Using Subcounty Estimates as Controls in Weighting for the American Community Survey

Keith A. Albright
U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract[1]
The American Community Survey (ACS) uses independent housing unit and population estimates, produced by the Census Bureau's Population Estimates Program (PEP), as controls. These controls are applied to county based weighting areas. The PEP also publishes estimates of totals for subcounty areas, which are incorporated places and minor civil divisions. The ACS estimates for totals in these subcounty areas are normally different from the PEP estimates, and in a few cases substantially different. These differences are disconcerting to many data users. In this paper we evaluate efforts to incorporate the PEP estimates for subcounty areas into the ACS weighting process. The main goal is to make ACS estimates of total housing units and population for subcounty areas closer to the PEP estimates, with minimal effect on distributions of housing and population characteristics.

Key Words: weighting, population controls, raking

1. Introduction

The Census Bureau's Population Estimates Program (PEP) produces estimates of demographic groups by age, sex, race, and Hispanic origin for counties, and in addition it produces estimates of the total population and housing units for subcounty areas. Current ACS methodology controls estimates by demographic groups to controls based on the PEP county estimates. These estimates are aggregated to county-based *weighting areas*, which are single counties or groups of small counties. However, no use is made of the PEP's subcounty estimates.

PEP subcounty estimates are produced for incorporated places and minor civil divisions (MCD) in 20 states where county MCDs have functioning governmental units. The subcounty areas for which PEP produces estimates include county/place parts, MCD/place parts, and unincorporated areas. A county/place part is the portion of an incorporated place that lies within a particular county, as incorporated places may cross county boundaries. Similarly, incorporated places may cross MCD boundaries and each portion is called an MCD/place part. When an incorporated place does not cross county or MCD boundaries, the county/place part or MCD/place part simply consists of the whole place. Unincorporated areas of a county are those that are not part of any incorporated place and are collectively referred to as the 'balance of county'. Similarly, portions of MCDs that are not part of an incorporated place are referred to as the 'balance

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

of MCD.' The PEP internet site (http://www.census.gov/popest/estimates.html) provides more information on the types of geography discussed in this paragraph, the methodology behind creating the estimates, as well as downloadable files containing the estimates.

In this paper, we evaluate the results of incorporating the PEP subcounty estimates into the ACS weighting methodology. For the study we use the 2005-2007 ACS data. We use several different strategies involving both the housing unit and population PEP estimates and different methods for defining subcounty areas to which we apply these estimates. This paper makes no judgments about whether ACS estimates (totals or distributions) are better or worse when using the PEP subcounty controls compared to the estimates that the ACS would produce without using the PEP subcounty controls.

1.1 Current ACS Weighting Methodology

The first step in the ACS weighting process is to assign an initial base weight to each housing unit in sample. These weights account for the differential nature of the sample selection system. In the second major step, a series of non-response adjustments are applied. These primary steps are not affected by controlling to subcounty PEP estimates and will not be discussed here. The ACS Design and Methodology Report (U.S. Census Bureau, 2009) gives a detailed description of all steps in the current ACS weighting procedure. The steps in the weighting process that are affected by the controlling to subcounty PEP estimates are briefly described below.

First is the Generalized Regression Estimation Adjustment (GREG). This adjustment is designed to reduce variability at the subcounty level, while having minimal effects on the expected value of the estimates of characteristics (Fay, 2006). It is a model-assisted step that makes use of sample frame counts and auxiliary information from administrative data sources. The models used to compute adjustment factors are specified and estimated at the subcounty level. Next is the housing unit post-stratification factor (HPF) adjustment. The HPF adjustment is done using a simple ratio adjustment that makes the weighted total of housing units equal to the PEP county based estimates at the weighting area level.

Following the HPF adjustment, sample persons are initially assigned the weight of their housing unit. Then a person post-stratification adjustment (PPSF) is applied. The current PPSF adjustment is computed using an iterative raking procedure first given by Deming (1940). A single iteration of the raking consists of three successive ratio adjustments. Each ratio adjustment is made at the weighting area level, with the total population controlled to the PEP county based estimate. The first ratio adjustment is spouse equalization and is applied to three categories of persons: householders in a married couple/unmarried partner relationship, the second person in this relationship, and all other persons. The adjustment makes the weighted total of householders in a relationship equal the weighted total of partners in these relationships. These totals are set equal to the weighted total of married couple/unmarried partner households. The next ratio adjustment is householder equalization and is applied to two categories of persons: householder and non-householder. The adjustment makes the weighted total of householders equal to the weighted total of occupied housing units. The final ratio adjustment is applied to race/age/sex groups. The number of categories varies by weighting area. It has a maximum of 156 cells defined by the combinations of six race and 26 age/sex groups. Cells are collapsed when necessary so that they meet minimum requirements for sample size and magnitude of adjustment. This adjustment makes the ACS weighted total for the collapsed race/age/sex groups equal to the PEP county based estimates. At the end of each iteration, a test for convergence is done. A weighting area passes the convergence test if the new estimates in each of the spouse and householder equalization categories are within a maximum threshold of the totals to which they are controlled. Raking for a weighting area stops when convergence is met. Note that the categories in the demographic ratio adjustment don't have to be tested since they are adjusted last in each iteration. The raking procedure results in an adjustment factor for each combination in the full cross classification of spouse, householder, and demographic equalization categories. As a final step Person weights are rounded using a controlled rounding process.

