Socio-Demographic Correlates of State-Level Coverage Rates for Young Children in the 2020 Census

By

Dr. William P. O'Hare[1],[2] and Dr. Susana Quiros[3]

September 2025

Abstract

Several past studies have examined various socio-demographic data elements and their correlations with geographic variations in young child census coverage rates. This study extends that stream of research by examining the relationship between 30 potential explanatory variables and state net coverage rates for young children in the 2020 Census. The 30 potential explanatory variables are clustered into five domains.

1. Race/ethnicity and immigration
2. Socioeconomic status
3. Family structure/living arrangements
4. Housing
5. Internet availability and response rates

Of the 30 potential explanatory variables examined here, 17 had statistically significant correlations with state young child coverage rates. Looking at the results across domains, six of the eight variables in the Race/ethnicity and immigration domain showed a statistically significant correlation and seven of the eight variables in the Family structure/living arrangement domain were highly correlated with variation in 2020 Census state young child coverage rates. Most of the findings are consistent with past research but a couple of intriguing results are examined in the Discussion section.

[1] Consultant to the Count All Kids Campaign.
[2] Dr. O'Hare's work on this Brief was supported by the Annie E. Casey Foundation, but they are not responsible for the contents of this publication.
[3] Adjunct Professor, University of Missouri.

Socio-Demographic Correlates of State-Level Coverage Rates for Young Children in the 2020 Census

1) Introduction

There is a growing body of research related to the undercount of young children in the U.S. Census (O'Hare 2015; U.S. Census Bureau 2019; Griffin and O'Hare 2020; Seltzer and Walker 2020; Johnson 2022; U.S. Census Bureau 2024; Quiros and O'Hare 2024; O'Hare 2023; Pascale 2024, Pascale 2025), but only a few studies have examined what explains geographic differences in state and county census coverage rates for young children. This study extends existing research by looking at correlations between state coverage rates for young (ages 0-4) children and 30 potential explanatory variables.

After providing some background information on the high net undercount of young children in the U.S. Census there is a review of the literature in this area. That is followed by an analysis of the correlations between 30 potential explanatory variables and state net census coverage rates for young children in the 2020 Census. The discussion section highlights a few of the most intriguing findings from this analysis.

2) Background

Figure 1 shows the net coverage rates in the 2020 Census for five-year age groups. The net undercount rate for young children (5.4 percent) is much worse than for any other age group. This is similar to the results in the 2010 Census (O'Hare 2015).

Note the net undercount of the population ages 0 to 4 is more than three times as high as that for children ages 5 to 9 (1.5 percent). This suggests there is something special about the population under age 5 even compared to slightly older children.

