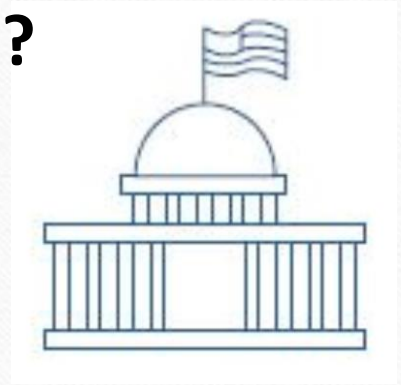
Methods Used by the Census Bureau to Measure the Accuracy of the 2020 Census Count

Data Science for the Public Good Seminar
June 26, 2023

Joseph Salvo
Institute Fellow
University of Virginia Biocomplexity Institute

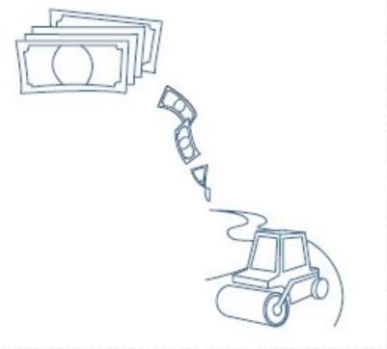
Why is a Good Census Count Important?



- **Reapportionment**
  Constitutional mandate to ensure appropriate delegation size for each state in the House of Representatives

- **Redistricting**
  Drawing Political Districts
  Based on data that reflect the actual population of our neighborhoods

- **Allocating Federal Funds**
  Money for programs and services that contain a population component -- more than 2.8 trillion dollars annually (as of FY 2021)



- **Local Planning**
  Foundation for Emergency management, school construction…and our public health response to epidemiological crises

How Do we Know Whether a Census Count is a Good Count?

Demographers calculate census "coverage" or the **completeness** of a census enumeration in two ways:

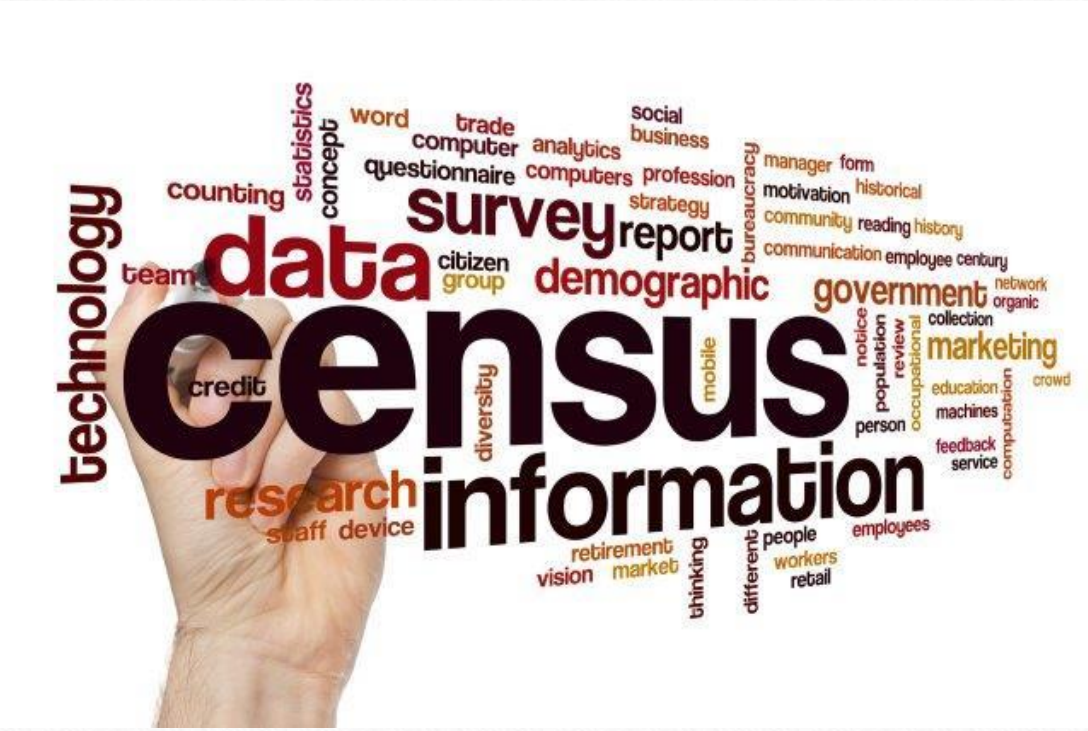
1. Demographic Analysis (DA)
2. Post Enumeration Survey (PES)

