Defining Historically Undercounted Populations and Populations at Risk of Coverage Error

Eric Jensen
Senior Advisor for Population Estimates
and Coverage Measurement
Population Division
U.S. Census Bureau

AAPOR 78th Annual Conference
May 10-12, 2023, Philadelphia, PA

This presentation is released to inform parties of ongoing research and to encourage discussion in progress. Any opinions and conclusions expressed herein are those of the author(s) and do not reflect the views of the U.S. Census Bureau.

United States® Census Bureau

"Taking a census is a major statistical undertaking but it is also a magical act and massive public relations task. How do you get the public to agree on a set of numbers designed to measure a moving target for which there is an infinite number of correct answers? The resulting census figure is a patchwork of counting, estimating, and guessing at many different levels. We play mental games in rationalizing what should be included in the census and what should be considered undercount or overcount."

—Jay Siegel (2005)

United States® Census Bureau

Demographic Analysis

- Demographic Analysis (DA) uses historical and current birth and death records, migration data, and Medicare enrollment records to produce population estimates
- The estimates are independent of the census being evaluated
- The DA results are compared to the census to produce estimates of net coverage error

Net Coverage Error =

(Census Count – DA Estimate) / DA Estimate * 100

United States® Census Bureau

Post-Enumeration Survey

- The Post-Enumeration Survey (PES) uses the dual-system estimation (DSE) approach to produce estimates of coverage error in the decennial census
- DSE is a capture-recapture method (akin to catch-and-release studies of wildlife biology)
- The census is one system and a survey conducted immediately after the census is the other system
- Survey respondents are matched to census observations

United States® Census Bureau

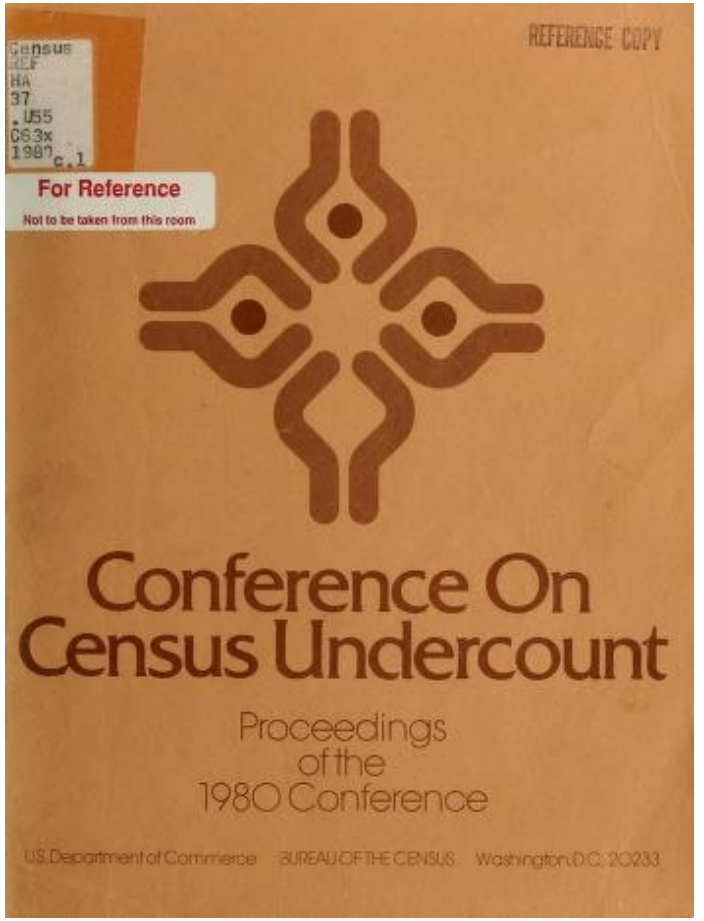
Post-Enumeration Survey (cont.)

		Enumerated in PES?	
		YES	NO
Enumerated in Census?	YES	Found in Both	Missing in PES Found in Census
	NO	Missing in Census Found in PES	Missing in Both

