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

9/20/11

2011 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES #ACS11-RER-06

DSSD 2011 AMERICAN COMMUNITY SURVEY RESEARCH MEMORANDUM SERIES ACS11-R-16

MEMORANDUM FOR	ACS Research and Evaluation Steering Committee
From:	David C. Whitford ***signed*** Chief, Decennial Statistical Studies Division
Prepared by:	Mark E. Asiala American Community Survey Estimation Branch Decennial Statistical Studies Division
Subject:	Evaluating Use of Family Equalization for American Community Survey Weighting Methodology

Attached is the final American Community Survey Research and Evaluation report for "Evaluating Use of Family Equalization for American Community Survey Weighting Methodology". Prior to 2006, the American Community Survey (ACS) had produced inconsistent estimates of households and householders and inconsistent estimates of husbands and wives in married couple households even though logically these estimates should be equal. With the Family Equalization Project, research was undertaken to remedy these differences using a three-dimensional raking methodology where, for two of the dimensions, the marginal control totals were derived from the survey itself rather than an independent source. The results from that research led to changes in the weighting methodology for the 2006 ACS. This paper evaluated the effects of that change in both the 2006 and 2007 ACS single-year estimates.

If you have any questions about this report, please contact Mark Asiala (301-763-3605).

Attachment

cc:
ACS Research and Evaluation Team
Stephanie Baumgardner (ACSO)
Phil Gbur (DSSD)
Steven Hefter
Anthony Tersine
Martin O'Connell (SEHSD)

September 20, 2011

Evaluating Use of Family Equalization for American Community Survey Weighting Methodology

FINAL REPORT

Mark E. Asiala
Decennial Statistical Studies Division

United States Census Bureau

Evaluating Use of Family Equalization for American Community Survey Weighting Methodology*

Mark E. Asiala
US Census Bureau, Room #4H057, 4600 Silver Hill Rd, Washington, DC 20233

Abstract
Prior to 2006, the American Community Survey (ACS) had produced inconsistent estimates of households and householders and inconsistent estimates of husbands and wives in married couple households even though logically these estimates should be equal. With the Family Equalization Project, research was undertaken to remedy these differences using a three-dimensional raking methodology where, for two of the dimensions, the marginal control totals are derived from the survey itself rather than an independent source. The results from that research led to changes in the weighting methodology for the 2006 ACS. This paper evaluates the effects of that change in both the 2006 and 2007 ACS single-year estimates. The focus of this paper is to assess the effects at the weighting area level where the raking was performed. We also looked at the effects on estimates for cities and towns.

Key Words: raking, iterative proportional fitting, weighting

1. Introduction

The American Community Survey (ACS) is part of the Census Bureau's plans for a re-engineered 2010 Census. The ACS will collect long-form (sample) data on an annual basis in order to produce single and multi-year estimates. Specifically, the 5-year estimates are comparable to the long-form estimates traditionally produced after each decennial census.

The ACS collects information on a wide variety of topics including housing, household, family, and person characteristics. The existing weighting methodology produces two weights to tally characteristics for all domains: a housing unit weight which is used for housing, household, and family characteristics and a person weight which is used for person characteristics including householder characteristics. Prior to the 2006 ACS, the weighting methodology produced these weights semi-independently with no method to try to ensure consistency between the weights. Because of this, some significant data inconsistencies were present in the ACS data prior to 2006. The largest was the inconsistency between the estimate of householders and the estimate of households. Logically, this should be a one-to-one relationship. However, in 2005 the ACS estimate for householders was 114.8 million but the estimate of households was only 111 million. Similarly, there should be a one-to-one relationship between the number of spouses in households and the number of married-couple households. Also from the 2005 ACS, there were 57.1 million spouses but only 55.2 million married-couple households. These inconsistencies did not exist in the unweighted data so clearly the weighting methodology was responsible for the observed differences in the estimates. In each of the examples given, the first estimate (householders, spouses) was tabulated from the person weights and the second estimate (households)

*This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau.

was tabulated from the housing unit weights. These differences were of great concern to data users both external and internal to the Census Bureau.

In 2007, the weighting methodology for the ACS was changed to remedy this inconsistency in the 2006 ACS weighting (see Asiala, 2007 for full details). Specifically, the design goals of the new methodology were to:

1. Reduce the inconsistency between the estimates of householders, households, and occupied housing units.
2. Reduce the inconsistency between the estimates of spouses and married-couple households.
3. Reduce the inconsistency between the estimates of unmarried partners and unmarried partner households.

The primary change to the weighting methodology was to replace the simple ratio estimation step with a three-dimensional raking process designed to address these inconsistencies.

