Developing Guidelines Based on CVs for when One-Year Estimates Can Be Used Instead of Three-Year Estimates in the American Community Survey (ACS) for Areas with Populations of 65,000 or More

Michael Ikeda, Statistical Research Division, U.S. Census Bureau

Abstract: A method is outlined for developing guidelines for acceptable use of 1-year estimates rather than 3-year estimates for areas with more than 65,000 total population in the American Community Survey (ACS). The method is based on the coefficents of variation of the estimates. It is applied to estimates from the 2006 ACS and the results are presented.

I. Introduction and Methodology

For areas with more than 65,000 total population both 1-year and 3-year estimates will soon be available. We want to develop simple guidelines, based on the coefficient of variation (CV), for when 1-year estimates can be used instead of 3-year estimates for these areas.

For estimates that are basically estimated counts of people, households, or housing units, we seem to be able to obtain reasonable guidelines by setting cutoffs based on the *percentage that the estimate is of the total*. For 1-year estimates, we can use the empirical bivariate distribution of "estimated percent of total" and "estimated CVs" to set cutoffs so that:
(a) most 1-year estimates which give percentages above the cutoff will meet a desired CV, and
(b) none or very few of these estimates will have CVs that exceed the desired CV by too much.
For a given target CV, if the percentage of total is greater than the cutoff, then the 1-year estimate can be used for any CV larger than the target; otherwise, the 3-year estimate can be used. We plan to divide the areas into several population categories and set separate guidelines for each category.

To assign cutoffs we will divide the "percent of total" distribution into ranges. For each target CV we find the "percent of total" range where the percentage of CVs greater than the target exceeds 10%. We then set the cutoff at the upper limit for that "percent of total" range.

For estimates that are not counts, we will subdivide by type of estimate (e.g. aggregate total, gini coefficient) and use the empirical CV distribution for each estimate type. We again expect to divide the areas into separate population categories and set separate guidelines for each category. However, there are some noncount estimates with high CVs even for very large areas, so any general guidelines would need to be used with caution.

II. Results

Table 1 gives cutoffs for single-year count estimates obtained by applying the methodology outlined above to estimates from the 2006 ACS. Cutoffs were calculated based on the minimum CV in each column. Estimates at the U.S., region, and division summary levels were excluded from the calculation, as were "components" within state. Tables 1a-1f expand the results in Table 1 to give explicit "Yes/No" recommendations for the use of single-year estimates for each population range. Based on feedback from Decennial Statistical Studies Division, Tables 1a-1f

would be the preferable format for illustrating the guidelines.

Cutoffs were also calculated for the combined category 65,000-300,000. The cutoffs for the combined category were generally between the cutoffs for the two subcategories. Cutoffs were also calculated separately for person estimates and household /housing unit estimates for each population category. The cutoffs for person estimates tend to be slightly lower and the cutoffs for household/housing unit estimates tend to be slightly higher than the overall cutoffs.

Table 1: Cutoffs for Percent of Total by Desired CV and Population Range for Single-Year Count Estimates

	Required Coefficient of Variation (CV)			
Population Range	5% to 9.9%	10% to 19.9%	20% to 34.9%	35% to 49.9%
65,000 to 150,000	75%	25%	7%	2.5%
150,001 to 300,000	40%	14%	3.5%	1.2%
300,001 to 1,000,000	16%	5%	1.2%	0.4%
1,000,001 to 3,000,000	7%	1.7%	0.4%	0.2%
3,000,001 to 15,000,000	2%	0.6%	0.2%	0.05%
15,000,001 to 40,000,000	0.6%	0.1%	0.04%	0.01%

Estimates for U.S., Region, and Division were not included in calculations.

Table 1a: Table of Guidelines on Use of One-Year Estimates for Geographies with 65,000 to 150,000 Population

	Required Coefficient of Variation (CV)			
Estimate as % of Total	5% to 9.9%	10% to 19.9%	20% to 34.9%	35% to 49.9%
less than 2.5%	No	No	No	No
2.5%-6.9%	No	No	No	Yes
7%-24.9%	No	No	Yes	Yes
25%-74.9%	No	Yes	Yes	Yes
75% and greater	Yes	Yes	Yes	Yes

Table 1b: Table of Guidelines on Use of One-Year Estimates for Geographies with 150,001 to 300,000 Population

	Required Coefficient of Variation (CV)			
Estimate as % of Total	5% to 9.9%	10% to 19.9%	20% to 34.9%	35% to 49.9%
less than 1.2%	No	No	No	No
1.2%-3.49%	No	No	No	Yes
3.5%-13.9%	No	No	Yes	Yes
14%-39.9%	No	Yes	Yes	Yes
40% and greater	Yes	Yes	Yes	Yes

Table 1c: Table of Guidelines on Use of One-Year Estimates for Geographies with 300,001 to 1,000,000 Population

