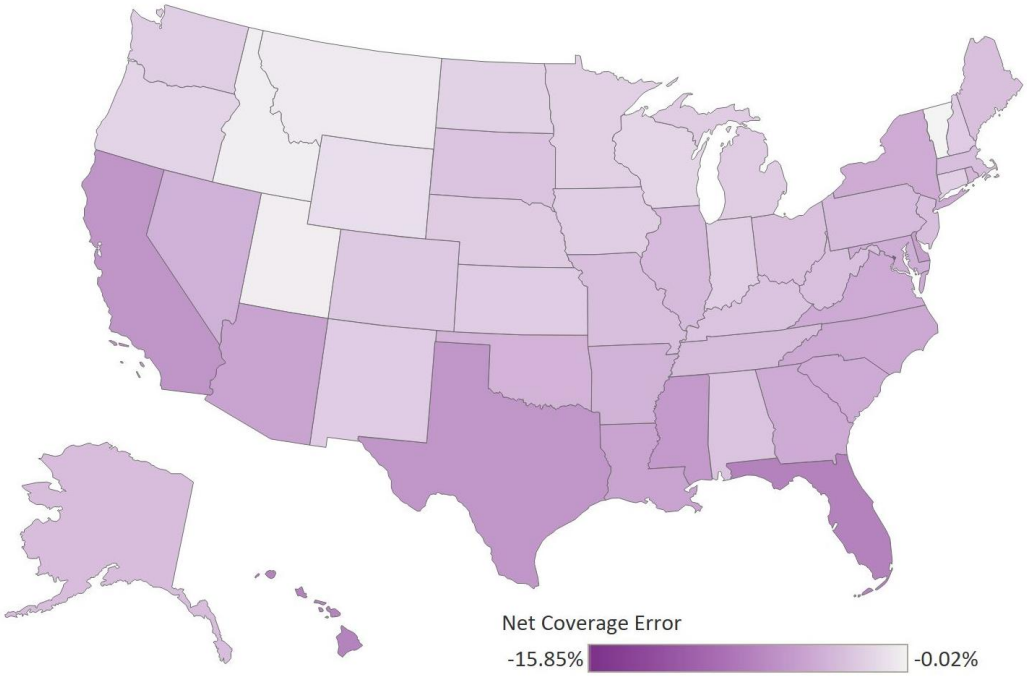
Summary of the U.S. Census Bureau's State and County-Level Net Coverage Error Findings for Young Children in the 2020 Census

Prepared by:

UMass Donahue Institute
Economic and Public Policy Research
Population Estimates Program

June 7, 2024

Figure 1. Net Coverage Error for Young Children (Aged 0-4) in the 2020 Census by State



UMass Donahue Institute. Source data: Demographic Analysis Estimates of Net Coverage Error in the 2020 Census for the Population Ages 0 to 4 by State (Table-1_0-4_ST). U.S. Census Bureau, Population Division. April 11, 2024

Background

On April 11th, 2024, the U.S. Census Bureau released net coverage error estimates of young children in the 2020 Census at the state and county-level. The newly released findings utilize the Census Bureau's Demographic Analysis population estimates to pinpoint undercounts of young children between the ages of 0 to 4 years old in areas across the country. The findings suggest that the largest undercounts were clustered along the West Coast, the South, and the Southwest. [1] Additionally, the Census Bureau highlighted a statistical relationship between net coverage errors and socioeconomic factors related to family structure, education, housing, and poverty. These specific characteristics were assessed by the Census Bureau as they cut across all race and ethnic groups.

---

[1] Jensen, Eric, and George M. Hayward. "Most Counties Had an Undercount of Young Children in the 2020 Census." census.gov, April 11, 2024. https://www.census.gov/library/stories/2024/04/children-undercount.html.

According to the data analysis, Massachusetts ranked 25th among all states by percentage, with an estimated net coverage error estimate of -4.15% for young children, equaling a total of 14,734 children not counted in the 2020 Census. The counties in Massachusetts with the largest undercounts of the young child population by percentage were Hampden (-6.4%), Suffolk (-5.71%), and Essex (-5.52%). The only counties in the state with a positive net coverage error estimate, and therefore an estimated overcount of the young child population, were Hampshire (2.62%) and Franklin (0.96%).

