

# Estimates Evaluation

## E<sup>2</sup>

Victoria A. Velkoff  
Jason Devine

Population Division  
U.S. Census Bureau

COPAFS Quarterly Meeting  
June 5, 2009

# Goal of Estimates Evaluation

To evaluate the performance of the methods we currently use to produce population estimates and to evaluate alternative methods relative to the Census 2010 counts.

# Outline

- Population estimates we produce and methods
- Background/overview of E<sup>2</sup>
- Estimates principles
- Accuracy measures
- Key points

# Annual Estimates

- Population
  - National by age, sex, race, and Hispanic origin
  - States by age, sex, race, and Hispanic origin
  - Counties by age, sex, race, and Hispanic origin
  - Incorporated places and minor civil divisions (total population only)
  - Puerto Rico Commonwealth and municipios by age and sex
- Housing units
  - States
  - Counties

# Current Method: ADREC

- National level

Cohort-component method (also called the Administrative Record Method or ADREC)

$$P_2 = P_1 + B - D + NIM$$

NIM = International Migration

- State and county level

Cohort-component method

$$P_2 = P_1 + B - D + NM$$

NM = Domestic and international migration  
(controlled to the national estimates)

- Sub-county level

Housing unit method

# Housing Unit Method

$$\text{Population}_t = \text{HU}_t * \text{O}_t * \text{PPH}_t + \text{GQ}_t$$

Where:

HU = Number of housing units

O = Occupancy rate

PPH = Persons per household

GQ = Group quarters population

# Past Research: HUBERT

1. Evaluations and basic comparisons of housing unit method and administrative records (ADREC) method for estimating population
2. Review of current methodology for estimating housing units
3. Research into ways to estimate persons per household (PPH) and occupancy

# This Estimates Evaluation (E<sup>2</sup>)

- Continues HUBERT research
- Incorporates lessons learned from HUBERT and past evaluations
- Formalizes the evaluations of the population estimates against decennial census results
- Continues collaboration with the Federal State Cooperative for Population Estimates (FSCPE) and other external researchers

# Project Overview

- Phase One
  - Develop estimates principles
  - Select accuracy measures
- Phase Two
  - Select alternative methods to test
  - Develop official evaluation estimates and alternative estimates
  - Evaluate official and alternative estimates

# Major Milestones

- Complete phase one – April 2009
- Determine alternative methods to test and agree on evaluation estimates – July 2009
- Produce evaluation estimates – December 2010
- Complete evaluations – October 2011
- Decide on post-2010 methodology – Spring 2012

# Evaluation Estimates

- Census Bureau
  - ADREC method
    - Nation, states, and counties
    - Total population estimates and demographic characteristics
  - Housing unit method with alternative sources of PPH and occupancy
    - Nation, states, counties, and sub-county
    - Total population estimates
- External researchers
  - Ratio correlation method
  - Other methodologies?

# Underlying Principles Document

- Documents the production requirements of the Census Bureau's Population Estimates Program
- Establishes basic principles for assessing alternative methods
- Builds on work from past evaluation projects
- Developed with input from the FSCPE and other external researchers

# Production Requirements

- The population estimates will reflect the estimated population change consistent with the census concept of usual residence.
- Because of their use in the distribution of Federal funds, producing an unbiased estimate of the distribution of the population is a priority.
- The most recent census count with geographic updates will serve as the estimates base.
- Products must be produced within deadlines established based on production requirements and user needs.
- Each vintage must include a time series from the last census.
- The latest available data will be used.
- Within any vintage, all products must sum to other products.
- One set of official products is developed within a vintage.
- The official population estimates must be developed by Census Bureau staff.
- Documentation of the method used must be available to data users.

# Methodological Principles

- Soundness
- Accountability
- Availability of Data
- Availability of Resources
- Robustness
- Comparability
- Adaptability
- Parsimony
- Reasonableness
  - Accuracy
  - Demographic Appropriateness
  - External Comparisons

# Accuracy Measures Document

- Documents in advance the measures of accuracy that will be used in the evaluations
- Provides justification for the selected measures of accuracy
- Builds on work from past evaluation projects
- Developed with input from the FSCPE and other external researchers

# Properties of “Good” Estimates

After a review of the estimates program in the 1970s by the Committee on National Statistics, four criteria were identified that a set of estimates should meet.

- 1) low average numeric error
- 2) low average percent error
- 3) few extreme percent errors
- 4) absence of bias for subgroups

Accuracy will be defined as the degree of closeness to the Census 2010 values.

# The Selected Measures of Accuracy

- Selected after a review of 19 different measures of accuracy
- Provide information on all four criteria
- Measure the accuracy of the estimated shares of the population
- Straightforward to compute and explain
- Commonly used
- Minimize duplication while providing a reasonably complete profile of accuracy

# The Five Selected Measures of Accuracy

1) Measure of Average Numeric Error:

Root Mean Squared Error

$$\text{SQRT}(\sum ((\text{Estimate}-\text{Census})^2)/N)$$

2) Measure of Average Percent Error:

Mean Absolute Percent Error (MAPE)

$$((\sum (|\text{Estimate} - \text{Census}| )/\text{Census}))/N)*100$$

3) Measure of Extreme Percent Error:

Number of percent errors greater than some established threshold

e.g., number of percent errors greater than plus/minus 10

# The Five Selected Measures of Accuracy

## cont'd

### 4) Measure of Bias:

Mean algebraic percent error (MALPE)

$$\left(\frac{\sum((\text{Estimate}-\text{Census})/\text{Census})}{N}\right)*100$$

### 5) Measure of Accuracy of Area's Share of Total Population:

Total absolute error of shares

$$\sum \left| (\text{Estimate}/\sum\text{Estimate}) - (\text{Census}/\sum\text{Census}) \right|$$

Measures will be calculated for predetermined size and growth categories and by geography to determine if there is bias for subgroups.

# Accuracy Profile: County Population Estimates for April 1, 2000

Measure of Accuracy	Method	
	ADREC	Housing Unit-Based
Root Mean Squared	10,341	13,464
MAPE (%)	3.2	5.9
MALPE (%)	-1.5	1.4
Extreme Percent Error (10%)	109	502
Accuracy of Shares	0.024	0.045

# Key Points

- Focus resources on most promising alternatives
- Use production requirements and underlying principles to guide our decisions
- Provide the most comprehensive evaluation possible given resource limitations
- Produce the evaluation estimates needed to conduct a comprehensive evaluation
- Provide datasets that will allow others to assess the accuracy of our estimates

# Contact Information

Victoria A. Velkoff  
(301) 763-7070

[Victoria.A.Velkoff@census.gov](mailto:Victoria.A.Velkoff@census.gov)

Jason Devine  
(301) 763-6070

[Jason.E.Devine@census.gov](mailto:Jason.E.Devine@census.gov)

Alexa Kennedy-Puthoff  
(301) 763-6030

[Alexa.K.Kennedy-Puthoff@census.gov](mailto:Alexa.K.Kennedy-Puthoff@census.gov)