The last step is to set the final housing unit weight. The housing unit weight for occupied housing units is then set equal to the final rounded person weight of the householder, making the estimate of occupied housing units equal to the estimate of householders. The weights of vacant housing units are rounded using a controlled rounding process. A consequence of this step is that the ACS estimates of total housing units is no longer equal to the PEP county based estimates, although they are still close.

2. Integrating the PEP Subcounty Estimates

This section describes changes to the ACS weighting made to make use of the PEP subcounty estimates. The subcounty estimates that will be used in weighting are estimated totals for housing units and population for incorporated places and MCDs (for the 20 states that have MCDs as functioning governmental units). These estimates will be incorporated into the housing unit post-stratification adjustment and the person raking process.

There are three alternative methods for incorporating subcounty estimates into the housing and person post-stratification adjustments that will be compared. The first is to not use PEP subcounty estimates, and is referred to as 'no controls'. The second is to only use PEP subcounty housing unit estimates and apply them to the HPF adjustment. In this alternative, we allow the results of the ACS to determine population estimates. The third is to use both housing unit and population PEP subcounty estimates. These three alternatives are summarized in the table at the end of this section.

2.1 Integrating PEP Subcounty Estimates into the Weighting Methodology

2.1.1 Revised Generalized Regression Estimation Adjustment

The only change to the GREG adjustment is in the creation of the subcounty areas for which models are created. We apply the adjustment to the same subcounty areas to which we will apply PEP estimates.

2.1.2 Revised Housing Unit Post-Stratification Factor

The HPF adjustment is made at the subcounty level when using subcounty housing unit estimates. The adjustment is made at the county level in counties where there is no subcounty area defined.

2.1.3 Revised Person Post-Stratification Factor

Incorporating the subcounty population estimates into the raking procedure requires modification of the ratio adjustments currently used. We still use three ratio adjustments, with the subcounty population total being one ratio adjustment. The spouse and householder equalization ratio adjustments are combined. As with the current method, the total population is kept controlled at the weighting area level after each adjustment. The first ratio adjustment is subcounty equalization. This adjustment makes the ACS estimate of total population of each subcounty area equal to the PEP estimates. In weighting areas where no subcounty areas are defined, the adjustment is applied at the county level. Note that this adjustment does not apply in the first two alternative weighting methods described in the beginning of Section 2. The second ratio adjustment combines the first two ratio adjustments of the current methodology and has four levels: Householders in a married couple/unmarried partner relationship, the second person in this relationship, householders not in a married couple/unmarried partner relationship, and all other non-householders. This adjustment simultaneously equalizes two pairs of estimates: householders in a relationship equal to the estimate of partners in these relationships, and the estimate of householders equal to the total of occupied housing units. As with the current methodology, this adjustment is applied at the weighting area level. The third ratio adjustment is the demographic equalization and is also be applied at the weighting area level. This adjustment is unchanged from the current methodology.

2.2. Defining Subcounty Areas for Weighting Adjustments

We use two different methodologies to define subcounty areas to which housing and population adjustments are applied. Both methods are evaluated and compared. The two methods that we will compare are what we refer to as the ‘whole’ and ‘parts’ methods. The distinction between the two is in how places and MCDs are treated in the 20 MCD states. There is no difference between the two methods in non-MCD states (and in several of the MCD states). Both methods feature a minimum population requirement for all subcounty areas and the entire county. All subcounty areas must have an estimated population of at least 8,000 population. This applies to all types of subcounty areas (county/place parts, MCDs, MCD/place parts, balance of county, and balance of MCD).

2.2.1 The ‘Whole’ Method

The “whole” method is the simplest. First, we describe its application in non-MCD states. The basic unit for forming subcounty areas is the county/place part for incorporated places. County/place parts with an estimated population of at least 8,000 are defined as a subcounty area to which PEP estimates are applied. All other places and unincorporated areas are grouped into a single subcounty area (the balance of county). If the estimated population in the balance of county is too small, it is combined with the smallest county/place part in the county. Combining with the smallest county/place part decreases the chance that the balance of county is combined with a place for which ACS estimates are published.

In the 20 MCD states, the basic unit for forming subcounty areas is either the place/county part or the MCD. In both cases, the method is very similar to what is used in the other states. MCD states are classified into those with ’strong‘ MCDs and ’weak’ MCDs. The MCD is used in the strong MCD states and the place is used in weak MCD states. The strong MCD states are Connecticut, Maine, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont and Wisconsin. The weak MCD states are Illinois, Indiana, Kansas, Missouri, Nebraska, North Dakota, Ohio, and South Dakota. In the strong MCD states, an incorporated place may be a subcounty area only if it is not part of an MCD with a population of at least 8,000. If a 'balance of county' area is too small, it will be collapsed with the MCD that has the smallest population. In the weak MCD states, an MCD may be a subcounty area only if it does not have any places in it that meet the population threshold. If a 'balance of county' area is too small, it will be collapsed with the place that has the smallest population.

