Evaluating Use of Alternative Population Controls for American Community Survey Weighting Methodology

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

Abstract
The objective of this research is to assess whether the use of independent population estimates produced by the Census Bureau's Population Estimates Program improves the estimates produced by the American Community Survey (ACS). Six alternatives to the current population controls were investigated by using controls either at higher levels of geography or with less demographic detail. An expected worst-case bias scenario for the population estimates is studied by applying the April 1, 2000 population estimates based on the 1990 Census as controls to the Census 2000 Supplementary Survey data. Various county-level demographic estimates were calculated for each of the seven methods. The county-level relative bias, variances, and relative mean square errors to the Census 2000 data were then calculated to compare the relative quality of the estimates for each method.

Key Words: population estimates, variance, coverage

1 Introduction

The purpose of this research is to investigate whether the use of population estimates as controls improve the quality of the American Community Survey (ACS) estimates in a measurable fashion. Nine alternative population controls are planned to be included in the study. Among the nine alternatives, both the current methodology and the use of no controls are included. At this time, results are available only for seven of the alternatives and only for the 2000 data. For each of the seven alternatives, estimates of bias, variance, and mean squared error for various characteristics were produced and a summary of those results are presented in this paper. A final report will include the remaining two alternatives and the 2005 data in full detail. While there are some exceptions, the general preponderance of the results show no clear indication that the current methodology presents serious flaws that would necessitate immediate action to change the weighting methodology for the ACS.

2 Scope and Purpose

As is the case for many of the Census Bureau demographic surveys, ACS estimates are controlled to agree with the population estimates. ACS weighted estimates of the number of persons are controlled to the population estimates by age, sex, race, and Hispanic origin at the weighting area level. An ACS weighting area is a county or a set of less populous counties that meet a minimum population or number of person interviews requirement. Adjustment of the survey estimates to agree with population controls is important for several reasons:

2.1 Reasons for use of controls

2.1.1 Correction for systematic undercoverage

Household surveys typically fail to locate every person in the target population, so that the weighted number of people using weights without the control adjustment tends to fall short of the census or population estimate. The causes for this phenomenon are a combination of missing whole households and leaving some persons off the household roster. Furthermore, the undercoverage is differential by age, Hispanic origin, race, and sex. The use of controls in the weighting helps to mitigate this differential.

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.

2.1.2 Reduction in sampling variance for survey estimates

The weighted estimate of the number of people using the pre-control weights varies from sample to sample, depending on the household size distribution and the vacancy rate measured by the sample. The weighted estimate of the number of persons by race, age, sex, and Hispanic origin also depends on which addresses are selected by sample. This sampling variance in the number of people contributes to the total sampling error in other estimates, such as the number of people with a high school diploma. But controlling to a fixed set of population controls, regardless of the sample selected, eliminates the sampling error for the controlled characteristics, and also helps to lower that sampling error of other estimates.

2.1.3 Smoothing of age distributions

Actual age distributions, even for single years of age, are usually fairly smooth, except for areas with very small population. In general, the populations of areas tend to grow or decline fairly smoothly with respect to the age distribution. However, even with fairly large samples, household surveys fail to capture these patterns accurately, and very often show large variations up and down over time, due to sampling error. These random jumps are often not "statistically significant", but the patterns displayed appear unrealistic. They also appear to be different in shape when compared to the pattern showed by the population estimates time series. The population estimates are not error-free, they tend to make the opposite error of showing a rather smooth trend line. In general, most data users are less disturbed by this than by "random variation" even if the "ups and downs" are declared to be not statistically significant.

2.1.4 Comparability with other estimates

The use of a single set of population estimates to control different surveys is advantageous because all of them give a consistent value for the population and brings more "credibility" to the survey estimates. Survey results can be compared without the need to account for inconsistencies in the population base. In practice, the consistency provided by the use of "controls" does not have to be perfect. It is common for many household surveys to target a different "total population", for example, people living in institutions or the military may be excluded or the "residence rules" may be deliberately different.

