Coverage and Agreement of Administrative Records and 2010 American Community Survey Demographic Data

Renuka Bhaskar
U.S. Census Bureau

Federal Committee on Statistical Methodology
Research Conference
November 4, 2013

Background

- Administrative Records (AR) provide high coverage of persons and addresses in censuses and surveys

- Demographic data in AR may be useful when demographic information are missing

- Comparisons between AR and 2010 Census reveal high agreement and coverage for some demographic variables

Motivation

- Expand on AR/2010 analysis by evaluating AR and the American Community Survey (ACS)

- There may be changes in the population and in the coverage of administrative records over time

- The ACS allows us to evaluate AR throughout the decade

Research Objectives

- *Coverage*: Evaluate the extent to which AR provide data on race, Hispanic origin, age, and sex for 2010 ACS respondents

- *Agreement*: Evaluate whether demographic responses in AR are in agreement with responses in the 2010 ACS

Data and Methodology

United States
Census
Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

Data

- 2010 ACS data compared to:

  - Centers for Medicare and Medicaid Services Medicare Enrollment Database
  - Housing and Urban Development (HUD) Public and Indian Housing Information Center
  - HUD Tenant Rental Assistance Certification System
  - HUD Computerized Homes Underwriting Management System
  - Indian Health Service Patient Registration File
  - Selective Service System Registration File
  - Health and Human Services Temporary Assistance for Needy Families
  - Social Security Administration (SSA) Supplemental Security Income Record
  - Commercial data (5 vendors, 9 datasets)
  - **SSA Numerical Identification File (Numident)**
  - **Previous Census Records (Census 2000 and ACS 2001-2009)**

Methodology

- Record Linkage - Unique identifiers called Protected Identification Keys (PIKs) were used to link persons in AR and the 2010 ACS

- Coverage

$$\frac{\textit{Individuals in the } 2010 \textit{ ACS for whom AR provides demographic data}}{2010 \textit{ ACS count}}$$

- Agreement

$$\frac{\textit{Individuals with matching AR and } 2010 \textit{ ACS demographic response}}{\textit{Individuals in } 2010 \textit{ ACS for whom AR provides demographic data}}$$
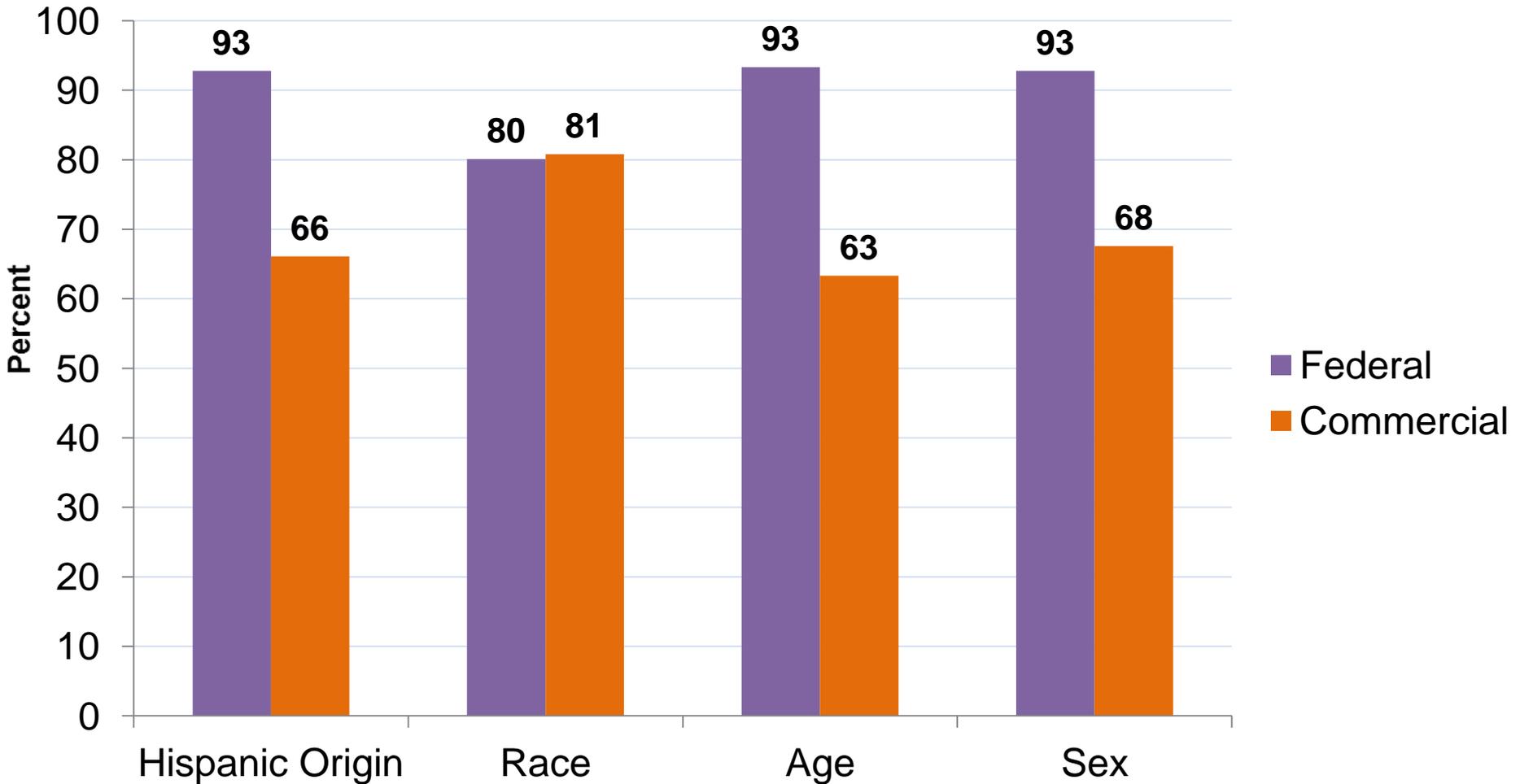
Results

