Investigation of Extreme Estimates of Census Coverage Error for Small Areas

Mary H. Mulry, Eric Schindler, Tom Mule, Nganha Nguyen[1], Bruce D. Spencer
U. S. Census Bureau, Washington, DC 20233
Department of Statistics and Institute for Policy Research, Northwestern University, Evanston, IL 60208

Abstract

Extreme estimates of census coverage error for a large number of small areas were a major problem for a revision of the Accuracy and Coverage Evaluation (A.C.E) Survey estimates of coverage of Census 2000, called A.C.E. Revision II. The existence of extreme estimates was a major reason the Census Bureau did not use the A.C.E. Revision II estimates in the intercensal estimates program. The estimation of coverage error used dual system estimation for poststrata, and synthetic estimation within the poststrata for small areas. The paper reports the results of an investigation into the cause and source of the extreme estimates. In particular, we examine the effect of having different poststratification variables in the enumeration (E) and population (P) samples used in the dual system estimation and the choice of the poststratification variables on the synthetic estimation for small areas.

Keywords: dual system estimation, undercount, synthetic estimation, Census 2000, Accuracy and Coverage Evaluation Survey

1. Introduction

Extreme estimates of census coverage error for a large number of small areas were a major problem for a revision of the Accuracy and Coverage Evaluation (A.C.E.) Survey estimates of coverage of Census 2000, called A.C.E. Revision II. The existence of extreme estimates was a major reason the Census Bureau did not use the A.C.E. Revision II estimates in the intercensal estimates program (U.S. Census Bureau 2003a). The estimation of coverage error used dual system estimation for poststrata, and synthetic estimation within the poststrata for small areas.

The A.C.E. Revision II estimated a net undercount rate of -0.5 percent, indicating an overcount of the population by the census. However, for more than 5 percent of places, the ratio of the adjusted number to the census enumeration for the area – what we will call the *area adjustment ratio* (AAR) – was -5.0 percent or lower. For about 0.5 percent of the places, the AAR was -10.0 percent or lower, which is considered unusually high overcounting by the census. In the other direction, the AAR for only about 0.5 percent of the places was 5.0 percent or higher. However, the focus in this study is on the areas with low AARs. Such low net undercount rate estimates of this size for a small area are considered an extreme, particularly relative to the estimated net undercount rate for the U.S. as a whole.

Intercensal estimates are important for a variety of reasons including fund allocations (Louis, Jabine, and Gerstein, 2003) and business decisions related to capital investment and public perceptions of vitality. The concern about the extreme estimates of census coverage error for some small areas arose since the intercensal estimates are derived as the sum of a census base number plus an estimate of net change since the census (Spencer and Lee 1980). When the census base changes by a large amount or proportion, the intercensal estimate will change by a similar amount. Large discrepancies between the census count and formal or informal independent estimates of population raise issues for local governments. The issues of concern are mostly related to overcounts in the census population estimates because adjusting the census base would decrease the size of the population. The concern is heightened when the declines are seen as being driven by data error and most especially by model error. In the case of A. C. E. Revision II small-area estimates, which are driven by synthetic estimation within poststrata, the issue is modeling error.

The research and preparations for 2010 Census coverage measurement is focusing on providing information useful for designing improvements in census-taking methodology. Therefore, the most important objective for the 2010 Census Coverage Measurement (CCM) program is to obtain separate estimates of erroneous census inclusions and census omissions. As a result, estimates of net error for small geographic areas are not important because the emphasis is not on adjusting census counts (Kostanich, Whitford,

[1]This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

and Bell 2004). Producing estimates of net error for the 2010 census continues to be an important objective. The plans include estimating the net undercount and the differential net undercount for demographic groups and by geography (Kostanich, Whitford, and Bell 2004).

This paper reports the results of an investigation into the cause and source of the extreme estimates for small areas by A.C.E. Revision II. In particular, we examine the effect of having different poststratification variables in the enumeration (E) and population (P) samples used in the dual system estimation and the choice of the poststratification variables on the synthetic estimation for small areas. The findings are relevant to the development of methodology for the estimation of the net undercount for demographic group and by geography and components of census coverage error.

