



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

MASTER FILE

March 4, 2003

DSSD A.C.E. REVISION II MEMORANDUM SERIES # PP-58

MEMORANDUM FOR Donna Kostanich
Chair, A.C.E. Revision II Planning Group

From: Dawn Haines *deh*
Chair, A.C.E. Revision II Estimation Subgroup

Prepared by: Thomas Mule *JTM*
Decennial Statistical Studies Division

Subject: A.C.E. Revision II Results: Change in Estimated Net Undercount

I. Introduction

This document shows an accounting of the A.C.E. Revision II estimation methodology. We did this analysis to better understand how the A.C.E. Revision II methodology changed the dual system estimate (DSE). We start with the March 2001 A.C.E. estimate of net undercount and look at the cumulative effect of incorporating each change one at a time. The A.C.E. Revision II DSE can be thought of as incorporating the following enhancements to a traditional DSE:

- New post-stratification to reflect different factors related to erroneous inclusions and omissions.
- Measurement corrections to the correct enumeration rate from the Further Study of Person Duplication (FSPD).
- Measurement corrections to the correct enumeration rate from the A.C.E. Revision II Measurement Coding Operation.
- Measurement corrections to the match rate from the FSPD.
- Measurement corrections to the match rate from the A.C.E. Revision II Measurement Coding Operation.
- Adjustment for correlation bias.

Table 1 shows the impact nationally. Table 2 shows the impact for the seven Race/Hispanic Origin domains. See Fenstermaker and Haines (2002) for a description of the domains.

II. Limitations

It is important to note that the change in the net undercount estimate shown in these tables reflect a specific ordering of incorporating the A.C.E. Revision II changes. If the order were rearranged, the estimates of change in the net undercount estimates for each incorporation would be different. However, the A.C.E. Revision II net undercount estimates at the bottom of each table would still be the same.

Futhermore, the net undercount change estimates are not equivalent to estimates of additional census erroneous enumerations measured by the A.C.E. Revision II. The tables show change in the net undercount estimates. For example, the table shows that after accounting for the new post-stratification and the additional erroneous enumerations, the net undercount estimate went from a net undercount of 3.3 million to a net overcount of 1.9 million. This is a change in the net undercount of 5.2 million people. This is *not* the change in erroneous enumerations.

The change in erroneous enumerations does not have a one-to-one relationship with the change in the net undercount estimates. The change in erroneous enumerations was slightly less than this. Using the aggregate formula for correct enumerations in the census (see Bell 2002), the March 2001 estimate of correct enumerations was 253 million. The estimate of correct enumerations for A.C.E. Revision II was 248.3 million. This corresponds to 4.7 million additional erroneous enumerations detected by A.C.E. Revision II. The only way this would be a one-to-one relationship was if the match rate for all post-strata was equal to 1. In our DSE formula, we divide by the match rate which is usually less 1. This leads to the change in the net undercount being larger than the change in erroneous enumerations.

Likewise, changes in net undercount estimates are not equivalent to the estimates of additional census omissions.

III. Explanation of Rows in Tables

The goal of the tables is to show the effect of incorporating each change. We start with the March 2001 A.C.E. data, post-stratification and estimator. Each row shows the effect of making one of the specific changes in our A.C.E. Revision II estimates. The changes due to duplication and coding errors in the E sample and P sample have some overlap. The results and magnitudes presented in Tables 1 and 2 are based on the specified order. We account for duplication first and then account for errors detected by coding error but not detected by duplication. If we flipped the order and presented coding error first, it would show a larger impact for coding error and a smaller impact for duplicates not detected by coding error.

1. March 2001 A.C.E. Estimate

This is the estimate of the net undercount from the March 2001 A.C.E. (Davis 2001).

2. New Post-stratification

This is the change in the net undercount estimate when we use the new post-stratification developed for the A.C.E. Revision II. We are only using the March 2001 data and the March 2001 estimator. Thus, the only change from the March 2001 estimate is the post-stratification. Though the effect of the new post-stratification is small at the national level, it has considerably more impact on subnational estimates, particularly for small areas.

3. E-sample Person Duplication Corrections

This is the change in the net undercount estimate when we account only for duplicates in the E sample detected by the FSPD outside the search area of the A.C.E. and use the post-stratification of the A.C.E. Revision II. We continue to use the March 2001 data to estimate the match rates for the P-sample post-strata.

4. E-sample Coding Corrections

This shows the change in the net undercount estimate after we account for the coding errors not identified by A.C.E. and not detected and accounted for by the computer matching in our estimate of correct enumerations. We use the post-stratification of the A.C.E. Revision II for these estimates. We continue to use the March 2001 data to estimate the match rates for the P-sample post-strata.