United States® Census Bureau

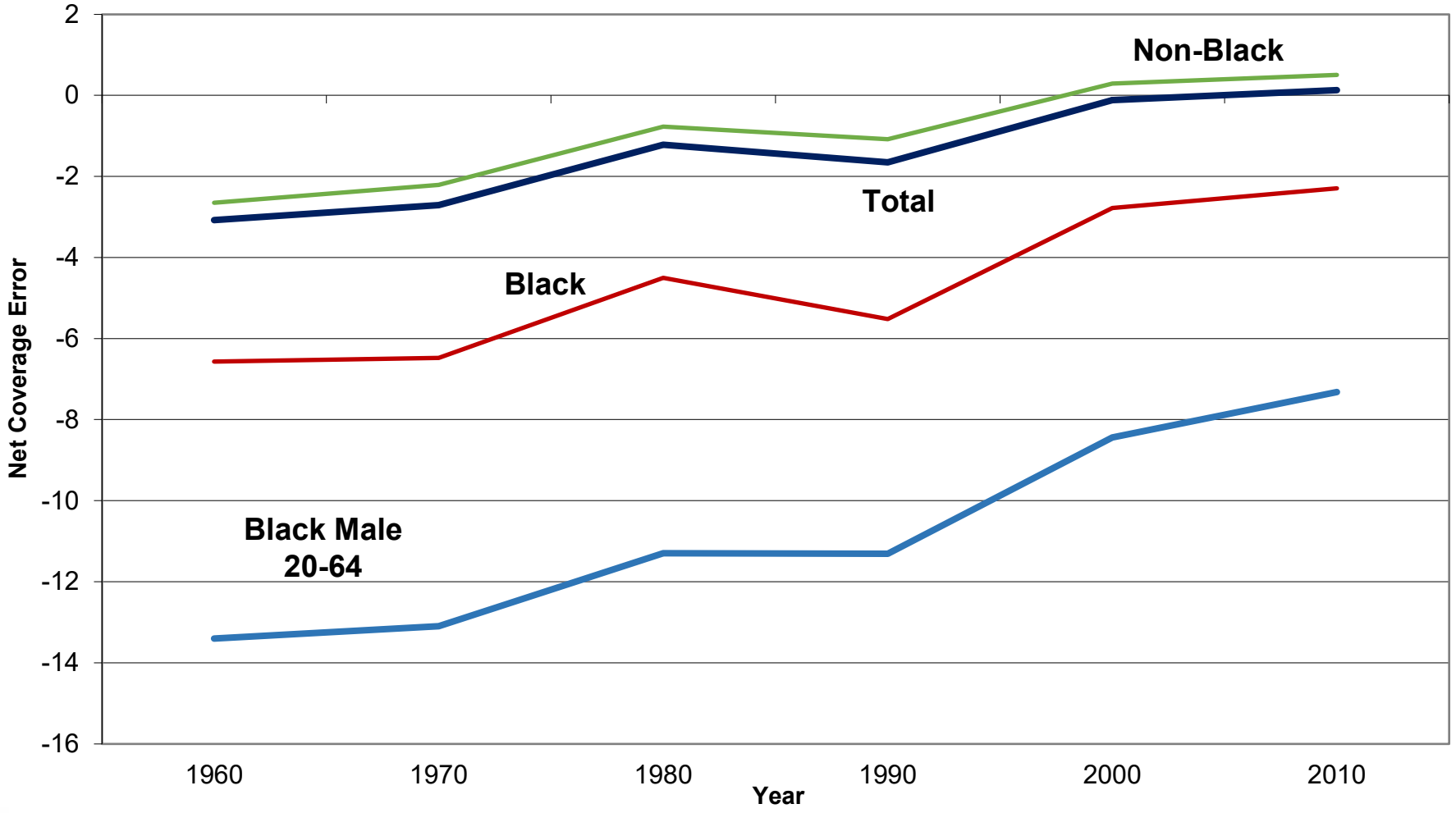
Historical Coverage Patterns

- Coverage error in the census has improved dramatically over the decades
- There are still large differences in coverage by age, sex, race, and Hispanic origin
- Renters consistently have lower coverage than owners



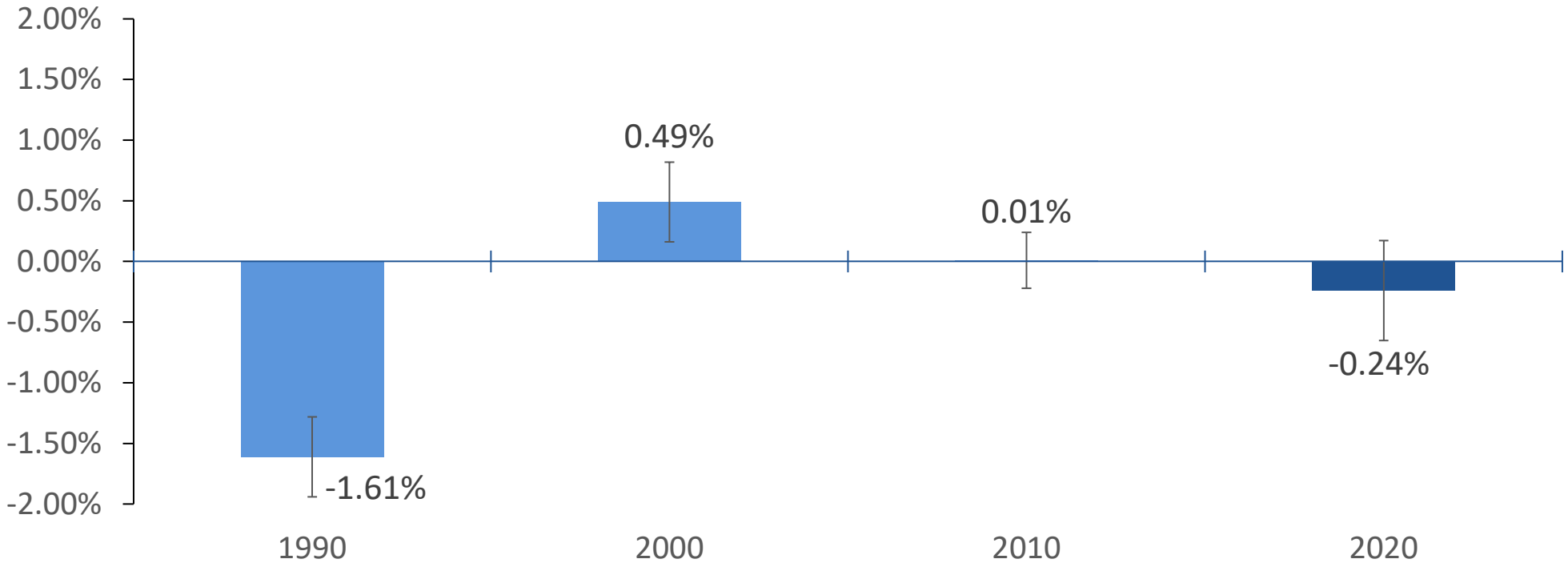
United States® Census Bureau

Demographic Analysis Estimates of Net Coverage Error for 1960 to 2010



Source: U.S. Census Bureau, Demographic Analysis Estimates, special tabulation.

U.S. Total Percent Net Coverage Error: 1990 to 2020



Source: U.S. Census Bureau, Post-Enumeration Survey Program.