1. Demographic Analysis (DA)

- Uses current and historical vital records, data on international migration, and Medicare enrollment to produce an independent population count
- Results are compared to the census counts to evaluate “net coverage error”

Demographic Analysis (DA)

- **Frequency:** Conducted every 10 years just prior to the release of the census count for apportionment
- **Geographic Detail:** National only
- **Subgroup Detail:** Single years of age, sex, limited race (black/non-black) and Hispanic origin for ages 0-29



Building the Demographic Analysis (DA) Count			
United States			
April 1, 2020 (Middle Series)			
		Number	
Total Population (000s)		332,601	
	Births	288,908	
	Deaths	(22,412)	
	International Migrants	44,256	
	Medicare Enrollees	21,849	
Source: U.S. Census Bureau. Population Division,			
	2020 Census Demographic Analysis		

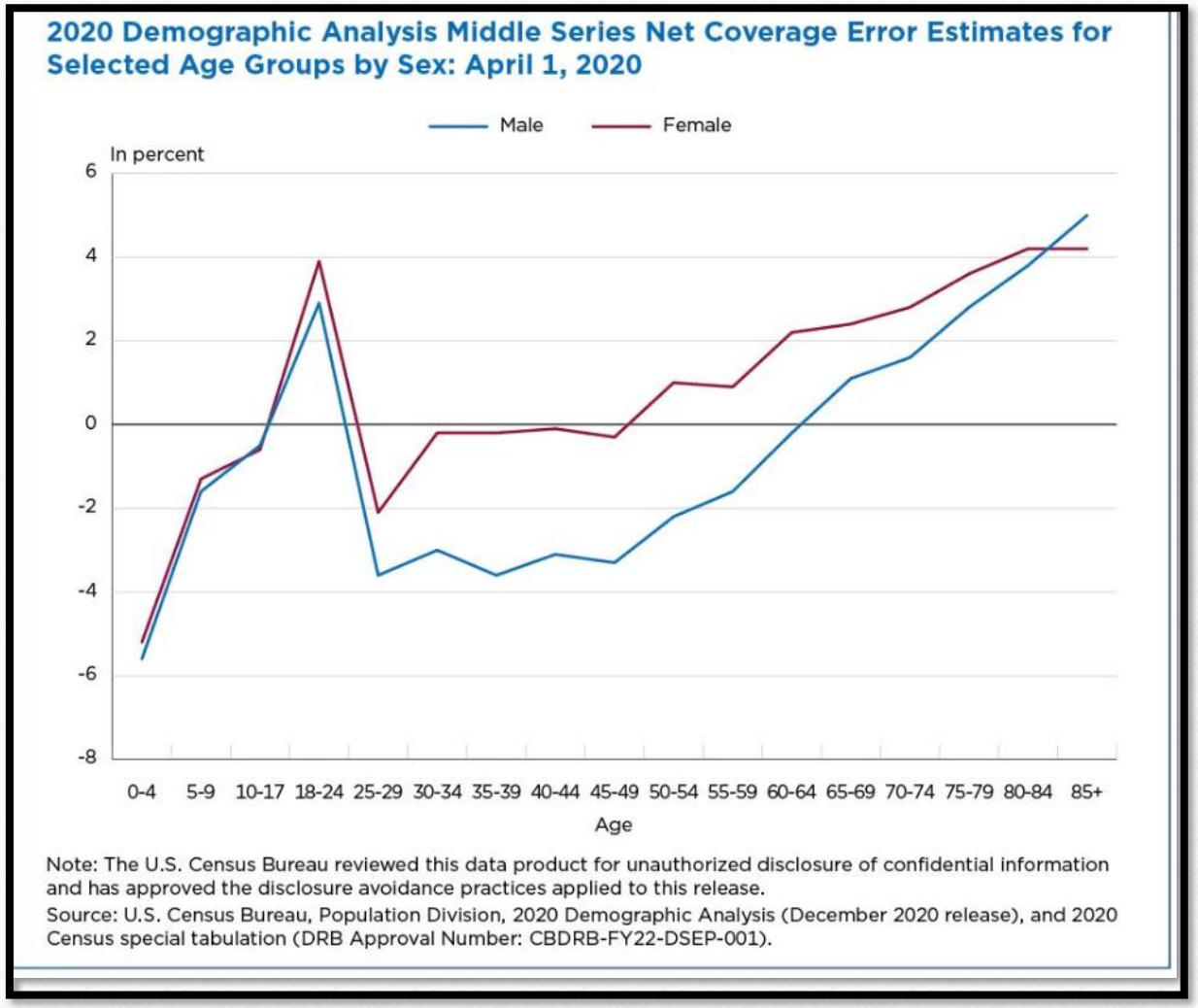
331,449,281

2020 and 2010 Census Coverage Results: DA

Year	DA Net Coverage Error (percents)
2010 Census	0.13
2020 Census	-0.35

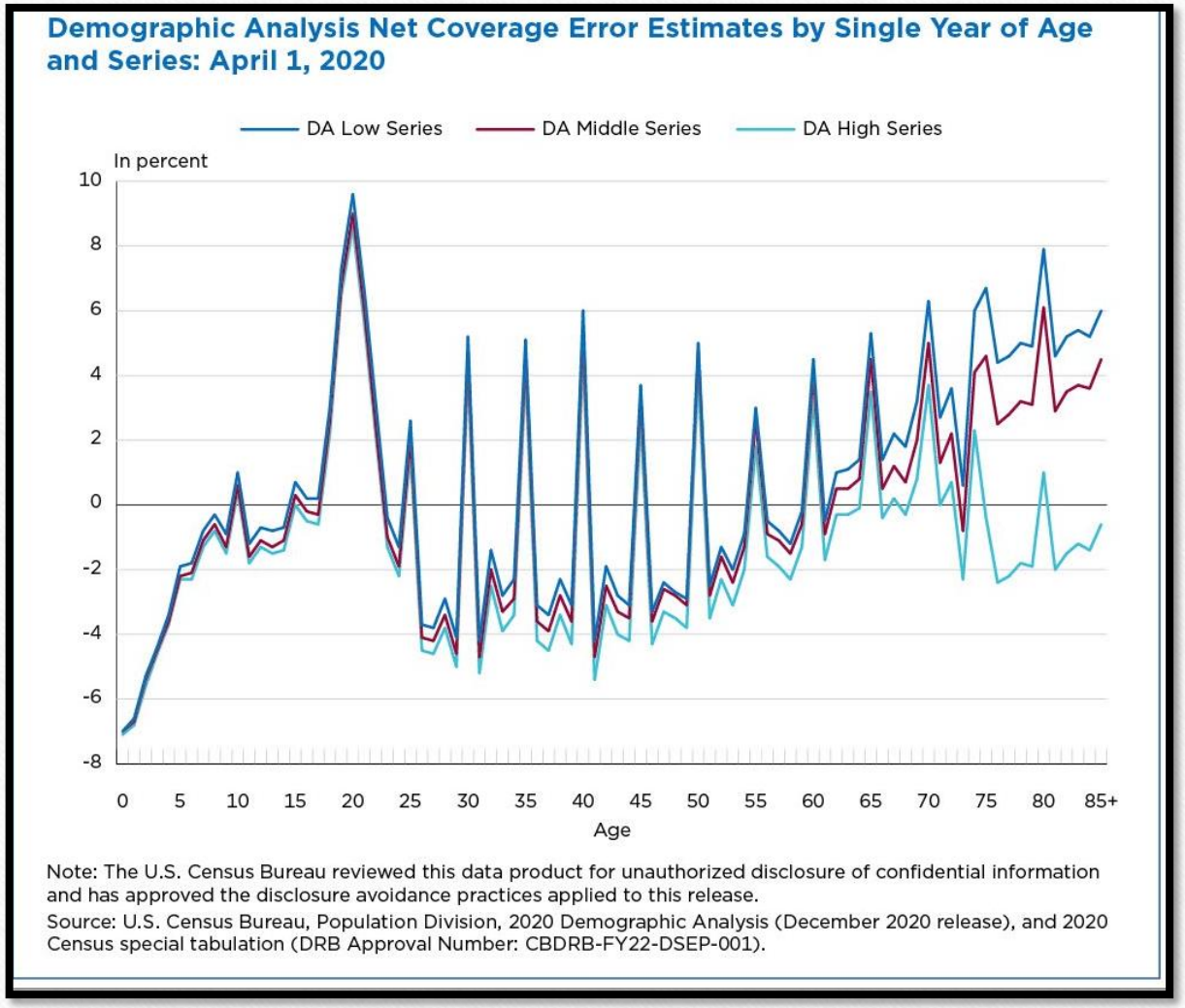
Source: U.S. Census Bureau, 2010 and 2020 Demographic Analysis Estimates, Middle Series.

