

## **Chapter 7: Assessing the Estimates**

Mary H. Mulry

The evaluations of the A.C.E. Revision II estimates may be divided into two categories. One category contains the evaluations that focus on individual error components. The other group consists of comparisons of the relative error between the Census and the A.C.E. Revision II estimator.

This section provides a brief description of the evaluation studies. The component errors examined by separate studies are sampling error, error from imputation model selection, error due to using in-movers to estimate out-movers in PES-C, synthetic error, error in the identification of the census duplicates as determined by administrative records, error in the identification of computer duplicates as determined by a clerical review, error from inconsistent poststratification variables, and potential error arising from the automated coding of some cases, called the at-risk coding, in the Revision Sample. The comparisons of relative error between the Census and the A.C.E. Revision II estimator include a comparison with Demographic Analysis, the construction of confidence intervals that account for bias as well as random error, and loss function analyses. Also in this category is an examination of the consistency of the estimates of coverage error measured by the A.C.E. Revision II estimator and the Housing Unit Coverage Study (HUCS). Although an adjustment for correlation bias is included in the A.C.E. Revision II estimates, no evaluations address the error in the level of correlation bias or the model used to distribute it across poststrata. The reason is that examining alternative models only would consider differences in models. Those differences would reflect the variations in how the several models correct the original DSEs for correlation biases, but would not reflect the presence or absence of correlation bias in these corrected DSEs.

### **7.1 Sampling error**

Sampling error gives rise to random error, quantified by sampling variance. The sampling variance is present in any estimate based on a sample instead of the whole population. The variance estimation methodology is a simplified jackknife with the block clusters being the primary sampling unit. The effect of in-cluster subsampling is implicitly captured in the weighting.

The 2001 experience with the A.C.E. data showed that the simplified jackknife method produces satisfactory variance estimates. Since an adjustment for correlation bias was incorporated, the correction was re-calculated for each replicate. An alternative variance estimation procedure assumed that the form of correlation bias adjustment was a scalar times the double sampling estimator. The replication method also accounts for the A.C.E. block cluster sampling.

## **7.2 Synthetic Error Evaluation**

The evaluation of synthetic error focuses on error in small area estimation. Synthetic estimation bias arises when the census from different areas but in the same poststratum have different coverage error rates but have the same census coverage correction factors. To assess synthetic estimation bias for a given area one needs to develop an estimate based on data from the area alone, called a direct estimate, but such an estimate is possible for only large areas. In lieu of direct estimates, we attempt to estimate synthetic estimation bias in undercount estimates from analysis of “artificial populations” or “surrogate” variables, whose geographic distributions are known. These surrogate variables are constructed as best as possible to have patterns similar to coverage error. Sensitivity analyses assesses the impact of synthetic estimation bias for these variables.

The evaluation of synthetic error uses an artificial population analysis similar to ESCAP I Report B-14 and the ESCAP II Report 23 (Griffin and Malec 2001a, 2001b) but this time compares the A.C.E. Revision II estimates and the 2000 Census. The study uses loss functions for assessing the effect of synthetic error. The major products are:

- (1) Estimates of the bias in the difference between census loss and A.C.E. Revision II estimator loss.
- (2) Indicator of whether the decision to use the A.C.E. Revision II estimator would have changed due to synthetic error.

## **7.3 Error Due to Using Inmovers to Estimate Outmovers in PES-C**

The error due to using inmovers to estimate outmovers is unique to the PES-C model for dual system estimation used in the original A.C.E. and the A.C.E. Revision II. For the PES-C model, the members of the P-sample are the residents of the housing units on Census Day. There is some difficulty in identifying all the residents of all the housing units on Census Day because some move prior to the A.C.E. interview. The A.C.E. interview relies on the respondents to identify those who have moved out, the outmovers. Since the outmovers are identified by proxies, many of the outmovers are not recorded. Therefore, the estimate of outmovers is too low. To avoid a bias caused by an underestimate of the number of movers, PES-C uses the number of inmovers to estimate the number of outmovers. The inmovers are those who did not live in the sample blocks on Census Day, but moved in prior to the A.C.E. interview. Theoretically the number of inmovers in the whole country should equal the number of outmovers. However, the number of inmovers may not equal the number of outmovers in a poststratum because of circumstances such as economic conditions causing more people to move out of an area than to move into an area.

The first step of the methodology consists of raking the number of outmovers to total inmovers. The distribution of the raked outmovers may better describe the outmovers than the distribution of the inmovers. Then we compare the A.C.E. Revision II estimates formed by using the number of inmovers with the A.C.E. Revision II estimates calculated using the raked number.