United States® Census Bureau

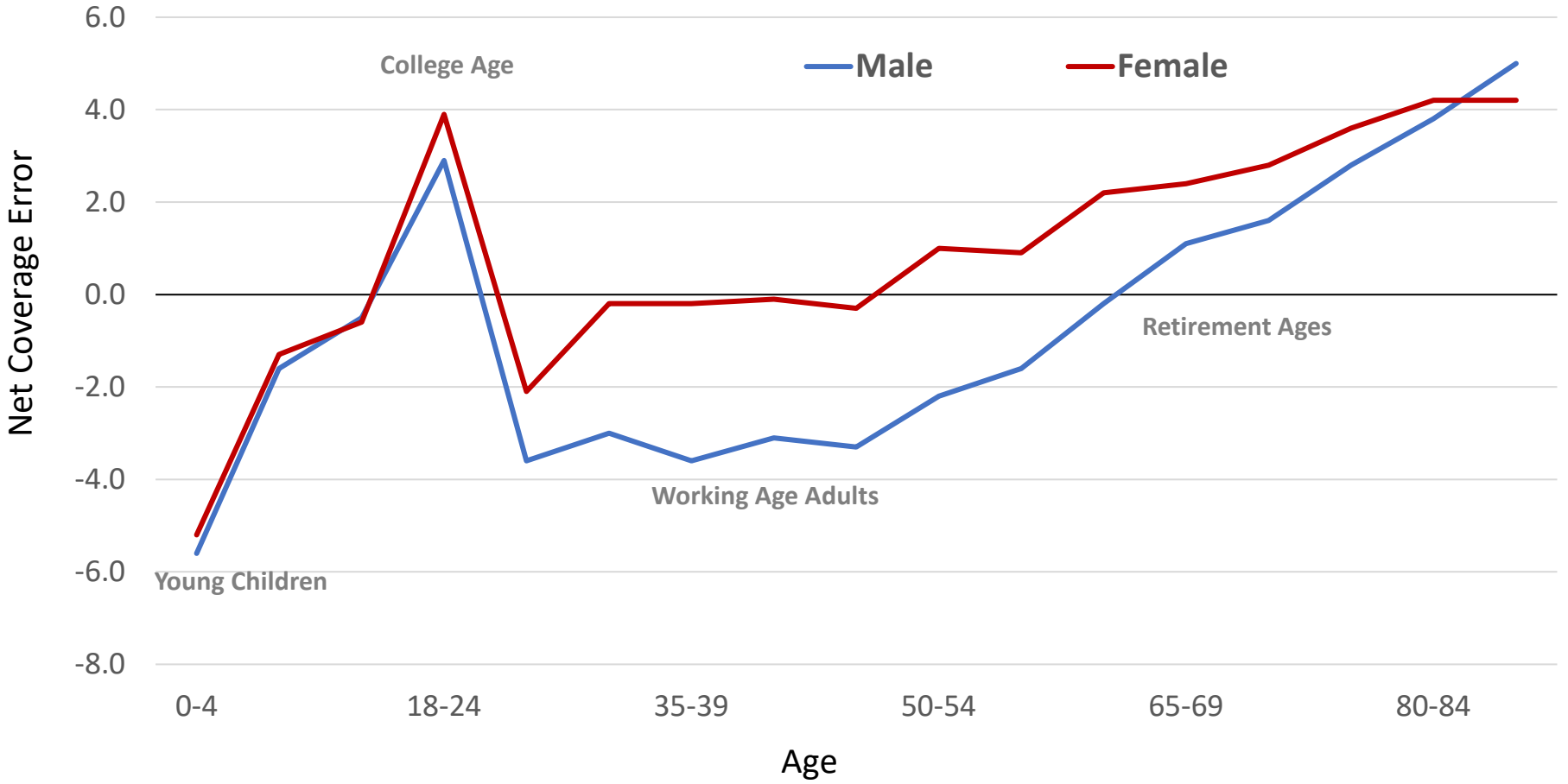
2020 Census Results

- The first PES and additional DA net coverage error estimates were released March 10, 2022
- Findings were shared at a press conference featuring the Census Bureau Director
- The results were some of the most comprehensive evaluations of the 2020 Census



United States® Census Bureau

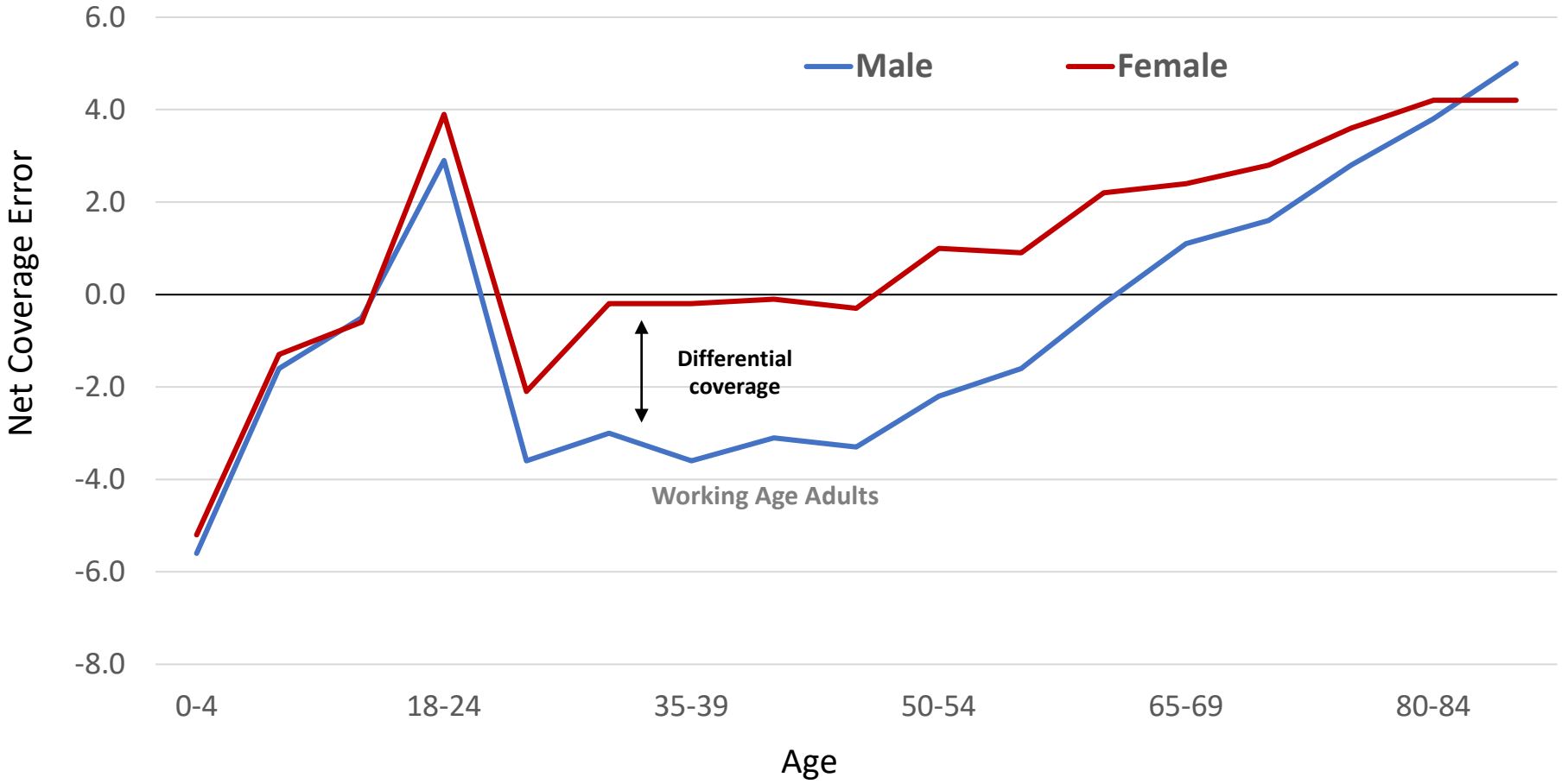
2020 Demographic Analysis Middle Series Net Coverage Error Estimates for Selected Age Groups by Sex: April 1, 2020



Source: U.S. Census Bureau, Population Division, 2020 Demographic Analysis (December 2020 release), and 2020 Census special tabulation (DRB Approval Number: CBDRB-FY22-DSEP-001).

United States® Census Bureau

2020 Demographic Analysis Middle Series Net Coverage Error Estimates for Selected Age Groups by Sex: April 1, 2020



Source: U.S. Census Bureau, Population Division, 2020 Demographic Analysis (December 2020 release), and 2020 Census special tabulation (DRB Approval Number: CBDRB-FY22-DSEP-001).

