Case Study 1:

How to Deal with Estimates with Low Reliability

2009 Population Association of America
ACS Workshop
April 29, 2009

What is Reliability?

Sampling Error is the uncertainty associated with an estimate that is based on data gathered from a sample of the population rather than the full population

Measures of sampling error give users an idea of how reliable, or precise, estimates are and speak to their fitness-for-use

Measures of Sampling Error

- Standard Error (SE) – foundational measure of the variability of an estimate due to sampling
- Margin of Error (MOE) – precision of an estimate at a given level of confidence
- Confidence Interval (CI) - a range (based on a fixed level of confidence) that is expected to contain the population value of the characteristic
- Coefficient of Variation (CV) - The relative amount of sampling error associated with a sample estimate

Calculating Measures of Sampling Error

At a 90 percent confidence level

MOE = SE x 1.645

SE = MOE / 1.645

CI = Estimate +/- MOE

CV = SE / Estimate * 100%

ACS Displays Margins of Error

Data Profile

You are here: Main ▸ Data Sets ▸ Geography ▸ Results

| Print / Download | Related Items

2007 Data Profiles:

- **Social**
- Economic
- Housing
- Demographic
- Narrative

2007 Comparison Profile

View this table...

- **from 2007**
- from 2006
- from 2005
- from 2004
- from 2003
- from 2002

View this table...

for other geographies (state, county, place...)

- Subject Definitions
- Quality Measures

Washington city, District of Columbia

Selected Social Characteristics in the United States: 2007

Data Set: **2007 American Community Survey 1-Year Estimates**

Survey: **American Community Survey**

Social - Education, Marital Status, Relationships, Fertility, Grandparents...
Economic - Income, Employment, Occupation, Commuting to Work...
Housing - Occupancy and Structure, Housing Value and Costs, Utilities...
Demographic - Sex and Age, Race, Hispanic Origin, Housing Units...
Narrative - Text profile with graphs for easy analysis...

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

For more information on confidentiality protection, sampling error, nonsampling error, and definitions, see Survey Methodology.

Selected Social Characteristics in the United States	Estimate	Margin of Error	Percent	Margin of Error
HOUSEHOLDS BY TYPE				
Total households	**251,039**	**+/-2,911**	**100%**	**(X)**
Family households (families)	108,181	+/-3,697	43.1%	+/-1.5
With own children under 18 years	45,089	+/-3,034	18.0%	+/-1.2
Married-couple family	55,790	+/-2,895	22.2%	+/-1.2
With own children under 18 years	18,752	+/-2,175	7.5%	+/-0.9
Male householder, no wife present, family	9,434	+/-1,574	3.8%	+/-0.6
With own children under 18 years	3,418	+/-1,105	1.4%	+/-0.4
Female householder, no husband present, family	42,957	+/-2,890	17.1%	+/-1.1
With own children under 18 years	22,919	+/-2,220	9.1%	+/-0.9

Example 1 – Calculating Sampling Errors

2007 ACS 1-year estimates for Washington, DC

Estimate of the percent of married couple families = 22.2% with a MOE of 1.2%

SE = MOE/1.645 = 1.2% / 1.645 = 0.729%

CI = Estimate +/- MOE = 22.2% +/- 1.2%
= 21.0% to 23.4%

CV = SE/Estimate * 100% = 0.729% / 22.2% * 100% = 3.28%

Interpreting Coefficients of Variation

CVs are a standardized indicator of reliability that tell us the relative amount of sampling error in the estimate

Estimates with CVs that are less than 15% are generally considered reliable, while estimates with CVs that are greater than 30% are generally considered unreliable

Distinguishing Between Reliable and Unreliable Estimates

There are no specific rules about acceptable levels of sampling error – the classification as “reliable” will vary based on the application

Some estimates warrant greater precision than others due to the consequences of their use

Reliability should always be considered when making comparisons

Example 2 – Assessing Utility

A mayor of a small town can receive funding to support a language program if the proportion of the population speaking Vietnamese exceeds 5 percent.

The 2007 ACS 1-year estimates shows the rate to be 1.2% with a MOE of 1.1%.

The CV of this estimate is over 50% and the estimates would be deemed unreliable, but the mayor can with confidence conclude that the Vietnamese-speaking population is less than 5%.