	Required Coefficient of Variation (CV)			
Estimate as % of Total	5% to 9.9%	10% to 19.9%	20% to 34.9%	35% to 49.9%
less than 0.4%	No	No	No	No
0.4%-1.19%	No	No	No	Yes
1.2%-4.9%	No	No	Yes	Yes
5%-15.9%	No	Yes	Yes	Yes
16% and greater	Yes	Yes	Yes	Yes

Table 1d: Table of Guidelines on Use of One-Year Estimates for Geographies with 1,000,001 to 3,000,000 Population

	Required Coefficient of Variation (CV)			
Estimate as % of Total	5% to 9.9%	10% to 19.9%	20% to 34.9%	35% to 49.9%
less than 0.2%	No	No	No	No
0.2%-0.39%	No	No	No	Yes
0.4%-1.69%	No	No	Yes	Yes
1.7%-6.9%	No	Yes	Yes	Yes
7% and greater	Yes	Yes	Yes	Yes

Table 1e: Table of Guidelines on Use of One-Year Estimates for Geographies with 3,000,001 to 15,000,000 Population

	Required Coefficient of Variation (CV)			
Estimate as % of Total	5% to 9.9%	10% to 19.9%	20% to 34.9%	35% to 49.9%
less than 0.05%	No	No	No	No
0.05%-0.19%	No	No	No	Yes
0.2%-0.59%	No	No	Yes	Yes
0.6%-1.9%	No	Yes	Yes	Yes
2% and greater	Yes	Yes	Yes	Yes

Table 1f: Table of Guidelines on Use of One-Year Estimates for Geographies with 15,000,001 to 40,000,000 Population

	Required Coefficient of Variation (CV)			
Estimate as % of Total	5% to 9.9%	10% to 19.9%	20% to 34.9%	35% to 49.9%
less than 0.01%	No	No	No	No
0.01%-0.039%	No	No	No	Yes
0.04%-0.09%	No	No	Yes	Yes
0.1%-0.59%	No	Yes	Yes	Yes
0.6% and greater	Yes	Yes	Yes	Yes

To illustrate how these tables can be used, suppose a researcher is interested in characteristics for Lake County, Illinois. Lake County currently has an estimated population in the range 300,001 to 1,000,000 (644,356 in the 2000 Census) so Table 1c would be used. Suppose the researcher is satisfied if the CV is less than 35%. Looking at Table 1c, any population estimate expected to be greater than 1.2% of the total population will usually meet the CV cutoff. Thus, for example, if the researcher is interested in an estimate of the foreign born population (14.8% of the total according to the 2000 Census) the one-year ACS estimate can be used, while for estimates that make up a small percentage of the county population (e.g., persons of Lithuanian ancestry) the one-year estimates would not be recommended.

Similar reasoning would be used for estimates of households, except that we would use the percentage of the total number of households. Thus, again using Lake County as an example, if the researcher wants the CV to be less than 35% then the one-year ACS estimate can usually be used for household estimates that are expected to be more than 1.2% of the total number of households.

For the noncount estimates we use a somewhat analogous method. We again break the areas into

population size categories. However, because "percent of total" is not really relevant for noncount variables, we set a single CV cutoff for each type of estimate within each size category. We consider four CV cutoff points: 0.05, 0.10, 0.20, 0.35 (the lower boundary points of the CV ranges in Table 1). For each type of estimate within each population size category we find the smallest cutoff point for which the percentage of CVs greater than the cutoff point is less than 10%. That cutoff point is our CV cutoff for the given type of estimate and population size category. Table 2 lists the CV cutoffs. The one-year estimates are usually acceptable for any target CV above the cutoff. As mentioned above, however, the cutoffs should be used with caution since there are noncount estimates with large CVs even for areas with large populations (assuming we even want to provide cutoffs for noncount estimates). As the Aggregate, Median, and Ratio columns behave somewhat similarly, a combined column is also shown for comparison.

Separate calculations were also done by breaking the 65,000 to 300,000 category into two subcategories (65,000 to 150,000 and 150,001 to 300,000). Dividing into two subcategories has little effect on the results.

Table 2: Single-Year Noncount Estimates--CV cutoff by Type of Estimate and Population Range

	Type of Estimate				
Population Range	Aggregate	Median	Ratio	A/M/R Combined	Gini
65,000 to 300,000	---	---	35%	---	10%
300,001 to 1,000,000	35%	35%	20%	35%	5%
1,000,001 to 3,000,000	20%	20%	20%	20%	5%
3,000,001 to 15,000,000	20%	20%	10%	20%	5%
15,000,001 to 40,000,000	5 %	10%	5%	10%	5%

CV cutoff not listed if 50% or greater. Estimates for U.S., Region, and Division were not included in calculations.

Again using Lake County as an example, for medians the one-year estimates are usually acceptable if the desired CV is greater than 35%. Thus, if the researcher is interested in median household income and has a desired CV of 20% the one-year estimate probably should not be used.

(For the above Lake County example, it does not matter whether the CV cutoff is calculated using just the medians, or after combining aggregates, medians, and ratios into a single category. Presumably we would not actually do both.)