

Data Show Young Children Are Missed in the Census for Different Reasons Than Adults Are Missed

By

Dr. William P. O'Hare^{1, 2}

April 2025

Executive Summary

In every U.S. Census, some people are missed and some are double counted. These errors are reflected in the Census coverage rate. It is less clear what causes these errors, and the errors are not always consistent across demographic groups. Are the net coverage rates of total population (all ages) and young children driven by the same set of factors or social forces? Several data points suggest they are not. After reviewing a couple of key data points from past research, this study adds new evidence on this issue by looking at correlation coefficients across states in the 2020 Census. Tables are developed showing correlation coefficients between state coverage rates (young child net coverage rates and total population net coverage rates) with each other and with 30 potential explanatory variables. If young children and total population census coverage rates are driven by the same forces, one would expect the correlation between total population net coverage rate and young child net coverage rates to be high and positive across states. One would also expect the correlations between total and young child state coverage rates and each of 30 potential independent variables (also called causal factors) to be similar in size and direction. In the 2020 Census, the correlations between state coverage rates for total population and state coverage rates for young children was $-.18$, which is not statistically significant and is in the wrong direction. If the major forces driving these two measures of census accuracy are the same

¹ Consultant to the Count All Kids Campaign

² This research was funded by the Annie E. Casey Foundation, but they are not responsible for the content of this publication

one would expect a high positive correlation across states. In looking at potential explanatory variables, it is clear that many other dissimilarities across correlations indicate the census accuracy for young children is driven by different forces than the accuracy for the total population. Only 3 of 30 correlations are in agreement. These results suggest It is important for the Census Bureau to develop operations, methods, and a communication outreach campaign for young children separate from those for total population in the 2030 Census.

Data Show Young Children Are Missed in the Census for Different Reasons Than Adults Are Missed

By

Dr. William P. O'Hare

Introduction

While 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 2022; Quiros and O'Hare 2024; O'Hare 2023; Pascale 2024), few studies have examined why the coverage results for young children might be different than the results for the total population.

This is important because there are some people who suggest that the census results for the total population and for young children are driven by the same forces. For example, in a review of activities undertaken in the 2020 Census to improve the count of young children, the Census Bureau listed several activities designed to improve the overall Census Count (Deaver 2021) thus suggesting a belief that improving the overall count would largely resolve the problems with the count of young children.

On the other hand, O'Hare (2024b) provides evidence suggesting the coverage of the total population and the coverage of young children are driven by different influences.