2.2.2 The 'Parts' Method

This is very similar to what was used to form subcounty areas for the GREG adjustment for 3-year estimates in the 2005-2007 ACS weighting. No distinction is made between strong and weak MCD states with this method. It attempts to give both MCDs and incorporated places the benefit of weighting adjustments at the subcounty level. The basic unit for forming subcounty areas is the MCD/place part. Place parts within an MCD that are large enough are made subcounty areas. Remaining portions of the MCD are combined into a 'balance of MCD' area. If this balance portion is not large enough, it is combined with the smallest place part in the MCD. So MCDs with a population of at least 8,000 are completely partitioned into one or more subcounty areas. MCDs and places that are not large enough are combined into a balance of county area. If this area is not large enough, it is combined with the smallest MCD.

2.2.3 Limitations of the Two Methods

Both methods share a limitation imposed by the 8,000 population threshold that is used when forming subcounty areas. The collapsing of subcounty areas that results from imposing the threshold can reduce the benefit of using subcounty controls for some places/MCDs because they get combined with other areas. When a place crosses county boundaries, for example, one or more of the county/place parts often has less than 8,000 population. These parts are placed in the balance of county, so only part of the place has its population controlled. Another situation that occurs is with counties that have the majority of the population in one place or MCD and the balance of the county's population is too small. This results in the county having no subcounty areas defined due to collapsing, so no part of the place or MCD has its population controlled.

2.2.4 Example

This section illustrates the application of the 'whole' and 'parts' methods to Porter County, Indiana. Porter County has 11 incorporated places and 11 MCDs for which PEP estimates are published. In this county, incorporated places are referred to as cities or towns and MCDs are referred to as townships. The intersections of these places and townships form 14 MCD/place parts, including balance of MCD portions, for which PEP estimates are also published. Table 1 shows the MCDs and MCD/place parts with their vintage 2007 population estimates. Note that the places Chesterton town, Portage city, and Valparaiso city have parts in more than one MCD. Also the MCDs of Jackson, Morgan Porter, and Union townships have no incorporated places in them.

Table 1. Incorporated Places and MCDs in Porter County, Indiana

MCD MCD/Place Part	**2007 Population Estimate**	**MCD** MCD/Place Part	**2007 Population Estimate**
Boone Township	**6,472**	**Portage Township**	**47,234**
Hebron town	3,646	Chesterton town (part)	3
Balance of MCD	2,826	Ogden Dunes town	1,276
Center Township	**40,165**	Portage city (part)	35,876
Valparaiso city (part)	29,412	Balance of MCD	10,079
Balance of MCD	10,753		
Jackson Township	**5,127**	**Porter Township**	**9,251**
Liberty Township	**7,593**	**Union Township**	**8,759**
Chesterton town (part)	925		
Balance of MCD	6,668		
Morgan Township	**3,362**	**Washington Township**	**4,151**
		Valparaiso city (part)	539
Pine Township	**3,144**	Balance of MCD	3,612
Beverly Shores town	720	**Westchester Township**	**20,612**
Town of Pines town	792	Burns Harbor town	1,073
Pleasant Township	**4,708**	Chesterton town (part)	11,557
Kouts town	1,811	Dune Acres town	228
Balance of MCD	2,897	Portage city (part)	629
		Porter town	5,344
		Balance of MCD	1,781

In applying the ‘whole’ method, note that Indiana is considered a ‘weak’ MCD state. Thus we give priority to incorporated places over MCDs when defining subcounty areas for weighting adjustments. The places Chesterton town, Portage city, and Valparaiso city have estimated populations over 8,000 so they are defined as subcounty areas. Since Porter and Union townships, both MCDs, have estimated populations over 8,000 *and* have no incorporated places in them that are large enough, they are also defined as subcounty areas. All other places and MCDs are grouped together into the single ‘balance of county’ area since they have estimated populations less than 8,000.

In applying the ‘parts’ method, we look at the MCD/place parts. Note that the MCDs Center, Portage, Porter, Union, and Westchester townships are the only ones with estimated populations greater than 8,000. So those are the only MCDs that will have any parts defined as subcounty areas. All other MCDs are grouped together into the ‘balance of county’. In Center township, the part of Valparaiso city and the balance of MCD portion are both large enough to be subcounty areas. In Portage township, the part of Portage city is defined as a subcounty area, while the parts of Chesterton town and Ogden Dunes city are added to the balance of MCD portion. Porter and Union townships are defined as subcounty areas like they were using the ‘whole’ method. In Westchester township, the part of Chesterton town is defined as a subcounty area, while the remaining places are all added to the balance of MCD portion.

3. Analyses and Results

The three options for using subcounty estimates and the two methods we use to define subcounty areas combine to create six alternative weighting methods. Each of these weighting alternatives were applied to the 2005-2007 ACS data. The six alternatives are summarized in Table 2. The nomenclature A1, A2, etc. in Table 2 will be used throughout the rest of this paper to refer to the six alternative weighting methods.

