Best Practices for Reporting American Community Survey in Municipal Planning

Chittenden County Regional Planning Commission, October 2018

Planners rely on various sources of data to understand the existing and changing conditions of a community's housing stock, economy, and demographics. These data are critical not only for telling the story of a community but also for informing policy change and implementation trajectories. The most extensive collection of data which planners turn to for developing comprehensive plans is the American Community Survey (ACS).

As stated in its name, the ACS is a survey that is conducted over time to provide current information about social and economic conditions of a community. **The ACS is NOT an official count of the population nor is it a point in time count.** The ACS provides an estimate of data on various topics and associated information on the reliability of the survey. Therefore, all ACS data topics include a margin of error (MOE) and such variation should be understood fully in planning practice before data is used to inform decision making. As the Journal of American Planning Association (JAPA) indicates "Since ACS data are estimates rather than actual counts, they contain a degree of statistical uncertainty—referred to as margin of error (MOE)—that planners must navigate when using these data" (Jurjevich et al., 2018).

To assist planners in using ACS data properly, JAPA recommends 5 guidelines when using ACS Data.

- Understand statistical reliability
- Report margin of errors
- Indicate when margins of error are not reported
- Consider alternatives for reducing statistical uncertainty
- Conduct statistical testing when comparing ACS data

This guide is intended to be an overview of appropriate guidelines to follow when using ACS Data. The level of effort needed to implement these best practices is high. If planners are not able to fully adhere to these practices it is recommend that any use of ACS data is accompanied by a disclaimer alerting the reader to the margin of error associated with the data and that statistical certainty and significance has not been tested. Additionally, the Vermont Housing Finance Agency (VHFA) is launching a new interactive data visualization tool that will include statistical certainty estimates to ensure proper use of ACS data.

Choosing ACS Data

- Context: Before applying the 5 guidelines listed above, planners need to understand the type of data available from the ACS. In choosing the appropriate data, think about the population size of your area because ACS data availability is based on the size of the population. For example, geographic areas with smaller populations sample too few housing units to provide reliable single-year estimates. The ACS provides period estimates derived from a sample collected over a period of time. The ACS offers 1-year estimates which includes data collected from a 12-month

period and 5-year estimates collected over a 60-month period. Each period estimate is associated with distinguishing features regarding availability, currency, and precision.

✓ Practice: Planners should use the table below to understand the differences between each ACS period estimate and choose the appropriate one for their needs considering a balance between currency, sample size, and reliability.

Table 1 Distinguishing Features of ACS Data

1-Year Estimate	3-Year estimates*	5-Year Estimates
12-months of collected data	36 months of collected data	60 months of collected data
Data for areas with populations of 65,000 and more	Data for areas with populations of 20,000 and more	Data for all areas
Smallest sample size	Larger sample size than 1-year	Largest sample size
Less reliable than 3-year or 5-year	More reliable than 1-year; less reliable than 5-year	Most reliable
Most current	Less current than 1-year; more current than 5-year	Least current
*** ACS 3-year estimates have been discontinued. ACS 3-year estimates will remain available to data users, but no new 3-year estimates will be produced.**		

Margin of Error

❖ Context: ACS estimates include a Margin of Error (MOE) at the 90% confidence level, which is the Census Bureau's standard. Higher confidence intervals can be calculated, if desired. The MOE measures the possible variation of the estimate around the population value. For example, as seen in the table below, the population of homes heated with utility gas is 37,073 (+/-1,650). With a 90% confidence level, there is a 90% chance that the true number of homes heated with utility gas lies somewhere within the range of 35,423 to 38,723 homes.

Example 1 – Margins of Error in ACS Data

	Estimate	Margin of Error
Total:	64,867	+/-1,151
Utility gas	37,073	+/-1,650
Bottled, tank, or LP gas	7,218	+/-1,119
Electricity	6,085	+/-1,286
Fuel oil, kerosene, etc.	9,751	+/-1,404
Coal or coke	0	+/-151
Wood	3,558	+/-752
Solar energy	166	+/-173
Other fuel	709	+/-406
No fuel used	307	+/-164

✓ Practice: Planners should first review the MOE against the estimate to determine reliability. If the MOE is larger than the estimate itself exercise caution with its use because it could signal that the sample size is too small to be reliable. For example, in block group 1 the 2012-2016 ACS 5-Year Estimate for the population of females under 5 years old is 28 with a MOE of +/- 33. The MOE for this population is higher than the estimate itself so data is not reliable, and the data should not be used.