2.2 Research Questions

The Census Bureau currently uses the full demographic detail of the Population Estimates Program's population estimates at the county level as controls in the weighting of the ACS data. While we have used the housing and population estimates at the county level since 2000, a comprehensive evaluation has not been conducted to determine if the current methodology makes optimal use of the estimates and whether this use of population controls results in a measurable improvement of the quality of the ACS estimates.

The purpose of this research is to determine if the use of controls in the ACS weighting improves the quality of the ACS estimates. To research if the current method makes the best use of the population estimates as controls, we consider several other reasonable alternative sets of population controls. We then make a determination across all methods as to whether the population estimates can be used in some fashion to improve the quality of the ACS estimates. The specific research questions we will address to make this determination are as follows.

(A) How much does controlling reduce the measure of variance for estimates at the county level? We will compare the variance of estimates before and after controlling to the population estimates. We conduct this study using data from both the Census 2000 Supplementary Survey (C2SS) and the 2005 ACS. The C2SS (and equivalent surveys in 2001–2004) used ACS methods to collect data at a national level in 2000 prior to full implementation of ACS in 2005. Even when we control at the state level, we still measure variance reduction for the county-based weighting areas. We maintain this geographical unit of analysis for comparability across alternatives and because the PSU sampling design of the C2SS does not allow the calculation of the estimates at the county-level for all counties.

(B) How close does controlling bring the ACS estimates to the Census 2000? How much does controlling actually reduce bias (in relation to Census 2000)? Ideally we would know the true population for 2000 and could directly measure bias in the 2000 ACS controlled and uncontrolled estimates. Since we don't know the true population we will compare the ACS, both controlled and uncontrolled, to the Census 2000. For the Census 2000 to be a useful proxy for truth it should be closer to truth than the ACS estimates either controlled or uncontrolled. We will measure differences from the Census 2000 for all variables that appear on the Census 2000 short form, namely, age, sex, race, Hispanic origin, relationship, and tenure. This question will not be addressed for the 2005 data.

(C) How much does controlling improve the overall quality of the ACS estimates? We will assess the improvement (or lack thereof) in the quality of the ACS estimates by comparing the mean squared error of each method of use of controls compared to using no controls. This combines the estimates of variance and squared bias in relation to the Census 2000. Thus the limitations of the mean squared error are the same as the calculation of the bias. Like the second research question, we only have the data to answer this question for the 2000 data and not the 2005 data.

The scope of the study is to make inferences about the reduction in variance and in error of decennial short-form characteristics and traditional sample characteristics on the ACS questionnaire at the county-based weighting area level (for reasons explained above). Due to practical considerations we make inferences about a subset of these characteristics. The subset of variables that we concentrate on will serve as a proxy for the complete set of ACS variables.

Note that comparing the Census 2000 to the population estimates and evaluating the quality of the population estimates are not within the scope of this study, though they are clearly relevant to the question of the efficacy of controlling. However, these questions are the focus of a separate Population Division study. Furthermore, for both studies we identify a common group of core variables which both will analyze, allowing us to obtain for those variables deeper insights into the efficacy of controlling.

3 Weighting Methodology

The weighting methodology used for the study was based on the methodology used for the Census 2000 Supplementary Survey (the name for the ACS in 2000). Thus the description of the weighting methodology presented here differs from the current methodology used for the ACS. Most notably is the absence of the group quarters weighting since no group quarters data collection occurred in 2000 for the ACS. There are also some enhancements that were put into place in 2006 to improve the consistency of certain key estimates from the survey that are also not present in the methodology used for the study. The decision not to include these enhancements was to improve the comparability of the alternatives studied to the production methodology that was used to publish the 2000 estimates.

The current weighting method has three primary steps: calculation of the baseweights, adjusting for non-response, and the application of controls.

3.1 Baseweights

The baseweights are defined as the inverse of the sampling probabilities. For housing units, these weights are the inverse of one of four different sampling rates used in the ACS sampling operation in the 36 ACS test counties (which utilized the full ACS sampling procedures and had a higher sampling rate) or the inverse of the single county-based sampling rate used in the 1203 national counties.

