# The 2020 Demographic Analysis Program

Eric B. Jensen

Population Division, U.S. Census Bureau

Southern Demographic Association

October 23-25, 2019

*This presentation is released to inform interested parties of research and to encourage discussion. Any views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.*

United States™ Census Bureau

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

# Census Coverage

- “Count everyone once, only once, and in the right place”
- Coverage error occurs when groups are overcounted or undercounted in the census
- Census Bureau uses two methods to estimate coverage:
  - Demographic Analysis (DA)
  - Post-Enumeration Survey (PES)

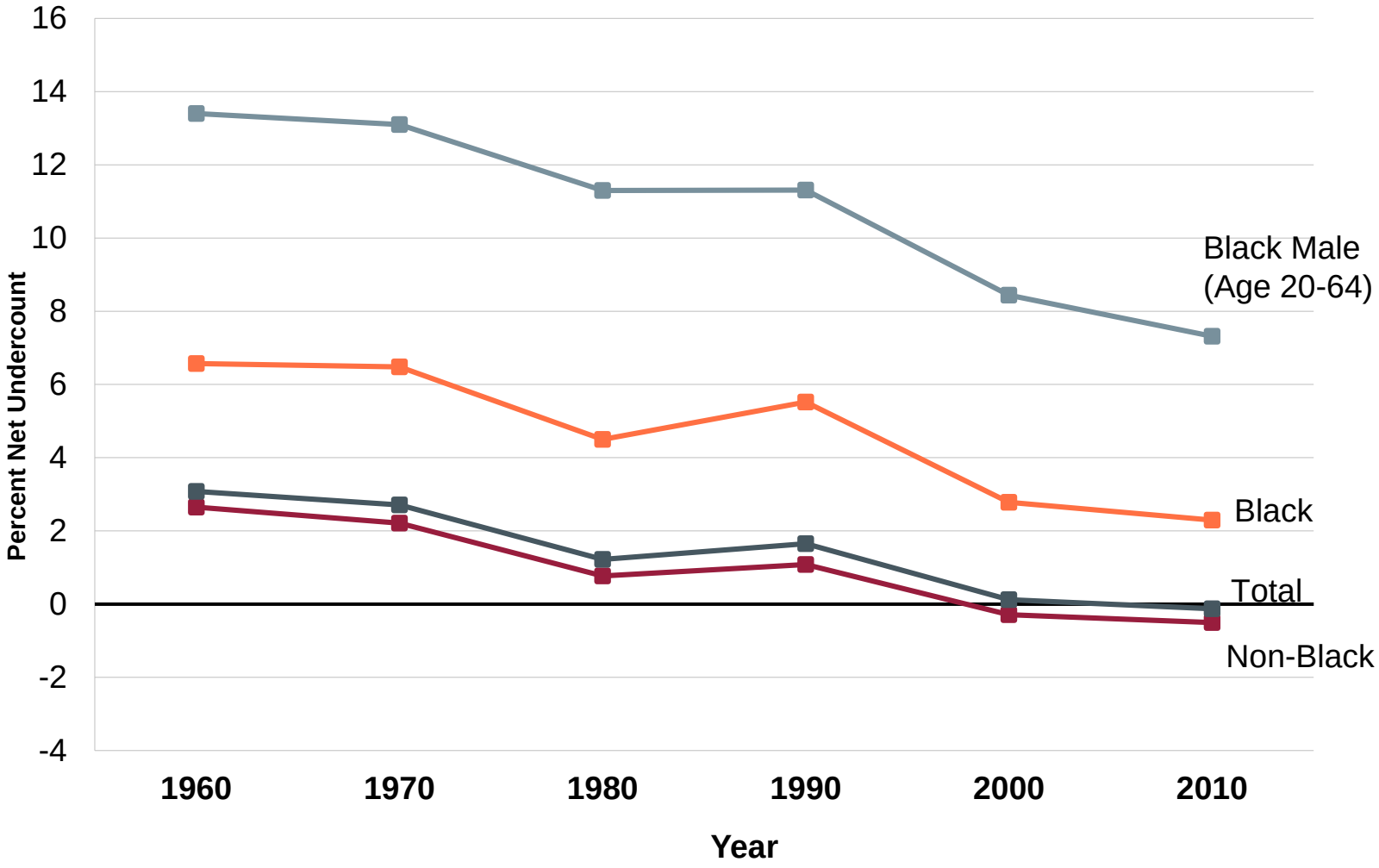
# Demographic Analysis

- The U.S. Census Bureau has used DA to measure net undercount and overcount in each decennial census since 1960
- DA techniques build an estimate of the size and composition of the population based on historical data on births, deaths, and international migration
- National-level estimates by age, sex, race, and Hispanic origin