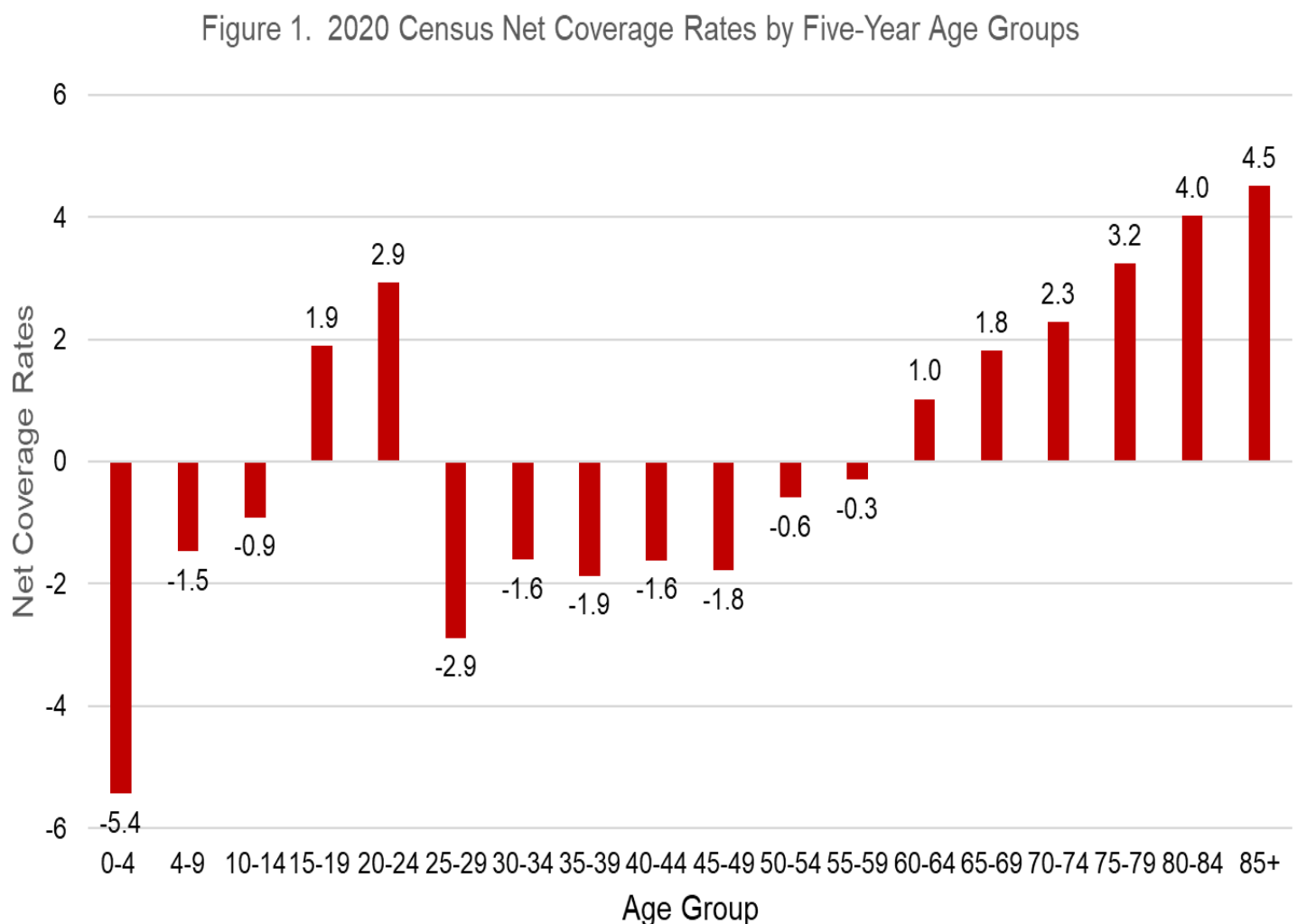
This is an important issue in terms of preparation for the 2030 Census. If the coverage of young children and the coverage of the total population are driven by different forces, it is important for the Census Bureau to develop operations, methods, and a communication outreach campaign for young children separate from the total population (or adults). This point is particularly significant because the net undercount of young children in the 2020 Census (5.4 percent) is higher than any other age group and (unlike other age groups) the net undercount of young children has been increasing steadily since the 1980 Census.

If the Census Bureau is going to reverse this negative trend, it is important to get a better understanding of what factors are driving the net coverage of young children. If the Census Bureau keeps doing the same thing it has done for the past forty years, there is no reason to believe the Census count of young children will improve in 2030.

After providing some background information from past research, new data on correlations across states reveal the extent to which census state net coverage rates for young children and the total population are dissimilar. That is followed by examination of whether state coverage rates (total population and young children) are consistently related to 30 potential explanatory variables in the same way.