To produce the net coverage error estimates, the Census Bureau compares the young child population across two datasets – the 2020 Census special tabulation (CBDRB-FY24-0218) and the 2020 Demographic Analysis (April 2024 release). The 2020 Demographic Analysis (DA) is a dataset produced by the Census Bureau that uses "birth and death records, data on international migration, and Medicare Enrollment data to produce population estimates that are compared to the census to estimate net coverage error." The DA is considered the Bureau's best benchmark when assessing the census count of young children as the birth records used as the primary source for the dataset are considered 100% complete in the United States. The net coverage error is then calculated as described in equation 1.

Equation 1. Calculation of Net Coverage Error

$$Net\ Coverage\ Error\ =\ 100\ x\ [(Census\ Count\ -\ DA\ Estimate)/DA\ Estimate]$$

Various data limitations and caveats should be noted when analyzing the net coverage error estimates. For example, data was only released for 1,927 of 3,143 U.S. counties as a population estimate of 1,000 or more for the 0-4 year old population was required. Births may also be incorrectly geocoded to an adjacent county rather than the actual county of residence. For example, the Census Bureau explains that the net coverage error estimate of -15.85% for the District of Colombia may be inflated due to mothers living in Virginia and Maryland giving birth in D.C. hospitals. Lastly, while undercounts of the young child population have been a recurring issue from census-to-census, the direct cause of the net coverage error is not known. Therefore, the statistical relationship between the socioeconomic factors outlined in this report and the net coverage error should be considered for their correlative association rather than causal effects.

State Comparisons

Massachusetts ranked 25th among all other states and the District of Columbia when comparing net coverage error figures. Overall, the net coverage errors at the state-level ranged from -15.85% (District of Columbia) to -0.02% (Vermont), suggesting that young children were undercounted, on average, in all states. [2] Fifteen states had estimated net coverage errors below the national average of -5.46%, meaning they had larger undercounts than the national average. More than two-thirds of these states are in the South, including the District of Columbia (-15.85%), Florida (-9.87%), Texas (-7.70%), Mississippi (-7.47%), Delaware (-6.84%), Louisiana (-6.44%), North Carolina (-6.06%), Virginia (-5.85%), Georgia (-5.78%), South Carolina (-5.78%) and Maryland (-5.47%). Three of the remaining states are in the West: Hawaii (-9.70%), California (-7.84%), and Arizona (-6.46%). New York (-5.75%) was the only state in the Northeast to make this list, while Massachusetts (-4.15%) and the remaining Northeastern states landed above the national average of -5.46%.

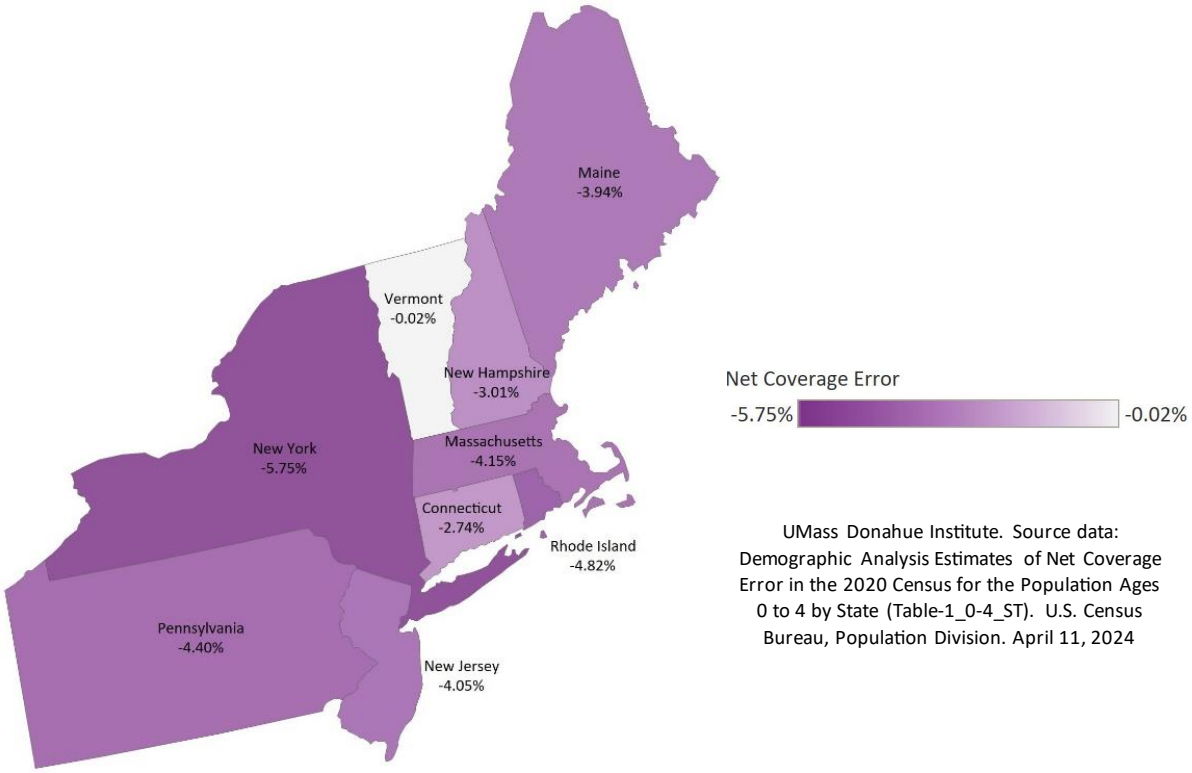
[2] US Census Bureau Population Division. "State and County Net Coverage Error Estimates for Children Aged 0 to 4 in the 2020 Census." census.gov, April 11, 2024. https://www.census.gov/library/visualizations/interactive/net-coverage-error-young-children.html.