United States® Census Bureau

Post-Enumeration Survey Net Coverage Error Rates for the Household Population in the United States by Race and Hispanic Origin (In percent)

Race or Hispanic Origin	2020		2010	
	Estimate	Standard Error	Estimate	Standard Error
Total	-0.24	0.25	0.01	0.14
Race alone or in combination with one or more other races				
White	0.66*	0.21	0.54*	0.14
Non-Hispanic White Alone	1.64*	0.21	0.83*	0.15
Black or African American	-3.30*	0.61	-2.06*	0.50
Asian	2.62*	0.77	0.00	0.52
American Indian Alaska Native	-0.91*	0.54	-0.15	0.71
On Reservation	-5.64*	2.72	-4.88*	2.37
American Indian Areas Off Reservation	3.06	2.72	3.86	2.99
Balance of the United States	-0.86*	0.47	0.05	0.58
Native Hawaiian or Other Pacific Islander	1.28	2.11	-1.02	2.06
Some Other Race	-4.34*	0.49	-1.63*	0.31
Hispanic or Latino	-4.99*	0.53	-1.54*	0.33

* Denotes a (percent) net coverage error that is significantly different from zero.
Note: A person can be included in more than one row. A negative (positive) estimate of net coverage error indicates an undercount (overcount).
Source: U.S. Census Bureau, Decennial Statistical Studies Division, 2020 Post-Enumeration Survey (March 2022 Release) and 2010 Census Coverage Measurement Survey.

United States® Census Bureau

Post-Enumeration Survey Net Coverage Error Rates for the Household Population in the United States by Race and Hispanic Origin (In percent)

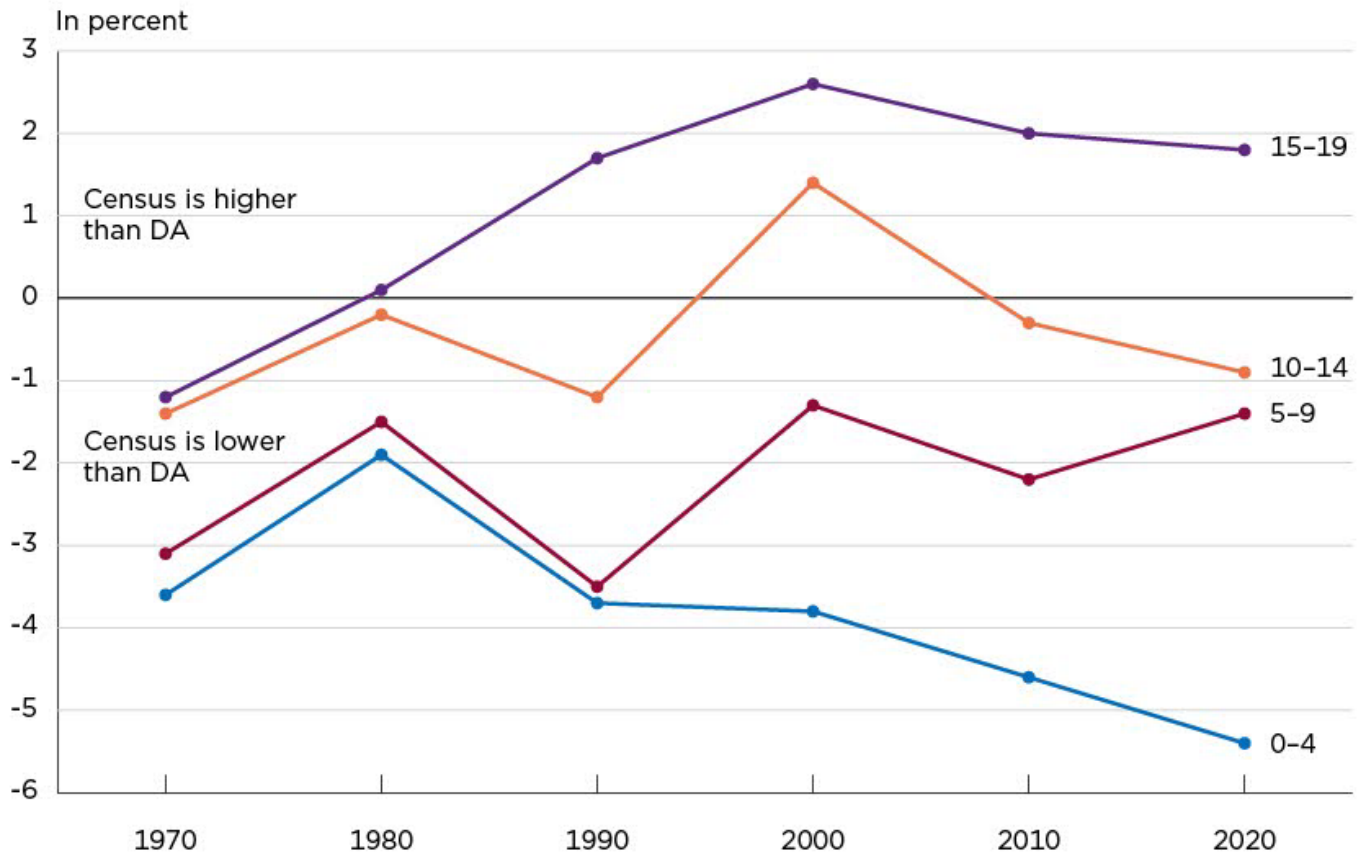
Race or Hispanic Origin	2020		2010	
	Estimate	Standard Error	Estimate	Standard Error
Total	-0.24	0.25	0.01	0.14
Race alone or in combination with one or more other races				
White	0.66*	0.21	0.54*	0.14
Non-Hispanic White Alone	1.64*	0.21	0.83*	0.15
Black or African American	-3.30*	0.61	-2.06*	0.50
Asian	2.62*	0.77	0.00	0.52
American Indian Alaska Native	-0.91*	0.54	-0.15	0.71
On Reservation	-5.64*	2.72	-4.88*	2.37
American Indian Areas Off Reservation	3.06	2.72	3.86	2.99
Balance of the United States	-0.86*	0.47	0.05	0.58
Native Hawaiian or Other Pacific Islander	1.28	2.11	-1.02	2.06
Some Other Race	-4.34*	0.49	-1.63*	0.31
Hispanic or Latino	-4.99*	0.53	-1.54*	0.33

* Denotes a (percent) net coverage error that is significantly different from zero.
Note: A person can be included in more than one row. A negative (positive) estimate of net coverage error indicates an undercount (overcount).
Source: U.S. Census Bureau, Decennial Statistical Studies Division, 2020 Post-Enumeration Survey (March 2022 Release) and 2010 Census Coverage Measurement Survey.