2. Estimates of Census Coverage Error

The A.C.E. Revision II estimation methodology used dual system estimates (DSEs) that incorporated corrections for measurement errors obtained from two sources: the re-coded cases from the A.C.E. evaluation data (Adams and Krejsa 2002) and census duplicates (Mule 2002). Both of these corrections affected the estimated E-Sample correct enumeration rates and the P-Sample match rates. The DSEs for adult males were also inflated by correlation bias adjustment factors estimated using Demographic Analysis sex ratios for the adult age groups (18-29, 30-49, 50+) at the national level by Black versus non-Black race groups. Also for the first time, the A.C.E. Revision II poststratification reflected one set of factors related to erroneous inclusions estimated from the E-sample and different factors related to omissions estimated from the P-sample. Previous estimates of census coverage error used the same poststratification for both samples. (U.S. Census Bureau 2003b)

The specific form of the A.C.E. Revision II DSE is given in equation (2) and discussed below. For a detailed discussion of the estimator, see U.S. Census Bureau (2004).

$$DSE_{ij} = Cen_{ij} \times r_{DD,ij} \times \frac{r_{CE,i}}{r_{M,j}} \times \phi \qquad (2)$$

where:

I and j denote the E- and P- Sample poststrata used to estimate the correct enumeration and match rates, respectively.

Cen_{ij} is the census count of the household population for the cross-classification of poststrata I and j. Includes the cases removed during the census because they potentially were duplicates but subsequently were reinstated.

$r_{DD,ij}$ is the census data-defined rate for the cross-classification of poststrata I and j. The reinstated cases are included in the denominator but not in the numerator.

$r_{CE,i}$ is the estimated correct enumeration rate for E-Sample poststratum I.

$r_{M,j}$ is the estimated match rate for P-Sample poststratum j.

ϕ is the correlation bias adjustment factor (for adult males, distinct for a given age-race group).

The construction of the A.C.E. Revision II estimates used synthetic assumptions. One type of synthetic assumption involved correcting the individual poststratum estimates for errors estimated at more aggregate levels, such as the corrections for correlation bias, duplicates, and measurement coding error. All involve estimates at a very aggregated level with little or no information available about how the effects being estimated truly affect correct enumeration rates and census inclusion probabilities for individual poststrata.

Another synthetic assumption was used in the application of poststratum coverage correction factors to persons with the poststratum in specific geographic areas. The assumption was that the net coverage error rate is constant within the poststratum. Equation (2) shows how the A.C.E. Revision II estimates are constructed for the cross-classified ij poststratum. To produce estimates for specific areas or population subgroups, the Census Bureau first defined coverage correction factors (CCFs) by dividing the dual system estimates from equation (2) by the corresponding census counts, i.e.,

$$CCF_{ij} = DSE_{ij} \,/\, Cen_{ij} = r_{DD,ij} \times \frac{r_{CE,i}}{r_{M,j}} \times \varphi \qquad (3)$$

To produce the estimate for any area or population subgroup a, the CCFs from equation (3) are applied synthetically:

$$\sum_{ij} Cen_{a,ij} \times CCF_{ij} = \sum_{ij} Cen_{a,ij} \times r_{DD,ij} \times \frac{r_{CE,i}}{r_{M,j}} \times \varphi \qquad (4)$$

where the summation is over all the cross-classified ij poststrata and $Cen_{a,ij}$ is the census count in poststratum ij for area or subgroup a. A.C.E. Revision II estimates for states, counties and places were estimated using Equation (4) and may be found in Kostanich (2003).

3. Cause and Sources of Extreme Estimates

Technical Assessment of A.C.E. Revision II (U.S. Census Bureau 2003b, Figure 1) shows that coverage correction factors (CCFs) for many poststrata were below .75, and most of those were for poststrata including proxy responses for the census numeration. The issue for extreme small area estimates is whether the AAR is far from 1.0. Unless corroborating evidence of a large amount of census error exists for a particular area, an AAR far from 1.0 is considered extreme.