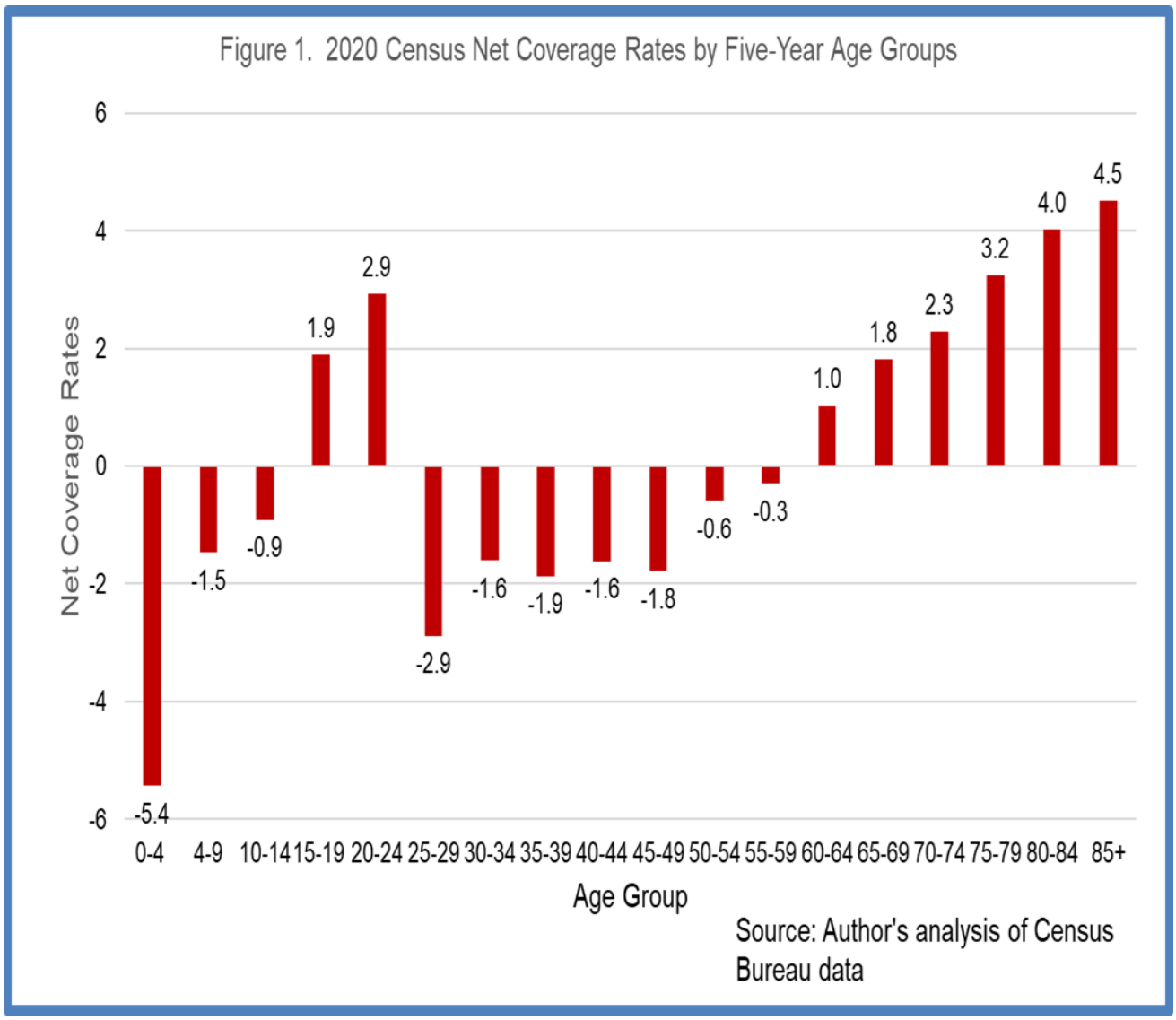
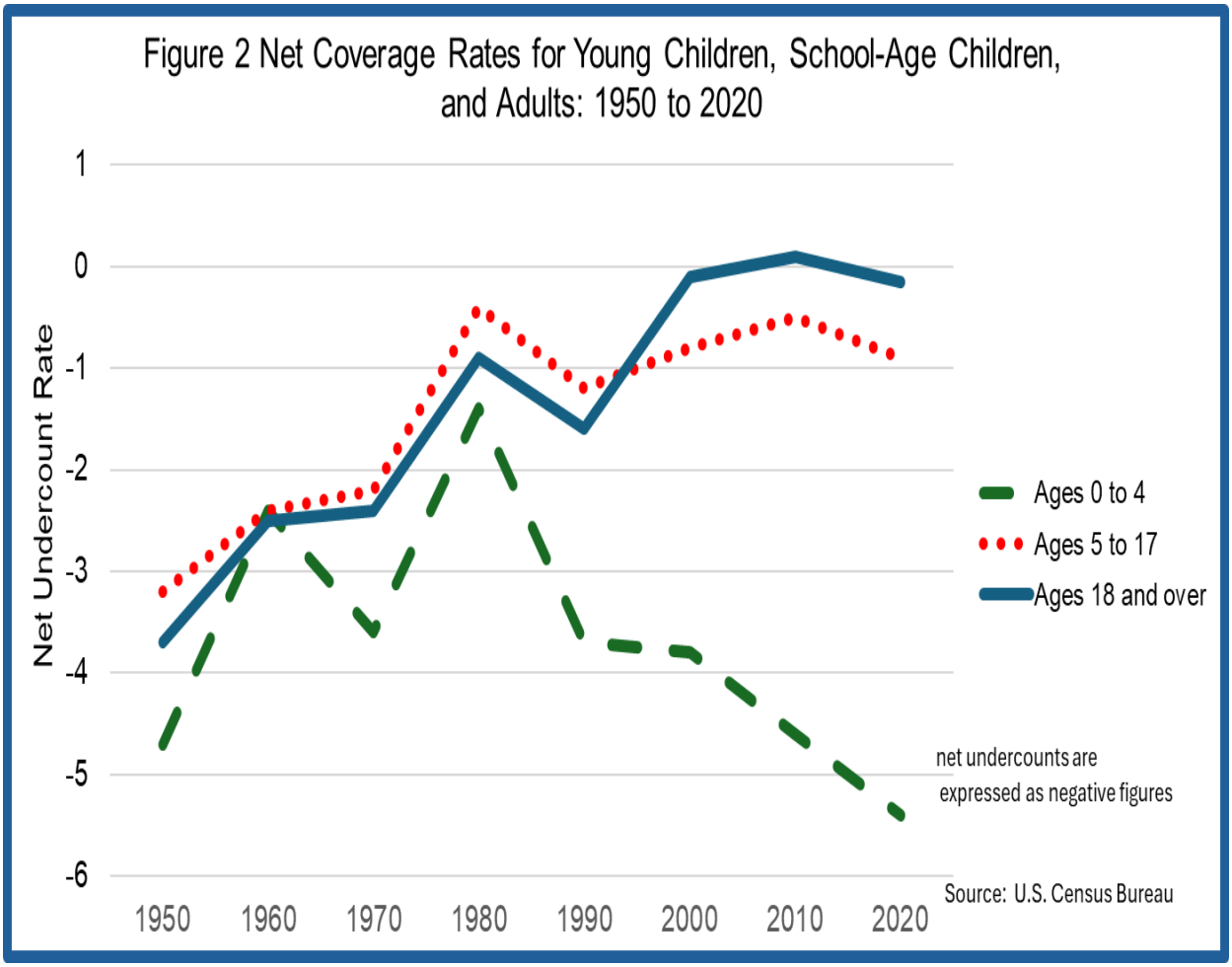


Figure 2 shows the coverage rates for young children (ages 0 to 4), school-age children (ages 5 to 17), and adults (ages 18 and older) from 1950 to 2020. After 1980, the coverage rate trend for young children is different than those for older children and adults. The coverage rate for young children goes from 1.4 percent in 1980 to 5.4

percent in 2020. Coverage rates for school-age children are relatively stable and the rates for adults improved a little after 1980. It is also noteworthy that young children are the only group that has worse coverage in the 2020 Census than they had in the 1950 Census.

These trends indicate that the changes over time in coverage of young children and the coverage of the total population are moving in different directions and are likely driven by different forces. O'Hare (2025) provides data showing the factors driving the census coverage of young children are different than those driving coverage rates for the total population.



The high net undercount of young children is geographically widespread. Children ages 0 to 4 were undercounted in every state in the 2020 Census (U.S. Census Bureau 2024). In addition, children ages 0 to 4 had higher net undercount rates than children ages 5 to 9 in every state in the 2020 Census (O'Hare 2024d). The persistent difference between the coverage of children ages 0 to 4 and those ages 5 to 9 raises an important question about why the youngest children have a much higher net undercount rate than slightly older children.

Analysis (Griffin 2014) suggests the high net undercount for young children is largely based on their high omissions rate rather than low erroneous enumeration rate and this begs two questions. First, why are some young children missed at a higher rate than other young children? For example, why are young Black and Hispanic children missed at a higher rate than young non-Hispanic white children? Second, what are the characteristics or factors that make omissions for young children higher than for other demographic groups?

We need to pinpoint the characteristics that make young children more likely than others to be missed in the Census.

3) Review of the Literature

In this section we review many of the studies that have examined the relationship between geographic variation in net young child coverage in the Census in relation to other socio-demographic variables. Most studies examined differences across states and counties.

O'Hare (2014) examined differences in state net coverage rates of young children in the 2010 Census in relation to a few other demographic measures. He found larger states (in terms of population size) had higher net undercounts of young children than smaller states. Four measures of race and ethnicity (Percent Black Alone or in Combination, Percent Hispanic, Percent Black Alone or in Combination plus Hispanic, Percent Minority) were found to be highly correlated with differences in state young child coverage rates, and all these correlations were found to be statistically significant.

O'Hare (2014, Table 6) also examined several hard-to-count factors. Those factors listed below were found to have a statistically significant correlation with state-level net census coverage rates of young children.