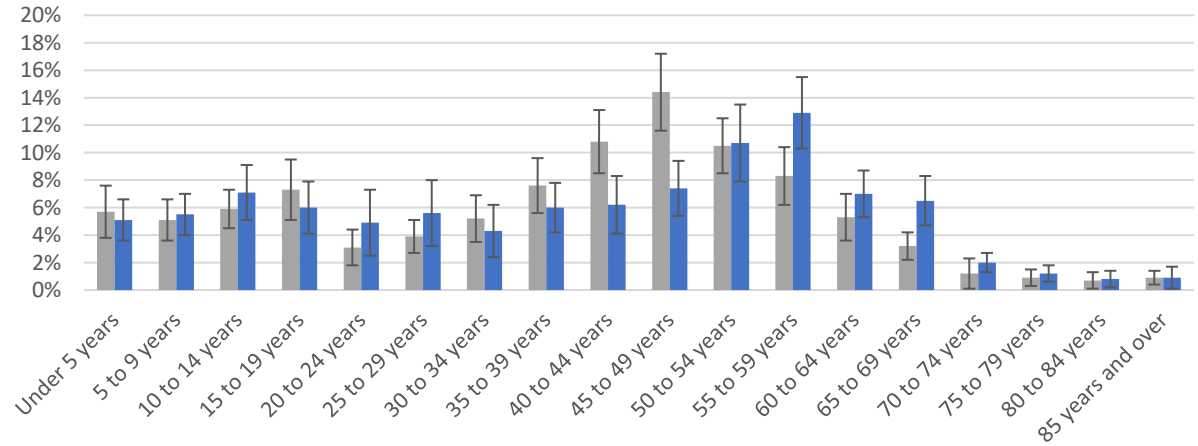
After comparing the MOE with the estimate, planners should include MOEs to alert readers of the variation that is possible with the ACS data. The MOE could be shown as another column in a table of raw values or as error bars in a graph (shown below). The error bars on a graph may be more desirable to achieve the look of a more concise and streamlined plan. If the margins of error are not reported, ensure that a disclaimer is included with the data notifying the reader as such.

Example 2 – Reporting Margins of Error in a Table

		2011		2016	
		Estimate	Margin of Error	Estimate	Margin of Error
Huntington	Proportion of the Population over 60	12.30%	+/- 2.30%	18.40%	+/- 3%
	Proportion of the Population over 65	6.90%	+/-1.50%	11.40%	+/- 2.50%
Chittenden County	Proportion of the Population over 60	16.30%	+/- 0.30%	19.00%	+/- 0.3%
	Proportion of the Population over 65	11.20%	+/- 0.10%	13.20%	+/- 0.01%

Source: 2012-2016 and 2007-2011 American Community Survey 5-Year Estimates (Table S0101: Age and Sex)

Example 3 – Reporting Margins of Error on a Graph



Source: 2012-2016 and 2007-2011 American Community Survey 5-Year Estimates (Table S0101: Age and Sex)

Statistical Reliability

- Context: ACS data users and planners should understand the statistical reliability or quality of the ACS data to determine the appropriate use for the data. This is especially important for small areas like census tracts, blocks, and municipalities with small populations.

- Practice: Planners should use the coefficient of variation (CV) test to assess fitness for use in planning efforts. The CV is a relative measure of uncertainty and expresses uncertainty as a percentage of the ACS estimate (Jurjevich et al., 2018). The larger the CV ratio, the less accurate the estimate. While various sources indicate there is no CV ratio that is universally accepted as "too high", generally, a CV of less than fifteen percent indicates a high reliability estimate. Data users can compute a CV directly from the MOE.

  - High Reliability – CV Less than 15%
  - Medium Reliability – CV between 15% and 20%
  - Low Reliability – CV > 30%, use with extreme caution

  To calculate a CV (at a 90% confidence level), use the following equation:

$$CV = \frac{Margin\ of\ Error/1.645}{Estimate} \times 100$$

  For example, if you have an estimate of 80 +/- 20, the CV for the estimate is 15.2%.

$$CV = \frac{20/1.645}{80} \times 100 = 15.2\%$$

  The sampling error represents slightly more than 15% of the estimate. The level of reliability should be footnoted according to the thresholds above. For this example, the level of reliability is medium.

Statistical Testing for ACS Data Comparison

- Context: Comparing ACS estimates involves more than just determining which estimate is higher or lower. Simply comparing an estimate from one year to another does not provide a reliable comparison. Statistical testing is needed to determine if a difference is unlikely to occur by chance known as statistical difference. To be statistically different, there must be statistical evidence that there is a difference between two estimates. Testing should be conducted for all comparisons of ACS data. Testing is done through a generic Z-score formula. Use the values shown in table 5 to follow the instructions below to test for significance[1].