3.1 Coverage Correction Factors

The first analysis we performed sought to identify the E-sample poststrata that are the source of the extreme downward CCFs under A.C.E. Revision II for a large number of small areas. We first set CCFs = 1 for all E-sample poststrata that specifically included proxy enumerations, and found that the number of places with AARs < 0.90 dropped from 91 to16. When we set CCFs = 1 for E-sample poststrata that specifically included proxy enumerations or late non-mail returns, not a single place had an AAR < 0.90. However, there were still 82 places (out of 19,269) with AARs < 0.95. Setting CCFs = 1 for E-sample poststrata that specifically included proxy enumerations or early or late non-mail returns reduced the number of places with AARs <0.95 to 4.

3.2 Offsetting Errors

It is noteworthy that the E-sample poststrata that are the source of the extreme downward AARs were all based on census operational variables (proxy response or non-mail return status) and therefore do not have counterparts in the P-sample. (They lack counterparts because P-sample enumerations that were not matched to census enumerations could not have determinate values for those variables.) The purpose of E-sample poststrata that are based on variables related to the data collection operation is to better allocate the erroneous enumerations geographically. This purpose can be undercut, however, if there are appreciable amounts of *offsetting errors* at the block cluster level that are handled by different E-sample and P-sample poststrata.

For an example of an offsetting error, note that census enumerations with *insufficient information for matching and follow-up*– what are called "KE" cases – are treated as census misses even if the person represented by the enumeration is found in the P sample, because a valid match cannot occur. An enumeration is classified as having insufficient information for matching and followup if the enumeration lacked a valid, complete name or if it did not have at least two other characteristics recorded. An enumeration is not data-defined if it does not have two or more characteristics reported. Of the proxy enumerations that were classified as erroneous, 73.7 percent were KE cases (Feldpausch 2001, Table A-7.A). Such enumerations do not contribute toward the count of correct enumerations, because they are classified as erroneous enumerations in the E-sample. Since a P-sample person represented by such an enumeration is being treated as a census miss, the two classifications cancel each other out at the aggregate level. However, when the synthetic estimates are calculated the canceling of errors may not occur at the small area level when the E- and P-samples have different poststrata. A type of enumeration that is excluded from the E-sample is those that are not data-defined. The people with enumerations that are not data-defined are considered census misses. The enumerations that are not data-defined are not a concern for offsetting errors in the DSE since they are not in the E-sample.

Note that if a poststratification is to be effective at reducing the bias associated with heterogeneity in the capture rates and the correct enumeration rates, then two conditions must hold: (I) E-sample correct enumeration rates vary across poststrata and the P-sample match rates vary across poststrata; (ii) distributions of census enumerations across the poststrata differ across areas.

That is, when (I) and (ii) hold, consistency of the E-sample and P-sample poststrata will account for offsetting errors from KE cases if those cases have sufficiently accurate demographic characteristics in the E sample (either reported or imputed) so that the person represented by the enumeration is classified consistently in both the E sample and P sample. If the person is not classified consistently, then having consistent definitions of E- sample and P-sample poststrata will not suffice to account for offsetting error (unless (I) or (ii) fails to hold).

When operational variables are used to define E-sample poststrata, and the correct enumeration rates vary greatly across poststrata with different statuses of operational variables (e.g., proxy versus nonproxy), then the synthetic estimates will be adversely affected by offsetting errors that are related to the operational variables. The match rates from the P-sample will not be able to capture the corresponding variation in match rates because the census operational variable is not well-defined for all members of the P-sample. This does not necessarily mean that operational variables should be avoided in defining poststrata, because if the operational variables are not used then the EEs (for example) are spread out more among the poststrata that are used, and cancellation of offsetting errors may still not occur. It does mean that potential improvements in accuracy for small-area estimates from using operational variables in defining poststrata can be eroded when the operational variables are related to offsetting errors.

4. Analyses

Our analyses explored whether the operational variables improve accuracy for small areas.

4.1 Insufficient Information for Matching and Followup

One approach examines the nature of the offsetting errors. In a recent study(Livermore Auer 2005), the Census Bureau's elite matching team attempted to link KE enumerations with P-sample people using more relaxed rules than permitted for A.C.E. matching. Each link was assigned a level of confidence that it was the same person - high, medium, or low.