Table 2. Six Weighting Alternatives Applied to the 2005-2007 ACS Data.

Method to Form Subcounty Areas	Applying Subcounty Estimates		
	No Controls	Housing Units Only	Housing Units and Population
Whole MCD or Places	A1	A2	A3
MCD/Place Parts	B1	B2	B3

Our analyses focus on the comparisons of total population and total housing units along with their variance estimates, and variance estimates for other population and housing characteristics. Variance estimates are analyzed using the coefficient of variation (CV). The CV of an estimate X is given by 100*SE(X)/X. We evaluated estimates for all incorporated places and MCDs for which 2005-2007 ACS estimates were published. Places and MCDs are classified into three size categories. Small areas are those with an estimated population between 20,000-65,000. Medium areas are those with an estimated population between 65,000-250,000. Areas with an estimated population greater than 250,000 are classified as large.

3.1 Estimates of Total Population and Total Housing Units

We first look at how consistent ACS estimates of total population and housing units are with PEP estimates for incorporated places and MCDs. Consistency is measured by the absolute relative percent difference between ACS and PEP estimates, which is given by $100*ABS(Est_j - PEP)/PEP$, where PEP is the PEP estimate and Est_j is the ACS estimate for weighting alternative j. We compared the absolute relative percent differences from different weighting alternatives by examining their distributions using the nonparametric Friedman test (Ostle and Malone 1988). The Friedman test compares dependent samples that are organized in a complete block design. In our application, the blocks are the places/MCDs and the treatments are the alternative weighting methods. This section compares estimates from places/MCDs in all states, using weighting alternatives B1, B2, and B3 (using the ‘parts’ method). In Section 3.3, we will see that the ‘parts’ method is preferable to the ‘whole’ method for forming subcounty areas.

3.1.1 Consistency Estimates of Total Population

Table 4 shows the distribution by percentile of the absolute relative percent differences between the ACS and PEP subcounty population estimates of total population. These distributions are shown for the weighting methods B1, B2, and B3 broken down by size category of place/MCD (there are 64 large, 479 medium, and 1,673 small place/MCDs). At the bottom of Table 4, Friedman test p-values are also given for each of three comparisons: B1/B2, B1/B3, and B2/B3.

The data in Table 4 shows that there is some improvement in consistency between ACS and PEP when we use only housing unit PEP subcounty estimates in the weighting (method B2). For example, the median absolute percent difference among the 479 medium size areas is for method B2 is 1.907% compared to 2.479% when no PEP estimates are used in weighting (method B1). Only in large areas is the difference between B1 and B2 not statistically significant. This may be because of the smaller sample size, 64, for the large areas. When we also use PEP population estimates in weighting (method B3), there is a large improvement in consistency. For example, the median absolute relative percent differences are reduced to near zero in all three size of area groups. To summarize, using only housing unit PEP subcounty estimates in the weighting provides some benefit to estimates of total population, but the best results are achieved by also including the population PEP estimates.

3.1.2 Consistency of Estimates of Total Housing Units

Results for comparison of ACS and PEP estimates of total housing units show the controls to PEP subcounty housing unit estimates make the ACS more consistent with the PEP. However, the controls to both the PEP subcounty housing unit and population totals make the ACS estimates of housing unit totals less consistent. Table 5 shows, for estimates of total housing units, the same information as in Table 4. Absolute relative percent differences are lower for method B2 than for B1. For example, the median absolute relative percent difference for medium size areas is 0.577% for method B2 compared to 1.29% for method B1. As an aside, the reason the absolute relative percent differences under method B2 aren't close to zero is due to setting the final housing unit weight equal to that of the householder as described in Section 1.1.4. But in method B3, the absolute relative percent differences actually increase and are larger than in method B1. For example, the median absolute relative percent difference for medium size areas increases to 1.806%. Only in large areas is there no significant difference between methods B1 and B3. In summary, using only housing unit PEP subcounty estimates in the weighting improves consistency for estimates of total housing units. But when PEP subcounty population estimates are also used in weighting, we get less consistency than when no PEP subcounty estimates are not used. This is likely due to differences between housing and population coverage at the subcounty level. Recall that the final housing unit weight is set equal to the weight of the householder, which results in ACS housing unit estimates no longer being equal to the values they were controlled to. The person weight of the householder is impacted by differential population coverage, and this in turn impacts the weights of the occupied housing units. Thus this leads the ACS estimates of total housing units to differ from the PEP when the subcounty population controls are also applied. Since we typically see greater variability in population coverage at the subcounty level than the county level, consistency between ACS and PEP subcounty housing estimates is more variable when subcounty population controls are used than when they are not used.

