Demographic Frame: An Overview

Jennifer M. Ortman

April 19, 2024

Presentation for the Annual Meeting of the Population Association of America

United States® Census Bureau

The Demographic Frame is a comprehensive database of person-level data that:

- Contains demographic characteristics and addresses associated with each person

- Is derived from administrative, third-party, census, and survey data sources

- Includes unique person-level identifiers used to link individuals across datasets

- Can be linked to information in other enterprise frames

- Is available only to approved, internal users in a secured computing environment
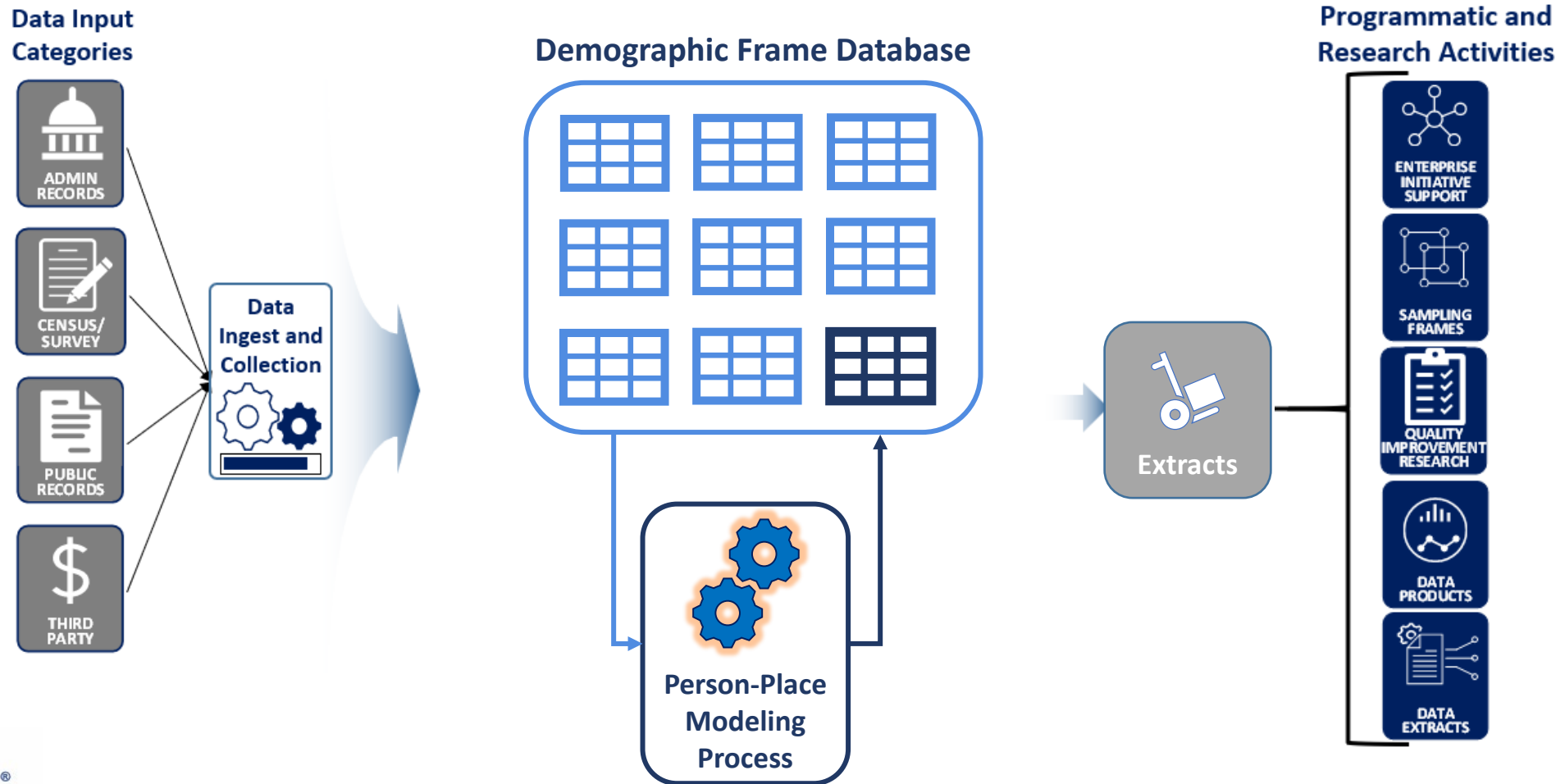

