American Community Survey Statistical Topics

Michael Beaghen and Daniel Sommers
U.S. Census Bureau
Washington, DC 20233

Prepared for the January 11, 2011 Cooperative Forecasting and Data Subcommittee Meeting of the Metropolitan Washington Council of Governments

Overview

- Use of independent estimates as controls for subcounty areas
- Break in series due to transition in the base for the controls
- Sample expansion and reallocation

Background – Subcounty Controls

- Two estimates of total population available for subcounty area
  - Specifically places, minor civil divisions
  - ACS and Population Estimates Program (PEP)
  - Typically did not agree prior to 2009
  - Disconcerting to data users
- Census Bureau investigated solutions to this problem for several years

Latest Research

- Three part research program
  - Documented the degree of inconsistency and looked for commonalities of the most egregious examples
  - Developed and tested several methodologies designed to reduce the degree of inconsistency
  - Evaluated the demographic impact of proposed methodologies on other characteristics besides total population

Results of the Research

- Places whose population greater than 250,000 tended to be underestimated by ACS compared to the PEP.
- Alternative weighting methodology is successful at reducing the inconsistencies for places and MCDs.
- Impact on other demographic characteristics tends to be small or in line with expectations.

Population Threshold for Subcounty Controls

- With the 5-year data product, only areas above a population threshold of 2,500 will be controlled.
- Not all places and MCDs with 2,500 will be controlled. It could be an aggregate of places and MCDs which is controlled.

Implementation for 2009

- We have implemented this new methodology for the data products released in 2010 including:
  - 2009 ACS 1-year estimates
  - 2007-2009 ACS 3-year estimates
  - 2005-2009 ACS 5-year estimates

Implementation Plans for 2010

- For the 2010 data release
  - Intercensal county-level estimates will be available to be used as controls for years 2006-2009.
  - No intercensal subcounty-level estimates will be available.
  - We will use 2000-based subcounty-level estimates as controls for 2006-2010 adjusted to agree with the county estimates.

Implementation Plans for 2011 Forward

- For the 2011 data release
  - Intercensal county- and subcounty-level estimates will be available for use as controls.
  - We will use these intercensals for years 2007-2009.
  - We will use postcensal (based on 2010) estimates as controls for years 2010-2011.

Summary – Subcounty Controls

- Goal of the new methodology is to reduce the inconsistencies but some will remain.
- Methodology does not ensure exact agreement between the ACS and PEP estimates.
- Multiyear products will be controlled to average of subcounty estimates over the period.
- 2010 will be a key transition year as we begin to incorporate updated population estimates.

Background – Break in Series

- The 2010 Census offers the opportunity to greatly improve the accuracy of the population estimates.
- As information is incorporated from the decennial census, the potential for significant shifts in the population estimates is present.
- Because we use the population estimates as controls, these shifts can also cause a break in the series of ACS estimates.

Education of Data Users

- The Census Bureau will need to alert ACS data users to this issue to help prevent misinterpretation of the year-to-year change.
- Outlets include:
  - Web site
  - Presentations
  - Partnerships with data user groups

Key Concepts for Handout

- Key concepts
  - Postcensal estimates are released annually except in the year of the census and are labeled by their year of release as their “vintage”.
  - Intercensal estimates are typically released once per decade, constructing a series from the previous census to the current census.
  - ACS multiyear estimates use the average of the population estimates over the period as controls.

Highlights from the Handout

- The use of 2010 Census information will be phased in over time.
- Release year 2011 (2010 ACS Data)
  - Simplified 2010 postcensal population estimates
  - Preliminary intercensal population estimates for 2006-2009
  - Subcounty controls will use internal postcensal estimates based on Census 2000 adjusted to agree with updated county controls

Highlights from the Handout (continued)

- Release year 2012 (2011 ACS Data)
  - From 2011 to 2012 ACS data has no break due to controls - the “steady state” is established.
  - Use postcensal controls for year 2010 forward, vintage 2011.
  - Use final intercensal controls for years prior to 2010.
  - Subcounty controls will use the same source as the county controls

Summary – Break in Series

- What to expect?
  - In release year 2011, the use of updated county controls will have an impact on most estimates.
  - In release year 2012, the use of updated subcounty controls should have a more focused impact on subcounty estimates.

Outline: Sample Expansion and Reallocation

- Design goals of the ACS
- Research into various expansion options
- Proposed expansion in President's 2011 budget
- Sample reallocation beginning in 2011

Key Design Goals of the ACS

- The ACS was designed to produce reliable estimates for similar characteristics and geographies as did the census long form
- Geographies include small areas such as tracts and small governmental units

Conceptual ACS Sample Design

- 3% sample of addresses per year
- 15% sample over five years
- Expectation was that the MOEs would be 1.33 times larger than the long-form

ACS Sample Design: 2005-2010

- Target sample is 2.9 million addresses per year (~2.2%)
- The sample size is a fixed number not a percentage
- Expectation is that the MOEs will be 1.75 times larger than the long-form.

Concerns for Select Areas

- Subsampling for personal visit reduces our effective sample size
- Particularly impacts area with:
  - High nonresponse
  - High nonmailable rates
  - Remote Alaska

Concerns for Select Areas

- Increased subsampling rates for areas with high mail/telephone nonresponse.
- High nonmailable rates especially impact
  - Hawaiian Homelands
  - Alaska Native Village Statistical Areas
  - Select American Indian (AI) areas
- Remote Alaska has special data collection procedures

Sample Expansion Plans for 2011

- Contingent on Congressional approval of the President's budget
- Increase the target sample to 3.54 million annually
- Improve the reliability of estimates for all areas
- Approximately a 9% reduction in the MOEs

Sample Expansion Plans for 2011

- Full person follow-up in select areas
  - Hawaiian Homelands
  - Alaska Native Village Statistical Areas
  - Remote Alaska
  - Select American Indian (AI) areas
- Restrict full follow-up to AI areas containing at least 10% American Indian based on Census 2000
  (approximately 75% of all AI areas)

Sample Reallocation

- Currently there is some variation in the reliability of tract estimates depending on the size
- For this reallocation, we wanted to determine how to better distribute our sample
- Goal is to reduce the differences in reliability across size categories

Sample Reallocation

- In 2005-2010, we used 7 different sampling rates determined by size of tract and governmental units
- Increased this to 16 different rates to better fine tune the sample
- Increased and added fixed rates for the smallest governmental units
- Shift sample from very largest tracts to smallest tracts

Sample Reallocation

- Not dependent on Congress approving the cost of the sample expansion
- Implemented beginning with the 2011 sample selection

Contact Information

- For questions regarding the weighting:
  Michael.A.Beaghen@census.gov
  Mark.E.Asiala@census.gov

- For questions regarding sample:
  Daniel.Sommers@census.gov
  Steven.P.Hefter@census.gov