3.1.3 CVs for Estimates of Total Population and Housing Units

The effects of the different weighting methods on these CVs largely mirror those for the estimates of consistency of total population and housing units that were observed in Section 4.1.2. Tables 6 and 7 show the distributions of the CVs for total population and housing units respectively. Using PEP subcounty housing unit estimates in weighting (method B2) results in a small reduction in CVs for total population and a larger reduction in CVs for total housing units compared to not any using PEP estimates (method B1). Using both PEP subcounty housing and population estimates in weighting (method B3) results in a large reduction in the CVs for total population. But the CVs for total housing units actually increase compared to not any using PEP estimates.

3.2 CVs for Other Population and Housing Characteristics

Our analysis of population and housing characteristics focused on comparing variance estimates that result from the different weighting methods. The characteristics that we computed estimates for come from the ACS 'Data Profiles' that are published on the Census Bureau's American Fact Finder website (http://factfinder.census.gov). We organized selected estimates into thirteen groups. These groups are listed in Table 3 below. For example, the category 'Sex' contains estimates of males, females, males/females over age 18, and males/females over age 65. The category 'Tenure' contains estimates of occupied, vacant, owner occupied, and renter occupied housing units.

Table 3. Groups for Population and Housing Characteristic Estimates

Population Characteristics	Number of Estimates	Housing Unit Characteristics	Number of Estimates
Age	13	Persons Per Household	4
Sex	6	Tenure	4
Race	8	Household Income	10
Relationship	6	Value	8
Marital Status	10	Vacancy Rate	2
Educational Attainment	9	Units in Structure	7
Occupation	10		

The coefficient of variation (CV) is used to compare variance estimates among the different weighting methods. As in Section 3.1, we make comparisons only among weighting alternatives B1, B2, and B3 (using the 'parts' method), using places/MCDs from all states. Within each subcounty area, the median CV for each characteristic group was computed. The Friedman test was conducted for each characteristic group to compare median CVs, within subcounty size category. P-values for the Friedman tests comparing the CVs are shown in Table 12. Tests are shown for the three comparisons B1/B2, B1/B3,and B2/B3 by size of area.

The comparisons of B1 vs. B2 show little difference between CVs, with the majority of tests not being significant 0.05 level. Most of the significant test results are for small areas, which is likely because of the larger sample size.

The comparisons of B1 vs. B3 show that method B3 produces lower CVs than B1 for most of the population characteristics. Significant test results are seen in the housing characteristics tenure, household income, value, and units in structure, except in large areas where only tenure is significant. The significant results among housing characteristic groups are due to the median CVs being *higher* for method B3 than method B1.

Comparisons of B2 vs. B3 show B3 producing smaller CV estimates than B2, with tests being significant for most characteristics. The exceptions for the housing characteristics noted in the previous paragraph also apply here.

In summary, using only housing unit PEP estimates in the weighting produces very little, if any, reduction in variance compared to not using the PEP. The most reduction in variance is achieved when both housing and population PEP estimates are used.

3.3 Comparison of 'Whole' and 'Parts' Methods

In this section we make comparisons for weighting alternative pairs A2/B2 and A3/B3. This allows comparisons of the 'whole' and 'parts' methods for forming subcounty areas while holding the method of applying subcounty estimates constant. As mentioned in Section 2.2, the 'whole' and 'parts' methods produce the same subdivisions of counties in many of the MCD states. The counties where the two methods subdivide the counties differently are mostly in Illinois, Indiana, Ohio, and New York, with a few more counties in Connecticut, Kansas, Missouri, and Vermont. The comparisons in this section only use places and MCDs in those counties where the ‘whole’ and ‘parts’ methods produce different results.

3.3.1 Comparison of Methods A2 and B2

Looking at weighting methods A2 and B2, we compare the performance of the ‘whole’ and ‘parts’ methods when only PEP subcounty housing unit estimates are used in weighting. Table 8 shows, for each method, the distributions of absolute relative percent difference between ACS and PEP subcounty population estimates for large, medium, and small size areas. The p-values for Friedman tests are also included. Table 9 shows the same information for subcounty housing unit estimates. For both population and housing unit estimates, the ‘parts’ method performs better, except for large areas where there were no significant differences.

3.3.2 Comparison of Method A3 and B3

Looking at weighting methods A3 and B3, we compare the performance of the ‘whole’ and ‘parts’ methods when both population and housing PEP subcounty estimates are used in weighting. Table 10 shows, for each method, the distributions of absolute relative percent difference between ACS and PEP subcounty population estimates for large, medium, and small size areas. The p-values for Friedman tests are also included. Table 11 shows the same information for subcounty housing unit estimates. For estimates of population, a significant difference is only seen in medium size areas where the ‘parts’ method performs better. But it should be noted that for small areas, the absolute relative percent differences at the upper ends of the distributions are much smaller for method B3 than A3, even though the statistical test was not significant. For housing unit estimates, no significant difference was found between the two methods, but for small areas the most extreme values of the absolute relative percent differences are smaller under the ‘parts’ method.

Based on the result discussed in the previous two paragraphs, we conclude that the ‘parts’ method produces better overall results than the ‘whole’ method in those counties where the two methods create different sets of subcounty areas to which PEP subcounty estimates are applied in the weighting process.

