

# ACS Group Quarters Small Area Estimation - What We have Accomplished and Future Work

Prepared for the State Data Center Annual  
Training Conference

October 14, 2010

Michael Beaghen

U.S. CENSUS BUREAU

U.S. DEPARTMENT OF COMMERCE

Washington, DC 20233

# Problem Statement

- ACS GQ sample is designed to produce state-level estimates of GQ population
- Estimates of GQ population contribute to the total resident population for substate areas
- As a consequence of sampling
  - Lack of representation of GQ sample for many small areas
  - Over-representation for other areas

# Solution

- Impute GQ person records to not-in-sample GQ facilities
- Donors are the sampled GQ person records
  - Copy the donor person's entire characteristics to imputed GQ

# Operational Goals

- Have evaluations completed this fall
- Implement for 2006-2010, 2008-2010 and 2010 ACS estimates

# GQ Types

- 27 specific types of GQs in ACS
- Grouped into seven major GQ types
  - Correctional institutions
  - Juvenile facilities
  - Nursing homes
  - Other long-term care facilities
  - College dorms
  - Military facilities
  - Other non-institutional GQs

# Technical Objectives

- If feasible – GQ representation, sample or imputed, by
  - Tract by major GQ type for 5-year data
  - County by major GQ type for 1- and 3-year data

# What We Have Accomplished - New GQ Estimation Methodology

# Large and Small GQ Sampling Strata

- Large stratum
  - Expected GQ pop of 16 or more according to ACS sampling frame
- Small stratum
  - Expected GQ pop of 15 or fewer according to ACS sampling frame

# Which GQs to Impute to?

- All large-stratum, not-in-sample GQs
- Selection of small stratum GQs to provide tract by major GQ type representation for 5-year estimates

# How Many GQ Person Records To Impute to Each GQ?

- **Large Stratum GQs:** 2.5% of the GQ's expected population
- **Small Stratum GQs:**
  - Expected pop/5 - if selected to give tract by major GQ type representation for the 5-year estimates.
  - Expected pop - if selected to give county by major GQ type representation for the 2010 estimates.

# Rules for Selecting Donors

- Period estimates - choose donors from same year that we impute into
  - Example - for a GQ on the 2007 sample frame but not-in-sample in 2007, choose as donors only 2007 ACS GQ data
- Expanding search
  - Select donors by geography and GQ type

# Expanding Search Algorithm for Selecting Donors

- 1) County and specific GQ type
- 2) County and major GQ type
- 3) State and specific GQ type
- 4) State and major GQ type

# Evaluation Based on Simulations with Census 2000 Data

- 25 sample selections, mimicking ACS sample selection
- Compute traditional ACS estimates
- Conduct imputation methodology
- Compare the two methods
  - For short form or 100% characteristics
  - By bias and variances

# Results of the Simulations

- Expanded search imputation superior to k-means clustering alternative
- Some donors were used repeatedly in small geographies
- Problem with single-sex GQ facilities
  - Males imputed to all-female facilities and vice versa

# Revisions to Imputation Methodology

- Limit donors to one use for each tract
- Identify single-sex and predominantly single-sex GQs from Census 2000, ACS, and Bureau of Prisons
  - Restrict donors by sex for these facilities
- We will continue to refine methodology as we learn more

# Current Work in Progress

- Evaluation based on 2006-2010 ACS data
  - Fabricated 2010 ACS data
- Compare imputation-based method with designed-based method
  - Sample or longform characteristics

# Production Activities

- In late 2010 conduct imputation
- In winter 2010-2011 Census subject matter experts review the ACS GQ data with imputations

# Conclusion

- Developed the basic imputation methodology
- First part of evaluations completed
- More work to do for 2010 estimates
- Continue to improve methods for future ACS estimates