With some assumptions, we can assess the implications of the KE enumerations that link to P-sample people on the correct enumeration rate for the proxy poststrata.

Using data in Feldpausch (2001, Table A-7.A) for the A.C.E., we can form an unpoststratified estimate of the correct enumeration rate for proxy responses. Feldpausch (2001, Table A-7.A) shows that of the proxy enumerations, 27.4 percent had insufficient information for matching, 7.8 percent were erroneous, and 1.9 percent had an unresolved enumeration status. If we assume that half of the unresolved cases were erroneous, the correct enumeration rate for proxy enumeration for the dual system estimator is 63.8 percent.

Table 4 in Livermore Auer (2005) shows that 14.8 percent of the enumerations with insufficient information that were proxy responses matched P-sample nonmatches with high confidence. In addition, 28.8 percent matched with medium confidence, and 6.2 percent were matched with low confidence. So, about 40 to 50 percent of the proxy responses with insufficient information could be matched to nonmatches with some level of confidence. Of the remaining approximately 50 percent, 49 percent definitely had no matches, and 1 percent consisted of duplicates of other census enumerations or some other classification.

Using the estimate that 50 percent of the proxy responses with insufficient information could be classified as correct enumeration, half of the 27.4 percent of the proxy responses with insufficient information, or 13.7 percent, could be classified as correct. The correct enumeration rate would rise to 77.5 percent.

Although 77.5 percent represents a 21 percent increase in the correct enumeration rate, it is still much lower than any of the census inclusion rate, or match rates, for the A.C.E. The match rates would increase if the ("with confidence") matches to the proxy responses with insufficient information were included although the effect on any P-sample poststratum would be small since the cases are distributed throughout all the P-sample poststrata. However, in a very small area with a high proportion of proxies, the synthetic estimate could increase by up to 21 percent since the full effect of including these cases on the match rates would be dispersed over a larger area.

The enumerations with insufficient information collected by proxy appear to contribute to the extreme estimates for small areas, but are not the only cause.

4.2 Coarser and Finer Poststratification Variables

Another approach is to examine the conditions when using a variable in the E-sample poststratification but not the P-sample poststratification is advantageous. For example, consider the variable that reflected the timing of the census enumeration. Although the variable could not be defined in the same way for the P-sample, the timing of the P-sample interview also is correlated with the match rate. The categories for the E-sample poststrata for census enumeration timing were early mail return, late mail return, early non-mail return, and late non-mail return. The P-sample interviewing had three phases, telephone phase conducted April 24 to June 13, 2000, personal visit phase conducted June 19 to September 11, 2000, and nonresponse followup conducted July 27 to September 11, 2000 (Petroni and Childers 2004). Telephone interviews were attempted only for households that responded to the census by mail although personal visit interviews were attempted when the telephone was not successful.

Childers et al (2001, Table 5o) show that the unweighted nonmatch rates before followup were 2.1 percent for the interviews collected in the telephone phase and 16.5 percent for those collected in the personal visit and nonresponse followup phases combined with an overall rate of 11.6 percent. For those that went to followup, Childers et al (2001, Table 9k) show the unweighted nonmatch rates after followup for those that went to followup were 1.4 percent for interviews from the telephone phase and 12.5 for those collected in personal visit and nonresponse followup phase with an overall nonmatch rate of 8.6 percent. The early mail return correct enumeration rates tended to be 4 to 8 percent higher than the early non-mail return and late non-mail return correct enumeration rates within domain, tenure, and family type.

Using a match rate averaged over all the P-sample collection phases with the early mailback correct enumeration rate probably tended to lead the DSE to overestimate the population in these poststrata. The reason for the tendency to overestimate is that many people with enumerations on early mailback returns were also telephone responses in the P-sample so the overall match rate was probably lower than the match rate for early mailback returns alone. In the late non-mailback poststrata, using the average match rate probably caused the DSE to underestimate the population and thereby measure an overcount. For small areas where the distribution of the timing of enumerations is not the same as the national distribution, this could cause a bias in the synthetic estimates.