DEMOGRAPHIC
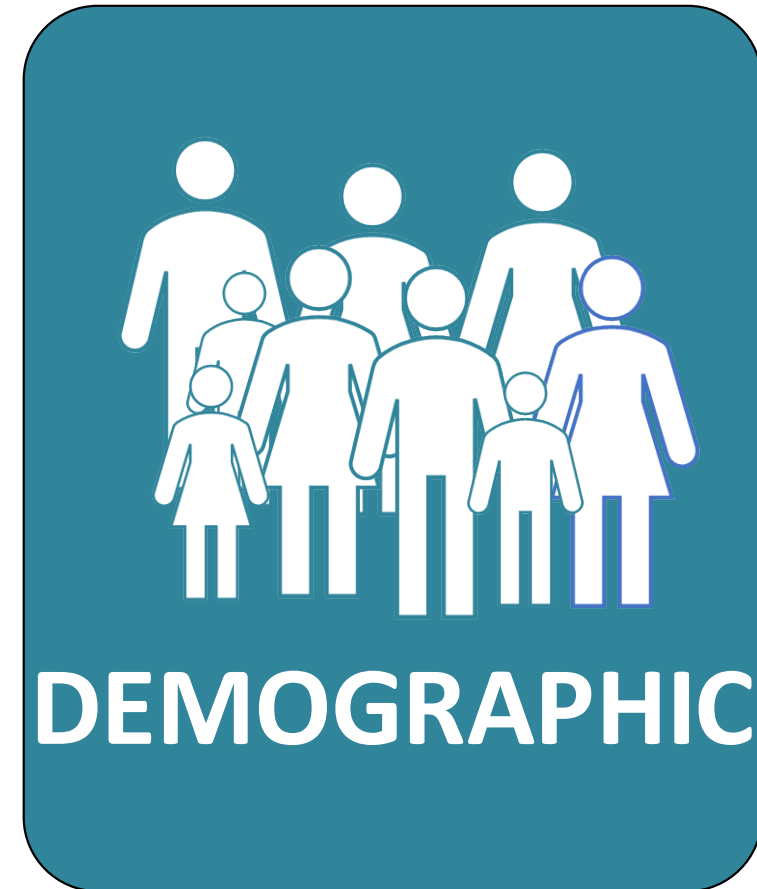
Capabilities under development include:

- Working with others across the enterprise to **improve record linkage**
  - Review and enhance the Person Identification Validation System (PVS) process to reduce the number of individuals without a Protected Identification Key (PIK)
  - Evaluate quality and determine how best to handle remaining individuals without a PIK

- Aggregating to **household and family units**

- Adding more years of data to facilitate **linking information about individuals across time and space**

- **Establishing a production team** to maintain the Demographic Frame for enterprise use

United States® Census Bureau

Building the Demographic Frame

Demographic Frame Extracts
Dates available: 2017, 2018, 2019, 2020, 2021, 2022

Protected Identification Key (PIK)	Master Address File Identifier (MAFID)	Modeled Probabilities	First and Last Date Observed in a Source	Date of Birth
Age	Date of Death	Sex	Hispanic Origin	Race
Block	Block Group	Tract	County	State

United States® Census Bureau

Applications of the Demographic Frame

- **Serve as a sampling frame for population surveys** where a focus on ensuring a representative sample around the characteristics of people is a better foundation than an address-based approach

- **Support enterprise initiatives such as census and survey taking**, including reducing the burden on our respondents by using information already available to the federal government instead of asking questions.

- **Creating blended data products**, including merging Frame data on topics not currently asked about on a survey or the census, to provide even more detailed information about each of America's communities

- **Improving data quality by drawing upon the Frame for data editing and imputation**, rather than just statistical approaches to assigning values

SAMPLING FRAMES

ENTERPRISE INITIATIVE SUPPORT

DATA PRODUCTS

QUALITY IMPROVEMENT RESEARCH

United States® Census Bureau
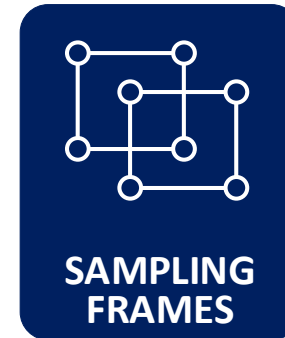
Current and Prospective Demographic Frame Customers

- 2030 Census Program
- Continuous Count Study (CCS)
- Self-Response Quality Assurance (SRQA) for Special Census
- Census Household Panel Survey
- American Community Survey (ACS)
- Community Resilience Estimates (CRE)
- Survey of Income and Program Participation (SIPP)
- Non-Employer Statistics – Demographics (NES-D)
- Longitudinal Employer-Household Dynamics (LEHD)