## **7.4 Error from imputation model selection**

This project estimates the uncertainty due to choice of imputation model by drawing on the analysis of reasonable alternatives to the imputation model conducted in 2001 (Keathley et al 2001). The ideal approach would be to repeat the very time-consuming analysis of reasonable alternatives for the A.C.E. Revision II estimator, but our limited resources do not permit it. Instead, we develop an estimate of the additional variance due to the choice of imputation model by using the previous work for the original A.C.E.

We form an estimate of the variance component for census coverage correction factors that accounts for the missing data error component due to the imputation of enumeration status, residency status, match status, and the P-sample noninterview adjustment. The replicates used to estimate the missing data variance are used in the loss function analysis to represent the random error due to the choice of the models imputation for missing data.

## **7.5 Examining the quality of the computer duplicates with administrative records**

Administrative records provide an opportunity to examine the quality of the estimates of duplicate enumerations used in the A.C.E. Revision II estimates. This study uses the Census NUMIDENT File and the Statistical Administrative Records System (StARS) 2000 to assess the effectiveness of the automated methodology used in the Further Study of Person Duplication (FSPD) to identify duplicate enumerations. Secondary goals are to provide data that can be analyzed to determine the nature of the census duplication so that the information may be used in reducing census duplication in 2010 and to aid in the evaluation of the methodology for the construction of StARS 2000.

The data for the study comes from three sources: the Census NUMIDENT file, the StARS 2000 database, and the FSPD duplicate file. First there is an attempt to assign a PIK (Personal Identification Key) to each census enumeration using the Census NUMIDENT file and the StARS database. The PIK is related to a person's SSN, but is used instead of the SSN for privacy protection. The next steps for the investigation are as follows:

- Use PIKs to identify FSPD duplicates who are not the same people. There are FSPD duplicates where the two members of the pair have different PIKs.
- Confirm FSPD duplicates by using all the addresses found for their assigned PIK. The StARS keeps all the addresses recorded in the different administrative records system. Both addresses shown for a set of FSPD duplicates may appear for the same PIK.
- Find additional census duplicates. Multiple census records might match to the same PIK, identifying potential census duplicates.

The study produces the following:

- Number of FSPD duplicates for the E and P samples that are not the same people
- Number of FSPD duplicates for the E and P samples that can be confirmed
- Number of FSPD duplicates for the E and P samples where the results are inconclusive, meaning that they can be neither confirmed nor denied.

- Number of additional E and P sample cases that have computer duplicates

## **7.6 Clerical review of computer duplicates**

The study examines accuracy of the FSPD computer identification of duplication in the census by having clerks review the enumerations that the computer designates as duplicates. The clerks determine whether the sets of two enumerations appear to be the same persons. In addition, census enumerations identified as duplicates by administrative records, but not by the computer, also have a clerical review. The potential census duplicates identified by administrative records are a by-product of the evaluation of the computer duplicates using administrative records.

The review is restricted to duplicates between enumerations in the E sample in the A.C.E. blocks and census enumerations outside the search area. Links between P-sample nonmatches and enumerations outside the search area also are reviewed.

The clerical review produces the following:

- the number of E-sample enumerations with false duplicate links identified by the computer
- the number of E-sample enumerations with missed duplicates identified by administrative records that are correct
- the number of P-sample nonmatches with false duplicate links identified by the computer
- the number of P-sample nonmatches with missed duplicates identified by administrative records that are correct

With these results, we compute the accuracy rate for the computer identification of duplicates in the census and between the P-sample nonmatches and the census.

## **7.7 At-Risk Coding**

The study assesses the amount of error at risk due to not having each and every case in the Evaluation Followup (EFU) sample reviewed clerically (Adams and Kresja 2001). The data collected in the Evaluation Followup of the A.C.E. found errors in the coding of E-sample census enumeration status and P-sample residence status and match status that needed to be corrected for the A.C.E. Revision II estimator. Ideally this would mean re-coding the entire A.C.E. sample, but that was not possible because the Evaluation Followup collected data in only 2,259 out of the 11,303 A.C.E. sample clusters. Even clerically re-coding the 70,000 cases in the Evaluation Followup sample was not feasible because of time constraints. A new strategy was devised to provide the most high quality data in the time allowed by restricting the clerical review to the more difficult cases. This strategy reduced the clerical workload to about 25,000, which could be done, and ensured the largest sample possible for the A.C.E. Revision II estimates.

Since the Production Followup (PFU) and the Evaluation Followup (EFU) questionnaires had been keyed and were now available in electronic form, we combined computer coding using an algorithm based on the keyed data and a clerical coding of the categories of cases where the

computer did not appear to do a good job.

The method compares the code assigned based on the PFU questionnaire to the code assigned based on the EFU questionnaire and then determines the best code. We assess the effectiveness of the computer algorithm by the agreement between the two new codes and a comparison with the re-codes assigned in Fall 2001 to a subsample of the EFU E-sample, called the Production Followup/Evaluation Followup Review (PFU/EFU). The PFU/EFU Review is believed to have been the best A.C.E. coding operation.