The ACS collects data from a monthly sample panel over a three month period of time and the data are tabulated according to which month the interview is conducted or the mail form is received. Variations in monthly response by mode of data collection could lead to some months having more interviews than others. The goal is for the ACS period estimates to represent a simple average of the monthly characteristics contained in the period. For that reason, an adjustment is made to the weights so that the total weight of the interviews for a given month is equal to the total weight of the sample cases allocated to that month's panel. Since the number of sample cases per month is relatively constant this makes each month contribute roughly an equal share to the single- or multi-year period.

3.2 Non-response Adjustment

While the level of unit non-response in the ACS is quite small, it is not negligible and hence an adjustment is done for non- response. In other surveys and censuses, characteristics that have been shown to be related to housing unit response include census tract, building type (single- versus multi-unit structure), and month of data collection (Weidman et al., 1995). This full cross-classification, however, produces too many cells (more than one million) than the sample can support so instead the non-response adjustment is conducted in two steps. The first step calculates a ratio estimate within the cross-classification of census tract and building type in each weighting area. The second step calculates a ratio estimate within the cross-classification of month of data collection and building type. In doing this two step process, information from all three characteristics are used in the adjustment in a manner that the sample can support. Because of the smaller sample in the 1203 (non-test) counties, only the second non-response adjustment is used.

3.3 Application of Controls

The nonresponse adjusted weights are then controlled to a set of independent housing unit estimates produced by the Census Bureau's Population Estimates Program (PEP) to adjust for housing unit coverage. The HU weights after the application of the housing unit controls then serve as the initial household person weights.

The household person records are placed into one of 154 cells defined by the cross-classification of race/ethnicity (6) and age/sex (26). The six race/ethnicity cells are: Hispanic and non-Hispanic crossed by the five major race groups, White, Black, American Indian and Alaskan Native (AIAN), Asian, and Native Hawaiian and Other Pacific Islander (NHOPI). The race categories of Some Other Race and multiple race responses are allocated among the five race groups for weighting purposes.

The person post-stratification cells are collapsed as necessary to ensure that there are at least 10 sample persons in each cell and that the ratio between the pre-controlled estimate and the control total is within a factor of 3.5. Once the collapsed cells have been defined, a ratio adjustment is applied to the sample records to control their weights to the cell controls.

As a final step in the HU weighting, the HU weight for all occupied HUs are set equal to the person weight of the principal person. This corrects for differential coverage of households. However, as this adjustment causes the estimate of total housing units to no longer be in agreement with the controls, the housing unit controls are again applied to the housing unit records.

4 Research Methodology

The three research questions will compare nine alternative methods of controlling the ACS estimates that include the current methodology and the use of no controls. All of the analysis will use statistics calculated at the county-based weighting area level even when we control at the state or national level for consistency purposes when doing the analysis.

4.1 Alternatives Studied

The research questions relating to bias and mean squared error will be analyzed only for the 2000 short-form type data where we have the Census 2000 short form to compare. For Census 2000 long-form characteristics (and in future research, 2005 data) only the research question relating to the reduction of variances will be studied since we have no proxy for truth (that does not also have sampling variability). The alternatives studied are as follows (where the term "demographics" refers to the full cross of race/ethnicity/age/sex):

1. County total population and demographics (current methodology)
2. State demographics, county population totals
3. County total population, no demographics
4. State demographics only, no county population controls
5. National demographics and state population totals

6. State demographics and county population totals for large counties (250,000 or more)
7. No controls (including none to housing unit totals)
8. National demographics and county housing unit totals [future]
9. State race / ethnicity and county age/sex totals [future]

With the exception of (8), housing unit controls would be applied at the broadest level of geography for which we controlled the population. For (1), (2), (3), (6), (8), and (9) they would be applied at the county level; for (4) and (5) at the state level. Only option (8) are the housing unit controls not applied at the same level (county) as total population (national). With the exception of alternative (8), person estimates are invariant to the housing unit counts and controlling to housing unit population estimates affects only estimates of housing unit and household variables.