Table 1. Largest Estimated Net Coverage Error for the Young Child Population by State

State Name	DA Population Estimate	2020 Census Count	Net Coverage Error	Rank
District of Columbia	44,083	37,095	-15.85%	1
Florida	1,143,120	1,030,284	-9.87%	2
Hawaii	85,659	77,352	-9.7%	3
California	2,319,173	2,137,439	-7.84%	4
Texas	1,971,128	1,819,260	-7.7%	5
Mississippi	185,510	171,647	-7.47%	6
Delaware	54,992	51,230	-6.84%	7
Arizona	419,488	392,370	-6.46%	8
Louisiana	300,610	281,257	-6.44%	9
North Carolina	611,557	574,468	-6.06%	10
Massachusetts	354,754	340,020	-4.15%	25
United States			-5.46%	-

UMass Donahue Institute. Source: Demographic Analysis Estimates of Net Coverage Error in the 2020 Census for the Population Ages 0 to 4 by State (Table-1_0-4_ST). U.S. Census Bureau, Population Division. April 2024

Figure 2. Net Coverage Error for Young Children (Aged 0-4) in the 2020 Census by Northeastern State



UMass Donahue Institute. Source data: Demographic Analysis Estimates of Net Coverage Error in the 2020 Census for the Population Ages 0 to 4 by State (Table-1_0-4_ST). U.S. Census Bureau, Population Division. April 11, 2024

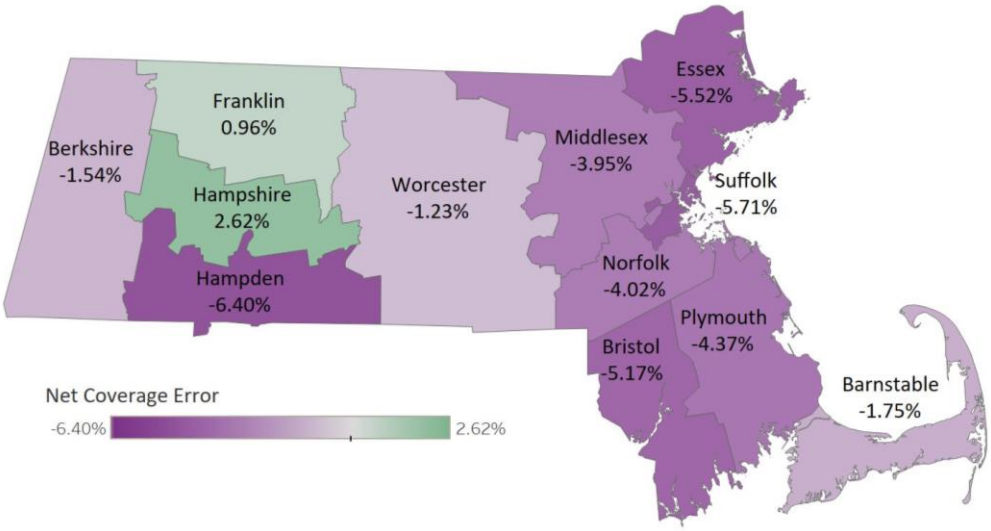
As mentioned, the states in the Northeast region of the U.S. demonstrated more moderate net coverage errors than several states in the South and West regions. For instance, the state with the greatest net

coverage error in the Northeast was New York at 5.75%, an error level that ranked 14th among all states. Additionally, the state of Vermont recorded the smallest net coverage error in the country at only 0.02%. Among the nine states in the Northeast, Massachusetts had the 4th largest negative net coverage error, behind New York (-5.75%), Rhode Island (-4.82%), and Pennsylvania (-4.40%).