For the P sample in the Evaluation Followup, a coding algorithm for the keyed data from the PFU and EFU questionnaires also was developed. Assessing the quality was not as easy for the nonmatches and unresolveds as for the matches. Although re-codes from the PFU/EFU Review were available for the matches in the P sample, none of the nonmatches or unresolveds were included.

The categories of cases not sent for clerical review had a high agreement rate between the codes for the PFU and EFU assigned by the computer algorithm. For the cases in these categories where the PFU and EFU disagreed, the selected code was the one from the form with more detailed information. Therefore, we have three types of cases in the estimation:

- 1) the PFU and EFU codes assigned by computer agree
- 2) the PFU and EFU codes assigned by computer disagree, but are in a category where there is high consistency between the PFU and EFU codes, and either the PFU form or the EFU form does not have answers to all the questions. The code for the form with complete data is selected.
- 3) clerically assigned codes

We call the first group the “at-risk cases.” The first group may be at a higher risk of error than the others because of the lack of clerical review, even though the codes assigned by the computer algorithm agree. However, the cases in the second group also may have error even though they are in a category with high consistency between the PFU and EFU, but we have no way of assessing the risk of error due to the lack of information on one of the forms.

To assess the potential for error, we assume the at-risk cases have the same error rate as found for the cases in their category in the PFU/EFU Review. We assess the potential impact by comparing the A.C.E. Revision II double sampling adjustment factors with the double sampling ratios under the assumption that incorporates the error rates. The double sampling adjustment factors are described in Chapter 6.

## **7.8 Inconsistency of poststratification variables**

Inconsistency in the E-sample and P-sample reporting of the characteristics used in defining the poststrata may create a bias in the dual system estimate (DSE). This bias affects the estimation of the P-sample match rate.

The analysis of the poststratification variables for the A.C.E. Revision II estimator follows a similar investigation as for the original A.C.E. The basic approach was to estimate the inconsistency in the poststratification variables using the matches and then assume that the rates also held for the nonmatches. The models used for the inconsistency of the original A.C.E. poststrata (Haberman and Spencer 2001) were fitted in two steps, first (i) models for inconsistency of basic variables, and then (ii) derivation of inconsistency probabilities for poststratification given the inconsistency probabilities of the basic variables. The inconsistency probabilities led to an estimate of the bias in the P-sample match rate that was used to estimate the bias in the DSE.

The approach we are taking for the ACE Revision II estimator is to calculate the proportions only for the collapsed poststrata used in the double sampling ratio. If the inconsistency in poststratification can only be based on the doubly-sampled cases, the models used in (i) must be simplified because they cannot be estimated with the reduced sample size under double sampling. We revise models in (i) and (ii) to reflect revisions to the P-sample poststratification and the subsample for the double sampling estimator.

To assess the bias due to inconsistency in the poststratification variables we calculate the A.C.E. Revision II estimates with a correction to the match rate for the inconsistency. Then we compare the estimates with and without the correction.

## **7.9 Consistency between the A.C.E. Revision II estimator and HUCS**

The study examines the validity of the A.C.E. Revision II estimates by assessing the consistency in the results from the A.C.E. Revision II estimates and the Housing Unit Coverage Study (HUCS) (Barrett, et al 2001). Since the A.C.E. Revision II estimates may be used in the Postcensal Estimates Program and that program makes use of the average household size in many calculations, considering the consistency between the A.C.E. Revision II estimator and the HUCS data is important.

Although the A.C.E. Revision II estimates census coverage for people and HUCS estimates census coverage for housing units, we expect to see similar patterns in the differential coverage for demographic and geographic groups. We also expect to see similar patterns in the measures of change in census coverage between 1990 and 2000 for demographic and geographic groups. However, if there is a substantial difference in the census coverage error caused by missing whole households and by missing people within households, the patterns of differential coverage of people and of housing units may not have similar patterns.

If there are demographic or geographic groups where the differential coverage from the A.C.E. Revision II estimator and HUCS is substantially different, the study attempts to describe whether the disagreement is a symptom of problems with the A.C.E. Revision II estimator or HUCS, or the result of legitimate differences in coverage.

## 7.10 Relative accuracy of the Census and A.C.E. Revision II estimator using Demographic Analysis

Demographic Analysis (DA) uses vital records, immigration statistics, and Medicare data to obtain an estimate of the population size. Since the methods are somewhat independent of the census, DA provides a method for assessing the relative quality of the Census and the A.C.E. Revision II. We assess the consistency of estimates of differential census coverage from the A.C.E. Revision II estimator and DA for demographic groups.