- Percent of households that were linguistically isolated
- Percent of the population aged 25+ not high school graduate
- Percent of people unemployed
- Percent of occupied housing units that are rental units
- Percent of units that are single detached units
- Percent of occupied housing units that are crowded
- Percent of people in poverty
- Percent of households that are not husband-wife households

O'Hare (2017b) examined differences in county-level coverage rates of young children in the 2010 Census in relation to a few selected socio-demographic measures. He found a big difference by county population size with the largest counties having much higher young child undercount rates than smaller counties. For example, the collective young child net undercount rate for counties with a half-million or more people was 7.8 percent compared to a net undercount rate under 1.1 percent for counties between 20,000 and 64,999 people and net overcount rates for counties with less than 19,999 people.

As with states, race and Hispanic origin metrics are closely related to differences in county-level net coverage rates for young children. When Black people plus Hispanic people were more than 50 percent of the county population, the net undercount rate for all young children was 8.9 percent compared to a net overcount of 1.9 percent for all young children when they were less than 5 percent of the county population.

The U.S. Census Bureau (2017a) studied the distribution of a couple of metrics related to potential problems with the count of young children in the 2010 Census across states and counties. They first examined the distribution of 611,606 positive responses to one of the child-specific undercount probes on the Census questionnaire. A positive response indicates uncertainty about including a child in the census and puts that child at risk of being missed. The positive response number was converted to a rate based on total number of young children in the state or county.

When states were ranked (U.S. Census Bureau 2017a, Table 5) based on positive response rates, most of the states at the top of the list were Southern or Southwestern states, and most of the states at the bottom of the list were in the Northeast or Midwest parts of the country.

The Census Bureau researchers also examined positive response rates for counties and Census tracts based on data in the Census Bureau's Planning Data Base. The researchers found the positive response rates was higher in larger (population wise) counties (those of 500,000 or more) compared to smaller counties. A county level map in the Census Bureau publication on this topic also shows the counties with the highest positive response rates are mostly in the South and Southwest areas of the country.

The Census Bureau researchers then examined the number of young children added through Coverage Followup Operation (CFU). An added child indicates the child would have been undercounted on the initial roster. The number was converted to a rate based on the total number of young children in the state or county.

There were substantial differences across states in the rate of potential errors in rostering young children. Again, the biggest problems appear to be in the South and Southwest regions of the country, and larger counties (those with 500,000 or more people) had higher rates of positive responses to the undercount count probes for young children than smaller counties.

O'Hare (2023b) examined variation in the 2020 Census state-level net coverage rates for young children derived by comparing Vintage 2020 Census Bureau population estimates to the Census counts. Most of the states with the highest net young children undercount rate were states with large racial/ethnic minority populations and states at the bottom of the ranking were states with little racial diversity. He also found an extremely high correlation (0.84) between the young child net coverage rates in 2010 and 2020.

The U.S. Census Bureau (2017b) also studied the geographic distribution of the 611,606 positive responses to the undercount probe question with respect to the Esri Tapestry segmentation system. (As noted above, a positive response means that household is uncertain whether to count their young child). The Esri Tapestry system divides the county up into many segments defined by combinations of demographic/social/economic characteristics. Some of those segments had higher positive response rates than other segments.

The Tapestry market segment with the highest rate of positive responses is labeled NeWest Residents (Census Bureau 2017b, no page) and is described as, “NeWest resident. This is a young Hispanic market. The population is new to America, new to their careers, with new young families. More than one-third of the households are linguistically isolated. The population is concentrated in larger metropolitan area in the South and West. More than half of the households have children, and the presence of young children is high compared with the U.S. average. Dependent children represent one-third of the total population.”

Associations between positive response rates and several census tract level characteristics from the Census Bureau’s Planning Database showed that higher positive response rates (a reflection of likely errors) were found in:

- Tracts with high level of language other than English spoken at home
- Tracts where Spanish spoken in the home at a high rate
- Tracts where educational attainment is low
- Tracts with higher rates of geographic mobility
- Tracts where unemployment is higher
- Tracts where poverty is higher
- Tracts where crowded housing is higher
- Tracts where census self-response rates were lower.

They conclude (U.S. Census Bureau, 2017b, no page number), “We find that the Esri Tapestry Segmentation and PDB categorizations can help us to identify neighborhoods and types of areas with coverage problems involving young children in the 2010 Census.”

King et al. (2018) examined subnational estimates of net coverage for young children based on state and county estimates of census coverage. They found the

highest net undercount of young children in the South and West. They also found race/ethnic composition of counties was related to net coverage rates.

The Population Reference Bureau (PRB) (O'Hare et al. 2019 and 2020) examined how differences in coverage rates of young children in the 2010 Census are related to other measures for a group of large counties (those with at least 5,000 children ages 0 to 4 in 2010). This set of counties was divided into two groups, one with more than 250,000 total population and one with less than 250,000 total population. Multiple regression results were more robust for larger counties. After extensive multi-variate analysis, they (O'Hare et al 2020, Table 6) found the combination of the variables listed below as the best predictors of young children coverages rates.

- Percent of minority (families with children)
- Percent of adults (age 18+) with less than a high school diploma, GED, or alternative
- Percent of persons ages 0-4 in poverty
- Percent of households that are linguistically isolated
- Percent of children ages 0-4 living with a grandparent householder
- Percent of households that have 7+ people
- Percent of children ages 0-4 who live with non-relatives or in groups quarters
- Percent of children ages 1-4 have moved in the past year.
- Percent of persons living in renter-occupied households

O'Hare (2023a) examined 2020 Census county-level net coverage rates for young children based on comparing Vintage 2020 Census Bureau population estimates to the 2020 Census counts. He found that 69 percent of the counties had a net undercount of young children. The counties with the highest net undercounts of young children were predominantly in the South and Southwest regions of the country. He also

found large counites (those with one million or more people) had much higher net undercounts for young children than smaller counties.

More than two-thirds of counties had a net undercount of young children in the 2020 Census (O'Hare 2023a). The 49 counties with more than one million people accounted for almost half (47.7 percent) of all net young child undercounts in the country.

O'Hare (2024c) examined changes in the county-level coverage of young children between the 2010 and 2020 Census. In general, the largest counties (1 million plus improved) had higher net young child net undercounts compared to smaller counties (less than 20,000 people). Across the four Census regions, the results do not differ much.

Quiros and O'Hare (2024) examined 30 factors thought to be related to state differences in the net coverage of young Hispanics in the 2020 Census and found 18 of the 30 correlations were statistically significantly correlated with state young Hispanic net coverage rates and in the predicted direction. The domain that appeared to be most highly related to variations in young Hispanic child coverage rates was race/ethnicity and immigration status. The domain of housing and the domain of family structure and living arrangements were also highly correlated.