Example 3 – Assessing Utility

Officials in Savannah city, GA, are considering an outreach program to the foreign-born population of the city using the public transportation system as advertising. Officials need to know how many foreign-born people use public transportation.

What do the 2007 ACS 1-year estimates show?

C08111. MEANS OF TRANSPORTATION TO WORK BY CITIZENSHIP STATUS - Universe: WORKERS 16 YEARS AND OVER
Data Set: 2007 American Community Survey 1-Year Estimates
Survey: American Community Survey

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is
Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states,
estimates of housing units for states and counties.

For information on confidentiality protection, sampling error, nonsampling error, and definitions, see Survey Methodology.

	Savannah city, Georgia	
	Estimate	Margin of Error
Total:	50,526	+/-3,821
Native		
Foreign-born:		
Naturalized U.S. c		
Not a U.S. citizen		
Car, truck, or van - dr		
Native		
Foreign-born:		
Naturalized U.S. c		
Not a U.S. citizen		
Car, truck, or van - ca		
Native		
Foreign-born:		
Naturalized U.S. c		
Not a U.S. citize	307	+/-257
Public transportation (excluding taxicab):	2,075	+/-892
Nativ	1,846	+/-813
Foreign-born:	229	+/-360
Naturalized U.S. citizen	43	+/-70
Not a U.S. citizen	186	+/-294
Taxicab, motorcycle, bicycle, walked, or other means:	2,935	+/-867
Native	2,897	+/-863
Foreign-born:	38	+/-70

Not a U.S. citizen	307	+/-257
Public transportation (excluding taxicab):	2,075	+/-892
Native	1,846	+/-813
Foreign-born:	229	+/-360
Naturalized U.S. citizen	43	+/-70
Not a U.S. citizen	186	+/-294
Taxicab, motorcycle, bicycle, walked, or other means:	2,935	+/-867
Native	2,897	+/-863

Example 3 – Assessing Utility

The 2007 ACS 1-year estimate of the foreign-born using public transportation is 229 with a MOE of +/- 360. This indicates a confidence interval of 0 to 589 and a CV of over 95%.

This is a highly unreliable estimate and shouldn't be used alone in an application such as this.

Example 4 – What to do with unreliable estimates

Officials in Cook County, IL are looking to improve the quality of life for the elderly population by identifying sub county areas with people over 65 who are poor or near poor.

An analyst finds a detailed table (B17024) from the 2007 ACS 1-year estimates that includes poverty data by age, providing a detailed series of income-to-poverty ratios.

Example 4 – Detailed Table

5.00 and over	2,432	+/-829
65 to 74 years:	6,198	+/-1,114
Under .50	0	+/-265
.50 to .74	355	+/-390
.75 to .99	0	+/-265
1.00 to 1.24	172	+/-215
1.25 to 1.49	620	+/-443
1.50 to 1.74	758	+/-527
1.75 to 1.84	0	+/-265
1.85 to 1.99	260	+/-254
2.00 to 2.99	799	+/-426
3.00 to 3.99	811	+/-325
4.00 to 4.99	888	+/-497
5.00 and over	1,535	+/-592
75 years and over:	3,963	+/-938
Under .50	155	+/-185
.50 to .74	93	+/-155
.75 to .99	222	+/-219
1.00 to 1.24	116	+/-112
1.25 to 1.49	403	+/-314
1.50 to 1.74	215	+/-212
1.75 to 1.84	186	+/-180
1.85 to 1.99	173	+/-180
2.00 to 2.99	671	+/-264
3.00 to 3.99	787	+/-534
4.00 to 4.99	232	+/-216
5.00 and over	711	+/-388

Source: U.S. Census Bureau, 2007 American Community Survey

In this table (B17024), data are available separately for people 65-74 years and 75 years and over and for 12 income-to-poverty ratios

