URBAN INSTITUTE

Reflections on the 2020 Census: Accuracy, Fairness, and Implications for the Decade Ahead

#LiveAtUrban

Simulating the 2020 Census

Miscounts and the Fairness of Outcomes

November 2, 2021
Reflections on the 2020 Census: Accuracy, Fairness, and Implications for the Decade Ahead
Diana Elliott, Principal Research Associate

URBAN INSTITUTE · ELEVATE · THE · DEBATE

Challenges to the 2020 Census Were Unprecedented

- Demographic change was going to make 2020 harder to count

- Politicization of a scientific agency and endeavor

- COVID-19 lockdown happened simultaneously with start of 2020 Census

    o Delays in fieldwork abutted hurricane and wildfire season

    o Reports of challenges counting apartment dwellers and college students in the pandemic

    o Uncertain residency and relocations make an address-based census a challenge

Accuracy and Fairness Matter in Every Census

- Challenges in 2020 have resulted in questions about accuracy (how well did we count the nation?) and fairness (how completely did we count all groups and places?)

- Some have studied early Census data quality indicators, but few definitive answers

- Post-enumeration Survey (PES) is the official Census measure of quality; unclear when this will be released

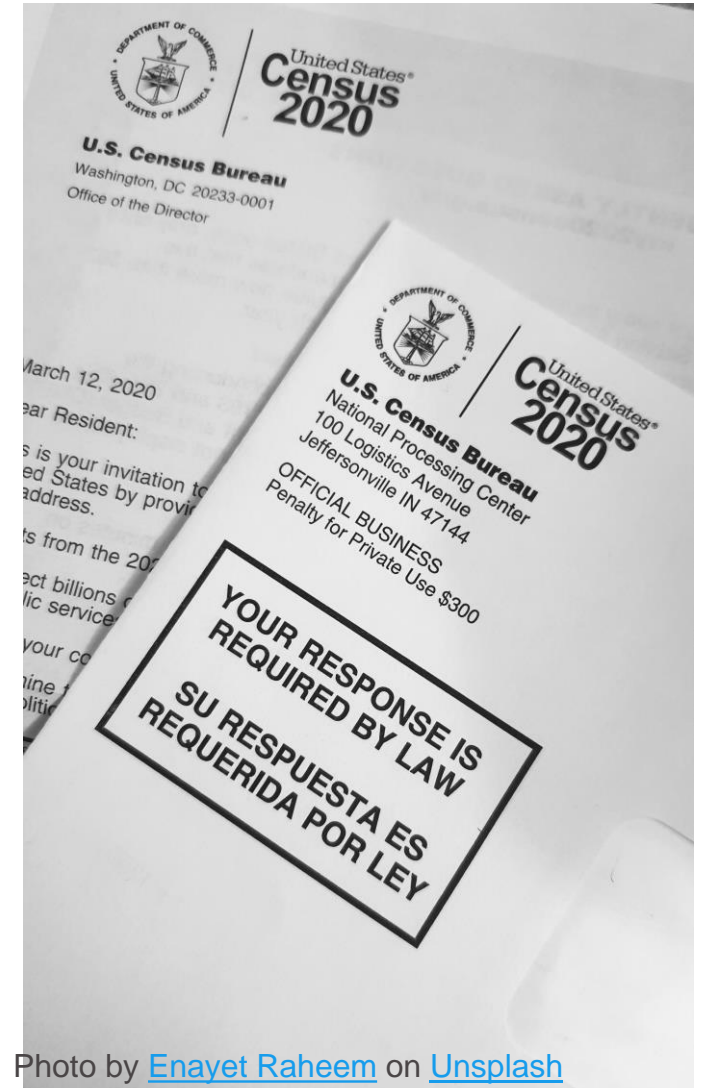- There is a need for external benchmarks that evaluate decennial censuses

Photo by Enayet Raheem on Unsplash

Goals of our Study

- **Develop a microsimulation model** to generate externally derived, yet plausible benchmarks for accuracy and fairness grounded in census data.

- **Produce estimates of miscounts** for demographic groups and geographies to help inform our understanding of the 2020 Census performance.