In order to make use of the 1990-based 2000 population controls, the Asian and "Native Hawaiian and Other Pacific Islander" race groups had to be combined to be comparable to the 1990 "Asian and Pacific Islander" category. The current methodology of assigning multi-race respondents to a single-race code for the purpose of weighting was used for all alternatives where demographic controls were applied.

4.2 Statistical Measures

We created estimates for several characteristics based on weights using each of the methodologies outlined above and based on the Census 2000 data. We then compare those estimates to answer each of the research questions relative to that characteristic.

For research question (A) regarding the reduction of variance, we calculate for each weighting area the variance of the estimate using each alternative and the percent reduction in variance of each alternative compared to the use of no controls (alternative #7). We then summarize the results by calculating the median percent reduction in variance across all weighting areas for that characteristic. A positive result represents a reduction in variance whereas a negative result represents an increase in variance. The largest positive result is 100% but there is no limit to the negative results. The Median Percent Reduction in Variance for Alternative $\#i$ is calculated as follows:

$$\text{Median}\left\{\frac{\text{Var Alt}\#7 - \text{Var Alt}\#i}{\text{Var Alt}\#7}\right\} \times 100\%$$

For research question (B) regarding the reduction in the squared bias, we calculate for each weighting area the squared bias compared to Census 2000 of the estimate for each alternative. We then calculate the percent reduction in the squared bias as compared to using no controls for each alternative and summarize the results by calculating the median percent reduction across weighting areas A positive result represents a reduction in bias whereas a negative result represents an increase in bias. The largest positive result is 100% but there is no limit to the negative results. The Median Percent Reduction in Squared Bias for Alternative $\#i$ is calculated as follows:

$$\text{Median}\left\{\frac{\text{Sq. Bias Alt}\#7 - \text{Sq. Bias Alt}\#i}{\text{Sq. Bias Alt}\#7}\right\} \times 100\%$$

where

$$\text{Squared Bias Alt}\#i = (\text{Est. Alt}\#i - \text{Census Est.})^2$$

The final statistic calculated is for research question (C) regarding the reduction in the mean squared error. The median percent reduction statistic for the mean squared error is calculated in a similar fashion to the two statistics above.

$$\text{Median}\left\{\frac{\text{MSE Alt}\#7 - \text{MSE Alt}\#i}{\text{MSE Alt}\#7}\right\} \times 100\%$$

where

$$\text{MSE Alt}\#i = \text{Squared Bias Alt}\#i + \text{Var Alt}\#i$$

4.3 Tables

The tables for analysis fall into two primary classifications: univariate and multivariate. The univariate tables have fewer cells overall which allows us to do the comparisons at either finer levels or by size of geographic area or both. The multivariate tables start with more cells than the univariate tables that requires us to limit the number of levels of a variable or our ability to break out the results by size of area.

All tables produced that involve race have been allocated to the 1990 Census race codes. The source data for the Census 2000 and ACS already had the Census 2000 Some Other Race allocated to the five Census 2000 single-race categories. For the purposes of the analysis, we allocated all muti-race respondents to the four 1990 single-race categories proportionally to the size of the single-race categories by weighting area, as determined by the population estimates for the ACS estimates and as determined by Census 2000 for the Census 2000 counts. We also combined the 2000 race categories Asian and "Native Hawaiian and Other Pacific Islanders" into the 1990 race category, "Asian and Pacific Islander".

4.3.1 Univariate Tables for Analysis

There are seven primary univariate tables for analysis.