According to DA, the 2020 Census had a net undercount of 0.35 percent, compared with a net overcount of 0.13 percent in 2010.



- If the census count exactly matched the DA count, all points on the line would be at zero, indicating no difference.
- **Net undercounts in the census fall below the line**, as is the case with those 0-4 years of age
- **Net overcounts in the census are above the line,** as with those 18-24 years and those in the oldest ages.

What are the reasons for these undercounts and overcounts?



- This chart compares the DA count to the Census count by **single years of age**.
- **Net undercounts in the census are below the line.**
- **Points above the line represents overcounts.**
- **Age heaping -** high overcounts in the census at ages ending in 0 and 5. ***Reflects the tendency of respondents to round age.***
- was much higher than in previous censuses.

Why do you think age heaping was higher in the 2020 census?

2. Post-Enumeration Survey

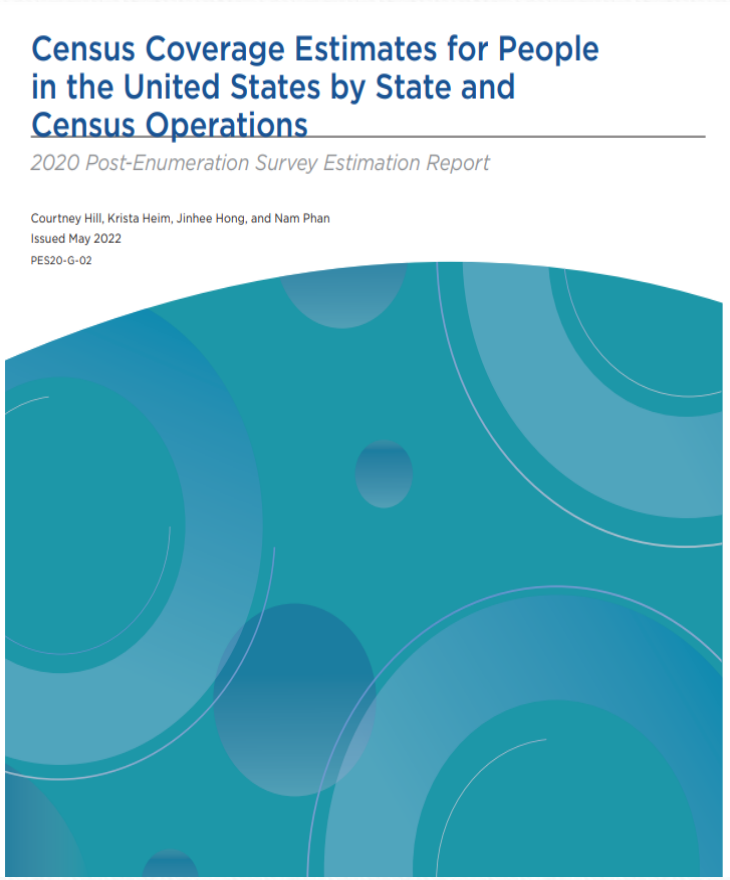
- Independently enumerating a sample of blocks in the census and comparing those results with the same blocks in the decennial census
- Interviews are conducted with all housing units in each of the PES blocks and occupants are matched to their respective records in the decennial census

Post-Enumeration Survey (PES)

- Household population only (323,200,000 census count in 2020)

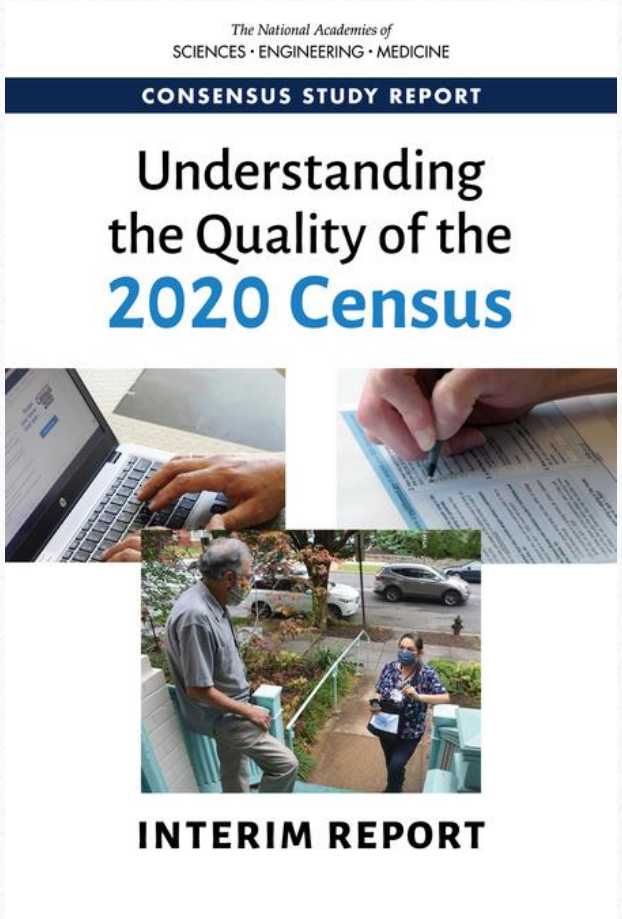
- Based on a sample of blocks, which makes it subject to sampling error