We compare estimates of differential census coverage by demographics, including race, sex, and age. There are some limited estimates by Hispanic ethnicity, but these are more problematic because data on Hispanic ethnicity has not been well included in vital records historically. The estimates of population size based on DA are not viewed with as high a confidence as the estimates of differential coverage. DA does a better job of measuring differences in coverage between groups than population size.

In addition, we compare the sex ratios from the A.C.E. Revision II estimates and DA. The sex ratio is the ratio of males to females and provides a measure of differential coverage of males and females, especially when calculated for race groups.

We also repeat these comparisons with 1990 PES and DA estimates to provide a context for viewing the comparisons with the 2000 data. We assess whether both methods measure the same change in differential net undercounts from 1990 to 2000.

## 7.11 Relative accuracy of the Census and the A.C.E. Revision II estimator using confidence intervals and loss function analysis

Two additional methods of assessing the relative accuracy of the Census and the A.C.E. Revision II estimates are using confidence intervals for the net undercount rate and a loss function analysis. We form the confidence intervals for net undercount rate using estimates of net bias and variance. Since most of the data available on the quality of the original A.C.E. is being incorporated in the A.C.E. Revision II estimates, the estimation of the net bias uses the data that were not included. In the loss function analysis, we use mean squared error weighted by the reciprocal of the census count to estimate loss for levels and shares for counties and places across the nation and within state.

Confidence intervals that incorporate the net bias as well as the variance for the net undercount rate  $\hat{U}$  provide a method for comparing the relative accuracy of the census and the A.C.E. Revision II estimates. We estimate the net bias in the census coverage correction factor for each poststratum. With the estimated bias and variance for each census coverage correction factor, we can estimate the bias  $\hat{B}(\hat{U})$  and variance  $V$  in the net undercount rate  $\hat{U}$  and form the 95% confidence interval for the net undercount rate by

$$(\hat{U} - \hat{B}(\hat{U}) - 2\hat{V}^{1/2}, \hat{U} - \hat{B}(\hat{U}) + 2\hat{V}^{1/2}).$$

Since  $\hat{U}=0$  corresponds to no adjustment of the census, one comparison of the relative accuracy of the census and the A.C.E. Revision II estimates is based on an assessment of whether the confidence intervals for the evaluation poststrata cover 0 and  $\hat{U}$ .

A loss function analysis for levels and shares compares the census and the A.C.E. Revision II estimator for counties and places across the nation and within state. The measure of accuracy used by the loss functions is the weighted mean squared error with the weights set to the reciprocal of the census count for levels and the reciprocal of census share for shares. The motivation for the selection of the groupings for the loss functions is the use of the postcensal estimates.

#### Levels

All Counties with population of 100,000 or less

All Counties with population greater than 100,000

All places with population at least 25,000 but less than 50,000

All places with population at least 50,000 but less than 100,000

All places with population greater than 100,000

#### Shares within state

All Counties

All places

#### Shares within U.S.

All places with population at least 25,000 but less than 50,000

All places with population at least 50,000 but less than 100,000

All places with population greater than 100,000

All states

#### **References**

Adams, Tamara, and Krejsa, Elizabeth A. (2001) "ESCAP II: Results of the Person Followup and the Evaluation Followup Forms Review." Executive Steering Committee For A.C.E. Policy II, Report No. 24. dated October 12, 2001. Census Bureau.

Barrett, Diane F., Beaghen, Michael, and Smith, Damon (2001) "ESCAP II: Census 2000 Housing Coverage Study", by Diane F. Barrett, Michael Beaghen and Damon Smith. Executive Steering Committee For A.C.E. Policy II, Report No. 17. dated October 17. Bureau of the Census.

Griffin, R. and Malec, D. (2001a) "Assessment of Synthetic Assumption." DSSD Census 200 Procedures and Operations Memorandum Series B-14\*. dated February 28, 2001. Census Bureau.

Griffin, R and Malec, D. (2001b) "Sensitivity Analysis for the Assessment of the Synthetic Assumption." ESCAP II Report 23 dated October 12, 2001. Census Bureau.



Haberman, Shelby J. and Spencer, Bruce D. “ Estimation of Inconsistent Poststratification in the 2000 A. C. E.” Paper prepared by Abt Associates Inc. and Spencer Statistics, Inc. under Task Number 46-YABC-7-00001, Contract Number 50-YABC-7-66020. Dated December 17, 2001.

Keathley, Don, Kearney, Anne, and Bell, William (2001) “ESCAP II: Analysis of Missing Data Alternatives for Accuracy and Coverage Evaluation.” Executive Steering Committee For A.C.E. Policy II, Report No. 12. dated October 11, 2001. Census Bureau.

Robinson, J. Gregory (2001) “ESCAP II: Demographic Analysis Results.” Executive Steering Committee For A.C.E. Policy II, Report No. 1. dated October 13, 2001. Census Bureau.