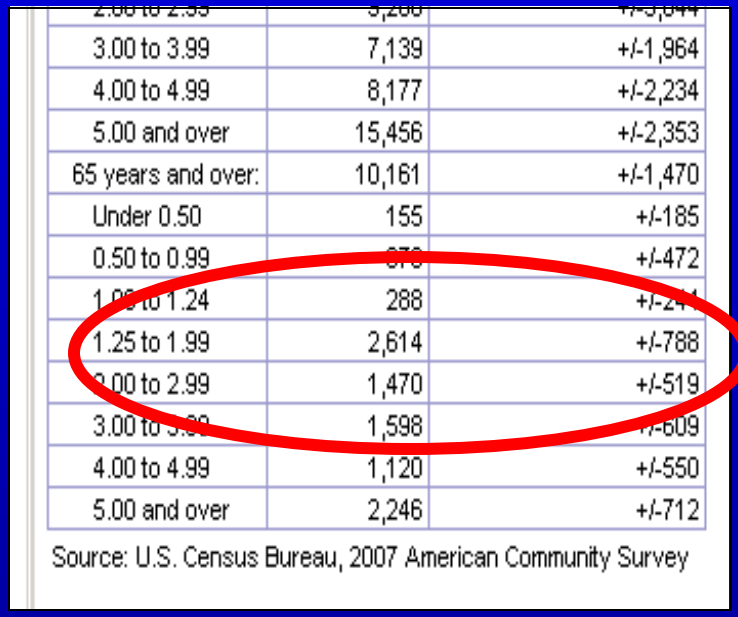
CVs are high – for example, the estimate of 403 persons 75 and over with a ratio of 1.25 to 1.49, has a MOE of 314 and a CV of 47.4%

Option 1
Consider the collapsed version of a table

You will find two versions of most detailed tables – one with full detail and another with detailed cells that have been collapsed

Collapsed tables include fewer estimates that are usually more reliable

Option 1
Check out the collapsed version of this table



3.00 to 3.99	7,139	+/-1,964
4.00 to 4.99	8,177	+/-2,234
5.00 and over	15,456	+/-2,353
65 years and over:	10,161	+/-1,470
Under 0.50	155	+/-185
0.50 to 0.99	[illegible]	+/-472
[illegible] to 1.24	288	[illegible]
1.25 to 1.99	2,614	+/-788
[illegible]00 to 2.99	1,470	+/-519
3.00 to [illegible]	1,598	[illegible]-609
4.00 to 4.99	1,120	+/-550
5.00 and over	2,246	+/-712

Source: U.S. Census Bureau, 2007 American Community Survey

In Table C17024 the two elderly age groups are combined and the 12 detailed income-to-poverty ratios are collapsed into 8 ratios

CVs are still high, but better; for example, the CV for the estimate of persons 65 and over with a ratio of 1.25 to 1.99 is 18.3%

Option 2
Consider additional collapsing of detail

In our example, we don't need the detail in the collapsed table. It is sufficient to identify the "poor and near poor" as including all people with an income-to-poverty ratio of less than 2.0.

We can collapse 4 detailed categories – under 0.5, 0.50 to 0.99, 1.00 to 1.24, and 1.25 to 1.99 to create a new category of "Under 2.00"

Option 2
Consider additional collapsing of detail

While summing estimates of people in poverty across four income-to-poverty ratios provides the combined estimate, summing MOEs will not produce the correct MOE.

The MOE of an aggregate estimate is determined by obtaining each component estimate's MOE, squaring it, summing these, and taking the square root of that sum.

Option 2 - Calculations

Age by Ratio of Income to Poverty Level – Bloom Township, Cook County, IL				
65 years and over	Estimate	MOE	MOE ^{2}	Square root of sum
Under 0.50	155	185		
0.50 to 0.99	670	472		
1.00 to 1.24	288	244		
1.25 to 1.99	2,614	788		
Under 2.00				

Source: 2007 ACS 1-year Estimates, Table C17024

Option 2 - Calculations

Age by Ratio of Income to Poverty Level – Bloom Township, Cook County, IL				
65 years and over	Estimate	MOE	MOE ^{2}	Square root of sum
Under 0.50	155	185	34,225	
0.50 to 0.99	670	472	222,784	
1.00 to 1.24	288	244	59,536	
1.25 to 1.99	2,614	788	620,944	
Under 2.00	3,727		937,489	968