- Big Advantage: PES estimates are available at subnational levels -- each of the 50 states and DC for 2020



PES Method in 4 Steps

1. **Prior to the beginning of the 2020 census**, the Census Bureau creates an independent list of addresses (on 10,000 blocks nationwide)
2. **Immediately after the census**, field staff interview housing units in these blocks, asking residents where they lived on April 1, 2020, about 161,000 housing units in total



PES Methods (continued)

3. **Information collected for the housing unit and occupants** is matched to 2020 Census data to determine whether people were or were not counted
4. **An estimate of the total U.S. population** is derived after matching and field follow-up to resolve unmatched cases

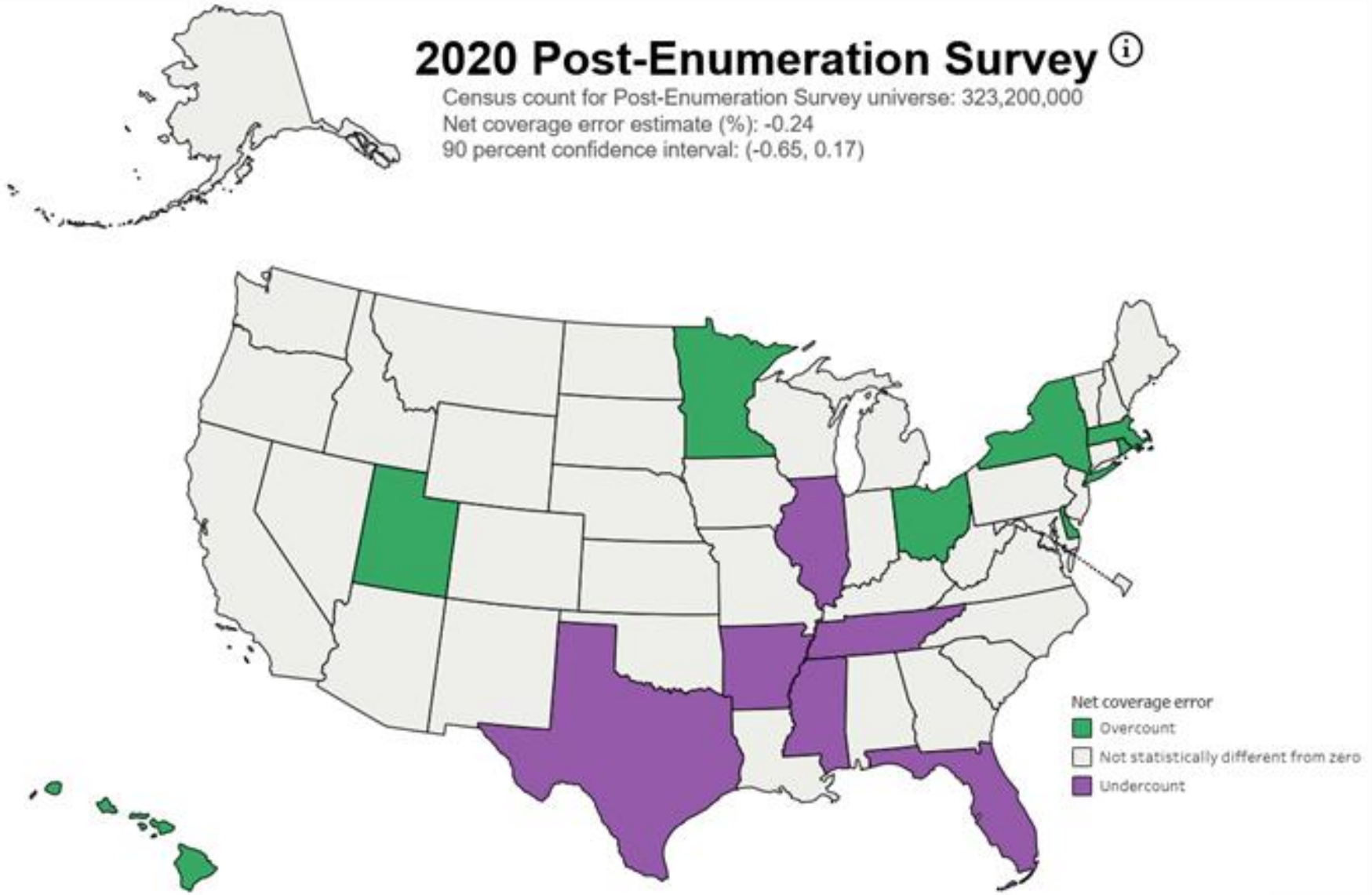
2020 Census Coverage Results: PES

Year	PES (percents)	
	Net coverage error	Standard error
2010 Census	0.01	0.14
2020 Census	-0.24	0.25

Source: U.S. Census Bureau, 2010 and 2020 Post-Enumeration Surveys and Demographic Analysis Estimates.

Overall differences between the PES and the Census were not significantly different from zero in both 2010 and 2020.

Differences for geographic areas and population subgroups are another matter.


2020 Post-Enumeration Survey
Census count for Post-Enumeration Survey universe: 323,200,000