United States® Census Bureau

Figure 1.

Demographic Analysis (DA) Estimates of Net Coverage Error by Select Age Groups: 1970–2020

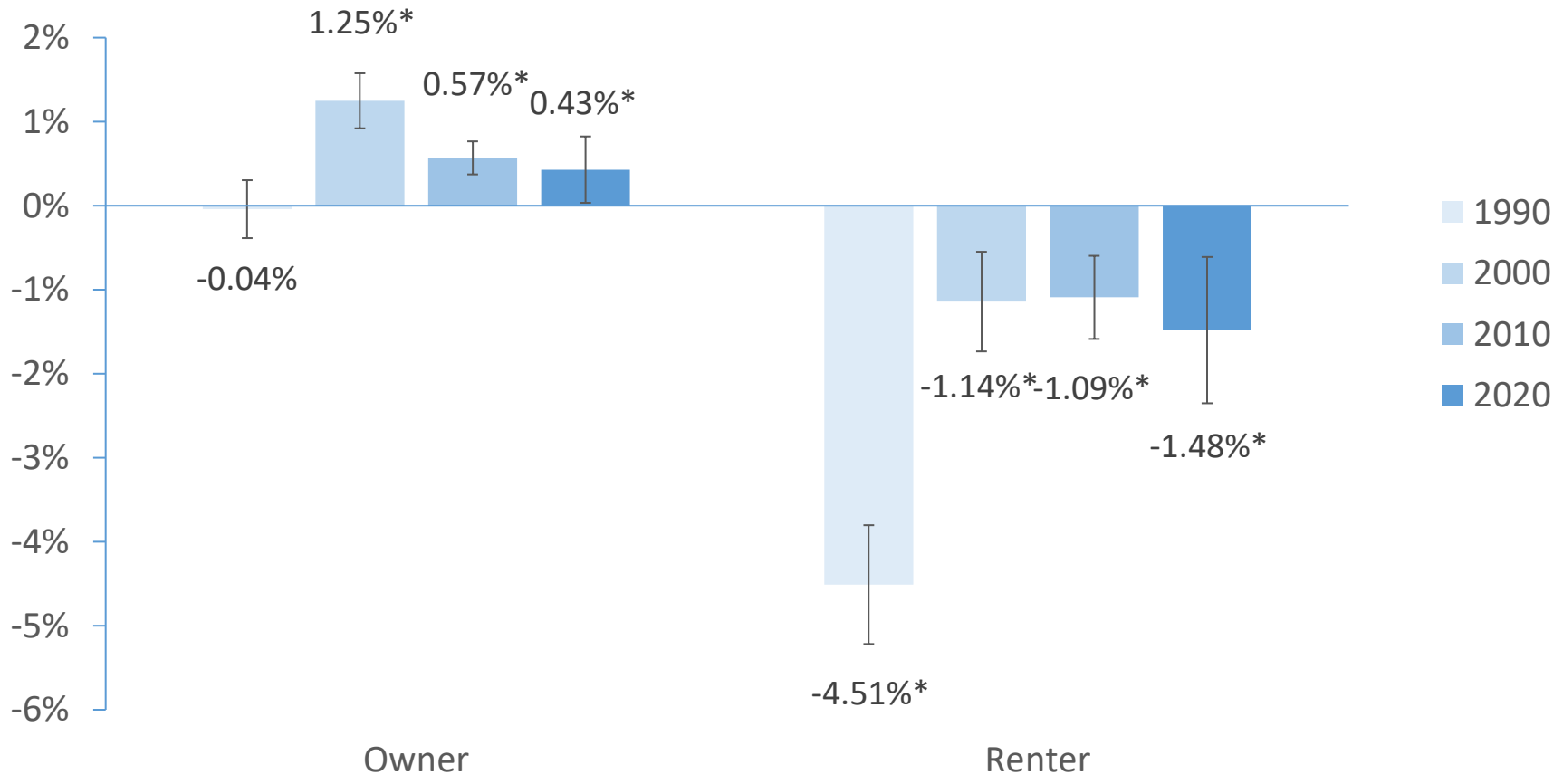


Note: The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. Demographic Analysis: CBDRB-FY22-DSEP-001.

Source: 1970, 1980, 1990, 2000, 2010, and 2020 Census and Demographic Analyses.

United States® Census Bureau

Percent Net Coverage Error for People by Tenure: 1990 to 2020



Source: U.S. Census Bureau, Post-Enumeration Survey Program.

Defining Historically Undercounted Populations

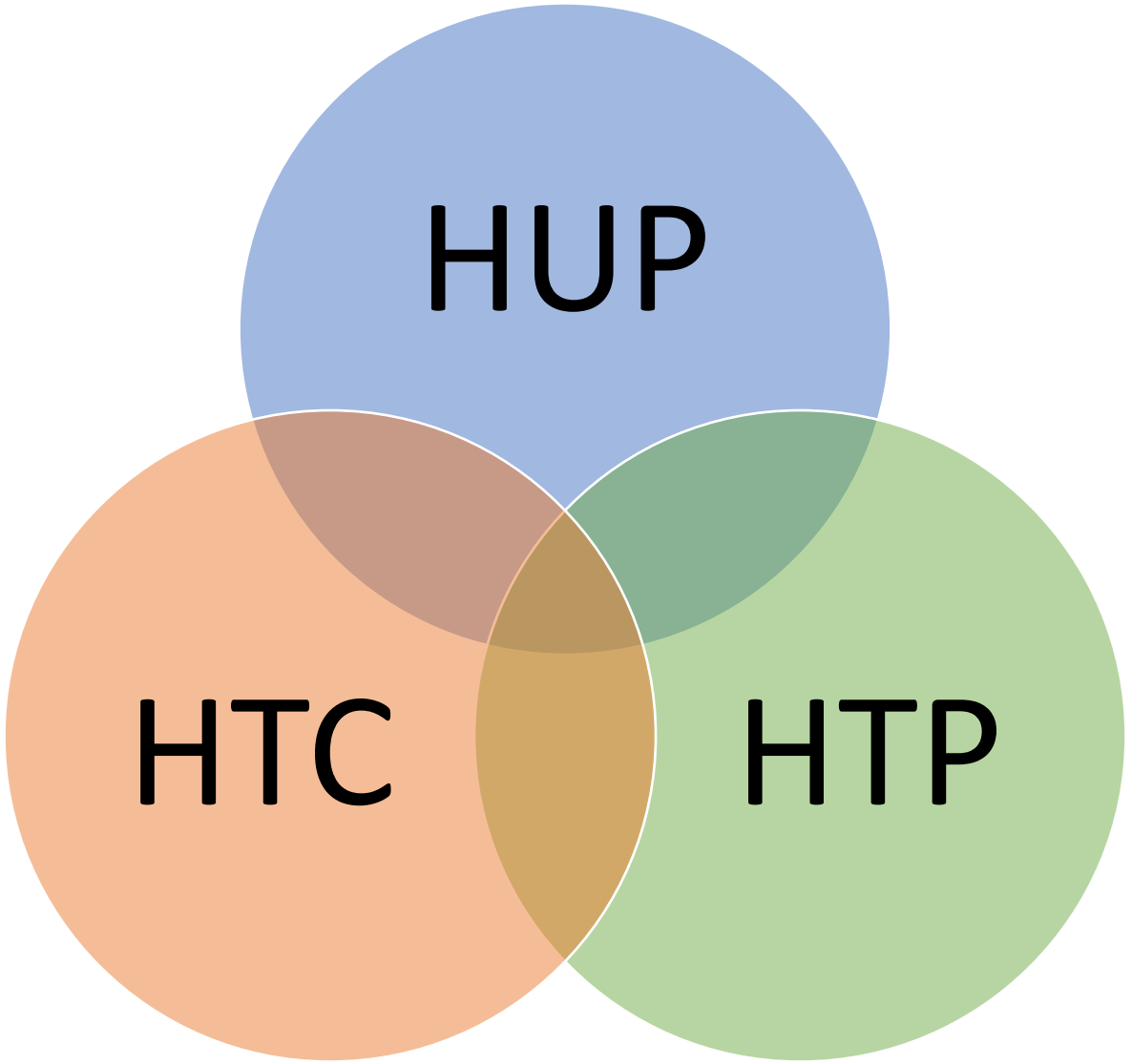
- We can only produce official coverage estimates for the characteristics in the full decennial census
  - Pre-2010, this means the short-form items
- We cannot produce official coverage error estimates by poverty status, English language ability, household structure, or other characteristics that people assume are related to the undercount
- Should our definition of Historically Undercounted Populations (HUPs) be limited to just the characteristics in the coverage measures?

