

American Community Survey (ACS): Data, Usage, & Guidelines

- Xuehao Chu
- Center for Urban Transportation Research



Outline

- The American Community Survey (ACS)
- Data
- Usage
- Precision
- Guidelines
- The ACS Statistical Analyzer



What is the American Community Survey (ACS)?

- Designed & administered by the U.S. Census Bureau
- Instrument is similar to the Census long form
 - Socio-demographic and housing
 - Economic and journey-to-work
- Administration is different from the Census long-form
 - Residence rule - current vs. usual
 - Frequency - annual vs. every 10 years
 - Measurement - continuous vs. snapshot (April 1st)
 - Sample size - 1 in 40 (2.5%) vs. 1 in 6 (16.7%)



What data are available?

- Type of data products
 - Pre-derived estimates
 - Sub-samples
- Pre-derived estimates
 - Published tables (at American FactFinder)
 - ACS Summary File (at ACS FTP site)
 - CTPP tables (FHWA's CTPP site)
- Sub-samples
 - 1% PUMS by state (at American FactFinder)
- Period of data products
 - 1-year, 3-year, 5-year



What forms do the estimates take?

- Frequencies (i.e., persons, households, housing units)
- Totals (i.e., total income, commute time, etc.)
- Averages
- Medians
- Ratios
- Percentages

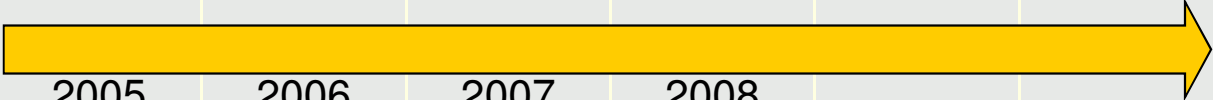
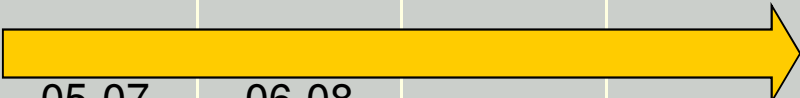
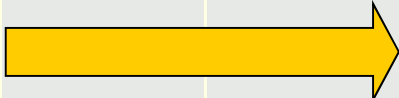


How do published estimates depend on their period and geography?

Population in the Geography of Estimates	Period of Estimates		
	1-Year	3-Year	5-Year
Large: 65,000+	X	X	X
Medium: 20,000+		X	X
Small: Under 20,000			X



When are published tables released?

Period of Data Products	Previous Year's Data Released in the Summer of					
	2006	2007	2008	2009	2010	2011
1-year						
3-year						
5-year						



What the data may be used for?


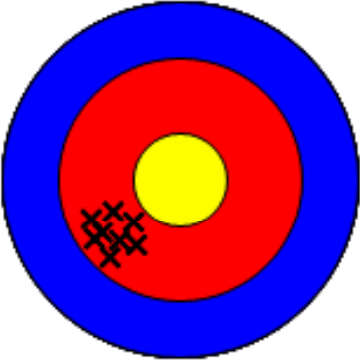

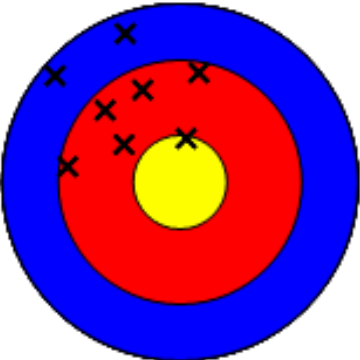
- Indicating conditions of a geography or a population group using individual estimates, e.g.:
 - Average commuting time
 - Median household income
 - % workers commuting to work by transit
- Indicating differences by comparing estimates
 - Over time
 - Across population groups
 - Across geographic areas
- Input data for multivariate analyses
- Input data for modeling



What are the critical issues in using ACS data?

- Recognize the nature of ACS estimates
 - They show average characteristics during a period
 - They show characteristics not totals
 - Control totals from Population Estimates Program
 - Non-institutional totals controlled at county level by sex, age, race, ethnicity
 - Institutional totals controlled at state level
 - Control totals reflect July 1st conditions
 - Totals \neq Census counts for 2010
- Consider the accuracy & precision of ACS estimates
- Impacts are most significant for
 - Smaller areas, especially those with rapid changes

What must be considered in using ACS Data?