Net coverage error estimate (%): -0.24
90 percent confidence interval: (-0.65, 0.17)
Net coverage error
Overcount
Not statistically different from zero
Undercount

Net Coverage Error Rates (Percents) for the Population by Race/Hispanic Origin					
United States					
2020 Census					
		2020	SE	2010	SE
Total (Household Population)		-0.24	0.25	0.01	0.14
White Non-Hispanic		*1.64	0.21	*0.83	0.15
Black or African American AOIC		*-3.30	0.61	*-2.06	0.50
Asian AIOC		*2.62	0.77	0.00	0.52
American Indian AOIC on Reservations		*-5.64	2.72	*-4.88	2.37
Hispanic or Latino		*-4.99	0.53	*-1.54	0.33
*Percent net coverage error is significantly different from zero					
AOIC=alone or in combination					
Source: Khubba, S, K. Heim, and J. Hong. (2022) “2020 Post-Enumeration					
Survey Estimation Report," PES20-G-01, U.S. Census Bureau, March					

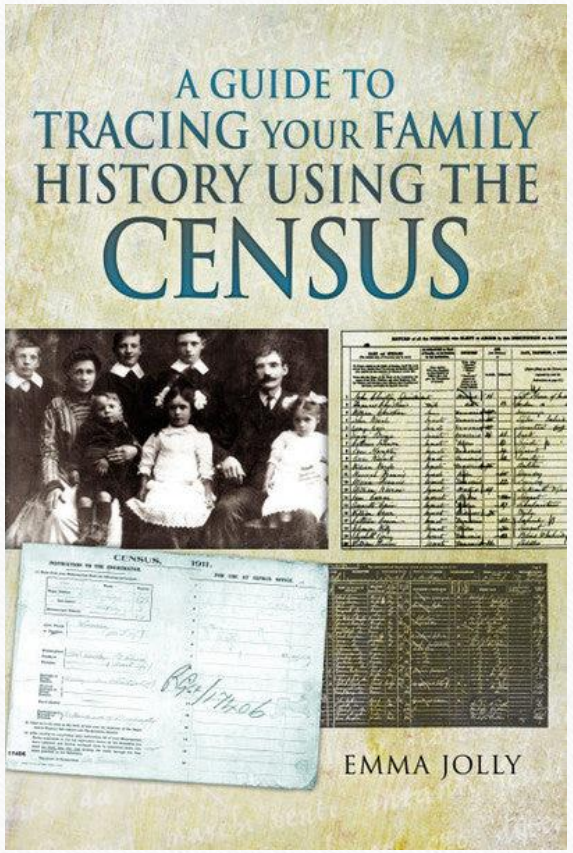
A Few Takeaways

- The Census is a civic ceremony mandated by the Constitution
- Although no census is perfect, it is still the “gold standard”
- Understanding the limitations of a census is important
- Changes are likely on the horizon as a result of understanding these limitations

Thank you!

United States Census 2020 Historical Population Data

1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020



UVA | BIOCOMPLEXITY INSTITUTE