United States® Census Bureau

Ongoing Evaluations

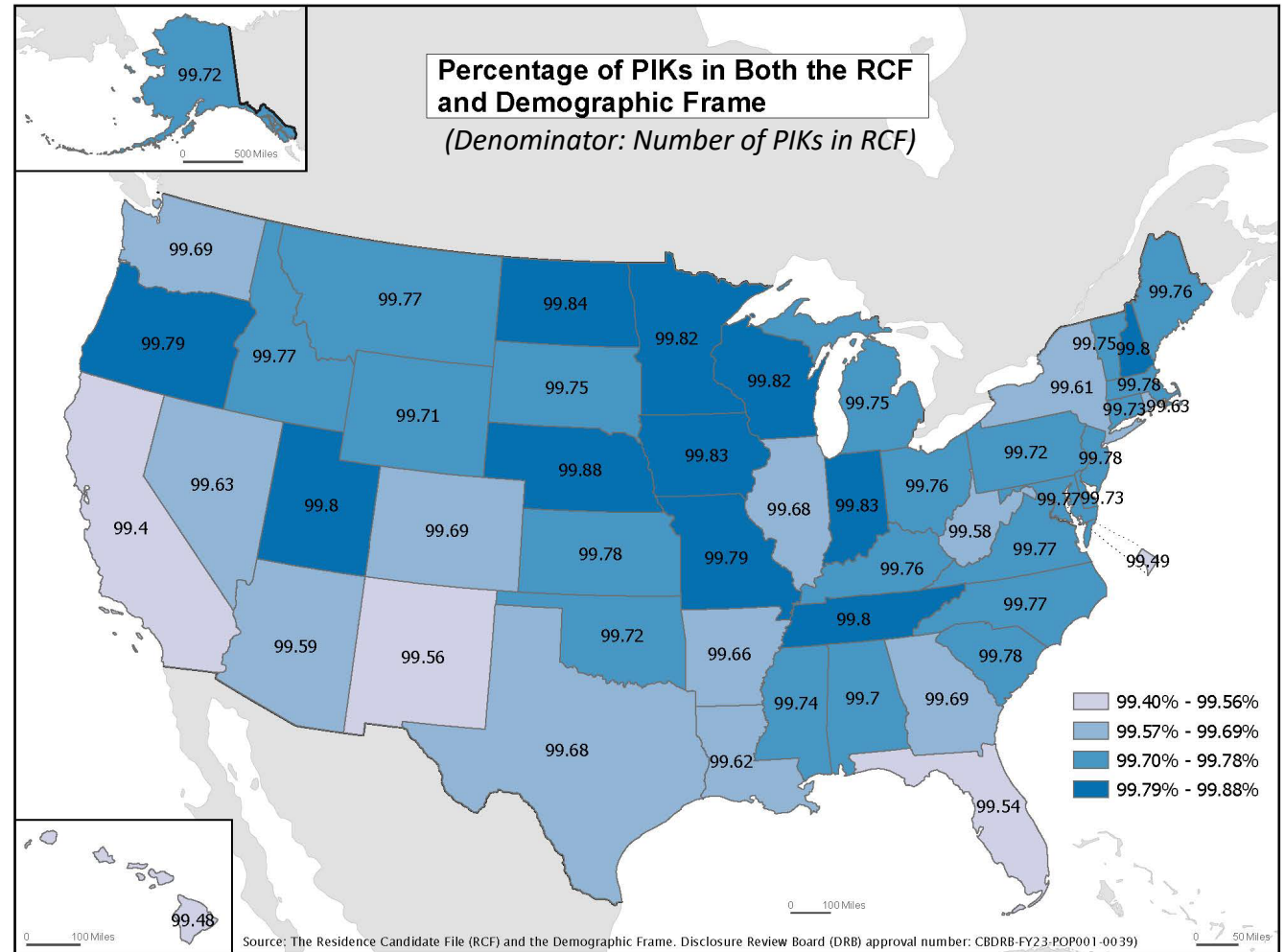| Person- and Address-Level Comparisons	Benchmarking (Aggregate Counts)	Person-Place Model
• Longitudinal Employer-Household Dynamics (LEHD) Residence Candidate File (RCF) • 2020 Census • Non-Employer Statistics by Demographics (NES-D)	• 2020 Census • Population Estimates Program • Administrative Records Population Estimates • American Community Survey (ACS)	• Evaluating 4 models • Establishing model metrics for ongoing evaluation

United States® Census Bureau

Highlights from Our Evaluations

- Does the Demographic Frame **find the same people and put them at the same location** as the LEHD's Residence Candidate File?

- Does the Demographic Frame **provide similar counts** compared to the 2020 Census or Population Estimates?

- Do **household counts** match between the Demographic Frame and the ACS?

*Does the Demographic Frame **find the same people** as the Residence Candidate File?*

Over **99 percent** of PIKs observed in the RCF were also found in the Demographic Frame at the national level and when looking by state.

Percentage of PIKs in Both the RCF and Demographic Frame
(Denominator: Number of PIKs in RCF)

99.40% - 99.56%
99.57% - 99.69%
99.70% - 99.78%
99.79% - 99.88%

Source: The Residence Candidate File (RCF) and the Demographic Frame. Disclosure Review Board (DRB) approval number: CBDRB-FY23-POP001-0039)

*Does the Demographic Frame **put people at the same address** as the Residence Candidate File?*

90.56 percent of PIKs were assigned to the same MAFID in the RCF and Demographic Frame.

Agreement ranges from **86 to 93 percent** by state.

Percentage of PIKs Assigned to Same MAFID in RCF and Demographic Frame

National Percentage = 90.56

Legend:
- 86.24%
- 86.25% - 89.88%
- 89.89% - 91.01%
- 91.02% - 92.55%

Source: The Residence Candidate File (RCF) and the Demographic Frame. Disclosure Review Board (DRB) approval number: CBDRB-FY23-POP001-0039