- **Explore outcomes like political representation and federal funding allocations** to understand the implications of an accurate and fair count.

- Have a **foundational model** to be refined over the decade and for 2030.

Snapshot of the Data

Step 1: Create a baseline population for Census Day:

➢ We start with ACS 2015-2019 5-year data and "mature" the population to April 1, 2020

Step 2: Make adjustments to simulate the 2020 Census environment:

➢ Use Census derived-estimates to adjust data for known 2020 Census factors (including 2020 self-response rates, historical accuracy rates by age, race/ethnicity, housing tenure; household citizenship status)

Step 3: Randomization to estimate over- and undercounts

➢ Nearly 16 million people in data set; each has probability of an under- and overcount

➢ Each case is assigned a random draw between 0 and 1 based on probabilities of under- and overcount; then assigned to 3 categories (omitted from the Census, accurately counted, or erroneously included)

➢ Results were calibrated to the official 2020 Census counts to derive miscounts among different groups

Scenarios Explored in the Study

Scenario 1: "Simulated 2020 Census"

"What will the 2020 Census count be if we replicate known factors that produce miscounts?"

- Project the person-level April 1, 2020 population and then adjust it for various factors

    - Housing tenure, age, race and ethnicity, presence of a noncitizen in household, 2020 self-response

- Allows us to understand 2020 Census miscounts

Scenario 2: "Hypothetical Full Count"

"What if the true resident population of the US was counted completely, accurately, and only once?"

- We assume no overcounts or undercounts among anyone in the US population

- Allows us to understand what a fair and accurate count would look like

Findings

A net undercount of the US population

In our "simulated 2020 Census" scenario:

- Average overcount of 3.6 percent

- Average undercount of 4.1 percent

- A net undercount of 0.5 percent

The undercount suggests the 2020 Census may not be as close to "net accuracy" as 2010 was, but is also not as dire as some had feared

Variation by geography is notable: States

In our "simulated 2020 Census" scenario:

- Mississippi and Texas have notable net undercounts (1.3 and 1.28 percents, respectively)

- Minnesota has a notable net overcount (0.76 percent)
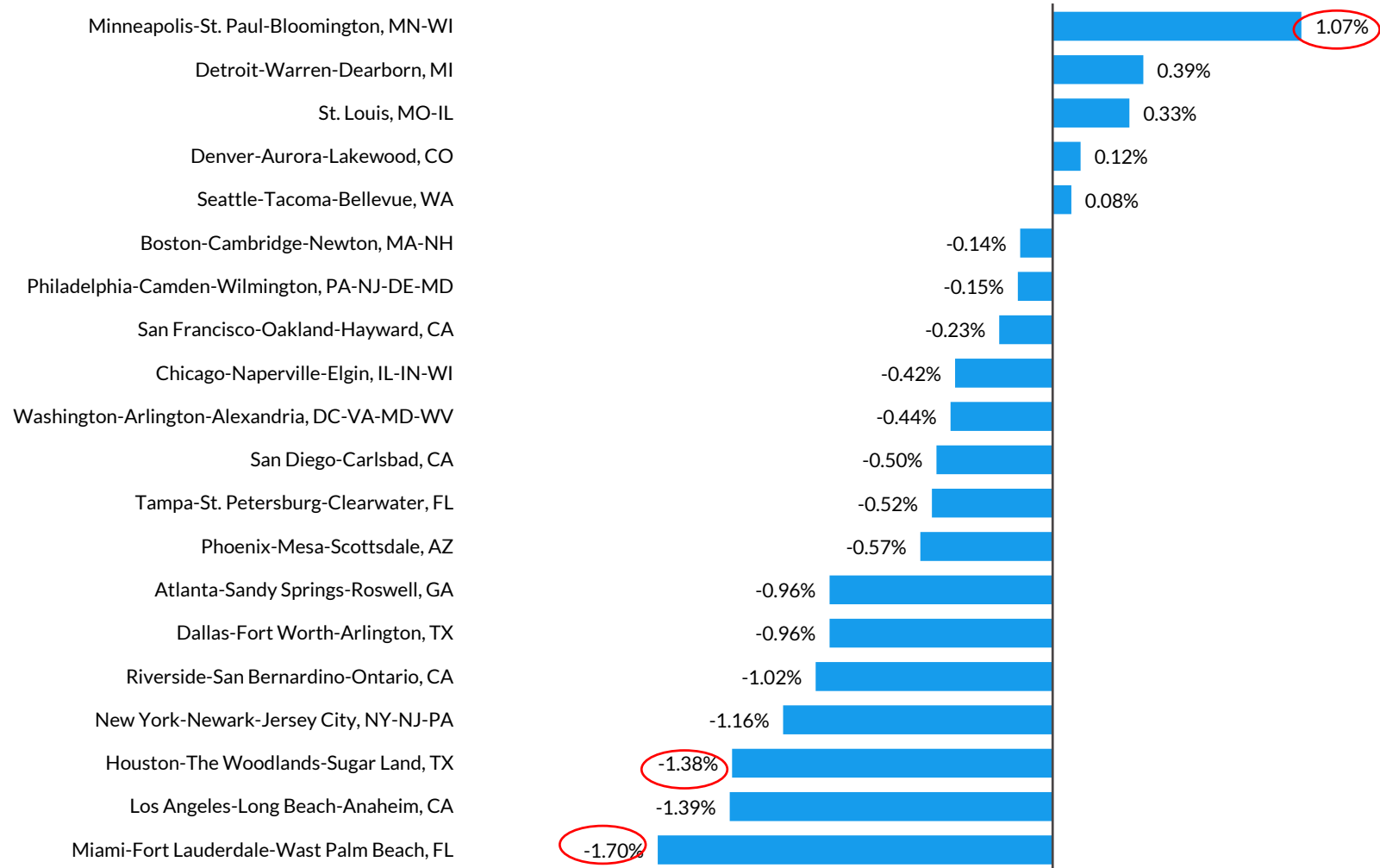
Has ramifications for funding:

- In our *"hypothetical full count"* scenario, Texas would receive over $247 million more in federal Medicaid reimbursements in 2021; Minnesota would receive $156 million less

- The *"hypothetical full count"* demonstrates that Mississippi and Texas residents will receive less of their fair share over the next decade; Minnesota residents will receive more

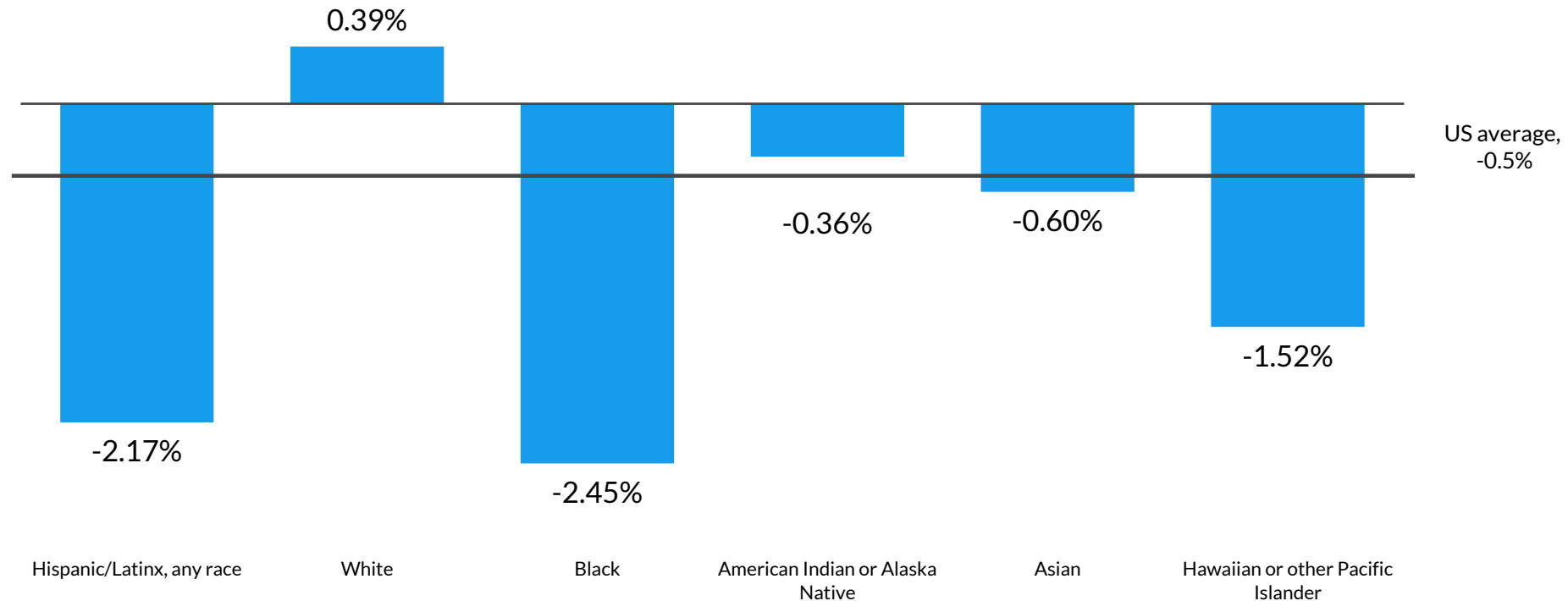# Variation by geography is notable: States (continued)