U.S. Census Bureau (2024) examined the relationship between in county level net coverage rates for young children for all the counties with at least 1,000 young children with four measures. Each of the four measures had a statistically significant collection with young child coverage rates.

- Female-headed households
- Renter-occupied housing units
- Children under 18 in poverty
- Adults aged 25+ without a high school degree

4) Data Sources for This Analysis

The state net coverage rates for the population ages 0 to 4 are from an experimental DA series released by the Census Bureau in 2024 (U.S. Census Bureau, 2024). There is no sampling error in the DA estimates like there is in the Census Bureau's Post-Enumeration Survey estimates, but there is undoubtedly some random error in the DA estimates. To reflect the idea of errors in the DA estimates, the Census Bureau issued a low, middle, and high series for the nation as a whole. The number of young children based on the three DA series varies from 19,427,000 to 19,480,000.

The potential explanatory variables used here are largely taken from past studies in this arena (Quiros and O'Hare 2024; National Academy of Sciences 2024; U.S. Census Bureau 2024a; O'Hare et al. 2019; O'Hare 2015). The specific data for the potential explanatory variables used in this study are largely taken from Quiros and O'Hare (2024).

The potential explanatory variables were sorted into five domains. These are the same domains used by Quiros and O'Hare (2024) and virtually the same as those used by the Population Reference Bureau (O'Hare 2019). Data for most of the potential explanatory variables are taken from the American Community Survey IPUMS files (see Quiros and O'Hare 2024 for more details on the data for the potential explanatory variables).

The five domains are:

1. Race/ethnicity and immigration status
2. Socioeconomic status
3. Family structure/living arrangements
4. Housing
5. Internet availability and response rates

The potential explanatory variables used here are not an exhaustive list, but they provide factors from five domains and are a reasonable set of factors to assess what types of factors are most highly correlated with the coverage of young children.

Keep in mind the net coverage of young children is constructed here such that a decrease in value is worse coverage. In other words, - 10 percent is lower (worse) than -5 percent. This is important in interpreting the correlation coefficients. If a net undercount rate goes down, that indicates a worsening in coverage.

5) Analysis

In this section of the paper, we examine how state-level variations in net coverage for young children are correlated with a set of potential explanatory variables. Table 1 shows the states ranked by net young child coverage rates in the 2020 Census. Even though every state has a young child net undercount in the 2020 Census, it is clear that there is a lot of variation across the states. Discounting the extremely high net undercount rate in the District of Columbia (15.85 percent) the coverage rates vary from a low of 0.02 in Vermont, to a high of 9.87 percent in Florida. The extensive variation in coverage rates should facilitate finding correlations related to the variation in net young child coverage rates.

Table 1. States Ranked by Young Child Net Undercount Rates Based on Demographic Analysis (DA) Estimates of Net Coverage Error in the 2020 Census for the Population Ages 0 to 4

Rank	State Name	DA Population Estimate	2020 Census Count	Net Coverage Error Estimate
1	District of Columbia	44,083	37,095	-15.85
2	Florida	1,143,120	1,030,284	-9.87
3	Hawaii	85,659	77,352	-9.70
4	California	2,319,173	2,137,439	-7.84
5	Texas	1,971,128	1,819,260	-7.70
6	Mississippi	185,510	171,647	-7.47
7	Delaware	54,992	51,230	-6.84
8	Arizona	419,488	392,370	-6.46
9	Louisiana	300,610	281,257	-6.44
10	North Carolina	611,557	574,468	-6.06
11	Virginia	511,295	481,405	-5.85
12	Georgia	651,900	614,218	-5.78
13	South Carolina	294,142	277,144	-5.78
14	New York	1,125,285	1,060,610	-5.75
15	Maryland	365,000	345,047	-5.47
16	Nevada	183,803	174,032	-5.32
17	Arkansas	189,309	179,575	-5.14
18	Oklahoma	253,946	241,242	-5.00
19	Rhode Island	54,534	51,903	-4.82
20	Illinois	738,282	705,616	-4.42
21	Pennsylvania	698,586	667,816	-4.40
22	Tennessee	411,470	393,767	-4.30
23	Alaska	50,255	48,104	-4.28
24	Missouri	370,450	355,024	-4.16
25	Massachusetts	354,754	340,020	-4.15
26	New Jersey	523,264	502,046	-4.05
27	West Virginia	92,944	89,207	-4.02
28	Maine	64,000	61,477	-3.94
29	Ohio	693,573	666,434	-3.91
30	Alabama	297,751	286,529	-3.77
31	South Dakota	59,967	57,742	-3.71
32	Kentucky	274,385	264,254	-3.69
33	Colorado	325,309	314,580	-3.30
34	Nebraska	130,620	126,605	-3.07
35	Kansas	185,068	179,446	-3.04
36	New Mexico	118,403	114,806	-3.04
37	New Hampshire	63,389	61,480	-3.01
38	Michigan	565,801	548,875	-2.99
39	Washington	450,442	437,231	-2.93
40	Iowa	195,743	190,064	-2.90
41	Connecticut	181,819	176,831	-2.74
42	Indiana	420,162	408,828	-2.70
43	Minnesota	349,568	340,357	-2.63
44	North Dakota	53,751	52,366	-2.58
45	Oregon	220,488	215,252	-2.37
46	Wisconsin	330,091	322,285	-2.36
47	Wyoming	34,505	33,955	-1.59
48	Montana	59,669	59,224	-0.75
49	Utah	241,052	239,780	-0.53
50	Idaho	114,638	114,128	-0.44
51	Vermont	28,561	28,555	-0.02

Source; Demographic Analysis Estimates of Net Coverage Error in the 2020 Census for the Population Ages 0 to 4 by State. Release Date: April 2024

Source: U.S. Census Bureau, Population Division, 2020 Demographic Analysis (April 2024 release), 2020 Census Demographic and Housing Characteristics File (DHC), and 2020 Census special tabulat**ion** (CBDRB-FY24-0218) for the District of Columbia only.

Several of the studies reviewed here found differences by region, so that issue is explored with the state data in Table 1. Figure 3 shows the collective young child coverage rates for the four Census regions. As was found in several previous studies, the South had the highest net young child undercount rate in the 2020 Census. The West region had a young child net undercount rate slightly above the national mean, but it should be noted that three states in the West (California, Arizona, and Hawaii) had exceptionally high net undercount rates and all the other states in the West had below-average net young child undercount rates.