Does the Demographic Frame *provide similar counts* compared to the 2020 Census or Population Estimates?
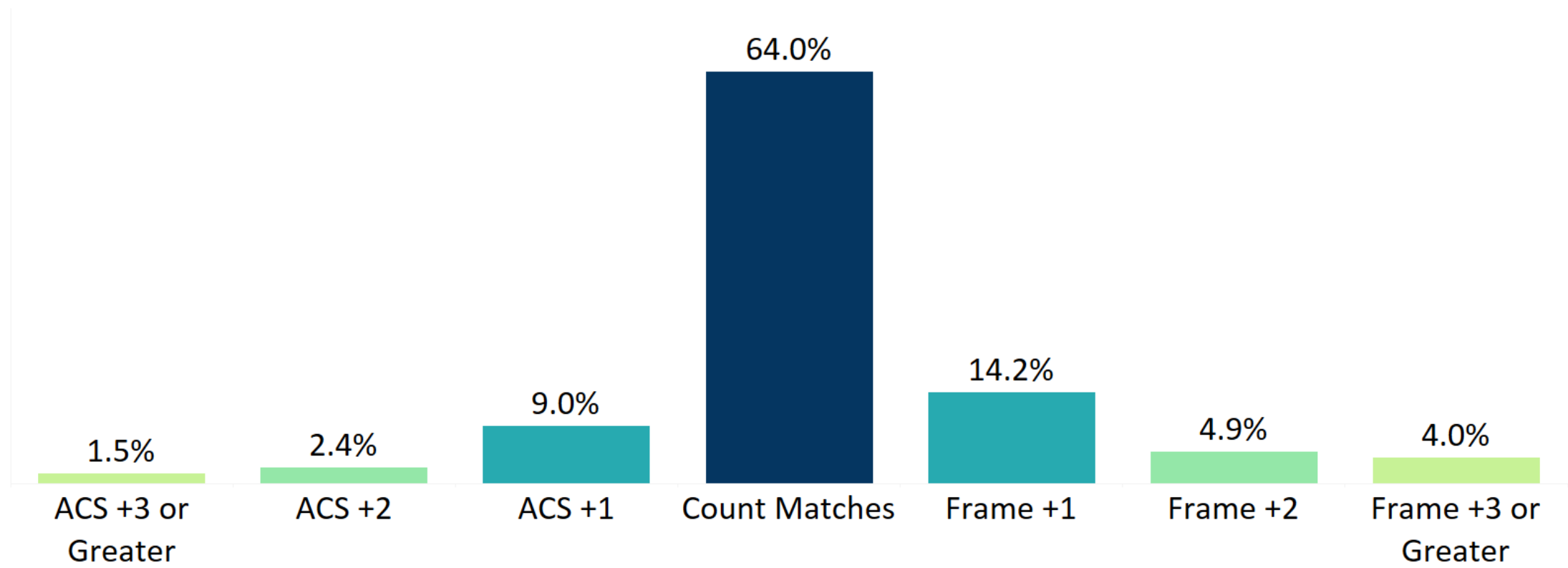
The Demographic Frame is:

2.45% less than the 2020 Census

1.87% less than the 2020 Population Estimates

| Geography | Demographic Frame | | 2020 Census | | 2020 Population Estimates | | Demo Frame Net Coverage Error (%) | |
	Number	Percent	Number	Percent	Number	Percent	Relative to Census	Relative to Pop Est.
United States	323,300	100.0	331,500	100.0	329,500	100.0	2.45	1.87
Numbers in thousands and rounded to 4-significant digits								

United States® Census Bureau

Do *household counts match* between the Demographic Frame and the ACS?

Bar chart:
- ACS +3 or Greater: 1.5%
- ACS +2: 2.4%
- ACS +1: 9.0%
- Count Matches: 64.0%
- Frame +1: 14.2%
- Frame +2: 4.9%
- Frame +3 or Greater: 4.0%

