Post-Enumeration Survey Overview

Presentation to the National Advisory Committee on Racial, Ethnic and Other Populations

May 6, 2021

Tim Kennel, Assistant Division Chief, Statistical Methods, Decennial Statistical Studies Division
1. Introduction to Census Quality
2. 2020 Post-Enumeration Survey Design
3. Schedule and Reports
Introduction to Census Quality
Post-Enumeration Survey
Measuring Census Quality

• Census process indicators
• Assessments and evaluations of census operations
• Demographic Analysis (DA) and other methods
• Post-Enumeration Survey (PES)
Post-Enumeration Survey
Coverage Measurement in the Past

Demographic analysis for decades

Coverage measurement surveys

• 1980 Post-Enumeration Program
• 1990 Post-Enumeration Survey
• 2000 Accuracy and Coverage Evaluation
• 2010 Census Coverage Measurement survey
Post-Enumeration Survey

2010 Census and Demographic Analysis (in Thousands)

2010 Census: 308,746

DA Low: 305,684
DA Middle: 308,475
DA Revised Middle: 308,346
DA High Middle: 310,038
DA High: 312,713
Net census coverage error (defined as)

\[ = \text{Census count} - \text{True population size} \]

- **Negative** coverage error means fewer people are included in the census than are in the true population (undercount)
- **Positive** coverage error means more people are included in the census than are in the true population (overcount)

**Note:** In the past, that is, up through 2010, documentation of results from the Census Bureau’s post-enumeration surveys has generally defined the net undercount using the opposite of the difference above. When providing results from 2020, and in these slides covering past results, we will use the definition above.
## Post-Enumeration Survey

### Net Undercount of US Household Population

<table>
<thead>
<tr>
<th>Year</th>
<th>Census Count</th>
<th>Net Coverage Error</th>
<th>Percent Coverage Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>300,700,000</td>
<td>36,000</td>
<td>0.01 %</td>
</tr>
<tr>
<td>2000</td>
<td>273,600,000</td>
<td>1,332,000*</td>
<td>0.49 %*</td>
</tr>
<tr>
<td>1990</td>
<td>248,700,000</td>
<td>– 3,994,000*</td>
<td>– 1.61 %*</td>
</tr>
</tbody>
</table>

* Statistically different from zero

Census counts from 1990 to 2010, exclude Remote Alaska.
Census counts for 2000 and 2010 exclude Group Quarters population.
Census count for 1990 included people in housing units and in noninstitutional and nonmilitary Group Quarters.
Post-Enumeration Survey
U.S. Total Percent Net Coverage Error from 1990 - 2010

Error bars indicate 90% confidence interval.
## Post-Enumeration Survey

### Components of Census Coverage for the 2010 United States Household Population (in Thousands)

<table>
<thead>
<tr>
<th>Component of Census Coverage</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Percent</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Census Count</td>
<td>300,700</td>
<td>0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Correct enumerations</td>
<td>284,700</td>
<td>199</td>
<td>94.7</td>
<td>0.1</td>
</tr>
<tr>
<td>Erroneous enumerations</td>
<td>10,040</td>
<td>199</td>
<td>3.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Due to duplication</td>
<td>8,521</td>
<td>194</td>
<td>2.8</td>
<td>0.1</td>
</tr>
<tr>
<td>For other reasons</td>
<td>1,520</td>
<td>45</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Whole-Person Census Imputations</td>
<td>5,993</td>
<td>0</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Estimate of Population</td>
<td>300,700</td>
<td>429</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Correct enumerations</td>
<td>284,700</td>
<td>199</td>
<td>94.7</td>
<td>0.1</td>
</tr>
<tr>
<td>Omissions</td>
<td>16,000</td>
<td>440</td>
<td>5.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Net Coverage Error</td>
<td>36</td>
<td>429</td>
<td>0.01</td>
<td>0.1</td>
</tr>
</tbody>
</table>
2020 Post-Enumeration Survey Design
Post-Enumeration Survey
What is the Post-Enumeration Survey?