# Benefits of DA

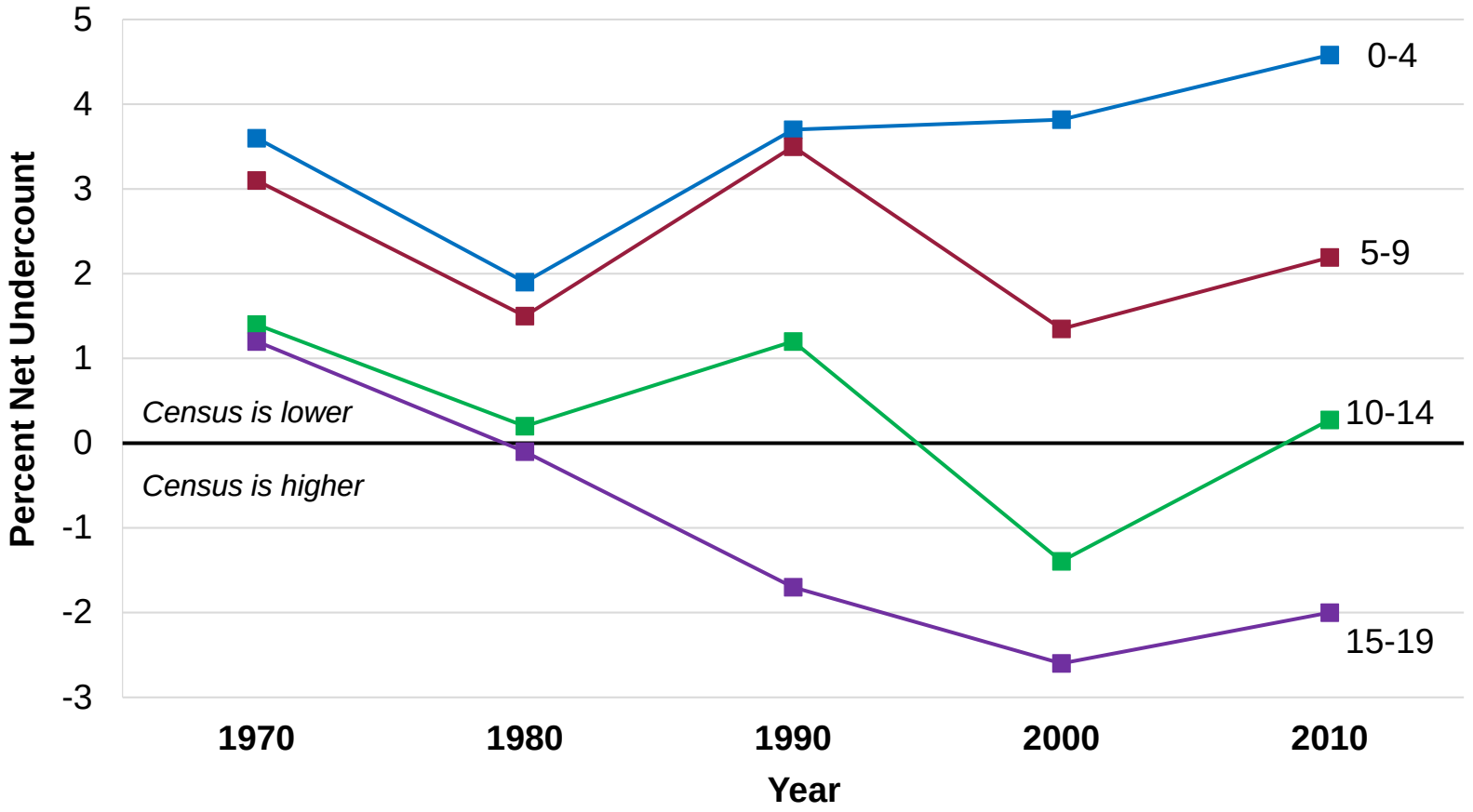
- The DA estimates are independent of the census that is being evaluated
- DA is a relatively inexpensive and non-intrusive evaluation of the decennial census
- DA sex ratios are used by the PES program to adjust for correlation bias in the dual system estimates
- DA highlights specific populations that are undercounted in the census (e.g., Black males and young children)

# DA Estimates of Percent Net Undercount: 1960 to 2010




Note: Positive values indicate an undercount.
Source: U.S. Census Bureau, Demographic Analysis Program, special tabulation.

# DA Estimates of Percent Net Undercount by Select Age Groups: 1970 to 2010




Note: Positive values indicate an undercount.
Source: U.S. Census Bureau, Demographic Analysis Program, special tabulation.