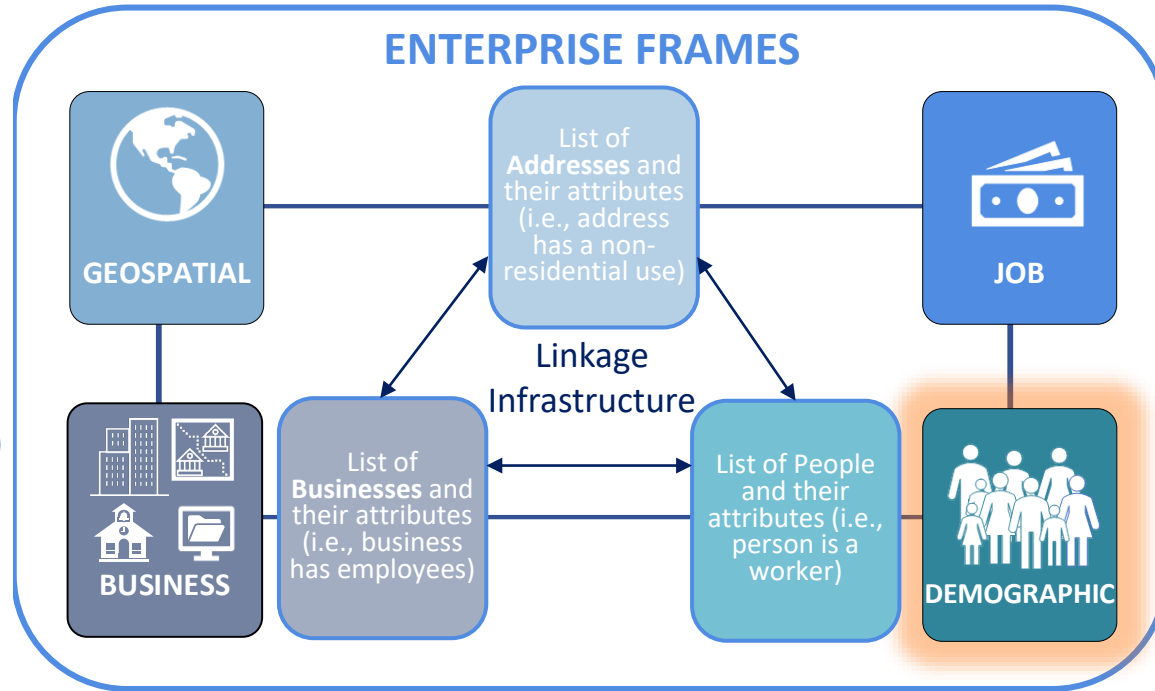United States® Census Bureau

For More Information

jennifer.m.ortman@census.gov
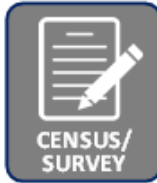
Census.gov/Frames

Appendix Slides

Frames Program

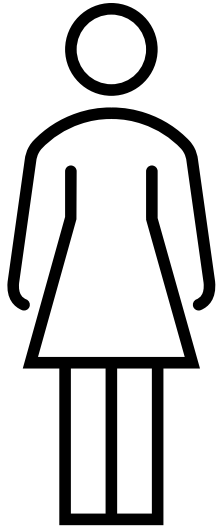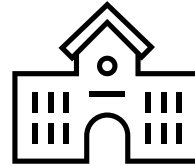*Creating an Infrastructure to Modernize the Census Bureau's Statistical Foundation*

Where does this person live?
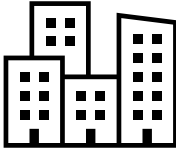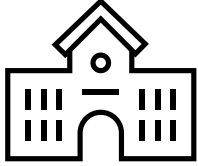
THIS HOUSE?   THIS DORM?   THIS APARTMENT?
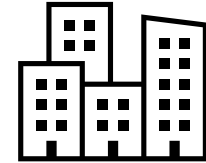
The Person-Place Model (PPM)

- A process for assigning probabilities to PIK-MAFID (i.e., person-address) records to determine a person's residence on a particular reference date

- Learns from "true" PIK-MAFID pairs (ACS) to make guesses for PIK-MAFID pairs from administrative records
  - Over 100 model features including info about: source file, IRS/income, address, and aggregate demographics

PIK	MAFID	Probability
		0.75
		0.20
		0.05

Additional Evaluation Results

RCF Analysis

*Does the Demographic Frame **find the same people** as the Residence Candidate File?*

Over **99 percent** of PIKs observed in the RCF were also found in the Demographic Frame at the national level and when looking by state.

Percentage of PIKs in Both the RCF and Demographic Frame
(Denominator: Number of PIKs in RCF)

	99.40% - 99.56%
	99.57% - 99.69%
	99.70% - 99.78%
	99.79% - 99.88%

Source: The Residence Candidate File (RCF) and the Demographic Frame. Disclosure Review Board (DRB) approval number: CBDRB-FY23-POP001-0039)

*Does the Demographic Frame **put people at the same address** as the Residence Candidate File?*

90.56 percent of PIKs were assigned to the same MAFID in the RCF and Demographic Frame.

Agreement ranges from **86 to 93 percent** by state.