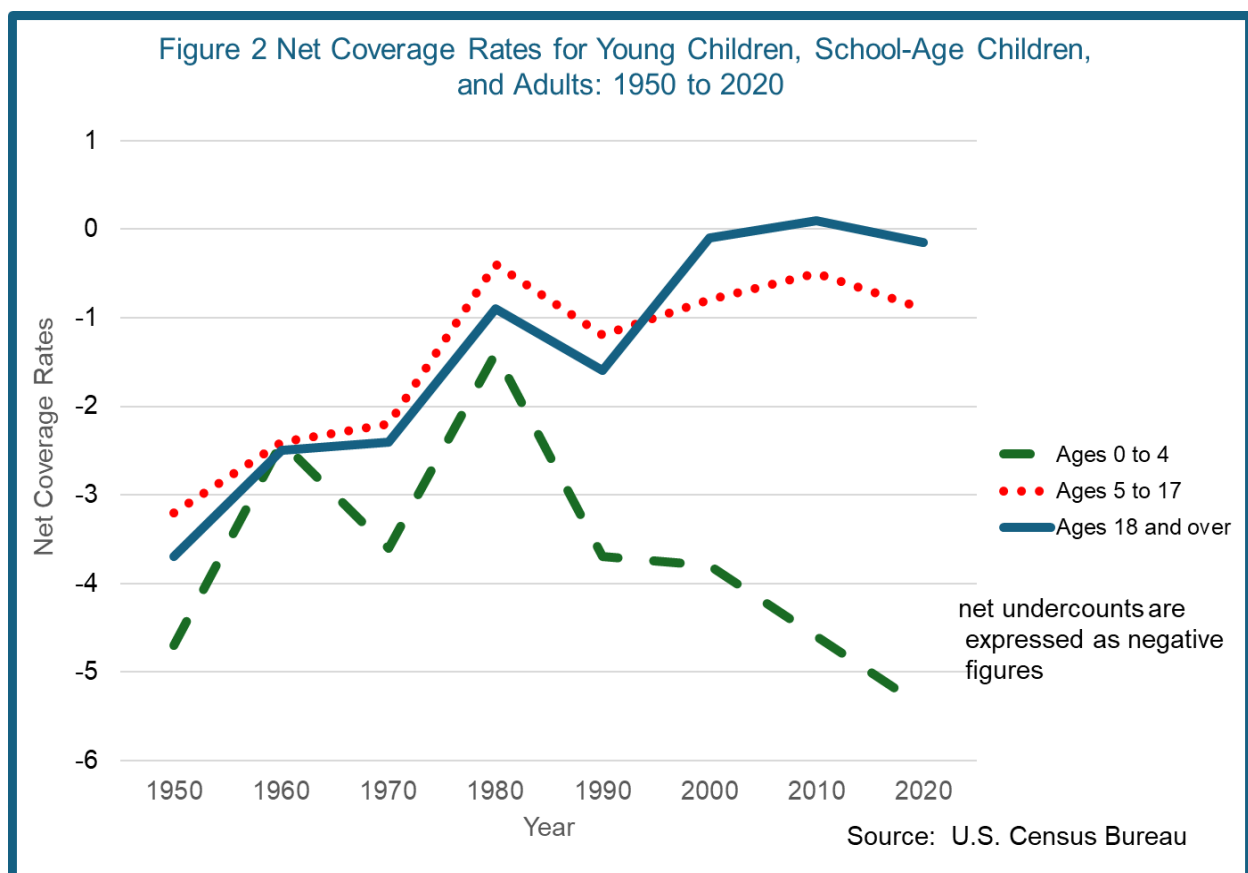
Background

Figure 1 shows the net coverage rates in the 2020 Census by five-year age groups. The undercount rate for young children (5.4 percent) is much worse than 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 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.



Source: Author's analysis of Census Bureau data

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 trends for young children are 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.



The high net undercount of young children is geographically pervasive. Children ages 0 to 4 were undercounted in every state in the 2020 Census (U.S. Census Bureau 2024). More than two-thirds of counties had a net undercount of young children (O'Hare 2023).

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 2024a). The persistent difference between the coverage of children ages 0 to 4 and those ages 5 to 9 raises an important question. Why do the youngest children have a much higher net undercount rate than slightly older children?

Data Sources for This Analysis

The state net coverage rates for the total population are taken from a publication from the U.S. Census Bureau (2022) based on the Census Bureau's Post-Enumeration Survey (PES)³. To some extent the coverage rates for the total population reflect coverage of the adult population since 78 percent of the total population are age 18 and over based on the 2020 Census. The PES state coverage estimates have sampling error

³ The District of Columbia is included as a state.

associated with them, but they are still the best estimates we have for census coverage at the state level.

The state net coverage rates for the population ages 0 to 4 are from an experimental Demographic Analysis (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 PES estimates, but there is undoubtedly some random error in the DA estimates. To reflect the idea of errors in the DA estimates of coverage, 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 a low of 19,427,000 to a high of 19,480,000, indicating that the DA data is likely to have little error.

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 2024; O'Hare et al. 2019; O'Hare 2015). The specific data for the potential explanatory variables used in this study largely are taken from Quiros and O'Hare (2024).

The potential explanatory variables used here are not an exhaustive list, but they provide factors from five different domains and are a

reasonable set of factors to assess the similarity of state-level correlations for the total population coverage and the coverage of young children.