United States® Census Bureau

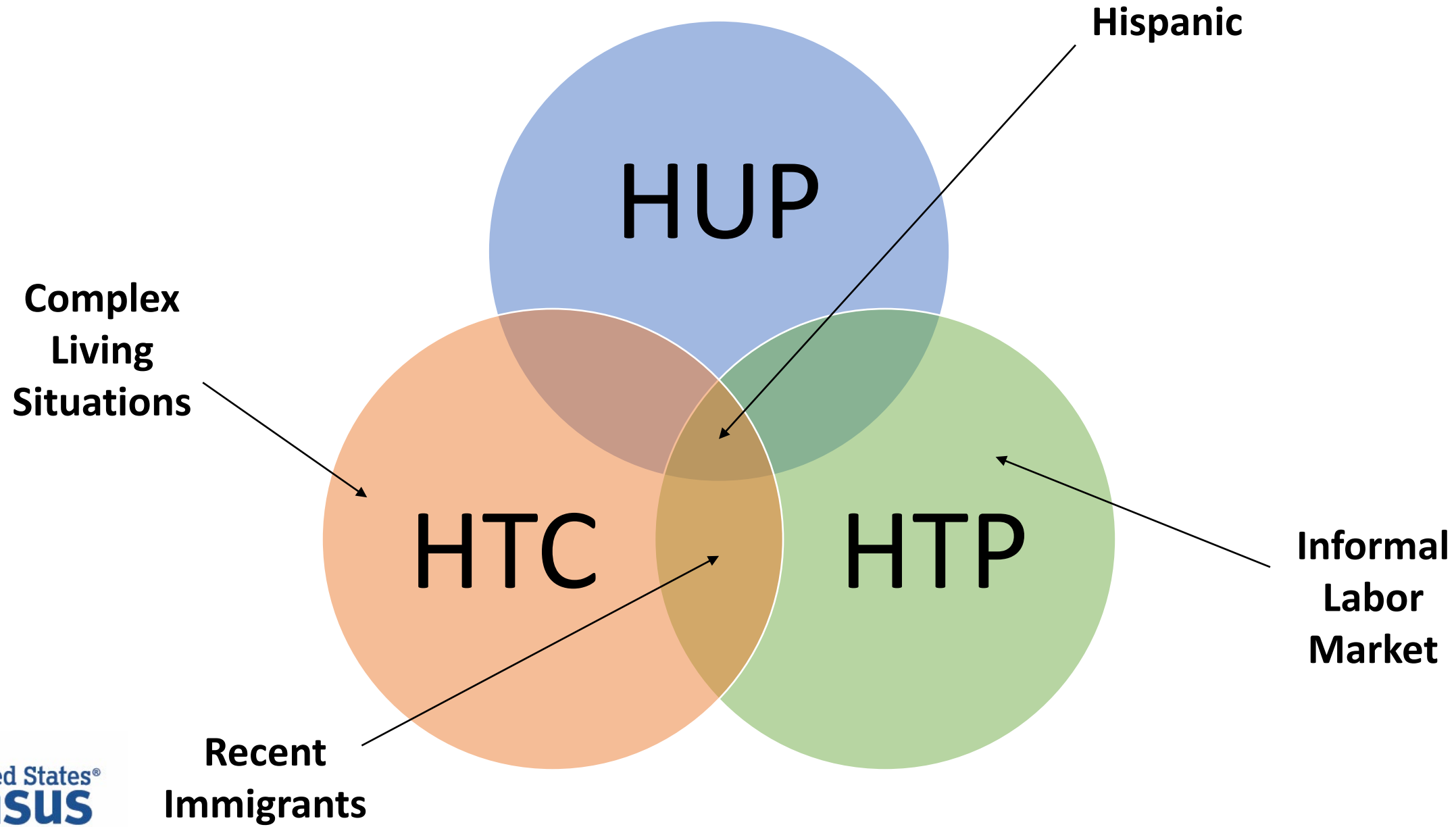
What about HTCs and HTPs

- Numerous qualitative studies have identified additional groups who may not be fully enumerated in the census
- In the 1990s, the Census Bureau started using the term “Hard to Count” to describe undercounted populations and groups that research has shown are difficult to enumerate or survey
  - Stakeholders have requested that we use different terminology
- In addition, some groups may not be found in Administrative Records, which makes them more challenging to assign a Personal Identification Key (PIK)
  - This makes them “Hard to PIK” (HTP)