Massachusetts Counties Overview

Among the 14 counties in Massachusetts, 12 counties had an estimated undercount of the young child population in the form of a negative net coverage error. Hampshire and Franklin counties, with net coverage errors of 2.62% and 0.96%, are the only counties in the state with a positive error – suggesting an overcount of the young child population occurred in these areas. Throughout the United States, there were 300 counties with overcounts recorded compared to 1,627 counties with undercounts. The average population size of young children in undercounted counties was 4,421, compared to an average of 10,017 total young children in overcounted counties. Based on these observations, it appears that less populated counties in the U.S. and Massachusetts were more likely to experience an overcount of the young child population than those with larger populations.

Figure 3. Net Coverage Error for Young Children (Aged 0-4) in the 2020 Census by MA County



UMass Donahue Institute. Source data: Demographic Analysis Estimates of Net Coverage Error in the 2020 Census for the Population Ages 0 to 4 by County (Table-2_0-4_CO). U.S. Census Bureau, Population Division. April 11, 2024

Table 2. Net Coverage Error Estimate and Population by MA County

MA County	DA Population Estimate	2020 Census Count	Net Coverage Error (Percentage)	Net Coverage Error (Numerical)
Hampden	24,961	23,364	-6.40%	-1,597
Suffolk	40,173	37,879	-5.71%	-2,294
Essex	43,898	41,476	-5.52%	-2,422
Bristol	29,710	28,175	-5.17%	-1,535
Plymouth	27,500	26,298	-4.37%	-1,202

Norfolk	38,131	36,599	-4.02%	-1,532
Middlesex	84,130	80,804	-3.95%	-3,326
Barnstable	7,612	7,479	-1.75%	-133
Berkshire	5,136	5,057	-1.54%	-79
Worcester	43,971	43,432	-1.23%	-539
Franklin	2,710	2,736	0.96%	26
Hampshire	5,265	5,403	2.62%	138
*Nantucket and Dukes counties are excluded due to the population of their young child cohorts being less than 1,000.				

Relationship with Socioeconomic Variables

The final component of the Census Buruea's data analysis for net coverage errors involved an assessment of county-level socioeconomic factors related to family structure, education, housing, and poverty. Specifically, the Bureau chose to review percentage figures for female-headed households, adults aged 25+ without a high school degree, renter occupied housing units, and children under 18 in poverty. Using a sample size of 3,854 U.S. counties, the data showed that the net coverage of young children tended to decrease as the percentage of the above-mentioned socioeconomic factors increased, suggesting an inverse relationship between the combined variables. While these findings suggest a degree of correlation between the variables, the Census Bureau emphasized that the data and methods are not available to calculate net coverage error estimates based on these county-level characteristics.

UMDI further assessed these socioeconomic variables for the state of Massachusetts to evaluate the potential relationship between county characteristics and the net coverage of young children. Two additional variables related to the 2020 Census self-response rates were also reviewed to determine whether census operational factors may also relate to net coverage errors. Table 3 summarizes the findings for Massachusetts counties.

Table 3. Summary of County-level Variables in Relation to the Net Coverage Error for MA Counties

County-level Variable		R-Squared	P-value
Socioeconomic Characteristics	Adults aged 25+ without a high school degree	0.4293	0.0207
	Female-headed households	0.4041	0.0263
	Renter occupied housing units	*0.1515*	*0.2111*
	Children under 18 in poverty	*0.1795*	*0.1700*
Census Self-Response Rates	Total Self-Response Rate	*0.0070*	*0.7954*
	Internet Self-Response Rate	*0.0171*	*0.6851*
UMass Donahue Institute. Source data: Census Interactive Gallery, "State and County Net Coverage Error Estimates for Children Aged 0-4" and "2020 Census: Tracking Self-Response Rates Map". U.S. Census Bureau, Population Division. April 11, 2024 and January 28, 2021			

Certain caveats should be noted when reviewing the findings from Table 3. First, the sample size of 12 counties is generally not sufficient when conducting a correlative study. The effects of a small sample size for such a study can lead to false relationships between variables being present. Second, the p-value

is generally viewed as significant if it is 0.05 or lower and can be seen as non-significant if higher than 0.05. Third, the R-squared value is generally considered a high level of correlation when above 0.7, a low level of correlation when below 0.4, and a moderate level of correlation when between these two points.