At the national level, all of these design goals were met. In the 2006 ACS, the estimated number of households, householders, and occupied housing units is now the same at all levels of geography. The estimate of the number of spouses was 55.48 million and the estimate of married-couple households was 55.52 million. The estimate of unmarried partners was 6.05 million compared to the estimate of unmarried partner households of 6.02 million. These sets of estimates remain statistically significant at the 90% confidence level due to the large correlation between them as the result of the revised weighting methodology. However, their difference is much smaller than the corresponding differences that existed in the 2005 data. One goal of this evaluation is to make similar comparisons for lower levels of geography as well.

One concern expressed by those who were briefed about the methodology change was that our convergence criteria was applied only at the national level even though the raking was performed at the weighting area level. In 2008, the methodology was tweaked to make the convergence criteria independent at the weighting area level for the 2007 ACS and the 2005–2007 ACS. A second purpose of this evaluation is to determine whether that change resulted in better consistency between estimates at the county-based weighting area level. We give our specific research questions below.

1.1 Research Questions

- How well is the equalization step meeting its design goals at the weighting area level?
- How well is the equalization step performing at the sub-weighting area levels?
- What was the benefit of moving, in 2007, to a weighting area level criterion versus a national level criterion?

2. Weighting Methodology

The weighting methodology for the ACS is a ratio-estimation method involving several steps. We present here a high-level overview with some additional detail given concerning the person weighting and the final steps of the housing unit weighting. For a more detailed description of the existing weighting methodology, see US Census Bureau (2009).

The weighting methodology produces two sets of weights: a housing unit weight and a person weight. The basic steps of the housing unit weighting are as follows:

1. Calculate the base weights defined as the inverse of the probability of selection.
2. Apply the Variation in Monthly Sample adjustment to smooth month to month response variation.
3. Apply a set of non-interview adjustments.
4. Apply a housing-unit post-stratification adjustment to control the estimate of total housing units.
5. Perform the person weighting [see below].
6. Apply the householder factor to make the housing unit weight for occupied units equal to the person weight of the householder.

2.1 Person Weighting

Prior to 2006, the person weighting was performed in one ratio estimation step. We used the intermediate housing unit weights from step (4) as the initial person weight. We then placed the sample persons into poststratification cells to apply demographic controls by race, ethnicity, age and sex. The methodology change for the family equalization was to introduce a raking-ratio estimation procedure by which we can reduce the data inconsistencies described previously. The Census 2000 sample weighting methodology used a similar approach but it also had the benefit of 100% Census counts for detailed demographics by householder/non-householder, spouses, and other characteristics. While the ACS population controls do not include this additional level of detail, we were able to design an expanded raking matrix that used ACS estimates in the marginals in order to carry out the raking methodology.

2.2 Raking Matrix

The raking matrix is defined with three dimensions. The first dimension carries out the spousal/unmarried-partner equalization. The second dimension carries out the householder/household equalization. The third, and final, dimension applies the demographic controls in the same fashion as the person post-stratification adjustment did in the original methodology. The geographic scope of the matrix is our weighting area which is defined to be a county or collection of counties. The order of the dimensions reflects the relative importance of each goal. Consistency with the controls has the highest priority and thus is the final step in the raking to ensure exact agreement with the controls as we have had in the past. The second to last step is the householder/household equalization step which was identified as a higher priority over the spousal equalization.

2.2.1 Spousal/Unmarried Partner Equalization

The spouse equalization dimension is defined as having 3 cells:

- Householder with spouse/unmarried-partner present
- Spouse/unmarried-partner of the householder
- Balance of sample persons

The first two cells are constrained to be equal to the estimate of married-couple or unmarried-partner households which is tabulated using the housing unit weight from step (4). The third cell is then constrained to be equal to the total population from the independent control minus the sum of the control totals for the first two cells. In this manner, total population is always controlled to the independent population control.

2.2.2 Household/Householder Equalization

The householder equalization dimension is defined as having 2 cells:

- Householders
- Non-householders (balance of sample persons)

The first cell is constrained to be equal to the estimate of occupied housing units and hence the estimate of households. The second cell is then constrained to be equal to the total population control minus the marginal control for the first cell.

2.2.3 Demographic Dimension

The demographic dimension is defined as having up to 156 cells per weighting area obtained from multiplying the number of cells for race/ethnicity (6) by age (13) by sex (2). In practice, collapsing of small cells typically makes the actual number of cells fewer than 156.