4. Conclusions and Recommendations

We compared three alternatives methods for incorporating PEP subcounty population and housing unit estimates into the ACS weighting methodology: not using them, using only the housing unit estimates, and using both housing and population estimates. We also compared two different methods, ‘whole’ and ‘parts’, for forming subcounty areas to which weighting adjustments using the PEP estimates were applied. These two methods only differ in counties where there are both incorporated places and minor civil divisions that are not coexistent.

Using only housing unit PEP subcounty estimates in the weighting process resulted in small improvement in consistency between ACS and PEP estimates of population for subcounty areas and greater improvement in consistency for the estimates of total housing units. Using both housing and population PEP subcounty estimates in the weighting resulted in very large improvement to consistency for estimates of population. However, estimates of housing units were actually less consistent than not using PEP subcounty estimates in the weighting. Similar results were observed for variance estimates of population and housing characteristics. This likely to due to differences between housing and population coverage at the subcounty level. We believe that the negative effect on housing unit estimates can be overcome by modifying the incorporation of subcounty population estimates into the raking process. This hypothesis will be tested in our next research effort on subcounty estimates. Based on the positive results for consistency in the population estimates, we would still recommend using both housing and population PEP subcounty estimates if the negative effect on ACS housing estimates is an acceptable trade off.

References

Deming, W. and Stephan, F. (1940). "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known." *The Annals of Mathematical Statistics,* Vol. 11, No. 4, 427-444.

Fay, R (2006). ''Using Administrative Records With Model-Assisted Estimation for the American Community Survey.'' *JSM Proceedings, Survey Research Methods Section.* Alexandria, VA: American Statistical Association. 2995−3001.

Ostle, B. and Malone, L. (1988). *Statistics in Research, Fourth Edition.* Iowa State University Press, Ames, IA.

U.S. Census Bureau (2009). *ACS Design and Methodology,* Washington, DC: U.S. Department of Commerce. http://www.census.gov/acs/www/Downloads/dm1.pdf

Table 4. Distribution of Absolute Relative Percent Difference between ACS and PEP Subcounty Population Estimates.

Percentile	Large (N=64) Pop >250,000			Medium (N=479) Pop 65,000-249,000			Small (N=1,673) Pop <65,000		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
100th (max)	12.193	9.543	0.804	22.797	21.621	3.853	87.645	53.670	10.284
99th	12.193	9.543	0.804	13.493	12.382	2.552	28.714	15.095	4.419
95th	8.825	7.879	0.178	9.414	7.986	0.356	12.948	9.923	1.251
90th	6.557	5.862	0.078	7.135	6.130	0.123	9.274	7.986	0.230
75th	4.118	3.027	0.017	4.662	3.799	0.020	5.935	4.882	0.041
50th (median)	2.281	1.542	0.005	2.479	1.907	0.010	3.096	2.659	0.019
25th	0.613	0.747	0.002	1.173	0.864	0.005	1.372	1.223	0.009
10th	0.176	0.288	0.001	0.463	0.317	0.003	0.518	0.483	0.003
5th	0.078	0.169	0.001	0.227	0.167	0.001	0.256	0.288	0.002
1st	0.043	0.038	0.000	0.015	0.026	0.000	0.065	0.050	0.000
0th (min)	0.043	0.038	0.000	0.001	0.012	0.000	0.002	0.000	0.000
Friedman Test P-value	B1 vs B2 0.1336			B1 vs B2 <0.0001			B1 vs B2 <0.0001		
	B1 vs B3 <0.0001			B1 vs B3 <0.0001			B1 vs B3 <0.0001		
	B2 vs B3 <0.0001			B2 vs B3 <0.0001			B2 vs B3 <0.0001		

Table 5. Distribution of Absolute Relative Percent Difference between ACS and PEP Subcounty Housing Unit Estimates.

Percentile	Large (N=64) Pop >250,000			Medium (N=479) Pop 65,000-249,000			Small (N=1,673) Pop <65,000		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
100th (max)	7.990	2.729	9.537	13.057	6.317	14.545	109.404	21.415	110.108
99th	7.990	2.729	9.537	9.624	4.174	10.183	23.673	5.886	14.899
95th	4.133	1.652	5.742	5.873	2.478	7.303	9.189	3.388	9.064
90th	3.012	1.193	4.694	4.405	1.868	5.914	6.602	2.566	6.957
75th	1.674	0.640	2.510	2.613	1.026	3.499	3.669	1.452	4.489
50th (median)	1.014	0.369	1.082	1.290	0.577	1.806	1.842	0.757	2.533
25th	0.560	0.149	0.513	0.599	0.231	0.867	0.825	0.354	1.117
10th	0.137	0.051	0.148	0.221	0.094	0.333	0.293	0.153	0.504
5th	0.085	0.033	0.063	0.136	0.041	0.147	0.126	0.070	0.228
1st	0.041	0.001	0.008	0.036	0.013	0.046	0.033	0.018	0.057
0th (min)	0.041	0.001	0.008	0.002	0.000	0.000	0.000	0.000	0.000
Friedman Test P-values	B1 vs B2 <0.0001			B1 vs B2 <0.0001			B1 vs B2 <0.0001		
	B1 vs B3 0.3173			B1 vs B3 0.0002			B1 vs B3 <0.0001		
	B2 vs B3 <0.0001			B2 vs B3 <0.0001			B2 vs B3 <0.0001		