Percentage of PIKs Assigned to Same MAFID in RCF and Demographic Frame

National Percentage = 90.56

	86.24%
	86.25% - 89.88%
	89.89% - 91.01%
	91.02% - 92.55%
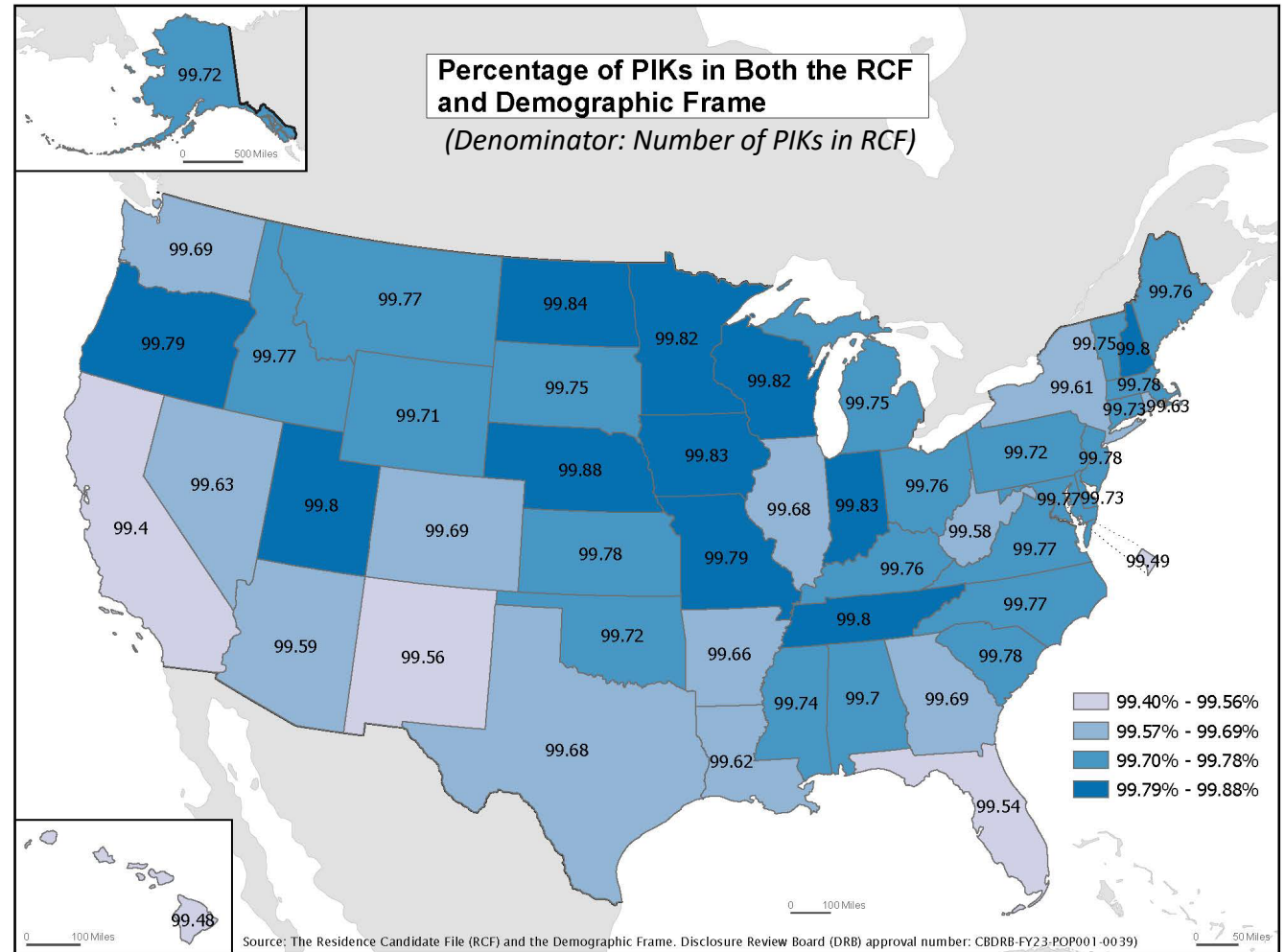
Source: The Residence Candidate File (RCF) and the Demographic Frame. Disclosure Review Board (DRB) approval number: CBDRB-FY23-POP001-0039)
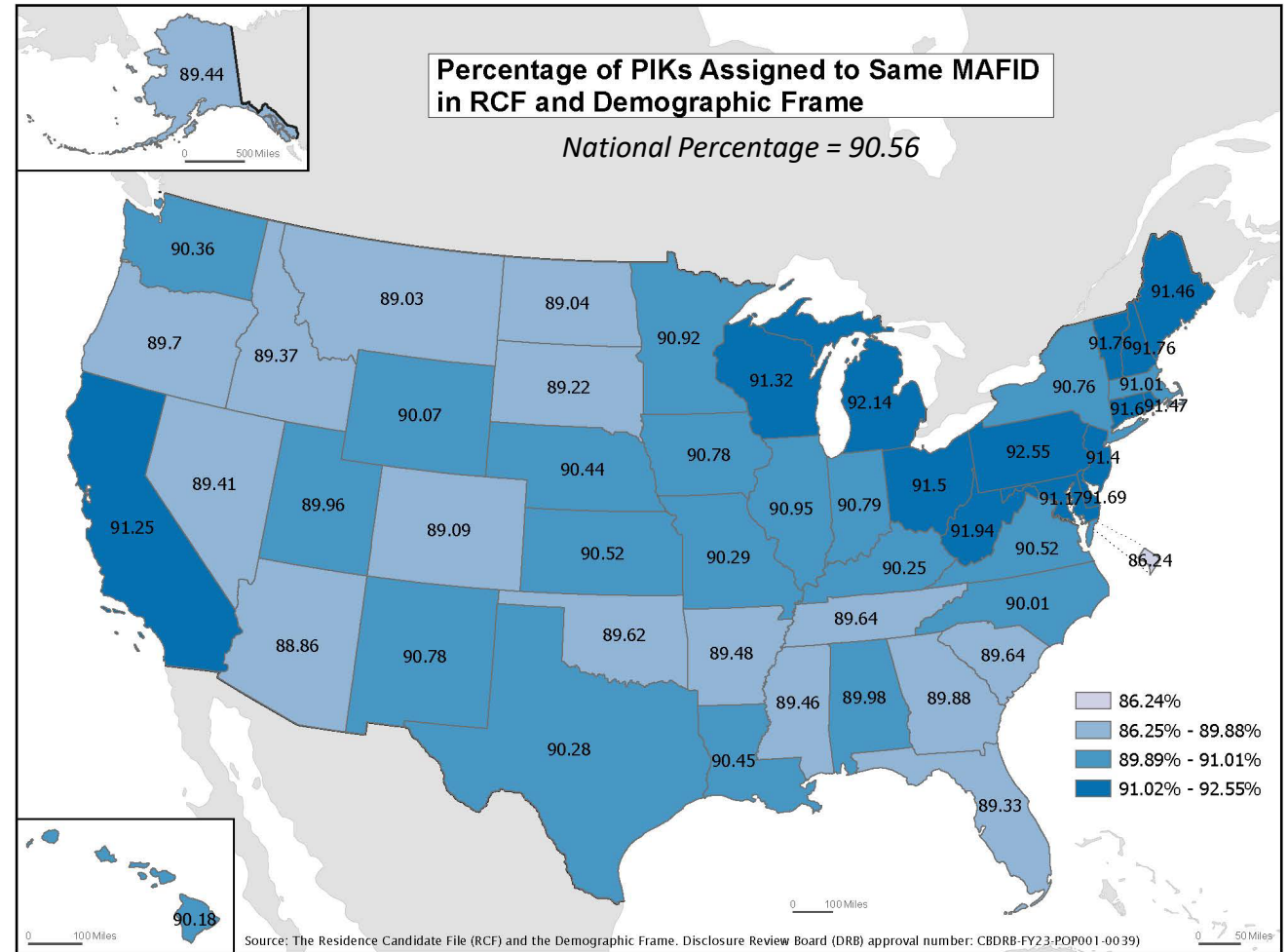
Nonemployer Statistics by Demographics (NES-D)