  1) Determine the standard error (SE) for each estimate. The SE is defined by the positive value of the MOE divided by 1.645.
  2) Square the resulting SE for estimate
  3) Sum the squared SEs
  4) Calculate the square root of the sum of the squared SEs

[1] U.S. Census Bureau, A Compass for Understanding and Using American Community Survey Data

5) Calculate the difference between the two estimates
6) Divide (5) by (4)
7) Compare the absolute value of the result of (6) with the critical value for the desired level of confidence (1.645 for 90 percent, the ACS standard).
8) If the absolute value of (6) is greater than the critical value, then the difference between the two estimates can be considered statistically significant at the corresponding confidence level.

In equation form:

$$If \left| \frac{Est_1 - Est_2}{\sqrt{\left(\frac{MOE_1}{1.645}\right)^2 + \left(\frac{MOE_2}{1.645}\right)^2}} \right| > 1.645,$$

the two estimates can be considered statistically different at a 90% confidence interval.

✓ Practice: Planners could consider determining statistical significance using the instructions above or by using the Census's Bureau's Statistical Testing Tool. This tool is an easy way to conduct statistical testing and determine if estimates are statistically different (higher or lower from each other). ACS data can be copied or downloaded into the spreadsheet to get instant results of statistical tests. Statistical testing can be used for any type of estimate (count, percent, median, etc.) and for comparing single years and **non-overlapping multi-year periods** (example: compare 2005-2009 ACS 5-year estimates to 2010-2014 ACS 5-year estimates). It is important to note that one should never compare overlapping multi-year periods. The statistical difference between time periods should be indicated when comparing data from two time periods (see example below).

Example 4 – Calculating Statistical Difference Between Two Time Periods

	2007-2011		2012-2016		Statistically Different?
	Estimate	**MOE**	**Estimate**	**MOE**	
Population	5,008	+/-23	5,052	+/-26	Yes

Consider Alternatives for Reducing Statistical Un-Reliability

❖ Context: Many small area geographies like tracts, block groups, and municipalities with small populations will have estimates with larger MOEs. Caution should be exercised when using these estimates because the reliability is questionable. Large MOEs and CVs of less than 15% can signal that the data are not accurate, often due to small sample sizes. To work around this, consider alternatives for reducing statistical unreliability.

✓ Practice: If data for small area geographies has poor reliability, consider using a larger geographic area. For example, use a county estimate instead of a town-level estimate, along with additional qualitative information known about the small area to bring context to the data

trends. Another method is to combine estimates across two small areas of geographies to form a larger estimate. If this is done, the MOE would need to be estimated for this larger geography.

To calculate estimates from two small areas to form a larger area:

1. Sum the estimate for the two small areas
2. Use the equation below to approximate the MOE:

$$MOE(Sum) = \sqrt{MOE_{Est1} + MOE_{Est2} \ldots}$$

3. Estimate the CV to determine reliability

Conclusion

Before including ACS data into plans, planners should apply the best practices described in this guide to determine the appropriate use of the data. A level of comfort should be gained from integrating statistical testing into the use of ACS data. However, if planners become less comfortable with the quality of the data because of statistical testing, planners could choose to not report the data with a note indicating that data on certain issues are not available due to reliability issues. Planners who use the ACS data should identify ways to integrate and communicate statistical uncertainty into their plans in a way that can be easily understood so that the underlying message of the data is not lost. While easier said than done, one final example is below.

Example 5 – Reporting on Data Reliability in a Town Plan

Most Huntington residents drive to work alone. American Community Survey data estimate that 812 workers drove to work alone in 2016. There was no statistical change between 2011 and 2016. No reliable data are available on the mode of transportation used by those workers who do not drive alone to work.[1] Though people working at home are not reflected in data about travel to work, anecdotal local evidence and national economic trends indicate that this is becoming more popular.

Figure – Transportation to Work, Huntington

Source: 2012-2016 American Community Survey 5-Year Estimates (Table S0802: Means of Transportation to Work by Selected Characteristics)

[1]Because of Huntington's small population, ACS estimates may have high margins of error. This plan uses the coefficient of variation (CV) for data points to determine whether margins of error are too high for a data point to be reliable. When the CV is below 15%, data are considered highly reliable, and these estimates are presented without caveat. When the CV is between 15% and 30%, the plan indicates that the data should be used with caution. Data with a CV over 30% are not reported to avoid unreliability. Methodology drawn from Jurjevich et al. "Navigating Statistical Uncertainty." Journal of the American Planning Association 84, no. 2 (Spring 2018): 112-126. Year-to-year changes between ACS estimates are statistically different unless otherwise reported.

References

Fuller, S. (n.d.). Using American Community Survey (ACS) Estimates and Margins of Error.

Jurjevich, J. R., Griffin, A. L., Spielman, S. E., Folch, D. C., & Meg Merrick, a. N. (2018). Navigating Statistical Uncertainity. *Journal of the American Planning Association*, 112-126.

U.S. Census Bureau, A. C. (n.d.). *https://www.census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018_ch03.pdf*.