UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

MASTER FILE

January 25, 2002

DSSD A.C.E. REVISION II MEMORANDUM SERIES PP-1

MEMORANDUM FOR Preston Jay Waite
Associate Director for Decennial Census

From: Howard Hogan
Chief, Decennial Statistical Studies Division

Subject Five Challenges in Preparing Improved Post Censal Population Estimates

We are now beginning the process of determining how to improve post-censal population estimates using the census, the Accuracy and Coverage Evaluation (A.C.E.), and the evaluation studies. This memorandum sets out five challenges which must be addressed to succeed in this task. Each challenge concerns a major facet or assumption in the dual system estimation (DSE) and adjustment process. I conclude by discussing a few related issues.

In planning a strategy, we must be mindful of the very small net national undercount. Demographic analysis implies, an "upper bound"on the net national undercount of only 0.37 percent. (Robinson, p. A-13). Thus a careful accounting of overcounts and undercounts by group is necessary.

Measurement of Erroneous Enumeration: We will need to estimate the level and distribution of census erroneous enumerations. Evidence clearly demonstrated that the A.C.E. process greatly underestimated the level of census erroneous enumeration. Fay (2001) produced lower bounds on the level of unmeasured erroneous enumerations of 2.9 million. (p.4) For the purpose of A.C.E. evaluation, this analysis is sufficient. However, it does not provide sufficient basis to produce alternative estimates of the population.

- As Fay clearly states: "The A.C.E. substantially underestimated erroneous enumerations. The actual effect is likely to be even greater than estimated here." (p.23) His estimates depend upon his estimate of computer matching efficiency, which he acknowledges is an upper bound (implying a lower bound on duplication) (p.3). To correct the census, we will need a point estimate rather than a lower bound. Using the lower bound, rather than a measure of central tendency, as a "point estimate" might result in adjusting up manyareas or groups that would and should be adjusted down if a proper accounting was made of census erroneous enumerations. The initial A.C.E. estimates failed because they

did not adequately account for all of these census errors. The revised estimates should not intentionally do the same.

- Fay's estimate depends, in part, on the Evaluation Follow-up field work. This sample cannot now be augmented. The follow-up sample for initially matched E-sample records was particularly small. The challenge will be how to deal with variance increases for any sub-national grouping.

- Much of the estimate of A.C.E. bias depends on inferences made at an aggregate national level. For example, the evaluation identifies a level of duplication that exceeds the level of erroneous enumerations currently measured by the A.C.E. For this analysis of erroneous enumeration at the national level, it did not matter which of the duplicate records in a pair was erroneous. However, since each member of the pair might be in different poststrata, for adjustment an assumption must be made about which is correct. We will need to address this issue in preparing sub-national estimates. The determination of which additional census enumerations were erroneous will also have implications for the coding of the P sample and for the treatment of missing-data.

Measurement of Census Omissions: We have a number of evaluations of the coverage rates estimated by the population sample (P sample). These include direct evaluations such as the matching error study, the evaluation follow-up and the Target Extended Search evaluations (TES-2, TES-3). In addition, the evaluations of the E sample has indicated that the P sample was allowed to match to erroneous census records. Converting a census record to erroneous necessarily changes the coding of any P-sample record linked to it. Some of these evaluations were based on relatively small samples. The TES samples were selected independently of the Evaluation Follow-up (EFU) and Matching Error Study (MES) samples. Fay (2000) indicates many additional census erroneous enumerations not identified by either the original A.C.E. process or by the EFU. It seems possible, indeed likely, that the P sample has matched to many of the implied erroneous enumerations. How these additional erroneous enumerations are allocated will imply changes to the P sample coding. We need to sort out the errors that have been discovered to be sure each error is accounted for once but only once. It may be that the net error in the P samples was small. However, this may not be true for sub-national groups and areas.

Missing Data: Each of the evaluation studies produced nonresponse and unresolved cases. Particularly problematic was the evaluation follow-up. Coded independently from the production results, the EFU produced a E sample unresolved rate of 9.4 percent (Adams and Krejsa 2001 Table 6). Coded jointly with the production interview, the EFU produced an unresolved rate of 4.8 percent, with an additional one percent coded as conflicting. The treatment of the unresolved cases will be critical.