2.3 Convergence Criteria

Each step, or dimension, in the raking is applied in series to make up one iteration, always ending with the demographic dimension. For the 2006 ACS weighting, the raking was performed a total of 19 iterations. It was at that number of iterations where the national estimate for each marginal ceased converging and began to diverge from the control. For the 2007 ACS weighting and the 2005–2007 weighting, a weighting area was allowed to continue until one of the following criteria were met for all relative differences between each marginal and its control satisfied:

1. Relative difference was less than 0.0001 if the number of iterations was fewer than 20.
2. Relative difference was less than 0.001 if the number of iterations was fewer than 40.
3. A maximum of 40 iterations was reached.

Thus, in some instances a weighting area could stop iterating earlier than before if it satisfied the first criterion or it could continue for more iterations if it stopped after satisfying criteria #2 or #3.

3. Research Methodology

The evaluation will make use of the 2005 ACS as a baseline for the ACS estimates prior to the family equalization methodology and use the 2006 1-year, 2007 1-year, and 2005–2007 3-year data to investigate the impact of the change in methodology. In addition, the 2006 1-year methodology used the fixed 19 iterations whereas the methodology used for the 2007 1-year and 2005–2007 3-year contained the three part convergence criteria described in the previous section.

Using the four sets of data, we constructed estimates of spouses and unmarried partners from the person data, complete with direct variances and also estimates of married-couple households and unmarried-partner households from the housing unit data. We then compared the corresponding estimates to tally the number of times the two corresponding estimates were statistically significant different at the 90% level by geography. The geography considered was nation, state, county, place, and minor civil division (MCD) and census county division[1] regardless of whether it met the publication threshold.

We also calculated the relative percent difference between each set of corresponding estimates and computed some basic distribution statistics for those. These data are then used to address the first two research questions.

To better capture the impact of moving from a national to weighting-area level convergence criterion, we reweighted the 2006 ACS 1-year data using the weighting-area level convergence criterion. This allowed us to isolate the impact of that change apart from normal year-to-year fluctuation. Since the convergence criterion is based on the maximum relative difference of the marginal to the control in the raking matrix, we summarized the distribution of that maximum across weighting areas by the iteration that the weighting area converged using the new convergence criterion.

4. Results and Discussion

The tables that follow are summary statistics for all geographies that contain non-zero estimates regardless whether they were published or not. The number of geographies actually published for the 1-year and 3-year estimates is actually much smaller.

4.1 Spouses and Married Couple Households

We computed the relative difference between the estimate of spouses and the estimate of married couple households for all geographies along with their assoicated variances. Table 1 and Table 2 show the median and mean of those relative differences. Both tables show a dramatic decrease in the relative differences between 2005 and 2006 that continues in both 2007 and 2005–2007. Those decreases occur at all levels of geography suggesting that the weighting-area based raking does filter down to the lower levels of geography (places and MCDs) as well.

We also note that the mean relative difference does show some improvement, at least nominally, in the 2005–2007 3-year data as compared to the 1-year data. This may be due to, in part, the fact that the 3-year weighting areas are smaller in size and thus more places and MCD's have the opportunity to benefit.

[1]For brevity, I will use the term MCDs to refer to both minor civil divisions and census county divisions for the balance of the paper.

Table 1: Median relative difference between the estimate of spouses and the estimate of married-couple households.

		Median Relative Difference			
Geography	N	2005	2006	2007	2005–2007
US	1	0.03365	-0.00075	-0.00077	-0.00082
State	52	0.03210	-0.00022	-0.00064	-0.00055
Place	25018	0.02500	0.00000	0.00000	0.00000
County	3219	0.02224	-0.00058	-0.00050	-0.00029
MCD	35274	0.01667	0.00000	0.00000	0.00000

Table 2: Mean relative difference between the estimate of spouses and the estimate of married-couple households.

		Mean Relative Difference			
Geography	N	2005	2006	2007	2005–2007
US	1	0.03365	-0.00075	-0.00077	-0.00082
State	52	0.03202	-0.00118	-0.00070	-0.00129
Place	25018	0.03006	0.00768	0.01175	0.00454
County	3219	0.02574	-0.00155	-0.00140	-0.00359
MCD	35274	0.02127	0.00480	0.00766	0.00056

We see in Table 3 that the number of geographies where the difference between the estimates of spouses and married-couple families is significant falls off dramatically between 2005 and 2006 (and also 2007). This is true not only for counties, states and nationally but for places and MCDs as well. We also note that across all geographies there is an increase in the number of each geography where the difference is statistically significant for the 3-year estimates. However, given the data from the previous two tables we can conclude that this increase is primarily due to the increase sample size used to form the 3-year estimates.

Table 3: Tally of number of each geography where the estimate of spouses is statistically different from the estimate of married-couple households.

