



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

MASTER FILE

December 31, 2002

DSSD A.C.E. REVISION II MEMORANDUM SERIES # PP-32

PRED CENSUS AND SURVEY MEASUREMENT STAFF MEMORANDUM SERIES:
CSM-A.C.E. REVISION II-07R

MEMORANDUM FOR: Donna Kostanich
Chair, A.C.E. Revision II Planning Group
Decennial Studies Statistics Division

From: Mary Mulry *signed 12/31/02* *MM*
Chair, A.C.E. Revision II Quality Indicators Group
Statistical Research Division

Through: David Hubble *signed 12/31/02* *RJP for*
Assistant Division Chief, Evaluations
Planning, Research, and Evaluations Division (PRED)

Prepared by: Donald Keathley
Mathematical Statistician, PRED

Subject: Study Plan for Comparing Dual System Estimates for the 2000
A.C.E. Revision II Using Inmover Counts vs Raked Outmover
Counts as Estimates for the Number of Outmovers

This document contains the study plan for comparing dual system estimates for the 2000 A.C.E. Revision II using inmover counts vs raked outmover counts as estimates for the number of outmovers.

Contact Don Keathley at donald.h.keathley@census.gov or at 301-457-8084 if you have any questions or comments.

cc. DSSD A.C.E. Revision II Memorandum Series Distribution List
Ruth Ann Killion (PRED)
David Hubble

Error Due to Estimating Outmovers Using Inmovers in the PES-C

Don Keathley, Planning, Research, and Evaluation Division

I. BACKGROUND

The error due to estimating outmovers using inmovers in the PES-C is unique to the PES-C model 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 it.

In an attempt to determine whether this has an effect on the DSEs, we're looking to rake the number of outmovers to total inmovers. The distribution of the raked outmovers may better describe the outmovers than the distribution of inmovers. Then we'll compare the A.C.E. Revision II DSEs that were formed by using the number of inmovers (DSE_{Inmover}) versus the A.C.E. Revision II DSEs that were calculated using the raked number (DSE_{Raked}).

II. QUESTIONS TO BE ANSWERED

- A. Are there any significant differences between the DSE_{Inmover} and DSE_{Raked} values?
- B. If yes:
 - 1. where are the differences?
 - 2. how many differences are there?

III. METHODOLOGY

A. Raking

Rake the number of outmovers to inmovers by the following two sets of marginals:

- 8 age-sex \otimes 7 domain groups (56 groups)
- 2 tenure \otimes 4 MSA/TEA \otimes 3 return rate groups (24 groups)

See Attachment A for the group definitions.

B. Estimation

Obtain $DSE_{Inmover}$ and DSE_{Raked} values for the following variables:

1. full P-Sample post-strata
2. FIPS state code
3. region

See Keathley (2002) for details on the post-strata.

C. Variance Estimation

Calculate variances for the $DSE_{Inmover}$ and DSE_{Raked} values at the post-strata, state, and region levels.

D. Error Estimation ($DSE_{Inmover}$ vs DSE_{Raked} comparisons)

Compare the DSEs for each post-stratum, state and region. Use the usual equation for the variance of a difference, where:

$$\text{Var}(DSE_{\text{difference}}) = \text{Var}(DSE_{\text{inmover}}) + \text{Var}(DSE_{\text{raked}}) - [2 \times \rho \times \text{SE}(DSE_{\text{inmover}}) \times \text{SE}(DSE_{\text{raked}})]$$

where ρ is the correlation coefficient. Use 0.9 and 0.99 as values for ρ . Then use t-tests to test for significant differences between the DSEs.