Table 6. Distribution of CVs for ACS Estimates of Total Population

Percentile	Large (N=64) Pop >250,000			Medium (N=479) Pop 65,000-249,000			Small (N=1,673) Pop <65,000		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
100th (max)	1.429	1.447	0.334	3.374	3.153	1.086	6.963	6.946	3.608
99th	1.429	1.447	0.334	2.970	2.863	0.749	4.851	4.778	2.364
95th	1.078	1.084	0.104	2.354	2.318	0.321	4.013	3.899	1.193
90th	1.007	0.973	0.077	2.140	2.094	0.115	3.594	3.528	0.266
75th	0.816	0.795	0.023	1.837	1.827	0.048	2.991	2.949	0.100
50th (median)	0.528	0.529	0.012	1.482	1.468	0.032	2.456	2.418	0.068
25th	0.383	0.369	0.008	1.199	1.172	0.024	2.000	1.981	0.051
10th	0.276	0.268	0.005	0.892	0.873	0.018	1.636	1.613	0.040
5th	0.219	0.209	0.004	0.729	0.707	0.015	1.464	1.443	0.035
1st	0.122	0.132	0.002	0.401	0.384	0.012	1.118	1.087	0.029
0th (min)	0.122	0.132	0.002	0.118	0.118	0.009	0.452	0.452	0.022
Friedman Test P-value	B1 vs B2 0.006			B1 vs B2 <0.0001			B1 vs B2 <0.0001		
	B1 vs B3 <0.0001			B1 vs B3 <0.0001			B1 vs B3 <0.0001		
	B2 vs B3 <0.0001			B2 vs B3 <0.0001			B2 vs B3 <0.0001		

Table 7. Distribution of CVs for ACS Estimates of Total Housing Units

Percentile	Large (N=64) Pop >250,000			Medium (N=479) Pop 65,000-249,000			Small (N=1,673) Pop <65,000		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
100th (max)	0.593	0.513	1.369	2.004	1.710	3.029	4.193	3.189	6.633
99th	0.593	0.513	1.369	1.532	1.334	2.586	2.893	2.443	4.511
95th	0.505	0.383	1.021	1.145	0.920	2.225	2.185	1.752	3.671
90th	0.416	0.342	0.837	1.006	0.790	2.011	1.924	1.454	3.274
75th	0.362	0.279	0.752	0.804	0.595	1.704	1.529	1.086	2.788
50th (median)	0.241	0.169	0.468	0.628	0.453	1.414	1.181	0.803	2.321
25th	0.169	0.117	0.318	0.489	0.351	1.103	0.899	0.598	1.908
10th	0.125	0.085	0.240	0.400	0.280	0.779	0.718	0.478	1.566
5th	0.106	0.066	0.178	0.351	0.242	0.633	0.611	0.425	1.359
1st	0.066	0.048	0.122	0.239	0.186	0.340	0.481	0.358	0.968
0th (min)	0.066	0.048	0.122	0.184	0.138	0.189	0.367	0.205	0.451
Friedman Test P-value	B1 vs B2 <0.0001			B1 vs B2 <0.0001			B1 vs B2 <0.0001		
	B1 vs B3 <0.0001			B1 vs B3 <0.0001			B1 vs B3 <0.0001		
	B2 vs B3 <0.0001			B2 vs B3 <0.0001			B2 vs B3 <0.0001		

Table 8. Distribution of Absolute Relative Percent Difference between ACS and PEP Subcounty Population Estimates in Counties where the ‘Whole’ and ‘Parts’ Methods Produce Different Results.

Percentile	Large (N=13) Pop >250,000		Medium (N=94) Pop 65,000-249,000		Small (N=382) Pop <65,000	
	A2	B2	A2	B2	A2	B2
100th (max)	8.794	8.777	13.207	9.149	38.424	17.358
99th	8.794	8.777	13.207	9.149	24.972	15.095
95th	8.794	8.777	9.215	6.560	13.383	9.751
90th	3.478	3.493	6.898	5.411	9.377	8.103
75th	2.810	2.846	4.421	3.827	5.916	5.206
50th (median)	2.341	2.352	2.787	2.272	3.255	2.898
25th	1.442	1.374	1.147	1.117	1.425	1.439
10th	0.786	0.610	0.430	0.469	0.584	0.598
5th	0.172	0.169	0.144	0.255	0.336	0.295
1st	0.172	0.169	0.052	0.040	0.026	0.056
0th (min)	0.172	0.169	0.052	0.040	0.007	0.000
Friedman Test P-value	0.7815		0.0488		0.0025	

Table 9. Distribution of Absolute Relative Percent Difference between ACS and PEP Subcounty Housing Unit Estimates in Counties where the 'Whole' and 'Parts' Methods Produce Different Results.