- Total Population
- Sex (male, female)
- Age (0–17, 18–34, 35–44, 45–64, 65+)
- Race/Ethnicity (Hispanic, non-Hispanic crossed by White, Black, Asian and Pacific Islander, and American Indian and Alaskan Native)
- Tenure (owner, renter)
- Relationship (householder, spouse, child, other relative, non-relative, unmarried partner)
- Household Income (ten levels)

4.3.2 Multivariate Tables for Analysis

For the cells defined by crossing Age, Sex and Race/Ethnicity we maintain the two Sex cells and five Race/Ethnicity cells but collapse the five Age cells to three groups: 0-17, 18-64, 65+. The 30 cells identified here are a result of crossing the four variables listed above. They are: 2 sex groups by 3 age groups by 5 race/ethnicity groups. In addition, to maintain sufficient sample in each cell we restrict the analysis to weighting areas that contain at least 10% of that race/ethnic group according to Census 2000 and have at least 20 unweighted persons in the C2SS.

4.4 Analysis by Size of Weighting Area

We will perform analysis in three weighting area size groups to achieve more homogeneity.

- small (household population $< 250,000$)
- medium ($250,000 \leq$ household population $< 750,000$)
- large (household population $\geq 750,000$)

There are two weighting areas with more than 5,000,000 which could be analyzed separately in some analysis. We stratify by size of weighting area because we suspect that the controls may be of lower quality for smaller areas. We would like to be able to measure this phenomenon if it is the case.

5 Results and Discussion

At this time, results are available only for the 2000 data and only for the first seven options. Given the volume of data produced, the data were summarized by scoring "winners" for each variance, bias, or MSE statistic for each characteristic. The winner was scored one point and ties (within 0.1%) resulted in both alternatives scoring a point. Tables 1 and 2 present tallies of total points. Table 1 tallies the scores by alternative for each statistic by table type (age, sex, etc.) in order to determine if certain alternatives perform better for different groups of characteristics. Table 2 tallies the scores by alternative for each statistic by size of weighting area to determine if there is a population size differential among the alternatives. Included in both tables is a maximum score column for context.

We expect that the weighting area variances for total population, sex, age, and race/ethnicity should be reduced dramatically if not completely through the use of (detailed) controls at the county level. For each of these characteristics, the current methodology shows the greatest reduction in variance compared to the alternatives for these items. The concern has largely been about the potential of introducing additional biases that may be present in controls at this level. Here we see that for total population, sex and age, the current methodology achieves the greatest or equal reduction of bias in approximately 92% of the table cells as well.

For race/ethnicity, the current methodology of using race demographics at the county level also yielded the greatest reduction in bias compared to the other alternatives. One race/ethnicity category where the alternatives consistently scored higher was for non-Hispanic American Indian and Alaskan Natives (AIAN) in all sized weighting areas. This is consistent with evaluations that have shown that this group also posed challenges to the 1990-based population estimates. In addition, alteratives other than the current methodology also performed better for Hispanics and non-Hispanic Whites (typically in the same areas) in the large weighting areas. This result is also consistent with evaluations that have shown that the estimate of Hispanics stemming from net foreign migration to be a challenge for the 1990-based population estimates. However, the current method was still the winner for all groups with respect to the greatest reduction in the MSE. In those categories where the current method did not perform the best with respect to bias, there was no clear-cut best alternative.

For tenure, the reduction in bias was generally best for the current method. For the estimate of owner occupied in small weighting areas, the current methodology demonstrated a reduction in bias compared to the use of no controls but not as much as the alternative that simply made use of the county population totals. The difference in the reductions was small, however, with the the top performing alternative achieving only 2–4% greater reduction in bias.

Finally, the multivariate table of race/ethnicity crossed by age and sex were also generally favorable to the current method. While the current method achieved the greatest variance reduction of all the alternatives as expected, it also achieved the greatest reduction in bias for about 70% of the cells. Non-Hispanic White and Blacks both consistently scored the highest for all three statistics using the current methodology. The exceptions were non-Hispanic AIAN females and males age 65 and older, non-Hispanic API age 18–64 and males 65 and older, and Hispanic females aged 0–7 and 18–64. The non-Hispanic API favored either state demographics with large county county totals or no controls. Non- Hispanic AIAN category favored state-level demographics with county totals included for males 65 and older and not included for females 65 and older which demonstrates the difficulty in identifying alternative methods to the current method. Because of the criteria for a weighting area to be included in the multivariate analysis, both of these race/ethnic categories have fewer weighting areas included in the results which suggests that there may be individual observations which are very influential in the results.