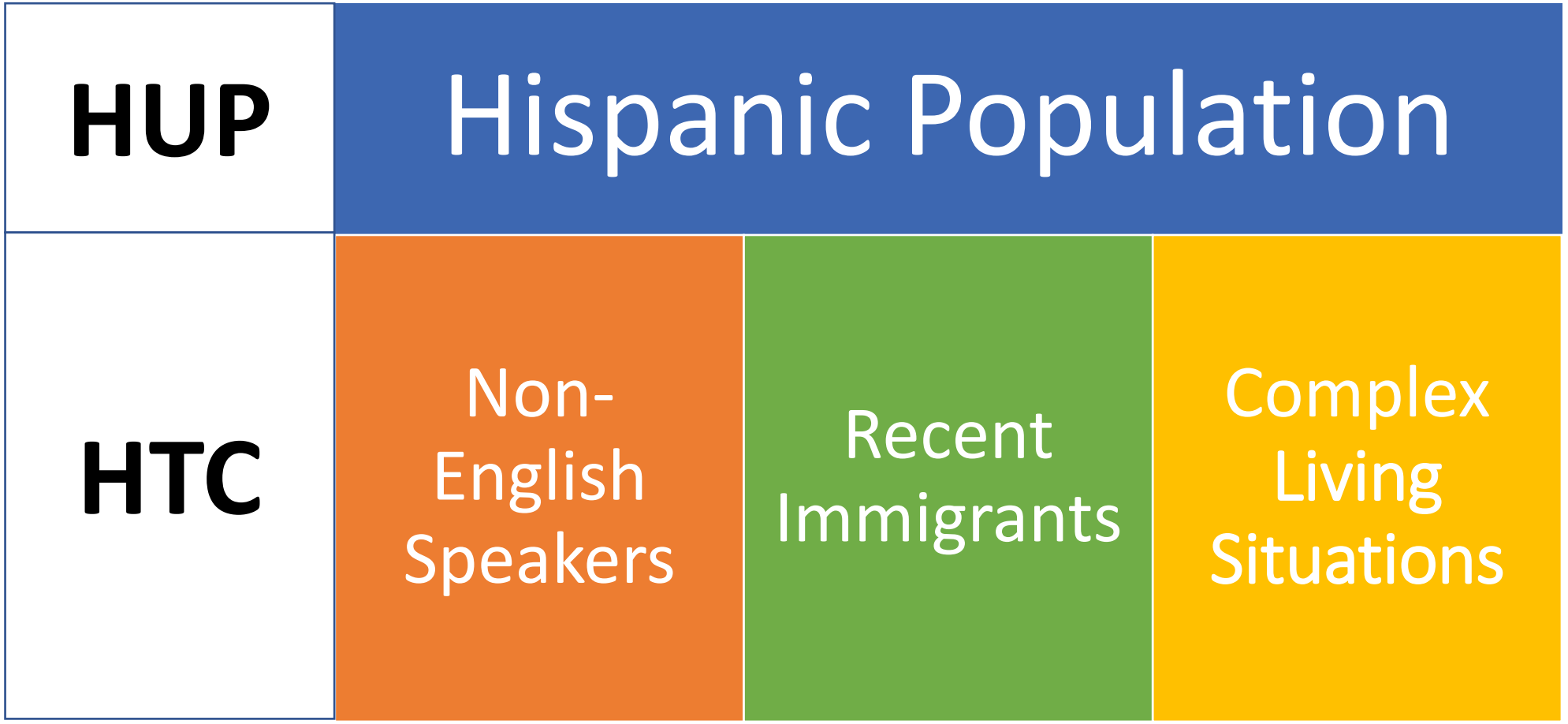
United States® Census Bureau



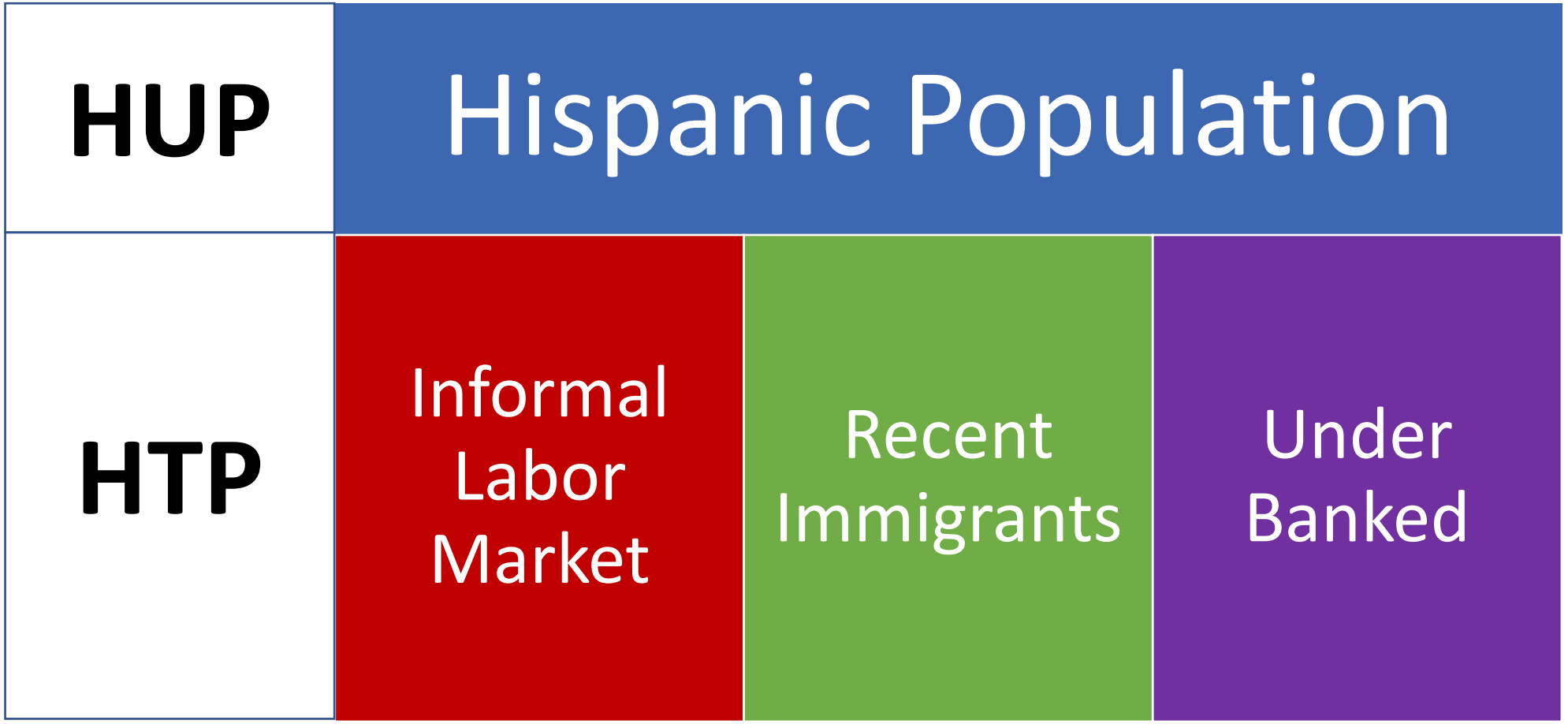
United States® Census Bureau



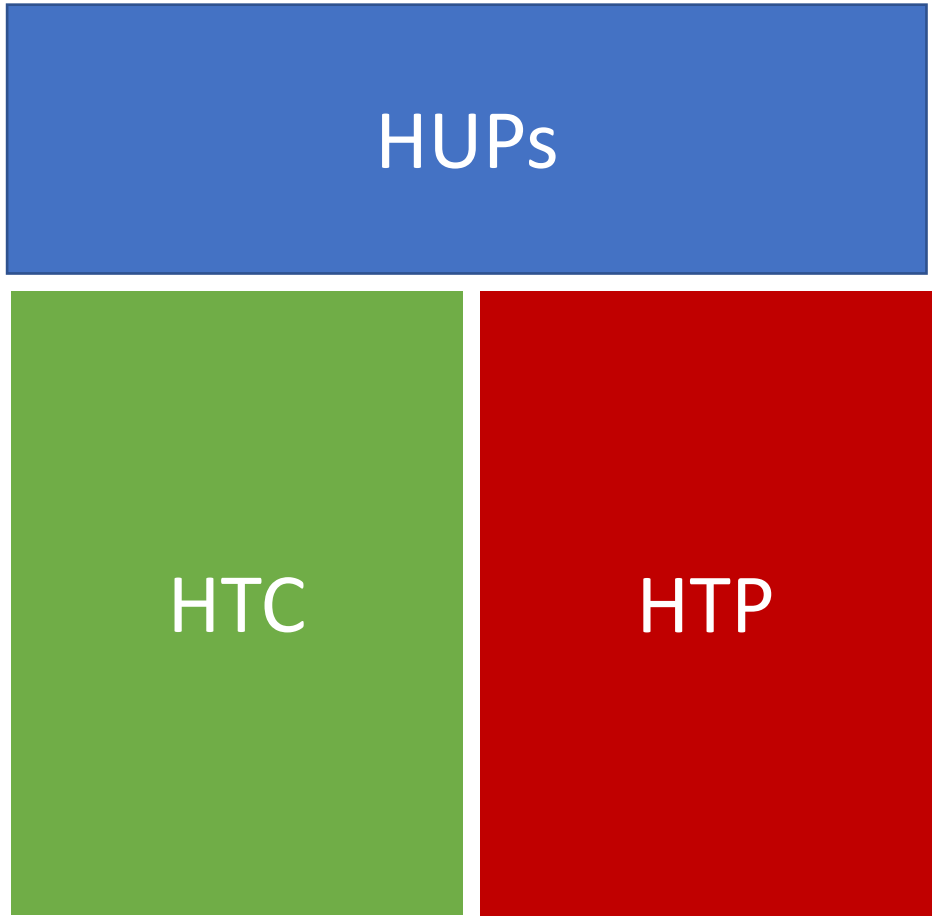
United States® Census Bureau



United States® Census Bureau

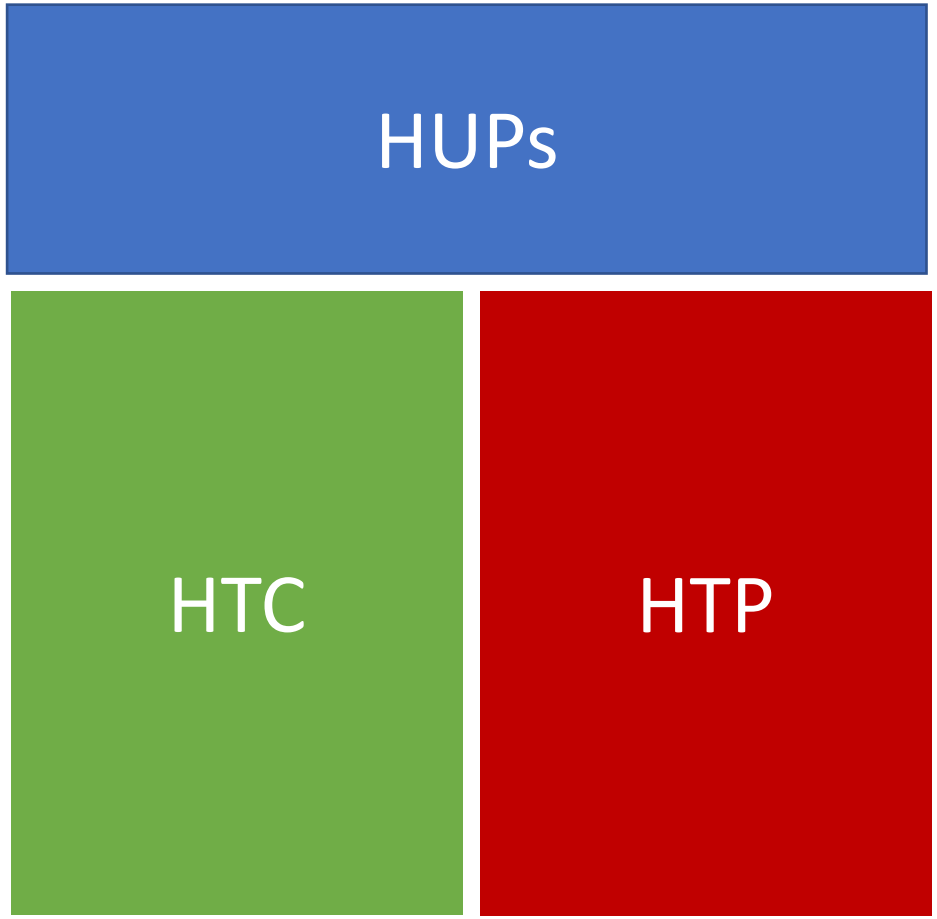


United States® Census Bureau



United States® Census Bureau

Populations at Risk of Coverage Error



United States® Census Bureau

Data Equity

Populations at Risk of Coverage Error

Disaggregation by Race and Ethnicity

Statistics on Hidden Populations

United States® Census Bureau

Conclusion

- How *should* we be defining HUPs?
- The definition will have implications for:
  - Research designed to improve coverage error
  - Outreach to communities and organizations
  - Planning for the 2030 Census
- There is currently a lot of interest in HUPs among staff at the Census Bureau
- Not having a clear definition is leading to a productive dialogue around why undercounts matter and what we can do to improve them

United States® Census Bureau

Thank you!

eric.b.jensen@census.gov

United States® Census Bureau