Based on four years of research on the undercount of young children in the 2010 Census, a Census Bureau report found several characteristics of young children associated with a higher risk of coverage errors (U.S. Census Bureau 2019, page 2).

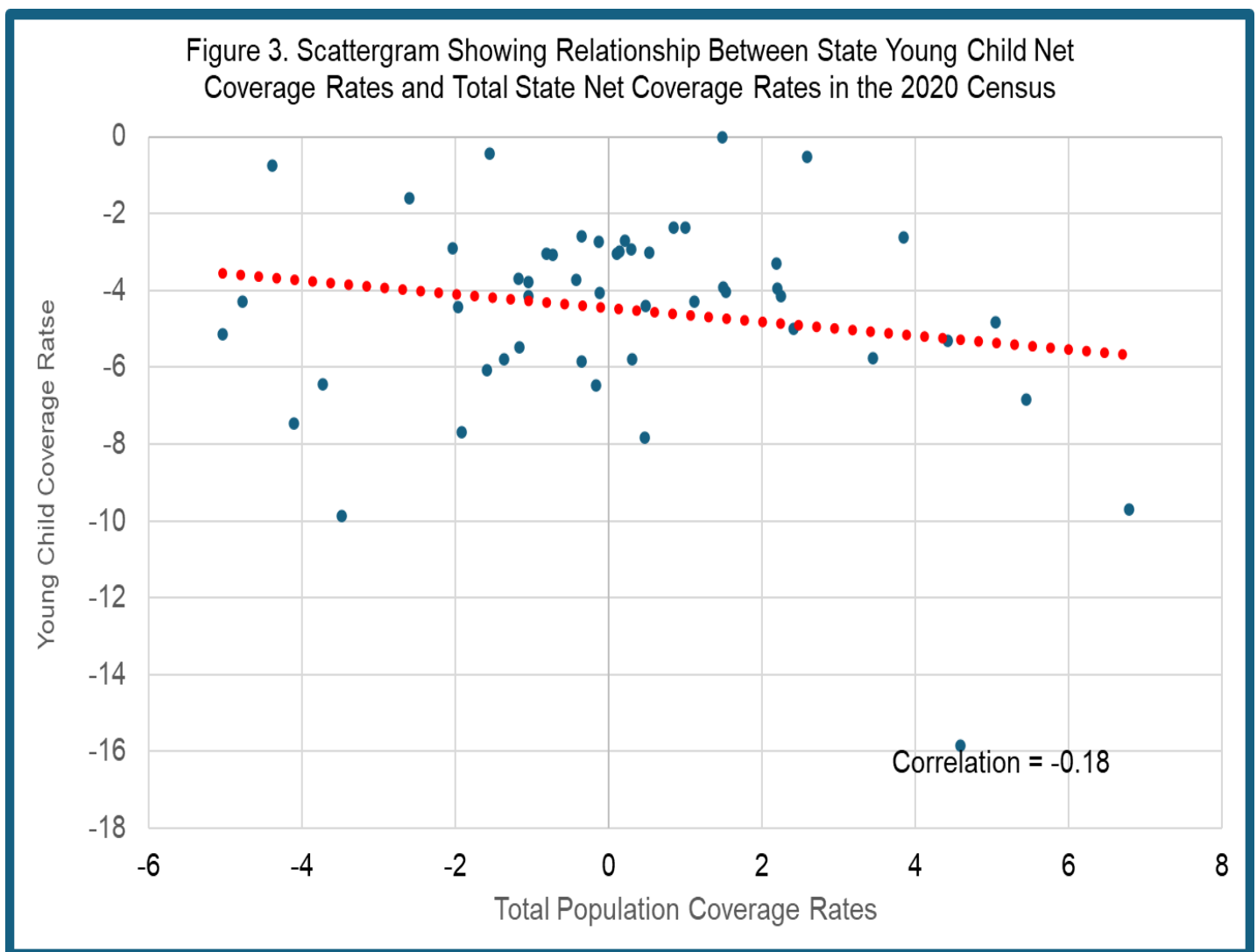
But many of the factors listed in that study such as being a renter, being Hispanic, or not be counted by self-response, apply to groups besides young children. It is important to identify factors that make young children more likely than other age groups to be missed or have a net undercount.