One can construct analysis of variance (ANOVA) models to illustrate, in theory, how offsetting errors could lead to a coarser poststratification based on consistent P- and E-sample poststrata being more accurate than a finer one based on separate P- and E-sample poststrata (Spencer 2005, Bell 2003). The relationship between the P- and E-sample poststratification variables parallels ways that a ratio estimator for the sample mean can be more or less accurate, depending on the strength of the correlation between the two variables in the ratio. The ANOVA models are not useful for decision making, but are only a formal scenario indicating general conditions under which the separate poststratification could be less accurate. The method applies to operational variables as well as other types of variables, such as relationship within the household.

4.3 Block Level Comparisons

Although estimates at the block level by themselves are not of interest, such estimates, especially for larger blocks, and aggregations of such estimates are useful in assessing various poststratified estimates for small areas. Along this line, a third approach in the analysis is to compare the synthetic estimates based on the alternative methods of poststratification to relatively accurate alternative estimates of small-area populations. The simplest method is to develop a "census-plus" type direct estimate of number of people in a block cluster(ignoring subsampling if used in a block-cluster, and ignoring people missed by both E- and P-samples),

$$\text{census-plus} = CE + P - M. \qquad (5)$$

A more complex method is to construct a direct DSE for a block cluster to attempt to account for people missed by both the census and the P-sample,

$$N = CE \times P/M.. \qquad (6)$$

Due to small sample sizes, this estimator N (Eq. 6) was unstable, and so we implemented caps to force its ratio to census-plus to be between 1 - k and 1 + k, for moderate k. Note that the cap does not apply to the undercount, but rather to the (smaller) number of people missed by both the census and the P sample. Consideration was also given to using the number of matches to E-sample people instead of more general matches, but the resulting national estimates were implausibly large. Formally, we define the "capped" direct DSE to be

$$\text{direct-DSE}(k) = \max\{(1 - k)\times\text{census-plus}, \min\{N, (1 + k)\times\text{census-plus}\}\}.$$

Choosing k = 0.05 would cap nearly 10 percent of the block clusters, and a value of k = 0.20 would cap about one to two percent of the clusters. The uncapped DSE can be represented as direct-DSE(∞).

Two synthetic estimates were computed, "938" which was similar to the A.C.E. Rev. II method and included proxy poststrata, and "648" which used consistent E-sample and P-sample poststrata, which were the P-sample poststrata in A.C.E. Revision II estimates. Direct estimates for block clusters were calculated from the A.C.E. data, based on the census-plus approach and on the direct-DSEs (i.e., direct-DSE(k) for k = ∞ , 0.20, 0.05). Empirical error estimates were computed based on the working assumption that the direct estimate was correct. If the A.C.E. data are correct for a block cluster, the error in the direct census-plus estimate is simply the size of the fourth cell (number missed by both census and ACE). The direct-DSE attempts to estimate the latter. There are also data issues concerning KE's, which the direct estimates excluded, and possibly other problems.

The empirical RMSE (root mean square error) and MAE (mean absolute error) for 2,163 block clusters with census counts of 100 or more (including non-data-defined persons but excluding late adds) were computed several ways. Table 1 shows the results when the estimates were weighted to account for sampling of block clusters and Table 2 shows unweighted results. (The weights were the average E-sample weight in the cluster.)

In addition, we computed the RMSE and MAE looking only at the block clusters where direct-DSE(∞) and direct-DSE(0.20) were the same, i.e., the block clusters where the direct-DSE and census-plus differed by less than 20 percent. The results are shown in Table 3, for the subset of 2,132 large clusters that satisfied this criterion (out of the original 2,163). Table 3 also contains the results for these blocks calculating the RMSE and MAE separately for the 105 block clusters where 10 percent or more of the census responses were given by a proxy and for the 2,029 block clusters where less than 10 percent of the census responses were given by a proxy.