5. P-sample Person Duplication Corrections

This is the change in the net undercount estimate when we account for the resident status of nonmover residents that were found by the FSPD computer matching to have a link to a census enumeration outside the A.C.E. search area. This estimate includes both of the E-sample revisions listed above. We use the post-stratification of the A.C.E. Revision II.

6. P-sample Coding Corrections

This is the change in the net undercount estimate when we account for residence and matching error in recoding not identified by the A.C.E. and not detected by the computer matching of the FSPD. This is our A.C.E. Revision II estimate without the correlation bias adjustment. We have implemented changes to the correct enumeration rate and match rate in the estimator and use the new post-stratification.

7. Correlation Bias

This shows the change in the net undercount estimate of adjusting the adult male estimates for correlation bias.

8. A.C.E. Revision II Estimate

This is the final estimate of the net undercount. Negative numbers are estimated net overcounts.

IV. Results

How did we go from a net undercount to a net overcount?

Table 1 starts with the March 2001 A.C.E. estimate of a national net undercount of just under 3.3 million persons. Each row shows the effect on the net undercount estimate of making one of the specific revisions. Using only the new post-stratification and not making any measurement error corrections would increase the estimated net undercount to 3.3 million, an increase of less than 39,000. When measurement error corrections are made to the correct enumeration rate, we see that if we first correct for those identified by the person duplication study, the estimated net undercount is reduced by 2.8 million. Next, adding in the correction identified by recoding reduces the estimated net undercount by another 2.4 million, resulting in an estimated net overcount of 1.9 million. Next we incorporate measurement error corrections into the match rate. First, adding in the corrections based on the person duplication study reduces the estimated net undercount by another 1.1 million. Adding in the corrections from the recoding causes the estimated net undercount to increase slightly by only 11,000. Making the final correction for correlation bias increases the estimated net undercount by 1.7 million, yielding the A.C.E. Revision II estimate of a 1.3 million net overcount.

What were the A.C.E. Revision II estimates of coverage for the Hispanic and Non-Hispanic Black domains?

Historically, the Hispanic and Black domains have had similar net coverage estimates, but the A.C.E. Revision II appears to depart from this coverage pattern. The A.C.E. Revision II estimated the net undercount for the Black domain is 1.84 percent but the percent net undercount estimate of 0.71 for the Hispanic domain is not statistically different from zero. The A.C.E. Revision II estimates of percent net undercount for the Black and Hispanic domains are not statistically different (Fenstermaker and Haines 2002).

A major reason for the departure is the adjustment for correlation bias of adult males for these two domains. From Table 2, we can see that the net undercount estimates before accounting for the adjustment for correlation bias for the Black and Hispanic domains are -177,000 and 146,000, respectively. However, the adjustment for correlation bias dramatically increases the level of the net undercount for Black to 628,000, but does not impact the Hispanics so dramatically. The net undercount for Hispanics only increases to 248,000.

What is the estimated coverage for the American Indian on Reservation domain?

Coverage measurement surveys, historically, have estimated a net undercount for American Indians on Reservations. The A.C.E. Revision II estimated a net undercount of -0.88 percent or -4,700 persons for the American Indian on Reservation domain. This is not significantly different than zero (Fenstermaker and Haines 2002).

Table 2 shows how the adjustments in the A.C.E. Revision II methodology produced this new estimate. The table shows reductions due to 1) duplicates not identified by A.C.E. in the E sample, 2) coding errors not identified by A.C.E. and not detected by computer matching and 3) nonmover nonresidents identified by computer matching.

The A.C.E. Revision II estimates of person duplication showed that 14,000 American Indians on Reservations were duplicated outside the search area: 11,000 were to another census unit, 1,000 were to a group quarters and 2,000 were to units deleted during the Housing Unit Duplication Operation. Of these 14,000 duplicates, 9,000 were to another enumeration on the same reservation. The computer matching operation also estimated that 15,000 American Indians on Reservations who were nonmover nonmatches in the P-sample linked to a census enumeration outside the search area. Of these 15,000 cases, 10,000 were to a census enumeration on an American Indian Reservation. These findings contributed to the results shown in Table 2.

References

Bell, W., "A.C.E. Revision II: Calculating aggregate data-defined, correct enumeration, and census inclusion rates (for groups that involve aggregation across post-strata)," DSSD A.C.E. Revision II Memorandum Series PP-40, U.S. Census Bureau, Washington, D.C., 2002.

Davis, P., "Accuracy and Coverage Evaluation: Dual System Estimation Results," DSSD Census 2000 Procedure and Operations Memorandum Series B-9*, U.S. Census Bureau, Washington, D.C., 2001.