		Tally of Signficant Differences			
Geography	N	2005	2006	2007	2005–2007
US	1	1	1	1	1
State	52	48	5	5	14
Place	25018	2981	250	240	456
County	3219	742	94	81	108
MCD	35274	3053	337	372	605

4.2 Unmarried Partners and Unmarried Partner Households

The relative differences between the estimates of unmarried partners and unmarried partner households follows the same pattern as was seen for the spouses and married couple households. For that reason, those data are included here.

Table 4 is similar to Table 3 and shows the number of geographies where the estimate of unmarried partners is statistically significantly different from the estimate of unmarried-partner households. The pattern is very similar to the results for spouses but, because of the much smaller population, there are fewer geographies where the differences are significant. We also see the higher tallies for the 3-year estimates.

Table 4: Tally of number of each geography where the estimate of unmarried partners is statistically different from the estimate of unmarried partner households.

		Tally of Signficant Differences			
Geography	N	2005	2006	2007	2005–2007
US	1	1	1	1	1
State	52	26	3	6	15
Place	25018	136	57	50	150
County	3219	137	56	64	133
MCD	35274	148	56	61	172

4.3 Convergence Criteria

Table 5 and Table 6 shows a comparison of the distribution of the maximum percent difference for the weighting area between the marginals and the controls for each cell in the 1st and 2nd dimension of the raking matrix. It is this maximum that determines whether a particular weighting will continue in the raking process or stop at its current iteration. Data for both the production 2006 ACS weighting and the revised raking convergence criterion methodology are shown. Note that both tables show the same data except for different ranges of iterations. There are two general observations to be made. The first is that those weighting areas which now stop in fewer than 20 iterations, generally achieved greater convergence if allowed to go to 19 iterations as was done in the 2006 production. On the flip side, most weighting areas that were allowed to proceed beyond 19 iterations achieved much better convergence than in production. Only a small number (74) failed to converge after 40 iterations. In fact, during testing we noted that those weighting areas that did not meet the weaker convergence criterion before 40 iterations were converging extremely slowly after approximately 30 iterations and showed little improvement in the final 10 iterations.

5. Conclusions

Overall, the family equalization is performing well at the design area intended, counties and geographies formed from counties. It also shows benefits below the design area as well. We do see that especially in the 3-year ACS estimates, some counties are producing less consistent estimates than the overall trend. Further evaluation is needed on these outlier counties.

We also note that while it appears that the more finely tuned convergence criterion is helping those counties that require 20 or more iterations to converge to the weaker criterion of 0.001, the counties which are able to converge to the stronger criterion of 0.0001 could benefit from simply allowing their respective weighting areas to continue to rake until 20 iterations to achieve a greater level of consistency. We would need to evaluate what potential negative impacts that may cause.

6. Future Research

As a followup to this base evaluation, we can investigate further into those counties that fail to converge in the 1-year and 3-year weighting. While we are aware of the reasons for some of these counties having difficultly in the raking process, a more thorough summary of these counties needs to be done.

We also could investigate further into whether there is any significant negative consequences to forcing all weighting areas to go through at least 20 iterations before stopping. By doing so, we could improve our overall level of consistency.

REFERENCES

Asiala, M. (2007), "Weighting and Estimation Research Methodology and Results From the American Community Survey Family Equalization Project," *Proceedings of the Federal Committee on Statistical Methodology Research Conference* available online at http://www.fcsm.gov/07papers/Asiala.IV-C.pdf

US Census Bureau (2009), *Design and Methodology: American Community Survey ACS-DM1*, US Government Printing Office, Washington DC.

Table 5: Distribution of max test for iterations 6–20.