In all three tables, the RMSE and the MAE appear to be slightly lower for the "648"estimates than for the "938" estimates, whose accuracy (as measured in the tables) tended to be comparable to that of the census. Although the "938" estimates empirically perform better for block clusters with larger amounts of proxy enumerations, they perform less well for the remaining clusters. The block clusters with a census proxy rate 10 percent or higher tend to have high overcount rates as measured by census-plus and direct-DSE(∞). The "938" estimates are not close to the standards of comparison but are closer than the "648" estimates because the low correct enumeration rates for proxy responses are applied to a high percentage of the census responses in the block clusters and not offset by low match rates. The block clusters with a census proxy rate lower than 10 percent tend not to have high overcount rates, and the "638" estimates are closer than the "938" to the standards of comparison. These results do not indicate any tangible advantage of the "938" methodology with its use of proxy poststrata.

Unweighted comparisons of accuracy tell much the same story, except that in comparisons of the estimates with the uncapped direct-DSE ("direct-DSE(∞)"), the census had the lowest RMSE by a very small margin. This pattern persists for aggregations of block clusters, as is shown in Tables 4-6, below. The unweighted comparison give more emphasis to block clusters selected with higher probabilities, which tended to be the larger block clusters. The apparent levels of error are largest for the unrestricted direct-DSE estimates ("direct-DSE(∞)"), and are next largest and quite similar for the census-plus and direct-DSE(0.20) standards, and are smallest by a fraction of a person for the direct-DSE(0.05) comparison. Which comparison gives the best read on the RMSE and MAE is not known. The levels of error of all three estimates appear larger than one would hope. Although our analysis could be refined, we do not anticipate a change in results. Possibly the direct estimates used as standards for comparison are highly inaccurate due to problems with the A.C.E. data, but if that is the case then both sets of synthetic estimates will be problematic as well. Of course, the census data could be problematic too.

Furthermore, we note that Spencer and Hill (2001) carried out an extremely careful construction of direct estimates based on the 1990 census and PES, restricting the analysis to blocks in the evaluation subsample. They found the distributions of empirical errors to have long tails, with (I) the mean greatly exceeding the median, (ii) mean relative absolute errors of about 15 percent, and (iii) root mean square relative errors of about 50 percent. Their results were generally similar to the current results.

Block clusters are not themselves of primary interest, but rather they are building blocks for areas such as states, counties, etc. In addition to the block clusters discussed above, three other aggregations of sample block clusters were considered:

- States (50 states and the District of Columbia)
- Counties (2,145 counties that included at least 1 sample block cluster)
- Large counties (152 counties with at least 1,000 enumerations in sample blocks, as well as the balance of the state for the 50 states)

Sums of squared errors (SSE) and sums of absolute errors (SAE) were computed for aggregates of block clusters, based on unweighted and weighted aggregations block-cluster sampling weights. The errors were computed based on deviations of the three estimates ("648", "938", and the census) from both sets of direct estimates (census-plus and direct-DSE(∞)). For every comparison except for unweighted aggregations against (the uncapped) direct-DSE(∞), the "648" estimates outperform both the "938" and the census. Aside from the unweighted comparisons against the uncapped direct-DSE(∞), the "648" uniformly outperforms the other two estimators census at the state, large county, and county levels. These analyses offer no support for concluding the separate P- and E-sample poststratification in the "938" estimator improved its accuracy compared to the "648" estimator. On the contrary, the comparisons indicate that the "648" estimator was more accurate.

5. Summary

The choice of poststratification variables for the E-sample or the P-sample has to consider the influence in the other sample. This applies when estimating components of coverage error and net coverage error. Even though small area estimation is not a priority for the 2010 Census Coverage Measurement program, the estimation methodology for net coverage error and components of coverage error has to include appropriate treatment of offsetting errors and consideration of the influence of variables in both the E- and P-samples to avoid introducing biases. The findings in our research contribute to the understanding of the implications of offsetting errors and of poststratification variables.

Logistic regression modeling, currently under consideration for 2010 Census Coverage Measurement (Kostanich.,Whitford, and Bell 2004), may have an advantage over poststratification when it comes to decisions on whether to use variables influential in both the correct enumeration rate and the match rate. Poststratification restricts the number of variables that may be used because high-order interactions are included causing the number of poststratification cells to increase dramatically with the number of variables. Logistic regression allows selective use of interaction terms. However, the same principles hold when selecting variables for logistic regression as when choosing poststratification variables.