Source: 2007 ACS 1-year Estimates, Table C17024

Option 2 - Results

Age by Ratio of Income to Poverty Level – Bloom Township, Cook County, IL				
65 years and over	Estimate	MOE	SE	CV
Under 0.50	155	185	112	72.6%
0.50 to 0.99	670	472	287	42.8%
1.00 to 1.24	288	244	148	51.5%
1.25 to 1.99	2,614	788	479	18.3%
Under 2.00	3,727	968	589	15.8%

Source: 2007 ACS 1-year Estimates, Table C17024

Option 2
Summary

The analyst should probably not directly use the estimates for each of the four income-to-poverty ratios to guide program planning (the CVs are very high for all but the last estimate)

Collapsing the four detailed ratios into one ratio with less detail results in a more reliable estimate

Option 3
Consider combining geographic areas

In our example, Bloom township is one sub county area in Cook County. It has two neighboring townships – Rich and Thornton

If the geographic detail isn't critical, estimates for these 3 areas could be combined

Option 3 - Calculations

Age by Ratio of Income to Poverty Level – Bloom, Rich, and Thornton Townships, Cook County, IL				
65 years and over	Estimate	MOE	MOE^2	Square root of sum
Under 0.50				
Bloom	155	185		
Rich	427	435		
Thornton	671	479		
Combined				

Source: 2007 ACS 1-year Estimates, Table C17024

Option 3 - Calculations

Age by Ratio of Income to Poverty Level – Bloom, Rich, and Thornton Townships, Cook County, IL				
65 years and over	Estimate	MOE	MOE^2	Square root of sum
Under 0.50				
Bloom	155	185	34,225	
Rich	427	435	189,225	
Thornton	671	479	229,441	
Combined	1,253		452,891	673

Source: 2007 ACS 1-year Estimates, Table C17024

Option 3 - Results

Age by Ratio of Income to Poverty Level – Bloom, Rich, and Thornton Townships, Cook County, IL				
65 years and over	Estimate	MOE	SE	CV
Under 0.50	1,253	673	409	32.6%
0.50 to 0.99	2,609	839	510	19.5%
1.00 to 1.24	2,684	1021	621	23.1%
1.25 to 1.99	6,635	1279	777	11.7%

Source: 2007 ACS 1-year Estimates, Table C17024

Option 3 - Calculations

Age by Ratio of Income to Poverty Level – Bloom, Rich, and Thornton Townships, Cook County, IL				
65 years and over	Estimate	MOE	MOE^2	Square root of sum
Under 0.50	1,253	673	452,929	
0.50 to 0.99	2,609	839	703,921	
1.00 to 1.24	2,684	1,021	1,042,441	
1.25 to 1.99	6,635	1,279	1,635,841	
Under 2.00	13,181		3,835,132	1,958

Source: 2007 ACS 1-year Estimates, Table C17024

Option 3 - Results

Age by Ratio of Income to Poverty Level – Bloom, Rich, and Thornton Townships, Cook County, IL				
65 years and over	Estimate	MOE	SE	CV
Under 0.50	1,253	673	409	32.6%
0.50 to 0.99	2,609	839	510	19.5%
1.00 to 1.24	2,684	1021	621	23.1%
1.25 to 1.99	6,635	1279	777	11.7%
Under 2.00	13,181	1958	1190	9.0%

Source: 2007 ACS 1-year Estimates, Table C17024

Option 3
Summary

Combining data for 3 neighboring areas improved the reliability of the detailed poverty data; collapsing this detail improved the estimate even more

Users need to consider the most important dimensions – geography or characteristic detail when considering collapsing

If both are critical, consider option 4

Option 4
Consider Multiyear Estimates

This will be covered in the next two case studies

Summary
Extrapolation to Large Data Sets

While these case studies referenced the use of a single set of estimates for a limited number of geographic areas, the underlying logic applies to analysts working with large data sets covering many areas

Be aware of the reliability limitations of the data before conducting your analyses, consider options to access or create more reliable estimates

What have we learned about dealing with ACS estimates with low reliability?

You should review the collapsed version of a detailed table to see if the collapsed values are sufficient for your needs

You can improve the reliability of ACS estimates by collapsing characteristic detail or combining geographies

Contact

Debbie Griffin
U.S. Census Bureau
deborah.h.griffin@census.gov