Percentile	Large (N=13) Pop >250,000			Medium (N=94) Pop 65,000-249,000			Small (N=382) Pop <65,000	
	A2	B2		A2	B2		A2	B2
100th (max)	1.698	1.652		10.341	3.854		29.568	8.175
99th	1.698	1.652		10.341	3.854		21.341	7.138
95th	1.698	1.652		5.657	2.014		6.767	3.873
90th	1.039	1.058		4.541	1.634		4.347	2.984
75th	0.831	0.827		1.787	1.072		2.483	1.642
50th (median)	0.556	0.560		0.871	0.609		1.025	0.770
25th	0.300	0.323		0.432	0.245		0.454	0.376
10th	0.037	0.101		0.191	0.101		0.178	0.160
5th	0.014	0.012		0.046	0.059		0.121	0.091
1st	0.014	0.012		0.015	0.014		0.018	0.027
0th (min)	0.014	0.012		0.015	0.014		0.004	0.014
Friedman Test P-value	0.7815			0.0006			0.0352	

Table 10. Distribution of Absolute Relative Percent Difference between ACS and PEP Subcounty Population Estimates in Counties where the 'Whole' and 'Parts' Methods Produce Different Results.

Percentile	Large (N=13) Pop >250,000			Medium (N=94) Pop 65,000-249,000			Small (N=382) Pop <65,000	
	A3	B3		A3	B3		A3	B3
100th (max)	0.420	0.452		13.093	3.853		38.647	10.284
99th	0.420	0.452		13.093	3.853		24.304	7.312
95th	0.420	0.452		6.731	1.519		8.760	3.250
90th	0.160	0.178		5.140	0.491		4.833	1.919
75th	0.017	0.026		1.498	0.076		2.186	0.357
50th (median)	0.005	0.005		0.407	0.010		0.046	0.033
25th	0.003	0.003		0.013	0.005		0.011	0.011
10th	0.002	0.001		0.004	0.003		0.004	0.005
5th	0.002	0.000		0.002	0.001		0.003	0.003
1st	0.002	0.000		0.001	0.000		0.000	0.000
0th (min)	0.002	0.000		0.001	0.000		0.000	0.000
Friedman Test P-value	0.7815			0.0006			0.2349	

Table 11. Distribution of Absolute Relative Percent Difference between ACS and PEP Subcounty Housing Unit Estimates in Counties where the 'Whole' and 'Parts' Methods Produce Different Results.

Percentile	Large (N=13) Pop >250,000		Medium (N=94) Pop 65,000-249,000		Small (N=382) Pop <65,000	
	A3	B3	A3	B3	A3	B3
100th (max)	5.607	5.622	10.040	9.125	29.718	20.077
99th	5.607	5.622	10.040	9.125	21.720	11.828
95th	5.607	5.622	6.668	6.101	10.740	8.643
90th	3.207	3.221	5.697	4.910	7.706	6.847
75th	2.193	2.187	3.830	3.360	4.997	4.958
50th (median)	1.204	1.225	2.375	1.850	2.769	2.836
25th	0.892	0.855	0.951	0.986	1.347	1.292
10th	0.611	0.585	0.498	0.524	0.623	0.569
5th	0.137	0.008	0.331	0.218	0.359	0.267
1st	0.137	0.008	0.026	0.100	0.079	0.022
0th (min)	0.137	0.008	0.026	0.100	0.011	0.000
Friedman Test P-value	0.7815		0.3023		0.2183	

Table 12. Friedman Test P-values for Comparing Median CVs of Population and Housing Characteristics

Group	Large (N=64) Pop >250,000			Medium (N=479) Pop 65,000-249,000			Small (N=1,673) Pop <65,000		
	B1 vs B2	B1 vs B3	B2 vs B3	B1 vs B2	B1 vs B3	B2 vs B3	B1 vs B2	B1 vs B3	B2 vs B3
Age	0.318	0.000	0.000	0.061	0.000	0.000	0.040	0.000	0.000
Sex	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Race	0.803	0.000	0.000	0.492	0.000	0.000	0.000	0.000	0.000
Relationship	0.318	0.000	0.000	0.819	0.000	0.000	0.002	0.000	0.000
Marital Status	0.803	0.046	0.134	0.001	0.000	0.000	0.000	0.000	0.000
Education	0.453	0.453	0.617	0.436	0.025	0.003	0.129	0.000	0.000
Occupation	0.617	0.001	0.006	0.049	0.000	0.000	0.000	0.000	0.000
Person per HH	0.803	0.001	0.006	0.131	0.031	0.061	0.303	0.170	0.141
Tenure	0.001	0.012	0.001	0.000	0.000	0.000	0.000	0.000	0.000
Household Inc.	0.803	0.318	0.453	0.000	0.000	0.000	0.000	0.000	0.000
Value	0.803	0.803	0.617	0.964	0.000	0.000	0.003	0.000	0.000
Vacancy Rate	0.617	0.000	0.000	0.964	0.000	0.000	0.094	0.000	0.000
Units in Structure	0.211	0.080	0.046	0.109	0.000	0.000	0.000	0.000	0.000