The use of direct estimates at the block cluster level is a potentially powerful way to compare alternative poststratification methods or alternative variables for logistic regression, provided that the direct estimates at the block cluster level are themselves accurate. If the direct block-cluster estimates could be made accurate, they could be used in real time in 2010 to compare alternative estimation methods. It would be useful to have more understanding concerning (I) why the direct block-cluster estimates are not accurate, (ii) just how accurate they are, and (iii) whether and how they could be made more accurate.

References

Adams, T. and Krejsa, E.(2002a) "A.C.E. Revision II Measurement Subgroup Documentation". DSSD A.C.E. Revision II Memorandum Series #PP- 6. U.S. Census Bureau, Washington, DC.

Bell, W. R. (2003) "Technical Issues and Developments in A.C.E. Revision II Estimation". Paper presented at the 2003 Joint Statistical Meetings, San Francisco, CA.

Childers, D., Byrne, R., Adams, R., and Feldpausch, R. (2001) "Person Matching and Follow-up Results". DSSD Census 2000 Procedures and Operations Memorandum Series B-6*. Census Bureau, Washington. Feldpausch, R. (2001) ESCAP II E-

Sample Erroneous Enumerations. Executive Steering Committee for A.C.E. Policy II (ESCAP II) Report 5. October 14, 2001. Washington, D.C. U.S. Census Bureau.

Kostanich, D. L.,Whitford, C., and Bell, W. R. (2004) "Plans for Measuring Coverage of the 2010 Census". *2004 ASA Proceedings.* American Statistical Association. Alexandria, VA. CD-ROM.

Kostanich, D. (2003) "A.C.E. Revision II: Adjusted Data for States, Counties, and Places." DSSD A.C.E. Revision II Memorandum Series #PP-60, U.S. Census Bureau . Washington, DC.

Livermore Auer, P. (2005) "Results of Feasibility Study to Match Census Enumerations Coded in A.C.E. as Insufficient Information for Matching and Followup". DSSD 2010 Census Coverage Measurement Memorandum Series. #2010-B01. Dated April 18, 2005.

Louis, T. A., Jabine, T. B., and Gerstein, M. (2003) *Statistical Issues in Allocation Funds by Formula.* Panel on Formula Allocations. Washington, D.C.: National Academy Press.

Mule, T. (2002) "Further Study of Person Duplication in Census 2000". DSSD A.C.E. Revision II Memorandum Series #PP- 51. U.S. Census Bureau, Washington, DC.

Petroni, R. J. and Childers, D. R.(2004) "Coverage Measurement from the Perspective of March 2001 A.C.E." Census Testing, Experimentation, and Evaluation Program, Topic Report Series, No. 4, TR-4. U.S. Census Bureau, Washington, DC.

Spencer, B. D. (2005) *Activity 1: Final Report. Statistical Analysis Research on Small Area Estimation of Census Coverage.* Report to the Bureau of the Census. Abt Associates Inc. And Spencer Statistics, Inc. Under contract No. 50-YABC-2-66036. March 22, 2005.

Spencer, B. D. and Hill, J. M. (2001) *Activity 3: Final Report on Evaluation of Block Level Estimates*. Report to the Bureau of the Census. Abt Associates Inc. and Spencer Statistics, Inc. under contract 50-YABC-7-66020. December 23, 1998; rev. July 12, 2001.

Spencer, B. D. and Lee, C.-F. (1980) Postcensal population estimation methods of the U.S. Bureau of the Census. Pp. 131-187 in Panel on Small-Area Estimates of Population and Income, Committee on National Statistics, *Estimating Population and Income of Small Areas,* Washington, D.C.: National Academy Press.

U.S. Census Bureau (2004) "Accuracy and Coverage Evaluation of Census 2000: Design and Methodology". DSSD/03-DM. Issued September 2004. U.S. Census Bureau, Washington, DC.

U.S. Census Bureau (2003a) "Decision on Intercensal Population Estimates". March 12, 2003. U.S. Census Bureau. Washington, DC.

U.S. Bureau of the Census (2003b) *Technical Assessment of A.C.E. Revision II.* Prepared for the Committee on National Statistics. March 12, 2003. Washington, D.C.: U.S. Census Bureau.