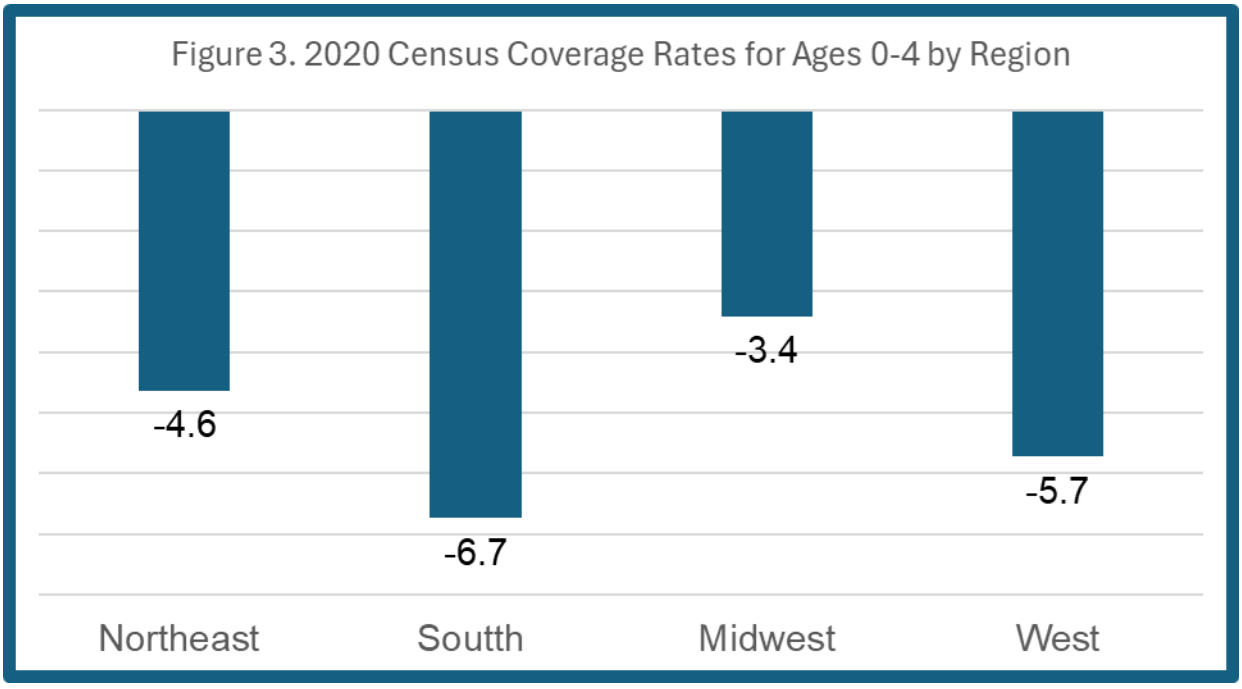


Table 2 shows the correlation coefficients between state young child coverage rates and 30 potential explanatory variables with the potential explanatory variables clustered by domain. The magnitude and direction of the correlation is provided along with an indicator of statistical significance.

Table 2. Correlation Coefficients Between 2020 Census State Young Child Coverage Rates and 30 Potential Explanatory Variables

	Correlation with State young child (ages 0-4) coverage rate	Statistical Significance*
Race/ethnicity and Immigration		
Percent Non-Hispanic Black alone (0-4)	-0.59	*
Percent Children of Color** (0-4)	-0.66	*
Percent non-Hispanic Black alone (families with related children)	-0.61	*
Percent Families of Color * (families with related children)	-0.67	*
Percent of adults ages 18+ who are foreign-born	-0.42	*
Percent of adults ages 18+ who are not US citizens	-0.42	*
Socioeconomic Status		
Percent of persons 0-4 in poverty	-0.11	
Percent of adults ages 18+ with less than HS or GED	-0.30	*
Family and Household Characteristics		
Percent of households linguistically isolated	-0.39	*
Percent of population ages 18 and older with limited English proficiency	-0.39	*
Percent of children ages 0-4 living in complex household arrangements	-0.34	*
Percent of children ages 0-4 living in multigenerational households	-0.42	*
Percent of children 0-4 who live with nonrelatives or in group- quarters	-0.37	*
Percent of children 0-17 who are not biological, adopted or step-children of householder	-0.54	*
Percent of children ages 1-4 who have moved in the past year	0.11	
Percent of all householders who are ages 15-34	0.27	*
Housing		
Percent of persons living in renter-occupied households	-0.65	*
Percent of housing units that are vacant	0.07	
Percent of household population living in multi-unit buildings	-0.52	*
Percent of housing units that are crowded (1.01 or more occupants per room)	-0.41	*
Percent of households that have 7+ people	0.01	
Percent of households without internet access	-0.09	
Percent of households without broadband internet	0.20	
Internet and Self-response rate		
Percent Households with Internet Access	-0.09	
Percent households without broadband access	0.20	
Self response rate (10/28/2020)	0.23	
* Statistically significantly different from zero at 10 level of significance		
** Persons of color are anyone other than Non-Hispanic White Alone.		

Table 2. Correlation Coefficients Between 2020 Census State Young Child Coverage Rates and 30 Potential Explanatory Variables

	Correlation with State young child (ages 0-4) coverage rate	Statistical Significance*
Race/ethnicity and Immigration		
Percent Non-Hispanic Black alone (0-4)	-0.59	*
Percent Hispanic (0-4)	-0.20	
Percent Children of Color** (0-4)	-0.66	*
Percent non-Hispanic Black alone (families with related children)	-0.61	*
Percent Hispanic (families with related children)	-0.23	
Percent Families of Color * (families with related children)	-0.67	*
Percent of adults ages 18+ who are foreign-born	-0.42	*
Percent of adults ages 18+ who are not US citizens	-0.42	*
Socioeconomic Status		
Percent of families who have children ages 0-4 and live in poverty	-0.03	
Percent of persons in poverty	-0.10	
Percent of persons 0-4 in poverty	-0.11	
Percent of adults ages 18+ with less than HS or GED	-0.30	*
Family and Household Characteristics		
Percent of households linguistically isolated	-0.39	*
Percent of population ages 18 and older with limited English proficiency	-0.39	*
Percent of children ages 0-4 living in complex household arrangements	-0.34	*
Percent of children ages 0-4 living in multigenerational households	-0.42	*
Percent of children 0-4 who live with nonrelatives or in group- quarters	-0.37	*
Percent of children 0-17 who are not biological, adopted or step-children of householder	-0.54	*
Percent of children ages 1-4 who have moved in the past year	0.11	
Percent of all householders who are ages 15-34	0.27	*
Housing		
Percent of persons living in renter-occupied households	-0.65	*
Percent of housing units that are vacant	0.07	
Percent of household population living in multi-unit buildings	-0.52	*
Percent of housing units that are crowded (1.01 or more occupants per room)	-0.41	*
Percent of households that have 7+ people	0.01	
Percent of households without internet access	-0.09	
Percent of households without broadband internet	0.20	
Internet and Self-response rate		
Percent Households with Internet Access	-0.09	
Percent households without broadband access	0.20	
Self response rate (10/28/2020)	0.23	
* Statistically significantly different from zero at 10 level of significance		
** Persons of color are anyone other than Non-Hispanic White Alone.		