Fenstermaker, D. And Haines, D., "A.C.E. Revision II: Summary of Estimated Net Coverage," DSSD A.C.E. Revision II Memorandum Series PP-54, U.S. Census Bureau, Washington, D.C., 2002.

Table 1: Change in Estimated Net Undercount of the Household Population

	Estimated Net Undercount	Change*	Cumulative
March 2001 A.C.E. Estimate	3,261,876		
New Post-Stratification		+38,618	3,300,493
E sample Person Duplication Corrections		-2,814,355	486,138
Coding Corrections		-2,427,198	-1,941,060
P sample Person Duplication Corrections		-1,103,805	-3,044,865
Coding Corrections		+11,032	-3,033,833
Correlation Bias		+1,702,176	-1,331,656
A.C.E. Revision II Estimate	-1,331,656	-4,593,532	

* Shows the effect of adding in one revision at a time. A different ordering of the revisions would result in slightly different intermediate effects, but yield the same overall net undercount estimate. Estimated changes in the net undercount is not the same as estimated additional census erroneous enumerations or additional census omissions.

Table 2: Change in Estimated Net Undercount of the Household Population by Race/Hispanic Origin Domain

		American Indian on Reservations			American Indian off Reservation			Hispanic		
		Estimated Net Undercount	Change*	Cumulative	Estimated Net Undercount	Change*	Cumulative	Estimated Net Undercount	Change*	Cumulative
March 2001 A.C.E. Estimate		26,895			52,991			1,013,988		
New Post-Stratification			-300	26,595		-2,496	50,495		-32,070	981,918
E sample	Person Duplication Corrections		-14,391	12,204		-21,388	29,107		-368,546	613,372
	Coding Corrections		-13,611	-1,407		-14,524	14,583		-335,229	278,143
P sample	Person Duplication Corrections		-6,647	-8,054		-10,377	4,206		-155,761	122,383
	Coding Corrections		+1,876	-6,178		+475	4,681		+23,852	146,234
Correlation Bias			+1,454	-4,724		+5,010	9,691		+101,801	248,035
A.C.E. Revision II Estimate		-4,724	-31,619		9,691	-43,300		248,035	-765,953	

* Shows the effect of adding in one revision at a time. A different ordering of the revisions would result in slightly different intermediate effects, but yield the same overall net undercount estimate. Estimated changes in the net undercount is not the same as estimated additional census erroneous enumerations or additional census omissions.

Table 2: Change in Estimated Net Undercount of the Household Population by Race/Hispanic Origin Domain (Continued)

		Non-Hispanic Black			Native Hawaiian and Pacific Islander			Non-Hispanic Asian		
		Estimated Net Undercount	Change*	Cumulative	Estimated Net Undercount	Change*	Cumulative	Estimated Net Undercount	Change*	Cumulative
March 2001 A.C.E. Estimate		740,809			28,490			96,405		
New Post-Stratification			+53,359	794,168		+1,110	29,600		+15,712	112,117
E sample	Person Duplication Corrections		-501,205	292,963		-7,238	22,362		-106,085	6,032
	Coding Corrections		-328,214	-35,251		-6,048	16,314		-91,400	-85,369
P sample	Person Duplication Corrections		-164,103	-199,354		-5,951	10,363		-25,944	-111,313
	Coding Corrections		+22,575	-176,779		+532	10,895		+1,101	-110,212
Correlation Bias			+804,777	627,997		+1,884	12,779		+35,793	-74,419
A.C.E. Revision II Estimate		627,997	-112,812		12,779	-15,711		-74,419	-170,824	

* Shows the effect of adding in one revision at a time. A different ordering of the revisions would result in slightly different intermediate effects, but yield the same overall net undercount estimate. Estimated changes in the net undercount is not the same as estimated additional census erroneous enumerations or additional census omissions.

Table 2: Change in Estimated Net Undercount of the Household Population by Race/Hispanic Origin Domain (Continued)

		Non-Hispanic White and Other	
		Estimated Net Undercount	Change*
			Cumulative
March 2001 A.C.E. Estimate		1,302,297	
New Post-Stratification			+3,303
			1,305,600
E sample	Person Duplication Corrections		-1,795,501
			-489,901
	Coding Corrections		-1,638,172
			-2,128,073
P sample	Person Duplication Corrections		-735,022
			-2,863,095
	Coding Corrections		-39,379
			-2,902,474
Correlation Bias			+751,458
			-2,151,016
A.C.E. Revision II Estimate		-2,151,016	-3,453,313

* Shows the effect of adding in one revision at a time. A different ordering of the revisions would result in slightly different intermediate effects, but yield the same overall net undercount estimate. Estimated changes in the net undercount is not the same as estimated additional census erroneous enumerations or additional census omissions.