Analysis

First, results are shown for the relationship between 2020 Census state coverage estimates for young children and the total population. Then correlations between 30 potential explanatory variables and state level coverage for young children and total population are examined.

Figure 3 is a scattergram showing the relationship between state net coverage rates for young children and state net coverage rates for the total population (all ages) in the 2020 Census.

If the forces driving the coverage of young children and the coverage of the total population are the same or similar, then one would expect a modest to high positive correlation. The correlation is -0.18, which is in the wrong direction if both coverage rates are driven by the same forces, and it is not statistically significant.



This evidence shows that the coverage rates of young children and the coverage rates of the total population are influenced by different factors.⁴

Correlations with Potential Explanatory Variables.

In this section of the paper, I examine how state-level variation in net coverage rates for the total population and for young children are correlated with a set of potential explanatory variables. The potential explanatory variables were derived from a host of past studies and have been sorted into five domains. 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).

Domains are:

1. Race/ethnicity and immigration
2. Socioeconomic status
3. Family structure/living arrangements

⁴ Data for individual states can be found on the Census Bureau website at <https://www.census.gov/library/visualizations/interactive/2020-post-enumeration-survey.html> and <https://www.census.gov/library/visualizations/interactive/net-coverage-error-young-children.html>.

4. Housing

5. Internet availability and response rates

Are the relationships between state level total population net coverage rates and potential explanatory factors the same as the relationship between young child coverage rates and the same potential explanatory factors? If the same forces are driving total population and young child coverage rates, one would expect the correlations with state young child coverage rates and total population coverage rates to be similar.

Keep in mind the net coverage of young children and the total population is constructed 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.

Table 1. Correlation Coefficients Between State Total Population Coverage Rates and Young Child Coverage Rates With 30 Possible Explanatory Variables in the 2020 Census

	Correlation with State Total Population (all ages) coverage rate	Statistical Significance *		Correlation with State young child (ages 0-4) coverage rate	Statistical Significance *	Correlations Agree on Sign and Statistical Significance
Race/ethnicity and Immigration						
Percent Non-Hispanic Black alone (0-4)	-0.18			-0.59	*	No
Percent Hispanic (0-4)	0.14			-0.20		Yes
Percent Children of Color** (0-4)	0.21			-0.66	*	No
Percent non-Hispanic Black alone (families with related children)	-0.17			-0.61	*	No
Percent Hispanic (families with related children)	0.13			-0.23		No
Percent Families of Color * (families with related children)	0.24	*		-0.67	*	No
Percent of adults ages 18+ who are foreign-born	0.32	*		-0.42	*	No
Percent of adults ages 18+ who are not US citizens	0.28	*		-0.42	*	No
Socioeconomic Status						
Percent of families who have children ages 0-4 and live in poverty	-0.41	*		-0.03		No
Percent of persons in poverty	-0.49	*		-0.10		No
Percent of persons 0-4 in poverty	-0.42	*		-0.11		No
Percent of adults ages 18+ with less than HS or GED	-0.28	*		-0.30	*	Yes
Family and Household Characteristics						
Percent of households linguistically isolated	0.24	*		-0.39	*	No
Percent of population ages 18 and older with limited English proficiency	0.18			-0.39	*	No
Percent of children ages 0-4 living in complex household arrangements	0.20			-0.34	*	No
Percent of children ages 0-4 living in multigenerational households	0.12			-0.42	*	No
Percent of children 0-4 who live with nonrelatives or in group- quarters	0.36	*		-0.37	*	No
Percent of children 0-17 who are not biological, adopted or step-children of householder	0.14			-0.54	*	No
Percent of children ages 1-4 who have moved in the past year	-0.36	*		0.11		No
Percent of all householders who are ages 15-34	-0.16			0.27	*	No
Housing						
Percent of persons living in renter-occupied households	0.30	*		-0.65	*	No
Percent of housing units that are vacant	-0.16			0.07		Yes
Percent of household population living in multi-unit buildings	0.48	*		-0.52	*	No
Percent of housing units that are crowded (1.01 or more occupants per room)	0.33	*		-0.41	*	No
Percent of households that have 7+ people	0.28	*		0.01		No
Percent of households without internet access	-0.43	*		-0.09		No
Percent of households without broadband internet	-0.51	*		0.20		No
Internet and Self-response rate						
Percent Households with Internet Access	-0.43	*		-0.09		No
Percent households without broadband access	-0.51	*		0.20		No
Self response rate (10/28/2020)	0.11			0.23		Yes
* Statistically significantly different from zero at 10 level of significance						
** Persons of color are anyone other than Non-Hispanic White Alone.						