Has ramifications for representation:

- In our hypothetical full count scenario:

    - Minnesota would lose a seat (from 8 to 7)

    - New York would gain a seat (from 26 to 27)

- Important to keep in mind that the *"hypothetical full count"* scenario is virtually unachievable for various reasons, but is a good thought exercise

Variation by geography is notable: MSAs

Minneapolis-St. Paul-Bloomington, MN-WI	1.07%
Detroit-Warren-Dearborn, MI	0.39%
St. Louis, MO-IL	0.33%
Denver-Aurora-Lakewood, CO	0.12%
Seattle-Tacoma-Bellevue, WA	0.08%
Boston-Cambridge-Newton, MA-NH	-0.14%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	-0.15%
San Francisco-Oakland-Hayward, CA	-0.23%
Chicago-Naperville-Elgin, IL-IN-WI	-0.42%
Washington-Arlington-Alexandria, DC-VA-MD-WV	-0.44%
San Diego-Carlsbad, CA	-0.50%
Tampa-St. Petersburg-Clearwater, FL	-0.52%
Phoenix-Mesa-Scottsdale, AZ	-0.57%
Atlanta-Sandy Springs-Roswell, GA	-0.96%
Dallas-Fort Worth-Arlington, TX	-0.96%
Riverside-San Bernardino-Ontario, CA	-1.02%
New York-Newark-Jersey City, NY-NJ-PA	-1.16%
Houston-The Woodlands-Sugar Land, TX	-1.38%
Los Angeles-Long Beach-Anaheim, CA	-1.39%
Miami-Fort Lauderdale-West Palm Beach, FL	-1.70%

The hardest to count will again be undercounted: Miscounts by race and ethnicity

0.39%

-2.17%

-2.45%

-0.36%

-0.60%

-1.52%

US average, -0.5%

Hispanic/Latinx, any race   White   Black   American Indian or Alaska Native   Asian   Hawaiian or other Pacific Islander