# DA Method

- Cohort component method

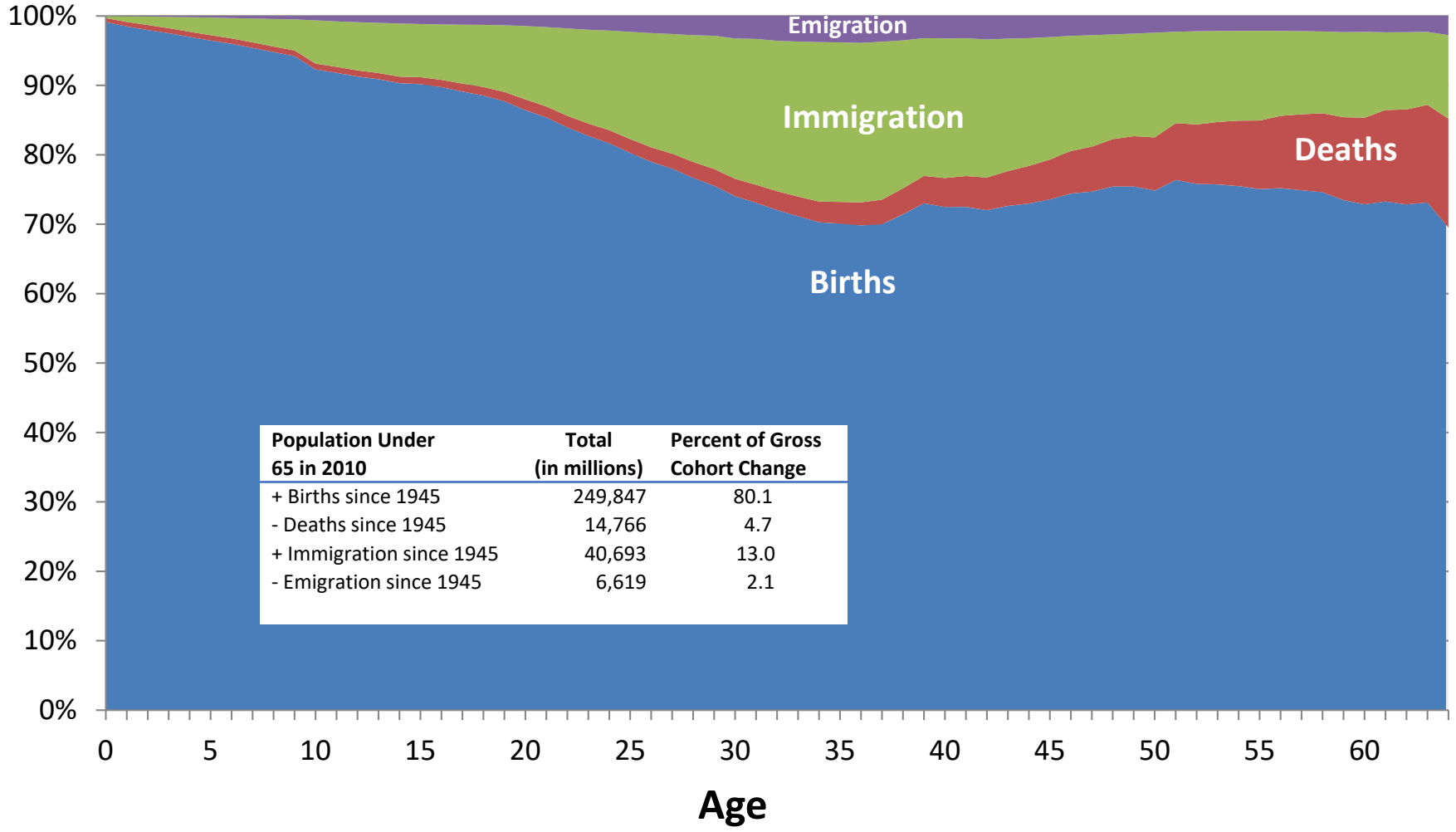
$$Pop_{total} = \sum_{i=1935}^{2020} Pop_i = \sum_{i=1935}^{2020} (Births_i - Deaths_i + Immig_i - Emig_i)$$

Where Pop is population, Immig is immigration, and Emig is emigration

- Medicare method for the oldest age groups
- Net Coverage Error

$$\frac{100 * (census\ count\ - DA\ estimate)}{DA\ estimate}$$

# Contribution of Births, Deaths, Immigration and Emigration to Gross Cohort Change by Age in 2010




| Population Under 65 in 2010 | Total (in millions) | Percent of Gross Cohort Change |
|---|---|---|
| + Births since 1945 | 249,847 | 80.1 |
| - Deaths since 1945 | 14,766 | 4.7 |
| + Immigration since 1945 | 40,693 | 13.0 |
| - Emigration since 1945 | 6,619 | 2.1 |

Source: U.S. Census Bureau, 2010 Demographic Analysis (May 2012 DA Release).