Of the 30 potential explanatory variables, correlation coefficients for 17 are statistically significant and all of those are in the expected direction. Statistically significant correlations are found in all of the domains except the internet and self-response rate variables. But variables in some domains are more highly correlated with young child census coverage than others. The results will be discussed domain by domain.

Several of the most highly correlated measures are related to race/Hispanic origin and immigration. Six of the eight potential explanatory variables in that domain have a statistically significant correlation. In the 2020 Census the net undercount rate for Black Alone or in Combination children ages 0 to 4 was 7.2 percent and for Hispanic children in this age range the net undercount was 8.6 percent so this finding is not surprising.

Many variables in the family and household characteristics domain are also highly correlated with young child census coverage across states. Of the eight potential explanatory factors correlations in this domain, seven are statistically significant. The only measure that was not significant was moving in the past year.

Of the four potential explanatory variables in the socio-economic status domain, only the education of adults is statistically significant. The lack of statistically significant correlations is surprising.

In the housing domain, only three out of the seven characteristics showed a statistically significant correlation.

None of the three measures in the internet and self-response rates domain showed a statistically significant correlation.

6) Discussion

The analysis here is largely correlational and it is widely understood that correlation is not the same as causation. Nonetheless, it is difficult to argue something is a causal factor if it is not correlated with the dependent variable.

It is noteworthy that of the 17 potential explanatory variables that show a statistically significant correlation with the state net young child coverage rates, only one (percent of adults ages 18+ with less than a high school or GED,) had a statistically significant correlation with total population net coverage rates from the PES (O'Hare 2025b). This underscores the extent to which the factors driving the high net undercount of young children in the Census are different than those driving the coverage of the total population and indicates that young children should be seen as a special population in the 2030 Census.

Most of the results found in this study were expected and this shows most of the results related to the 2010 Census are also found in the 2020 Census. However, there are a couple of results that deserve further comment.

Surprisingly, the percent of young children who are Hispanic does not have a statistically significant correlation with state net coverage rates. That may be because Hispanics are highly concentrated in a few states (such as California, Texas, and Florida) and there are a large number of states where the Hispanic population is a small share of the total population. In many states the Hispanic child population is not large enough to drive the overall state coverage rate. In about two-thirds of the states (34 out of 51) Hispanic children make up less than 20 percent of the population ages 0 to 4.

Interestingly, Quiros and O'Hare (2024) found that there was a statistically significant correlation of 0.32 (the way the coverage was measured in this study this means the higher percent Hispanic the worse the coverage rate) when the percent of Hispanic children ages 0 to 4 was correlated with the state net coverage rate for young

Hispanic children. In 2010, the correlation between percent Hispanic and Hispanic young child coverage rates was 0.13 and not statistically significant.

In the housing domain, one of the most intriguing correlations is that between the share of multi-unit buildings and the young child undercount rate. Moreover, it is one of the higher correlations (-0.52) seen here. O'Hare (2014) found a relationship between state young child coverage rates and the percent of young children living in single detached units. The higher the percent of young children living in single detached units, the lower the net undercount. Percent living in single detached units is kind of the complement of those living in multi-unit structures.

Quiros and O'Hare (2024) found the percent of population living in multi-unit building was correlated with the state-level net coverage of young Hispanic children in both 2010 and 2020 at a statistically significant level.

This finding is important because the connection between living in a multi-unit building and being missed in the Census has not been explored in the past. It is also a noteworthy finding because it helps planning for the 2030 Census by alerting us to the kinds of housing units most likely to miss young children.

In conducting the 2020 census, census enumerators were sent to housing units that did not respond to the Census Bureau's invitation to participate sent out in March. There has been some concern about how well these enumerators could gain access to certain housing units such as gated communities and large apartment complexes.

It was discovered in the summer of 2020 that the information Census Bureau enumerators were given to hand out to landlords and property managers to gain access

to the structure was not as strong as it could have been. The information on the enumerators card asked the landlords and property managers to please let the Census enumeration their property. In fact, the laws says landlords and property managers must let the enumerators onto their property.

This issue was discussed at a couple of meetings of the U.S. Census Bureau's National Advisory Committee. The U.S. Census Bureau's National Advisory Committee (2024a, no page number) summarized the concern this way,

> "The NAC recommends that the Bureau, as a means to securing access to restricted access properties, research effective communication (e.g. awareness campaigns) regarding the legal requirement for landlords and property managers to provide access for enumerators and relevant Census employees—and any relevant penalties for landlords, property managers, and agents who do not respond or comply – as well as the benefits to the community of full Census count."

In another related document, the Census Bureau's U.S. Census Bureau's National Advisory Committee (2024b), stated,

> "The NAC recommends that the Census Bureau's communication team communicate clear messages to the public regarding all legal obligations to cooperate with the Census, for example, by landlords and their employees to allow census enumerators onto the property.

Table 3 shows that young Black and young Hispanic children are highly concentrated in multi-unit buildings compared to their non-Hispanic White counterparts. Young Black children and young Hispanic children are two demographic groups that also have remarkably high net undercount rates. This suggests that living in a multi-unit building may be related to being missed in the 2020 Census. To the extent the lack of access to multi-unit buildings was a problem in the 2020 Census it was likely to impact the count of young Black and Hispanic children disproportionately.

Table 3. Percent of Group Living in Multi-Unit Structures *	
	Percent Living in Multi-Unit Structures
All Children ages 0 to 4	18
Non-Hispanic White ages 0 to 4	8
Hispanic ages 0 to 4	24
Black or African American ages 0 to 4	35
Source: 2023 ACS tabulation on IPUMS system	
*Multi-unit structures are defined here as those with 3 or more units in one structure.	