	High Accuracy	Low Accuracy
High Precision		
Low Precision		



What guidance does the U.S. Census Bureau has?

- "As the ACS estimates are based on a sample survey of the U.S. population, information about the sampling error [i.e., precision] associated with the estimates must be taken into account when analyzing individual estimates or comparing pairs of estimates across areas, population groups, or time periods."



What can be used to measure precision?

- Measures of precision, independent of confidence
 - Standard Error (SE)
 - Coefficient of Variation (CV)

$$CV = 100 * SE / Estimate$$

- Measures of precision, dependent on confidence
 - Margin of Error (MOE)

$$MOE_{90\%} = SE * 1.645$$

- Confidence Interval (CI) =

$$(Estimate - MOE_{90\%}, Estimate + MOE_{90\%})$$



How can measures of precision be obtained?

- Published ACS estimates come with an MOE but not other measures of precision
- Most CTPP ACS estimates come with an MOE but not other measures of precision
- Users need to derive measures of precision
 - Census 2000 published
 - CTPP 2000 estimates
 - Census 2000 PUMS data
 - User-derived estimates from published ACS estimates
 - User-derived estimates from an ACS PUMS



Are my estimates usable?

- MOE is an absolute indicator of precision & ineffective
- CV is a relative indicator & effective
- General guidelines for using CV
 - 10% or less is desirable
 - Cautious to use if $10\% < CV \leq 50\%$
 - Avoid if $CV > 50\%$
- Consider the importance of an estimate
 - Use lower CV for more important usage (e.g., funding)



What strategies may be used to improve precision?

- For a given period
 - Combining smaller geographies into a larger one
 - Combining smaller population groups into a larger one
- For a given geography or population group
 - Using estimates for a longer period



What are some guidelines for using individual estimates?

- Trade off between currency and precision
- Use single-year estimates for
 - Larger geographies
 - Larger population groups
 - Examining year to year changes if CV is small
 - Geographies experiencing rapid changes
- Use multiyear estimates
 - When single-year estimates have a large CV
 - For smaller geographies
 - For smaller population groups in larger geographies



What are some guidelines for comparisons?

- Comparing ACS estimates
 - Avoid overlapping areas or population groups in a period
 - Avoid 1-year period with multiyear period
 - Avoid two overlapping multiyear periods
 - Avoid pre-2006 and post-2006 (including 2006)
- In comparing ACS and 2000 estimates, consider:
 - Effects of different residence rules
 - Effects of different reference periods
 - Effects of seasonal variation
- Use higher confidence for comparisons if
 - The conclusion is important (e.g., funding related)
 - Multiple pairs are compared



How can two estimates be compared statistically?

- Take into account the precision of estimates
- Conduct a test of significance
- Select a confidence level for the test
- Calculate a test statistic based on the estimates and their precision levels
- Determine the critical value for the selected confidence
- If test statistic $>$ critical value, the difference is statistically significant at the selected confidence level
- If test statistic \leq critical value, the difference is not statistically significant



How does the *ACS Statistical Analyzer* help?

- It derives measures of precision
- It tests if the difference between two estimates is statistically significant at a given confidence level
- It does these for six different forms of estimates
- It focuses on ACS but deals with Census 2000 too
- It is a 2003 Excel-based template
- It has 4 functions (A, B, C, D)
- It has 15 sub-functions
- Users just need to enter the required data
- It takes care of the statistical procedures and formulas involved



What are the four functions?

- Focuses on published ACS estimates at American FactFinder
- Derives measures of precision for estimates without any measure of precision
- Derives measures of precision for estimates obtained from two or more other estimates with an MOE
- Tests whether a difference between two estimates is statistically significant at a given confidence level



Summary

- Both accuracy and precision are important
- Users have some control on precision
- Precision affects the selection of period & geography
- Precision determines usability and comparisons
- The *ACS Statistical Analyzer* takes care of the statistical procedures and formulas
- Focus on ACS estimates but cover 2000 estimates
- Allows six forms of estimates