# 2020 DA Program

- DA is a sub-operation under the Evaluations and Experiments Integrated Project Team (IPT)
- DA is conducted by staff in the Estimates and Projections Area of the Population Division
  - Similar input data to the official population estimates
  - However, DA is not built from the prior census counts
- Work with internal and external research partners to develop the methodology for the 2020 DA
  - Universities, research institutes, other federal agencies

# Research Partners

**Demographic Analysis**
Gregg Robinson, *U.S. Census Bureau (retired)*
Jason Devine, *U.S. Census Bureau*
Bill O'Hare, *O'Hare Data and Demographic Services, LLC*

**Births**
Brady Hamilton, *National Center for Health Statistics*
Luke Rogers, *U.S. Census Bureau*

**Mortality**
Irma Elo, *University of Pennsylvania*
Elizabeth Arias, *National Center for Health Statistics*
Jana Johnson, *University of Minnesota*

**Net International Migration**
Katharine Donato, *Georgetown University*
Jenny Van Hook, *Penn State University*
Bryan Baker, *Office of Immigration Statistics*
Jeff Passel, *Pew Research Center*
Jason Schachter, *U.S. Census Bureau*

**Aging**
Wan He, *U.S. Census Bureau*
Kirsten West, *U.S. Census Bureau (retired)*

**Race and Hispanic Origin**
Richard Alba, *City University of New York*
Carolyn Liebler, *University of Minnesota*
Roberto Ramirez, *U.S. Census Bureau*
Nicholas Jones, *U.S. Census Bureau*

# 2020 DA Program (cont.)

- For 2020, we plan to produce 6 series of estimates
  - Black alone / Non-Black
  - Black alone or in combination
  - Hispanic / Non-Hispanic
  - Extended Hispanic
  - Full race and Hispanic origin detail (0-17)
  - State and county estimates for young children

- A range of estimates—high, medium, and low—to measure variation in the estimates

| Proposed Series for the 2020 DA Estimates | | | |
|---|---|---|---|
| **Series** | **Type of Estimates** | **Characteristics** | **Cohorts** |
| Black alone / Non-Black | Official | Age, sex, race | 0-85+ |
| Black alone or in combination / Non-Black alone or in combination | Official | Age, sex, race | 0-39 |
| Hispanic / Non-Hispanic | Official | Age, sex, Hispanic origin | 0-29 |
| Full race and Hispanic origin detail | Experimental | Age, sex, race, Hispanic origin | 0-17 |
| Extended Hispanic / Non-Hispanic | Experimental | Age, sex, Hispanic origin | 0-39 |
| Subnational DA[1] | Experimental | Age | 0-4 |

[1]The Subnational DA series will not be produced until the final birth data are available in 2021.

# 2020 DA Challenges

- Assigning race and Hispanic origin to birth records
  - Kid-Link method
- Producing estimates with race and Hispanic origin detail that is consistent with the census
- Developing estimates of the oldest-aged population
- Estimating the foreign-born population
- Implementing differential privacy

# 2020 Research

## Births

- Incorporate provisional birth data for most recent years where final microdata are not available
- Account for differences between race assignment approaches used in historical birth records (before 1980) and in more recent birth records (1980 and after)

## Race and Hispanic Origin

- Standardize DA race assignment methods across all cohorts for all decades
- Improve race assignment methods to reflect increasing multiracial births
- Account for individuals who change their self-identification of race over time

# 2020 Research (cont.)

## Mortality and Aging

- Addressing misclassification of age, race, and Hispanic origin on death certificates
  - We are working on a proposal to use the Mortality Disparities in American Communities (MDAC) database, which links responses from the 2008 ACS to death records
- Looking at the validity of death data for the oldest ages
  - Supplement the vital statistics data with data from Medicare and other sources
- Working with people at the Center for Medicare and Medicaid Services (CMS) to filter out records that shouldn't be included in our estimates
  - Medicare database currently has a lot of people aged 115 or more

# 2020 Research (cont.)

**Net International Migration**

- Developing methods to estimate the total stock foreign-born population rather than estimating migration flows, which was done in the past
- Improving the estimates of the native population living abroad, especially for young children
  - Using Mexico's 2010 Census data and the National Survey of Occupation and Employment (ENOE), we found that approximately 200,000 children aged 0 to 4 that were born in the United States were living in Mexico at the time of the 2010 Census
  - For the 2020 DA, we are monitoring trends for this population in the ENOE survey and the 2018 Mexican Survey of Demographic Dynamics (ENADID)
- Estimating Puerto Rico migration after Hurricane Maria

# 2020 Research (cont.)

**Coverage of Children**

- Develop a series for the 2020 DA with full race and Hispanic origin detail for the population aged 0-17
- Develop DA state and county estimates of the population aged 0 to 4 in 2020

# Thank you

**Eric Jensen**

Population Division

Population Evaluation, Analysis & Projections Branch

301-763-3723

eric.b.jensen@census.gov