More information about the relationship between multi-unit buildings and census young child undercounts could be extremely helpful in thinking about the 2030 Census. If counting young children in large multi-unit buildings has been a hidden problem in the past, the Census Bureau has a chance to focus more attention on it in the 2030 Census.

Another result that merits discussion is the fact that the internet and self-response rate domain had no statistically significant correlations with state net coverage rates for young children. There are other analyses that show self-response rates are linked to census accuracy for the total population (O'Hare 2020)

The fact that the internet availability and self-response measures were not related to the variations in the net undercount of young children underscores the idea that many young children are missed in households that respond to the Census. There is growing evidence that young children are often missed because respondents do not think they are supposed to be included on the Census form (Griffin and O'Hare 2020; U.S. Census Bureau 2023; Pascale 2024). The 2020 Census Tracking study repeatedly

surveyed the public about their views of the census between January and May 2020, and one of the questions asked respondents about including a young child in the Census. Based on the 2020 Census tracking survey, the Census Bureau (2023, page 37) concludes, “Notably for each of these three weeks we also observed a sizable percentage of respondents who said they were unsure about whether to include a child on their census questionnaire – between 13 and 19 percent across all three time periods and for both modes.” In addition, when Mulry et al. (2024, Figure 1) compared individuals on the 2020 Census roster to administrative records roster for the same household, they found children were often the group left off.

If a large share of young children are missed in the Census because they are left of returned census forms, that would explain why there is no statistically significant correlation between self-response and state level census accuracy for young children.

7) Summary

Census Bureau data show that every state had a net undercount of young children in the 2020 Census. Nonetheless, there was substantial variation across states. Discounting the extremely high net undercount rate in the District of Columbia (15.85 percent) the coverage rates vary from a low of 0.02 in Vermont, to a high 9.87 percent in Florida.

The analysis presented in this paper will help us determine which characteristics are more closely related to the net undercount of young children and which are not. Most of the results seen here were consistent with past research. Variables from five different domains were examined in terms of their correlation with the variation in net

coverage rates of young children in the 2020 Census. This study found 17 of 30 different potential explanatory variables had a statistically significant correlation with state-level variation in the net coverage of young children in the 2020 Census. We found measures of race/ethnicity and immigration highly correlated with variation in net undercounts. Similar results were found with family and household structure characteristics. Of the 17 correlations found to be statistically significant, 6 are from the Race/Ethnicity and Immigration domain and 7 are from the Family and Household Characteristics domain.

Two results deserve further attention. There was a high correlation between the share of the population in a state living in multi-unit buildings and net young child Census coverage. More research on that topic is needed before the 2030 census.

The lack of correlations between self-response rates and young child net coverage supports the idea that many young children are missed because respondents do not list them on the roster of a returned census form. More analysis of the 2020 Census Tracking survey (for example breakdown by race, location, family structure, and other demographic characteristics known to be linked to differences in the undercount of young children.) would help develop compelling messages for the 2030 Census. This demographic research should be combined with communications research to find out what message(s) work best with which groups.

8) References

Griffin D. H. (2014). “Final task force report: Task Force on the undercount of young children.” Memorandum for Frank Vitrano, Washington, DC, U.S. Census Bureau February 2. http://www.census.gov/library/working-papers/2014/demo/2014-undercount-children.html

Griffin, D., and O’Hare, W.P. (2020). “Counting Young children in the U.S. Census – Important Differences Between 2010 and 2020,” Posted on the Leadership Conference Education Fund website. March 29, 2022. https://civilrightsdocs.info/pdf/reports/Counting-Young-Children-in-the-Census-Report.pdf

Johnson, J. (2022). “Does the Census Miss the Native-Born Children of Immigrant Mothers? Evidence from State-Level Undercount by Race and Hispanic Status,” Population Research and Policy Review Vol 41, pp 139-195. https://link.springer.com/article/10.1007/s11113-021-09651-w

King, H., Ihrke, D. and Jensen, E., (2018). “Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census,”” paper present the 2018 Population Association of American Conference, April 25-28 , Denver Colorado, Downloaded May 6, 2018. https://paa.confex.com/paa/2018/meetingapp.cgi/Paper/21374.

Mulry, N., Tello-Trillo, C.J, Mule, V. T., and Keller, A. (2024). “Comparisons of administrative rosters to census self-response and NRFU household member responses,” Statistical Journal of the International Association of Official Statistics, pp 41-52.

National Academy of Sciences (2023). Assessing the 2020 Census : Final Report, National Academy of Sciences Press. https://nap.nationalacademies.org/catalog/27150/assessing-the-2020-census-final-report

O’Hare, W.P. (2014). “State-Level 2010 Census Coverage Rates for Young Children,” *Population Research and Policy Review,* Volume 33, no. 6, pages 797-816. State-Level 2010 Census Coverage Rates for Young Children | SpringerLink

O'Hare, W.P. (2015). *The Undercount of Young Children in the U.S. Decennial Census.* Springer Publishers. https://www.springer.com/gp/book/9783319189161#

O'Hare, W.P., Mayol-Garcia, Y., Wildsmith, E., and Torres, A., (2016). "The Invisible Ones: How Latino Children Are Left Out of Our Nation's Census Count," Child Trends Hispanic Institute & National Association of Latino Elected Officials, Child Trends, Washington DC. https://www.childtrends.org/publications/the-invisible-ones-how-latino-children-are-left-out-of-our-nations-census-count

O'Hare, W.P. (2017). "Geographic Variation in 2010 U.S. Census Coverage Rates for Young Children: A Look at Counties," *International Journal of Social Science Studies,* Vol. 5, No. 9 Sept. Redframe Publishing. 295180286.pdf (core.ac.uk)

O'Hare, W.P. (2020). "Are Self-Participation Rates Predictive of Accuracy in the U.S. Census?" International Journal of Social Science Studies, Vol 8. No. 6 November, pp 23-34. https://doi.org/10.11114/ijsss.v8i6.4967

O'Hare, W.P, Jacobsen, L.A., Mather, M. VanOrman and Pollard, K. (2019a). "What Factors Are Most Closely Associated with the Net Undercount of Young Children in the U.S. Census?" Population Reference Bureau, Washington, DC. https://www.prb.org/wp-content/uploads/2019/03/net-undercount-children-acs.pdf

O'Hare, W.P., Griffin, D., and Konicki, S. (2019b). "Investigating the 2010 Undercount of Young Children - Summary of Recent Research," 2020 Census Memorandum Series, U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-2010-undercount-children-summary-recent-research.pdf