With these caveats in mind, we can cautiously evaluate the relationships between the county-level variables and the net coverage error for Massachusetts counties. The self-response rates demonstrate no direct relationship with the net coverage error, suggesting that this critical portion of the enumeration process did not impact the count of young children. Furthermore, two of the socioeconomic variables, renter occupied housing units *and* children under 18 in poverty, appear to have no impact on the net coverage error for Massachusetts counties. This leaves two socioeconomic variables, adults aged 25+ without a high school degree *and* female-headed households, that show a moderate level of correlation at a significant level with the net coverage error.

The following scatter plots depict the relationship between the two socioeconomic variables most correlated with the net coverage error in Massachusetts counties. Both charts describe an inverse relationship between the variables in which the net coverage error decreases as the percentage of adults aged 25+ without a high school degree *and* female-headed households increases.

Figure 4. Percent of Female-headed Households

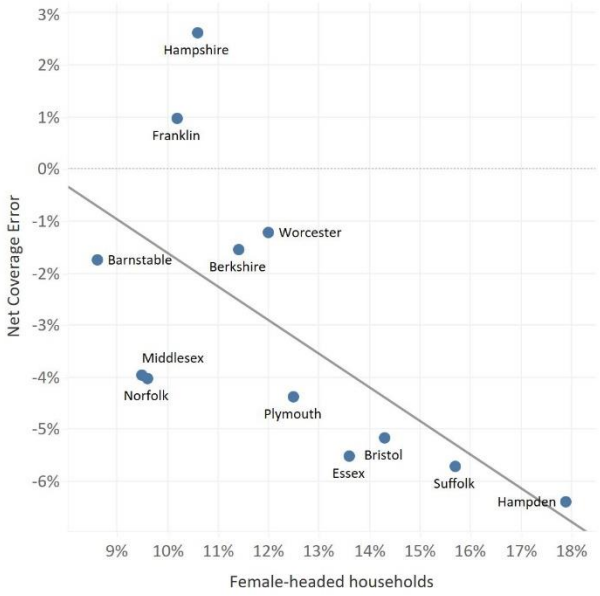
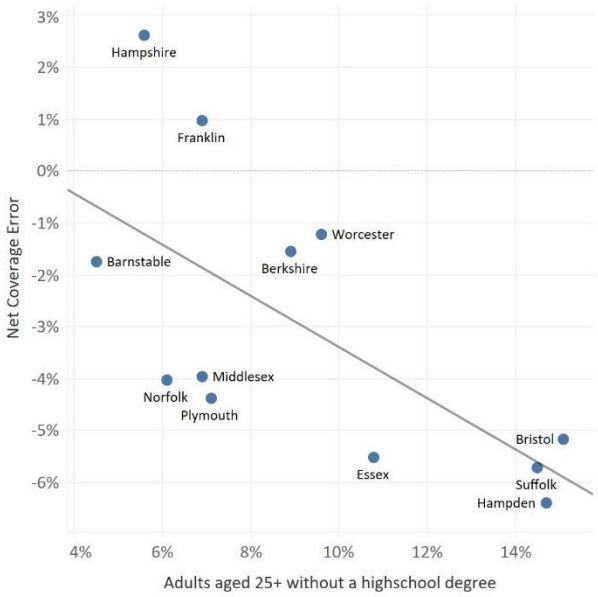


Figure 5. Percent of Adults Aged 25+ without a Highschool Degree



UMass Donahue Institute. Source data: Census Interactive Gallery: State and County Net Coverage Error Estimates for Children Aged 0-4. U.S. Census Bureau, Population Division. April 11, 2024

The data reviewed in this section allows researchers to point towards potential socioeconomic causes of net coverage errors in Massachusetts, although it does not paint a clear method for calculating net coverage errors based on socioeconomic characteristics alone. Perhaps a similar correlative study produced with a larger sample size at the city and town-level, or a weighted correlative approach based on R-squared values, could produce a viable calculation method for net coverage errors to be factored into future population estimates and projections.

Summary prepared by:

Christopher Diaz, Research Analyst
Susan Strate, Senior Research Manager