The hardest to count will again be undercounted: Miscounts by age

0.96%

0.63%

-0.13%

US average, -0.51%

-0.87%

-1.85%

-4.86%

Birth to age 4   Ages 5 to 9   Ages 10 to 17   Ages 18 to 29   Ages 30 to 49   Ages 50 to 99
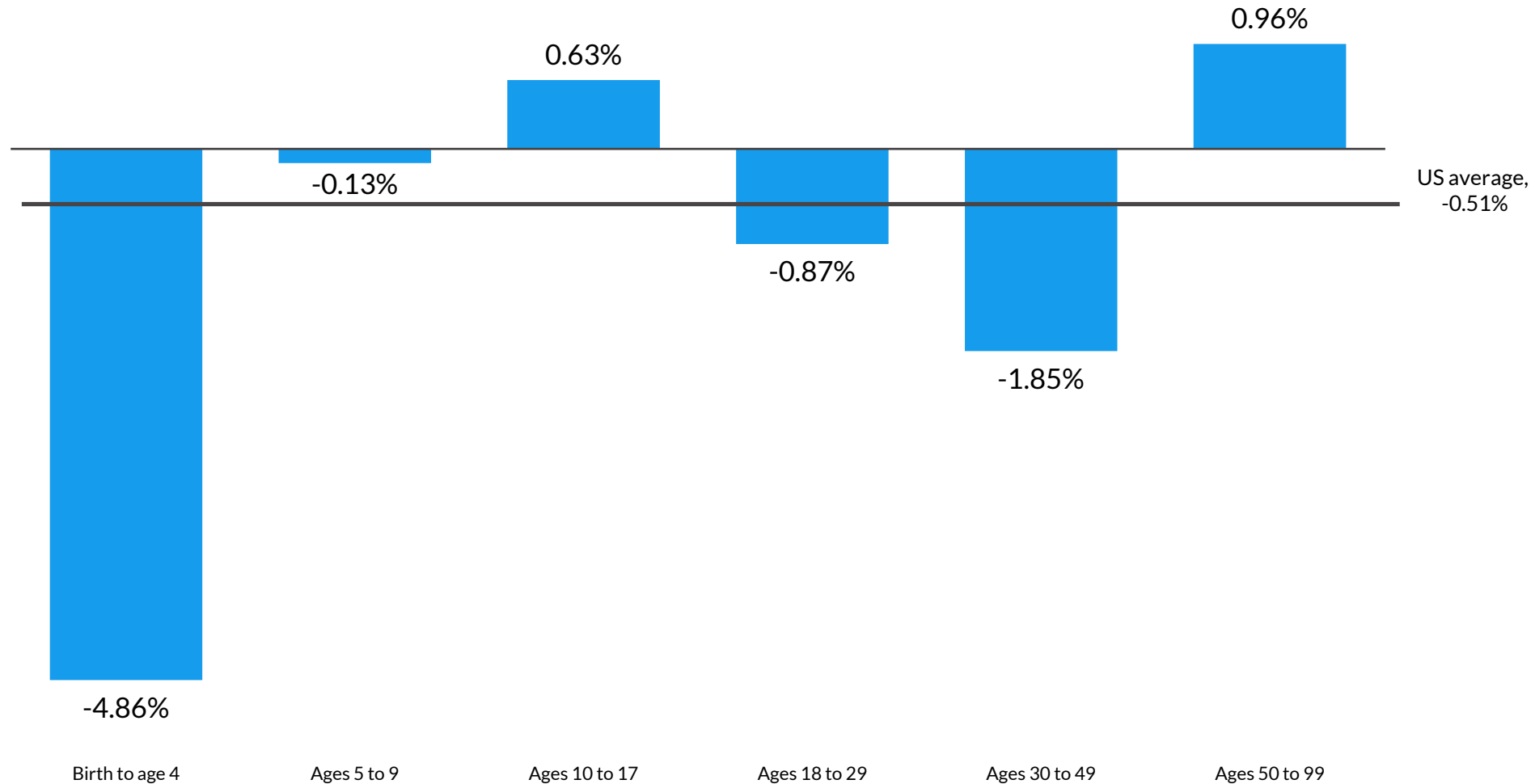
The hardest to count will again be undercounted

- Black and Hispanic/Latinx identified people have a net undercount of over 2.45 and 2.17 percent, respectively, in our simulated 2020 Census.

- Young children, or those under age 5, are likely to be net undercounted by 4.86 percent.

- Nationwide, renters are likely to be undercounted by 2.13 percent overall.

- People in households with a noncitizen present in them are anticipated to be undercounted by 3.36 percent overall.

Conclusion

- Counting the US population is an extremely difficult task, becomes harder each decade

- Urban's model offers a novel way to check decennial census quality; can be refined as new data become available

- Some communities and their residents will be shortchanged due to an incomplete count

- Underscores the importance of accuracy and fairness, with long lasting implications and outcomes over the next decade