Nonemployer Statistics by Demographics (NES-D)

- Annual series of nonemployer business counts, receipts, and business owner characteristics based on administrative records and Census Bureau data

    https://www.census.gov/programs-surveys/abs/data/nesd.html

- NES-D is investigating ways to use **demographic characteristics of business owners** from the Demographic Frame to reduce existing efforts to consolidate data sources

Nonemployer Business Owners

Match Rates between 2018 NES-D and 2020 Demographic Frame by Type

Legal Form of Organization	Percent Matched
Partnerships: *Two or more people who combine their resources to form a business*	99%
Individual proprietorships: *Unincorporated business owned and run by one person or couple*	97%
S-Corporations: *Smaller corporation with board of directors, corporate officers, bylaws, and management structure*	97%

Results from Marquette (2023)
Disclosure Review Board (DRB) approval number: CBDRB-FY23-ESMD002-027

Demographic Consistency between 2018 NES-D and 2020 Demographic Frame

Demo Frame Characteristics	Agreement Rate with NES-D
Overall Race*	**90%**
White	96%
Black or African American	92%
American Indian or Alaska Native	83%
Asian	73%
Native Hawaiian or Pacific Islander	71%
Missing	48%
Overall Hispanic Origin	**93%**
Hispanic	92%
Non-Hispanic	98%
Missing	18%
Overall Sex	**99%**
Male	99%
Female	99%
Missing	77%
Overall Age	**88%**

*Following OMB's rules for business statistics and surveys, NES-D does not publish the race category "Some Other Race" or "Two or More Races."
Results from Marquette (2023). Disclosure Review Board (DRB) approval number: CBDRB-FY23-ESMD002-027

ACS Household Size Evaluation

Does the Demographic Frame *put people at the same address* as the ACS?

	Respondent Address Matchs Person Place Model's Best Address	76.8%
ACS Person Record Found on the Frame	Respondent Address is not the Best Address from Person Place Model	9.4%
	Respondent Address Doesn't Match Any Frame Addresses for Respondent	10.3%
ACS Respondent not Found on the Frame		3.5%

~96.5% of ACS respondents are found on the Demographic Frame.
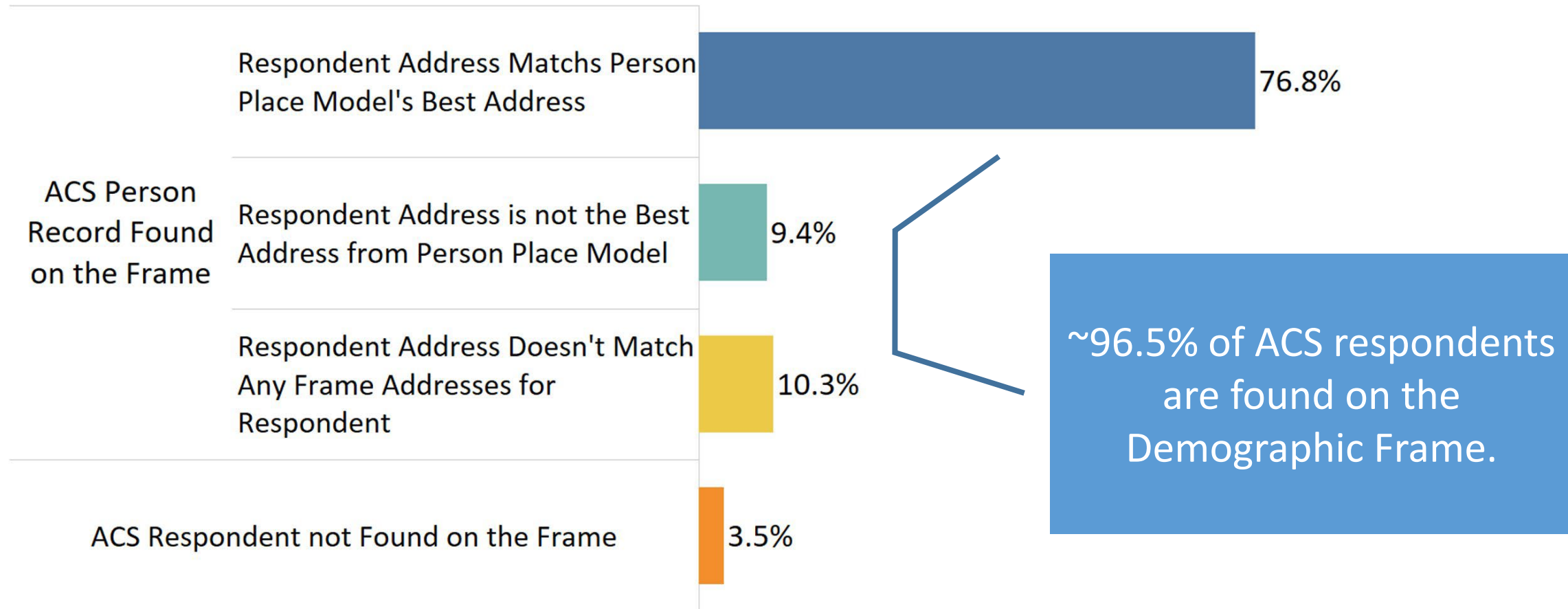
Do Household Counts Match Between the ACS and Demographic Frame?