Table 1. Weighted Estimates of RMSE and MAE for synthetic DSEs with "938" and "648" poststratifications and for census, for block clusters with census count 100 or more, with four alternative standards for comparison.

	standard for comparison is							
	census-plus		direct-DSE(∞)		direct-DSE(0.20)		direct-DSE(0.20)	
Estimator	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**
"938"	4676	2412	8102	2621	4818	2460	4686	2409
"648"	4536	2288	8016	2505	4676	2344	4536	2296
census	4637	2429	8045	2609	4735	2447	4581	2399

Table 2. Unweighted Estimates of RMSE and MAE for synthetic DSEs with "938" and "648"poststratifications and for census, for block clusters with census count 100 or more, with four alternative standards for comparison

	standard for comparison is							
	census-plus		direct-DSE(∞)		direct-DSE(0.20)		direct-DSE(0.05)	
Estimator	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**
"938"	19.52	10.37	24.75	11.68	19.70	10.45	18.99	10.13
"648"	19.50	10.25	24.56	11.49	19.53	10.26	18.92	9.97
census	19.62	10.47	24.54	11.56	19.58	10.34	18.99	10.05

Table 3. Estimated RMSE and MAE for synthetic DSEs with "938" and "648" poststratifications and for census, for block clusters with census count 100 or more, where the direct-DSE differed from census-plus by 20 percent or less, by proxy rate.

	standard for comparison							
	census-plus		direct-DSE(∞)		census-plus		direct-DSE(∞)	
	Weighted				Unweighted			
Estimator	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
"938"	4455	2328	4617	2374	17.50	9.62	17.22	9.51
>10% proxy	4931	2720	5030	2732	25.62	14.20	25.31	13.45
<10% proxy	4429	2308	4595	2355	16.98	9.39	16.70	9.31
"648"	4335	2208	4480	2259	17.46	9.50	17.08	9.34
>10% proxy	5626	3149	5389	2924	28.01	16.37	25.82	14.50
<10% proxy	4259	2160	4429	2225	16.75	9.15	16.51	9.07
census	4449	2352	4543	2364	17.65	9.72	17.26	9.44
>10% proxy	5586	3063	5387	2858	27.45	15.91	25.84	14.33
<10% proxy	4384	2316	4496	2339	17.00	9.40	16.70	9.19

Table 4. Weighted and unweighted estimates of SSE and SAE for synthetic DSEs with "938" and "648" poststratifications and for census, for block clusters aggregated to state level.

	Weighted				Unweighted			
	comparison based on census-plus		comparison based on direct-DSE(∞)		comparison based on census-plus		comparison based on direct-DSE(∞)	
Estimator	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)
"938"	194	2.2	270	2.2	4.5	8.6	3.3	9.8
"648"	137	1.8	169	1.9	3.6	7.8	2.9	8.9
census	365	3.5	677	2.8	7.9	11.1	2.6	8.2

Table 5. Weighted and unweighted estimates of SSE and SAE for synthetic DSEs with "938"and "648" poststratifications and for census, for block clusters aggregated to level of large county or balance of state.

	Weighted				Unweighted			
	comparison based on census-plus		comparison based on direct-DSE(∞)		comparison based on census-plus		comparison based on direct-DSE(∞)	
Estimator	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)
"938"	149	3.6	163	3.7	2.1	14	2.7	15.7
"648"	117	3.3	137	3.3	1.8	13.2	2.4	14.7
census	215	4.3	159	3.8	2.2	15.2	1.7	13.3

Table 6. Weighted and unweighted estimates of SSE and SAE for synthetic DSEs with "938" and "648" poststratifications and for census, for block clusters aggregated to counties.

	Weighted				Unweighted			
	comparison based on census-plus		comparison based on direct-DSE(∞)		comparison based on census-plus		comparison based on direct-DSE(∞)	
Estimator	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)	SSE (millions)	SAE (1000's)
"938"	129	7.7	124	7.8	1.8	27.1	1.8	28.3
"648"	107	7.2	106	7.3	1.6	26.2	1.7	27.2
census	145	8.3	122	8.3	2.0	27.8	1.7	27.0