For housing units identified by the EFU as containing additional erroneous enumerations, we will have the EFU interview results to make a determination about the status of the housing unit on Census Day that can guide the P-sample missing data proceedures. Again, it seems possible,

indeed likely, that the P sample has matched to many of the implied erroneous enumerations, that is, the additional census erroneous enumerations not identified by either the original A.C.E. process or by the EFU. These will probably include P-sample whole households currently matched to census enumerations now determined (or implied) to be erroneous. A logical inference is that the P-sample household was also included erroneously. Removing these records creates a P sample nonresponse housing unit arising through a specific set of circumstances.

Our past approach to housing unit nonresponse was to assume that the units were occupied and that the residents were as likely as their neighbors to be Census Day residents and matches. This treatment reflected our understanding of the circumstances that created the non-response. Many of the converted units will be, in fact, vacant on Census Day, often seasonal vacants. However, if these units were, in fact, occupied on Census Day by another household, then that other household was most likely missed by the census. Because of this, the assumption that the residents these occupied units are as likely to be correctly enumerated as their neighbors is highly questionable. We know that the census counted the wrong people in these housing units. We may need to develop nonresponse methods that impute these housing units from only housing units with erroneously enumerated whole housing units in the census. Failure to address this issue properly could increase correlation bias.

Correlation Bias: Correlation bias is always present in dual system estimates. We have allowed a partial correction for correlation bias in the 1990 and the 2000 total error models. However, in neither of these censuses has the "production" estimates allowed for correlation bias. This was justified, at least in part, with the argument that such an estimate was "conservative." Because of correlation bias, the DSE will underestimate the population. Thus if the true undercount of a group was, say, four percent, but we only adjusted up three percent, the correction was in the right direction and better than doing nothing. It was just not as big as it should have been. However, when we are dealing with possible overcounts, the situation differs. If the true overcount is two percent, but because of correlation bias we underestimate the population and adjust down three percent, the adjustment is too far, although, in this example still closer to the truth. Ignoring correlation bias is no longer conservative. Indeed, for a group with a small net undercount, we could be adjusting down when we should have been adjusting up. Because of this, the treatment of correlation bias needs to be re-considered in light of the current situation. Correction for correlation bias would necessarily be model based and would take us away from the "empirical dual system estimate". Ignoring correlation bias could cause us to underestimate the true differential undercount, and could result in adjusting down further minority groups with true undercounts.

Poststratification: Our current post strata were developed in reference to our experience in all previous censuses, but especially the censuses of 1980 and 1990. In those censuses, gross omissions was the dominant error, with gross erroneous inclusions being smaller. Our knowledge led us to conclude that the determinants of net undercount would follow largely socio-economic groupings. However, these strata may aggregate and thus wash out many of the factors that are driving differentials in Census 2000. A postratification that averaged undercounts and overcounts might cause unacceptable synthetic error. This problem may be

made worse if the A.C.E. Revision II relies on aggregations of poststrata, such as the evaluation poststrata.

The differential errors by group and area are most probably driven as much by the factors that caused overcounts as those that caused undercounts. Bill Bell has previously suggested poststratifing separately for gross omission (P sample) and gross erroneous enumerations (E sample). We rejected this approach because of the tight schedule we were previously under. I think that this approach should be carefully investigated, especially to reduce and control synthetic error. We should investigate alternative post strata to reflect the new situation as was done after 1990 and was quite successful.

Other Issues: Beyond these global challenges, several important issues need to be addressed. First, we will clearly need models to optimize our data. Are there any policy restriction on the use of models? Second, should we attempt to include in the processing the "Re-instated cases"? This might reduce variance, but would necessitate more missing data modeling. Should we attempt to remove duplicates from the census files before computing the estimates and before performing the synthetic estimation? This could potentially reduce synthetic bias. Finally, what criteria will be used to judge success? It will be hard to plan a successful project without knowing the "acceptance criteria."

I think that with time, work imagination and the proper commitment of resources, this can be a very rewarding project. However, the challenges are real and success is far from assured.

References

Fay, Robert, Evidence of Additional Erroneous Enumerations from the Person Duplication Study, Executive Steering Committee for A.C.E. Policy II, Report 9, Preliminary Version October 26, 2001.

Robinson, J.G., Demographic Analysis, Executive Steering Committee for A.C.E. Policy II, Report 4, October 13, 2001