Percent of Households by Response where Count Matches in 2019 ACS and 2020 Demo Frame

Response	Match %
Internet	64.4%
Mail	70.2%
Personal Interview	48.1%

Results from Sawyer (2023). Disclosure Review Board (DRB) approval number: CBDRB-FY23-ACSO003-B0056

United States® Census Bureau
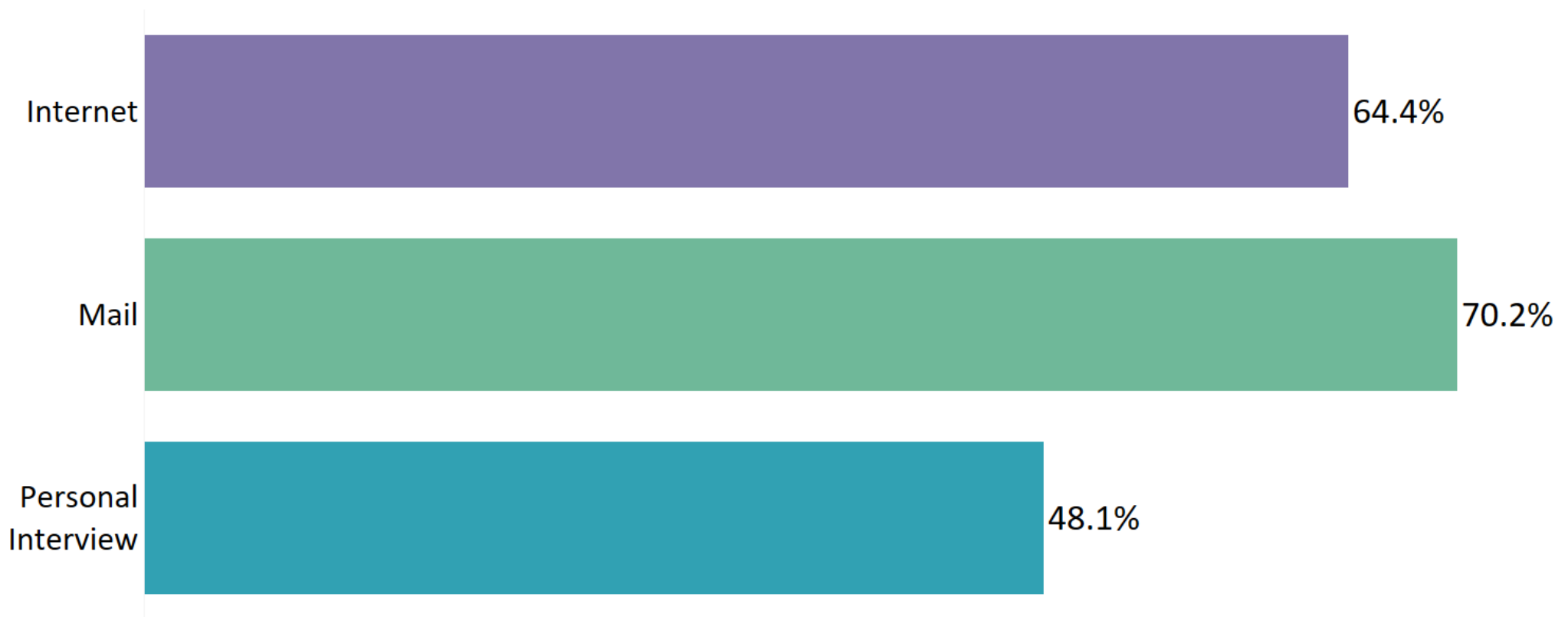
Does the Demographic Frame put people in the same address they are found in the ACS?

	March 2019	July 2019	October 2019
Respondent Address Matchs Person Place Model's Best Address	76.9%	79.3%	80.5%
ACS Person Record Found on the Frame — Respondent Address is not the Best Address from Person Place Model	10.7%	9.2%	8.3%
Respondent Address Doesn't Match Any Frame Addresses for Respondent	9.0%	8.5%	8.3%
ACS Respondent not Found on the Frame	3.4%	3.0%	2.8%

United Census Bureau