Table 1 shows the correlations between coverage rates (total population and young child) and 30 potential explanatory variables. A positive correlation means as the potential explanatory variable increases, the coverage rate also increases (gets better). A negative correlation means as the potential explanatory variable increases, the coverage rate

decreases (gets worse). If there is no statistical significance it means the relationship is not strong and might even be zero.

The last column of Table 1 indicates where the correlations for the total population coverage and the young child coverage agree. When they agree, it indicates that this force affects both the count of young children and the count of adults. Where they do not agree, it suggests that this force effect is not consistent in how it affects the coverage of young children and the total population.

Both the sign of the correlation coefficients (positive and negative correlation) and magnitude (statistical significance) are examined to determine agreement. If both correlations are in the same direction and statistically significant, I determined it was an agreement (that is, the factor affects the census accuracy of children and adults similarly) and if both coefficients were not statistically significant, I determined it was an agreement. The degree of statistical significance is not shown in the tables but correlation coefficients with a magnitude above .23 are statistically significant.

If the forces affecting coverage of young children and the total population are the same, one would expect a high level of agreement in the 30 sets of correlations.

In summary, Table 2 shows that young children and the total population have different relationships with the vast majority of potential explanatory variables. Of the 30 potential explanatory variables in Table 1, there are only 3 where the correlation coefficient for state total population net coverage rates and the correlation coefficient for state young child net coverage rates agree on sign/direction and magnitude (statistically significance) and in two of three both are not statistically significant. The only factor that seems to be statistically significant in the same way for both total population and young children is the percent of adults with less than a high school degree or GED, which suggests that improving efforts to reach households with low literacy could improve the count of both the total population and young children.

Table 2. Summary Table of Agreement Between State-Level Total Population Coverage Correlations and Young Child Correlations with 30 Possible Independent Variables by Domain		
	Number of Characteristics	Number of correlations where there is agreement on direction and statistical significance
Race/ethnicity and immigration status	8	1
Socioeconomic status	4	1
Family structure /living arrangements	8	0
Housing	7	0
Internet availability and response rates	3	1
Total	30	3

Summary and Conclusions

All of the data examined here support the idea that most factors or forces affecting census coverage of young children are different than those affecting coverage of the total population.

Figures 2 and 3 show the net coverage of young children and the total population are dissimilar over time and across states. Over the past 40 years, the census accuracy trend of young children has been quite different than that of adults and even different than older children. Figure 2 shows what has been done for the past forty years to improve the count of young children in the Census is not working. Figure 3 shows us that there is no statistically significant correlation between the undercount of young children and the count of the total population across states.

This study shows very few (only three out of 30) potential explanatory variable correlations between the state net coverage rates for the total population and state net coverage rates for young children in the 2020 Census were similar. This is another sign that the factors driving the coverage of young children are different than those driving the coverage of the total population.

Collectively, this provides convincing evidence that the variations of coverage for young children and the total population are driven by different factors or forces. General improvements in counting the total population are unlikely to improve the coverage of young children.

Given the results of the past 40 years, the Census Bureau needs to try new bold strategies specifically focused on counting young children in the 2030 Census. Tweaking past approaches is not likely to work (O'Hare 2024c). Unless we do something quite different in the 2030 census from what we have done in the past forty years, we should expect the coverage of young children to continue to deteriorate. Many such ideas have been proffered in the public comments at Census Bureau advisory committee meetings in the past few years and some are summarized by Stein (2025).

Because the Census operational plans take years to develop, the Census Bureau needs to assess substantial and innovative methods soon in order to embed successful approaches into the 2030 Census operations. Child advocates would be encouraged if the Census Bureau could identify major new elements in their plan for 2030 that tests have shown will improve the count of young children significantly. The Census Bureau used some new approaches for counting young children in the 2020 Census (Walejko, G and Konicki, S.2018; Deaver 2021) but as far as I can tell they

did not evaluate any of them and the overall result was the net undercount of young children increased between 2010 and 2020.