IV. DATA REQUIREMENTS

A. Raking

1. final weights, incorporating missing data procedures
2. in-mover counts, by the two sets of marginals in III.A.
3. out-mover counts, by each of the combinations of marginals in III.A. ($56 \times 24 = 1,344$ combinations, or cells)
4. post-stratum codes
5. FIPS state code
6. region

B. Error Estimation

1. $DSE_{Inmover}$ values by post-stratum, state, and region
2. DSE_{Raked} values by post-stratum, state, and region
3. Variance estimates for all $DSE_{Inmover}$ and DSE_{Raked} values
4. Variance estimates for all differences ($DSE_{Inmover} - DSE_{Raked}$)
5. t-statistics for each difference

V. DIVISION RESPONSIBILITIES

A. Raking (out-movers to in-movers)

1. specifications (PRED)
2. software (program) development (DSSD)
3. software implementation (DSSD)
4. verification (PRED)

B. Produce $DSE_{Inmover}$ and DSE_{Raked} values (DSSD)

C. Variance Estimation

1. specification memo (DSSD)
2. software development (DSSD)
3. software implementation (DSSD)

D. Compare $DSE_{Inmover}$ vs DSE_{Raked}

1. specifications (PRED)
2. software development (DSSD)

- 3. software implementation (DSSD)
- E. Analysis (PRED)
- 1. Analyze the results from V.D.3..above
 - 2. Form conclusions

F. Report (PRED)

Write a report exhibiting the results (V.D.3.), analysis (V.E.1.), and conclusions (V.E.2.)

VI. MILESTONE SCHEDULE

ITEM #	WHO	ACTIVITY	WHEN
1	PRED	Develop specifications for raking outmovers to inmovers	12/16
2	PRED	Develop specifications for comparing DSE_{Raked} and DSE_{Raked} values	12/16
3	DSSD	Develop and write variance estimation methodologies and specifications	12/16
4	DSSD	Write raking program(s)	12/17
5	DSSD	Write variance estimation program(s)	12/18
6	DSSD	Run the raking program(s)	12/18
7	DSSD	Produce DSE_{Raked} and DSE_{Raked} values	12/26
8	DSSD	Run the variance estimation program(s)	12/26
9	DSSD	Create output file(s) from items 6, 7, and 8	12/26
10	PRED	Obtain output file(s) from item 9	12/26
11	PRED	Analyze DSE comparisons	12/31
12	PRED	Write a report, exhibiting the results and conclusions	12/31

VII. LIMITATIONS

We're assuming that DSE_{Raked} values more closely reflect true DSEs than $DSE_{Inmover}$ values.

VIII. RELATED STUDIES

None known.

IX. REFERENCES

Keathley, Don (2002), "*Procedure for Raking Outmovers to Inmovers in the A.C.E. Revision II for the Evaluation of using Inmovers to Estimate the Number of Census Outmovers,*" DSSD A.C.E. Revision II Memorandum Series # PP-13, dated December 31, 2002

Kostanich, Donna (2003) "*A.C.E. Revision II Design and Methodology,*", DSSD A.C.E. Revision II Memorandum Series # PP-30, dated January, 2003

Raking Marginals

- A. 8 age-sex \otimes 7 domain groups (56 cells), where we defined the individual groups on:
1. Age-sex groups
 - a. age 0-9, both sexes
 - b. age 10-17, both sexes
 - c. age 18-29, male
 - d. age 18-29, female
 - e. age 30-49, male
 - f. age 30-49, female
 - g. age 50+, male
 - h. age 50+, female
 2. Domain groups
 - a. American Indian or Alaska Native on a reservation
 - b. American Indian or Alaska Native not on a reservation
 - c. Hispanic
 - d. Non-Hispanic Black
 - e. Native Hawaiian or Pacific Islander
 - f. Non-Hispanic Asian
 - g. Non-Hispanic White or 'some other race'
- B. 2 tenure \otimes 4 MSA/TEA \otimes 3 return rate groups (24 cells), where we defined the individual groups on:
1. Tenure
 - a. Owner
 - b. Non-owner
 2. MSA/TEA
 - a. Large MSA MO/MB
 - b. Medium MSA MO/MB
 - c. Small MSA & Non-MSA MO/MB
 - d. All other TEAs

Raking Marginals

3. Return Rate
 - a. Domains 3, 4, and 7, from section B.1.b. above
 - (1) High - tracts with the top quarter of census form return rates
 - (2) Low - all other tracts
 - b. Domains 1, 2, 5, and 6 - no return rate stratification