O'Hare, W.P, Jacobsen, L.A., Mather, M. VanOrman (2020). "Predicting Tract-Level Undercount Risk for Young Children," Population Reference Bureau, Washington, DC. https://www.prb.org/wp-content/uploads/2020/02/us-census-undercount-of-children.pdf

O'Hare, W.P. (2023a). "County-level Coverage Rates of Young Children in the 2020 Census: The National-Level Data Do Not Tell the Full Story." October, Posted on the Count All Kids website. https://countallkids.org/resources/county-level-coverage-rates-of-young-children-in-the-2020-census-the-national-level-data-do-not-tell-the-full-story/

O'Hare, W.P. (2023b) "State Undercount Rates for Young Children in the 2020 Census" August 2023. https://countallkids.org/resources/state-undercount-rates-for-young-children-in-the-2020-census/

O'Hare, W.P. (2024a) .WHAT THE NEW CENSUS BUREAU DEMOGRAPHIC ANALYSIS (DA) EXPERIMENTAL YOUNG CHILD COVERAGE ESTIMATES FOR STATES AND COUNTIES TELL US ABOUT METHODOLOGY, May 16, 2024.https://countallkids.org/wp-content/uploads/2024/05/WHATTH1.pdf

O'Hare, W.P. (2024b). WHAT THE NEW CENSUS BUREAU DEMOGRAPHIC ANALYSIS (DA) EXPERIMENTAL YOUNG CHILD COVERAGE ESTIMATES FOR STATES AND COUNTIES TELL US ABOUT METHODOLOGY, May 16, 2024. https://countallkids.org/wp-content/uploads/2024/05/WHATTH1.pdf

O'Hare W. P. (2024c). "Counties Where the Coverage for Young Children Deteriorated Between 2010 and 2020." January, Posted on Count All Kids website. https://countallkids.org/resources/counties-where-the-coverage-for-young-children-deteriorated-between-2010-and-2020/

O'Hare, W.P. (2024d). THE NET COVERAGE OF CHILDREN AGES 5 TO 9 IN THE U.S. CENSUS, MARCH 12, 2024. https://countallkids.org/wp-content/uploads/2024/03/The-Net-Coverages-of-Children-age-5-to-9-in-the-2020-Census.pdf

O'Hare W.P. and Quiros, S. (2025a). "Analysis of 2020 Census Net Undercount of Young Hispanic Children by County," January, Posted on the Count All Kids website. https://countallkids.org/wp-content/uploads/2025/01/Analysis-of-2020-Census-Net-Undercount-of-Young-Hispanic-Children-by-County.pdf

O'Hare, W. P., (2025b). "DATA SHOW YOUNG CHILDREN ARE MISSED IN THE CENSUS FOR DIFFERENT REASONS THAN ADULTS ARE MISSED, April 17, 2025.

https://countallkids.org/resources/data-show-young-children-are-missed-in-the-census-for-different-reasons-than-adults-are-missed/

O'Hare, W.P. (2025c). Analysis Shows There Was an Extremely High Net Undercount of Young Black Children in the 2020 Census. https://countallkids.org/resources/analysis-shows-there-was-an-extremely-high-net-undercount-of-young-black-children-in-the-2020-census/

Pascale J. (2024). "The Undercount of Young Children in the Census: A Synthesis of the Literature and an Examination of a New Data Source," Paper delivered at the Southern Demographic Association, Savannah, GA Oct 16-18.

Pascale, J. (2025). Targeting At-Risk Households to Understand and Reduce the Undercount of Young Children: A Research Agenda, U.S. Census Bureau, Research Report Series (Survey Methodology #2025-04). https://www.census.gov/library/working-papers/2025/adrm/rsm2025-04.html

Quiros, S. and O'Hare W.P. (2024) Towards a More Accurate Count: Identifying State-Level Predictors of the Undercount of Young Hispanic Children in the 2010 and 2020 U.S. Censuses," *Population Research and Policy Review*, Vol. 43, No 75, November 11, 2024. https://link.springer.com/article/10.1007/s11113-024-09918-y

Seltzer, J., and Walker, D. K. (2020). "Counting Children in the U.S. 2020 Census: Assure Our Future is Represented" Global Alliance for Behavioral Health and Social Justice and Boston University School of Public Health March 23, 2020.https://nam.edu/wp-content/uploads/2020/03/Counting-Children-in-the-US-2020-Census.pdf

U.S. Census Bureau (2017a). Investigating the 2010 Undercount of Young Children – Geographic Distribution of Coverages Followup Results.; FINAL REPORT Deborah Griffin and Scott Konicki, July. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-2010-undercount-children-coverage-followup-results.pdf

U.S. Census Bureau (2017b). Investigating the 2010 Undercount of Young Children – Analysis of Coverage Follow up Results Using the Esri Tapestry Segmental and Planning Data Base, Deborah Griffin and Scott Konicki, July. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-2010-undercount-children-coverage-followup-analysis.pdf

U.S. Census Bureau (2017c). Investigating the 2010 Undercount of Young Children – A Comparison of Demographic, Housing and Household Characteristics of Children by Age. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-2010-undercount-children-characterisitcs-by-age.pdf

U. S. Census Bureau (2019). "Investigating the 2010 Undercount of Young Children - Summary of Recent Research," O'Hare, W.P., Griffin, D., and Konicki, S., 2020 Census Memorandum Series, U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-2010-undercount-children-summary-recent-research.pdf U.S. Census Bureau, Washington, DC.

U.S. Census Bureau (2024). "Exploring Undercount of Young Children in the 2020 Census by County. Eric Jensen and George M Hayward. https://www.census.gov/library/stories/2024/04/children-undercount.html

U.S. Census Bureau's National Advisory Committee (2024a). "Memo from Cherokee Bradley to Robert Santos, Recommendations and Comments to the U.S. Census Bureau from the National Advisory Committee 2023 Special Session Meeting" January 29. https://www2.census.gov/about/partners/cac/nac/meetings/2023-07/response-nac-special-session-2023-recommendations.pdf

U.S. Census Bureau's National Advisory Committee (2024b). "Memo from Cherokee Bradley to Robert Santos, Recommendations and Comments to the U.S. Census Bureau from the National Advisory Committee 2023 Fall Meeting" April 30. https://www2.census.gov/about/partners/cac/nac/meetings/2023-11/nac-fall-2023-recommendations.pdf