Coverage of Each Demographic Item

- There are 4.47 million people in the 2010 ACS

- Of the 2010 ACS sample, AR provides:
    - Hispanic origin data for 4.16 million, or 93 percent
    - Race data for 3.61 million, or 81 percent
    - Age data for 4.17 million, or 93 percent
    - Sex data for 4.16 million, or 93 percent

Demographic Characteristics

- AR provides higher coverage of Hispanic origin data for non-Hispanics in the ACS (94 percent) compared to Hispanics (83 percent)

- Coverage of race data varies by race group
  - White alone and Black alone: 83 percent
  - American Indian or Alaska Native alone: 82 percent
  - Asian alone and Native Hawaiian and Other Pacific Islander alone: 76 percent
  - Two or More Races: 67 percent
  - Some Other Race alone: 46 percent

- Coverage of age and sex data high (over 90 percent) for all age groups and both sexes

Coverage by Federal or Commercial Administrative Records Source File

Source: 2010 American Community Survey and Administrative Records Data

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

Federal AR Files

- Numident
  - Highest coverage of sex and age data
  - Highest coverage of race data for certain groups (Black alone, Asian alone, Native Hawaiian Pacific Islander alone)

- Previous Census Records
  - Highest coverage of Hispanic origin data
  - Second highest coverage of sex and age data
  - Highest coverage of race data for certain groups (White alone, Two or More Races, Some Other Race alone)

- Medicare Enrollment Database
  - High coverage of age data for older population

Agreement

- Coverage results indicate that AR provides substantial data on certain demographic items

- Next, we look at whether demographic data in AR and in the 2010 ACS are in agreement

- Restrict to cases whose 2010 ACS demographic response was not allocated

Agreement of Hispanic Origin Response

Administrative Records Source File	Percent in Agreement: Non-Hispanic	Percent in Agreement: Hispanic
Federal Files		
Previous Census Records	99.5	91.6
Numident	N/A	52.9
HUDCHUMS	98.6	82.0
HUDPIC	98.1	85.6
HUDTRACS	98.5	77.9
TANF	98.2	68.9
MEDB	N/A	26.9
Commercial Files		
Commercial File 1	98.6	78.4
Commercial File 2	98.2	75.3
Commercial File 3	98.2	78.4
Commercial File 4	98.2	75.5

Agreement of Race Response

2010 ACS Race	Range in Percent Agreement by AR Source Files	
	Federal	**Commercial**
White Alone	96.9 to 99.0	94.9 to 98.0
Black Alone	88.7 to 98.0	38.8 to 63.2
American Indian Alaska Native Alone	30.4 to 97.9	7.9 to 16.4
Asian Alone	55.7 to 94.3	73.9 to 85.0
Native Hawaiian or Other Pacific Islander Alone	33.6 to 72.9	13.0 to 18.2
Some Other Race Alone	12.3 to 53.6	1.3 to 3.3
Two or More Races	2.5 to 35.0	N/A

Agreement of Age Response

- Age in the ACS calculated as age at the time of survey (January 1 – December 31, 2010)

- Age in AR files calculated as age as of Census reference date: April 1, 2010

- Agreement: AR age is within one year of ACS age

- Future work will compare AR and ACS date of birth

Agreement of Age Response

2010 ACS Age Group	Range in Percent Agreement by AR Source Files	
	Federal	**Commercial**
0 to 2	96.4 to 99.8	96.2 (only one file)
3 to 17	92.5 to 99.7	22.7 to 83.6
18 to 24	98.6 to 99.2	85.3 to 99.2
25 to 44	98.3 to 99.2	89.4 to 99.3
45 to 64	97.1 to 99.2	88.8 to 99.3
65 to 74	94.4 to 99.6	91.3 to 99.2
75 and older	91.8 to 99.3	94.0 to 98.0

United States™
Census
Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

17

Agreement of Sex Response

2010 ACS Sex Response	Range in Percent Agreement by AR Source Files	
	Federal	**Commercial**
Males	97.9 to 100.0	94.8 to 98.6
Females	98.0 to 99.7	95.7 to 98.9

Discussion

- Sex: high coverage, high agreement

- Age: high coverage, agreement to be re-calculated using date of birth

Discussion

- Hispanic origin and race: coverage and agreement more variable but high for some groups

- AR may be able to assist when demographic data are missing

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

Future Research

- How to determine a single demographic response from the various AR files

- Additional AR data sources
    - Supplemental Nutrition Assistance Program
    - Medicaid Statistical Information System

- Agreement of allocated responses

- Continue analysis throughout the decade

Thank you!

renuka.bhaskar@census.gov