Iter	N	Method	Min	1st Q	Median	3rd Q	Max
6	2	Production	0.000000	0.000000	0.000001	0.000003	0.000003
		Revised	0.000057	0.000057	0.000072	0.000087	0.000087
7	2	Production	0.000002	0.000002	0.000004	0.000006	0.000006
		Revised	0.000043	0.000043	0.000062	0.000081	0.000081
8	3	Production	0.000001	0.000001	0.000006	0.000014	0.000014
		Revised	0.000037	0.000037	0.000081	0.000081	0.000081
9	4	Production	0.000002	0.000002	0.000002	0.000007	0.000011
		Revised	0.000079	0.000081	0.000085	0.000093	0.000099
10	5	Production	0.000005	0.000009	0.000009	0.000011	0.000025
		Revised	0.000079	0.000088	0.000090	0.000092	0.000094
11	4	Production	0.000010	0.000010	0.000015	0.000024	0.000029
		Revised	0.000087	0.000089	0.000090	0.000092	0.000093
12	4	Production	0.000005	0.000009	0.000016	0.000020	0.000021
		Revised	0.000078	0.000086	0.000095	0.000099	0.000100
13	7	Production	0.000004	0.000009	0.000014	0.000022	0.000025
		Revised	0.000063	0.000072	0.000077	0.000095	0.000097
14	18	Production	0.000006	0.000015	0.000021	0.000028	0.000042
		Revised	0.000075	0.000086	0.000090	0.000095	0.000099
15	25	Production	0.000007	0.000012	0.000021	0.000033	0.000040
		Revised	0.000063	0.000075	0.000081	0.000092	0.000100
16	31	Production	0.000012	0.000025	0.000036	0.000043	0.000048
		Revised	0.000059	0.000081	0.000087	0.000096	0.000098
17	41	Production	0.000028	0.000039	0.000050	0.000055	0.000067
		Revised	0.000074	0.000084	0.000088	0.000093	0.000098
18	49	Production	0.000043	0.000054	0.000063	0.000067	0.000080
		Revised	0.000066	0.000079	0.000086	0.000091	0.000099
19	56	Production	0.000067	0.000079	0.000087	0.000092	0.000099
		Revised	0.000067	0.000079	0.000087	0.000092	0.000099
20	1231	Production	0.000100	0.000263	0.000440	0.000740	0.001316
		Revised	0.000070	0.000196	0.000338	0.000582	0.000997

Table 6: Distribution of max test for iterations 21–40.

Iter	N	Method	Min	1st Q	Median	3rd Q	Max
21	82	Production	0.001182	0.001295	0.001360	0.001432	0.001650
		Revised	0.000754	0.000843	0.000884	0.000931	0.001000
22	70	Production	0.001454	0.001557	0.001630	0.001725	0.001970
		Revised	0.000781	0.000847	0.000874	0.000929	0.000991
23	52	Production	0.001633	0.001846	0.001953	0.002148	0.002692
		Revised	0.000807	0.000846	0.000897	0.000929	0.000993
24	40	Production	0.001702	0.002122	0.002283	0.002539	0.003053
		Revised	0.000822	0.000862	0.000905	0.000945	0.000996
25	32	Production	0.001903	0.002238	0.002613	0.002803	0.003403
		Revised	0.000829	0.000873	0.000918	0.000955	0.000994
26	30	Production	0.002185	0.002668	0.002907	0.003124	0.003856
		Revised	0.000828	0.000903	0.000930	0.000958	0.000996
27	25	Production	0.002449	0.002946	0.003348	0.003909	0.004334
		Revised	0.000822	0.000874	0.000900	0.000940	0.000990
28	18	Production	0.002708	0.003034	0.003485	0.004131	0.004745
		Revised	0.000855	0.000885	0.000920	0.000948	0.000998
29	18	Production	0.002983	0.003645	0.004325	0.004957	0.005722
		Revised	0.000857	0.000907	0.000935	0.000948	0.000991
30	15	Production	0.003265	0.004009	0.004702	0.005759	0.006937
		Revised	0.000849	0.000882	0.000918	0.000962	0.000970
31	14	Production	0.003121	0.004196	0.004888	0.005766	0.005911
		Revised	0.000872	0.000902	0.000930	0.000962	0.000989
32	9	Production	0.004257	0.004732	0.005457	0.005714	0.006049
		Revised	0.000882	0.000921	0.000958	0.000980	0.000981
33	6	Production	0.003137	0.004800	0.005566	0.006448	0.006612
		Revised	0.000916	0.000917	0.000925	0.000978	0.000998
34	3	Production	0.005118	0.005118	0.005610	0.007326	0.007326
		Revised	0.000924	0.000924	0.000936	0.000994	0.000994
35	9	Production	0.006463	0.006669	0.007006	0.007275	0.007681
		Revised	0.000897	0.000938	0.000977	0.000980	0.000998
36	4	Production	0.004973	0.006289	0.007747	0.009422	0.010953
		Revised	0.000904	0.000910	0.000930	0.000964	0.000986
37	6	Production	0.005982	0.006326	0.007578	0.009201	0.009653
		Revised	0.000903	0.000919	0.000926	0.000955	0.000969
38	8	Production	0.005401	0.008032	0.008934	0.009437	0.010766
		Revised	0.000926	0.000929	0.000942	0.000981	0.000998
39	8	Production	0.006751	0.007785	0.008476	0.009130	0.010525
		Revised	0.000912	0.000931	0.000938	0.000969	0.000995
40	74	Production	0.007225	0.011641	0.016433	0.025521	0.172413
		Revised	0.000918	0.001971	0.003155	0.007097	0.171476