The bottom line is summed up well by a recent recommendation from the National Census Bureau Advisory Committee to the Census Bureau (2024):

“The current methods employed by the U.S. Census Bureau to identify and count young children have not worked. The problem is getting worse. The Census Bureau needs to test new methods, operations, and strategies to improve the undercount of young children. Such changes need to be well-tested before they are used in the 2030 census.

11. Recommendation: the NAC Recommends that the U.S. Census Bureau report to the NAC and the public on what they plan to do differently in the 2030 census to improve the count of young children and what research and testing will take place to inform operations,”

References

Deaver, K (2021). "Thank You To Our Partners, Hosted by Partnership for America's Children," April 1, 2021

Griffin, D., and O'Hare W. P. (2020). "Are Census Omissions of Young Children Due to Respondent Misconceptions about the Census?" *International Journal of Social Science Studies*, Vol 8. No. 6 November 59-72. <http://redfame.com/journal/index.php/ijsss/article/view/4994/5223>

Johnson J., E., (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*, 2022, 41(1) pp 139-195, <https://link.springer.com/article/10.1007/s11113-021-09651-w>

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

National Census Bureau Advisory Committee (2024). Memo to Robert Santos Director U.S. Census Bureau From: Cherokee Bradley National Advisory Committee (NAC) Chair Recommendations and Comments to the Census Bureau from the National Advisory Committee – 2024 Spring Meeting May 2, 2024

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, Jacobsen, L.A., Mather, M. VanOrman and Pollard, K (2019). "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. (2023). "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. (2024a). "THE NET COVERAGE OF CHILDREN AGES 5 TO 9 IN THE U.S. CENSUS," *March 12, 2024*, posted on the Count All Kids website, <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. (2024B). "WHAT PAST RESEARCH TELLS US ABOUT HOW TO PREPARE FOR THE 2030 U.S. CENSUS COUNT OF YOUNG CHILDREN." *June 11, 2024*, <https://countallkids.org/wp-content/uploads/2024/06/What-past-research-tells-us-about-2030-Census-count-of-young-children.pdf>

O'Hare, W. P. (2024c). "No Time for Tweaking," October 9, Posted on Count All Kids Website, <https://countallkids.org/resources/no-time-for-tweaking/>

Quiros, S. and O'Hare W.P. (2024). Identifying State Level Predictors of the Net Undercount of Young Hispanic Children in the 2010 and 2020 U.S. Census. Forthcoming, Population Research and Policy Review,

Pascale, J. (2024). "The Undercount of Young Children in the Census: A Synthesis of Literature and an Examination of a New Data Source: Paper presented at the Southern Demographic Association Conference. October 17, Savanna, GA.

Seltzer, J.A., and Walker, D.K., (2020). "Counting Children in the 2020 Census: Assure Our Future is Represented," National Academy of Medicine, <https://nam.edu/counting-children-in-the-u-s-2020-census-assure-our-future-is-represented/>

Stein, D. (2025), RECOMMENDATIONS TO THE CENSUS BUREAU FOR IMPROVING THE COUNT OF YOUNG CHILDREN IN 2030 April 3, 2025 , <https://countallkids.org/resources/recommendations-to-the-census-bureau-for-improving-the-count-of-young-children-in-2030/>

U.S. Census Bureau (2019) "Investigating the 2010 Undercount of Young Children - Summary of Recent Research," 2020 Census Memorandum Series, O'Hare, W.P., Griffin, D., and Konicki, S. 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 (2022b). "Census Coverage Estimates for People in the United States by State and Census Operation, PES20-G-02RV, by Courtney Hill, Kristen Heim, Jinhee Hong, and Nam Phan, U.S. Census Bureau, <https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/census-coverage-estimates-for-people-in-the-united-states-by-state-and-census-operations.pdf>

U.S. Census Bureau (2024). "Census Bureau Releases Experimental Estimates of State and County Undercounts and Overcounts of Young Children in the 2020 Census, April 11, 2024. Press Release Number: CB24-CN.12, <https://www.census.gov/newsroom/press-releases/2024/undercounts-overcounts-children-2020-census.html>

Walejko, G and Konicki, S. (2018) Census Efforts to Reduce the Undercount of Young Children, Presentation at the 2018 Joint Statistical Methods conference.