Goal
• Measure coverage of people and housing units in the decennial census
• Measure net coverage error and components of coverage
• No intent to adjust the census counts

How is coverage estimated?
• Conduct a survey in sample blocks
• Match people in the PES independent survey with enumerations in the census to determine who was missed or counted in error

How are the data used?
• To assess the success of the current census
• To provide information to improve the coverage of future censuses

History
• Used in the United States census since 1950
Post-Enumeration Survey
2020 Post-Enumeration Survey Universe

Household Population and Housing Units
Excludes
• Group Quarters and Group Quarters Population
• Remote Alaska
Post-Enumeration Survey
Sample Sizes

First sample (Basic Collection Units)
- Selected 9,800 Basic Collection Units in 50 states and District of Columbia
- Selected 400 Basic Collection Units in Puerto Rico

PES listing
- 510,000 housing units in 50 states and District of Columbia
- 18,000 housing units in Puerto Rico

After subsampling and review
- 161,000 housing units in 50 states and District of Columbia
- 7,100 housing units in Puerto Rico
Capture (Enumeration sample, E sample)
• A sample of census records (enumerations) in the PES sample blocks, called E sample
• Measure correct enumerations in census
• Determine if enumeration is correct or erroneous; duplicate? correct location?

Recapture (PES sample, population sample, P sample)
• A sample of blocks, independent of the census, in the post-enumeration survey (PES), called P Sample; interview people in units
• Match people in P Sample to Census
## Post-Enumeration Survey

### How PES Estimates are Created

<table>
<thead>
<tr>
<th>Enumerated in Census?</th>
<th>Enumerated in PES?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>YES</strong></td>
<td>YES</td>
<td><strong>Found in Both</strong></td>
</tr>
<tr>
<td></td>
<td>NO</td>
<td><strong>Missing in PES</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Found in Census</strong></td>
</tr>
<tr>
<td><strong>NO</strong></td>
<td></td>
<td><strong>Missing in Census</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Missing in PES</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Missing in Both</strong></td>
</tr>
</tbody>
</table>
Post-Enumeration Survey  
Current version of Dual System Estimator (DSE)

\[
\overline{\text{TOT}}_D = \sum_{j \in D} \left[ \pi_{dd,j} \pi_{ce,j} / \pi_{m,j} \right]
\]

**Where**

- \(j\) is a census enumeration, and
- \(D\) is the estimation domain (total, tenure status, age/sex groupings, etc.).
- \(\pi_{dd,j}\) is the predicted probability census case \(j\) is data-defined
- \(\pi_{cd,j}\) is the predicted probability that census case \(j\) is a correct enumeration
- \(\pi_{m,j}\) is the predicted probability that census case \(j\) was matched
Post-Enumeration Survey

Schedule and Reports
## Post-Enumeration Survey
### Revised field milestone schedule

<table>
<thead>
<tr>
<th>Operation</th>
<th>Original Dates</th>
<th>New Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Start</td>
<td>Finish</td>
</tr>
<tr>
<td>Independent Listing</td>
<td>1/16/2020</td>
<td>3/20/2020</td>
</tr>
<tr>
<td>Initial HU Followup</td>
<td>5/6/2020</td>
<td>6/19/2020</td>
</tr>
<tr>
<td>Person Followup</td>
<td>2/3/2021</td>
<td>3/26/2021</td>
</tr>
<tr>
<td>Final HU Followup</td>
<td>5/19/2021</td>
<td>6/18/2021</td>
</tr>
</tbody>
</table>

Operations in **blue** are complete.

Dates include reinterview.
Three major groups:
- Demographic or housing unit characteristics
  - Age Group
  - Age Group by Sex
  - Sex
  - Race and Hispanic Origin (alone or in-combination and mutually exclusive groups)
- Relationship to Householder
- Tenure (owner or renter)
- Geographic areas
  - Region
  - State
- Census operations
  - Coverage Improvement Universe
  - GSS Partnership County
  - In-Field Address Canvassing Outcome
  - Nonresponse Followup Respondent Type
  - Nonresponse Followup Workload
  - Response Mode
  - Response Rate Decile
  - Type of Enumeration Area
Post-Enumeration Survey
Release Coverage Reports and Tables

First Release
• National demographic estimates of net coverage of people
• National demographic estimates of components of coverage for people
• Tables can be downloaded on Census Bureau website

Second Release
• State coverage estimates for people
• More detailed national coverage estimates for people
• All housing unit coverage estimates
• Coverage estimates for Puerto Rico
Post-Enumeration Survey
Schedule to Release Methodology Reports

First Release

• Source and Accuracy Statement – National Person

Second Release

• Estimation Design
• Estimation Methods: Characteristic Imputation
• Estimation Methods: Missing Data
• Estimation Methods: Net Coverage Estimation
• Estimation Methods: Component Estimation
• Source and Accuracy – All of PES
Post-Enumeration Survey

References


How have you used census post-enumeration survey estimates in the past?

How do you think the Census Bureau should use the estimates of net coverage error and components of coverage in planning for the future?
Questions