6 Conclusions

The trend in these preliminary results show a fairly consistent pattern of alternative #1, the current method using the population estimates full detail at the county level, achieving the greatest reduction in bias, variance and mean squared error. There are exceptions to this trend that warrant further study of extreme values in order to better understand the data. However, there is no clear indication that the current methodology presents serious flaws that would necessitate immediate action to change the weighting methodology for the ACS.

7 Limitations and Future Research

We will continue to produce and analyze additional data from this study to confirm if this trend continues. We fully recommend that this study be repeated after the 2010 Census data and the 2010 population estimates projected forward from 2000 are internally available. The reasons for this are two-fold. The first is that due to improvements in the methodology used to produce the population estimates, our expectation is that the quality of the population estimates has improved since 2000. The second is that this study was limited to studying relatively large weighting areas due to the smaller sample sizes for the C2SS as compared to the full implementation ACS. Using full implementation data, we will be able to conduct the study for finer categories of counties that are significantly smaller than the weighting areas used for C2SS. This could reveal different results than the current study once smaller counties are studied.

8 References

Weidman, L., Alexander, C., Diffendahl, G., and Love, S. (1995), "Estimation Issues for the Continuous Measurement Survey," Proceedings of the Survey Research Methods Section, American Statistical Association, pp. 596-601, 1995, www.census.gov/acs/www/AdvMeth/Papers/ACS/Paper5.htm.

		Maximum Score	Status Quo	State Demographics, County Totals	County Totals Only	State Demographics Only	National Demographics, State Totals	State Demographics, Large County Totals	No Controls
Total Population	Variance	3	3	3	3	0	0	2	0
	Bias	3	3	3	3	0	0	2	0
	MSE	3	3	3	3	0	0	2	0
Sex	Variance	6	6	0	0	0	0	0	0
	Bias	6	6	0	0	0	0	0	0
	MSE	6	6	0	0	0	0	0	0
Age	Variance	15	15	0	0	0	0	0	0
	Bias	15	13	1	0	1	0	0	0
	MSE	15	15	0	0	0	0	0	0
Race	Variance	15	13	0	0	0	0	0	0
	Bias	15	10	1	2	0	0	1	2
	MSE	15	15	0	0	0	0	0	0
Tenure	Variance	6	6	0	0	0	0	0	0
	Bias	6	4	1	2	0	0	0	0
	MSE	6	6	0	0	0	0	0	0
Relationship	Variance	18	18	0	0	0	0	0	0
	Bias	18	10	2	1	2	0	3	1
	MSE	18	16	0	0	0	0	1	1
Household Income	Variance	30	23	1	2	0	0	1	2
	Bias	NA	NA	NA	NA	NA	NA	NA	NA
	MSE	NA	NA	NA	NA	NA	NA	NA	NA
Race / Ethnicity / Age / Sex	Variance	30	29	0	1	0	0	0	0
	Bias	30	21	1	3	2	0	1	2
	MSE	30	27	1	0	2	0	0	0

Table 1: Summary of Scores for Each Alternative by Characteristic

		Maximum Score	Status Quo	State Demographics, County Totals	County Totals Only	State Demographics Only	National Demographics, State Totals	State Demographics, Large County Totals	No Controls
Small Weighting Areas	Variance	21	20	1	1	0	1	0	0
	Bias	21	16	3	4	0	0	1	0
	MSE	21	21	1	1	0	0	0	0
Medium Weighting Areas	Variance	21	20	1	1	0	1	1	0
	Bias	21	19	1	2	0	0	1	1
	MSE	21	21	1	1	0	0	1	0
Large Weighting Areas	Variance	21	21	1	1	0	0	1	0
	Bias	21	11	4	2	3	0	4	2
	MSE	21	19	0	1	0	0	2	1

Table 2: Summary of Scores for Total Population, Sex, Age, Race, Tenure and Relationship for Each Alternative by Size of Weighting Area