What Factors Are Most Closely Associated With the Net Undercount of Young Children in the U.S. Census?

By
William P. O'Hare,* Linda A. Jacobsen,** Mark Mather,** Alicia VanOrman,** and Kelvin Pollard**

*O'Hare Data and Demographic Services, LLC
**Population Reference Bureau

Acknowledgments: We would like to acknowledge support from the 2020 Census Project, a funders collaborative to promote a fair and accurate census, hosted by New Venture Fund. We would also like to thank the U.S. Census Bureau for making county-level net undercount data available for this analysis, and Paola Scommegna, Heidi Worley, Jessica Woodin, and Lillian Kilduff at PRB for their assistance in editing this report.

Table of Contents

Executive Summary

In the 2010 U.S. Census, young children (ages 0 to 4) were more likely to be missed than any other age group, Census Bureau research shows. More than one in 10 young children were not counted in 2010, and the net undercount rate for young children (the percent of children who were missed minus the percent who were double counted) was nearly 5 percent. The net undercount for young children has been increasing while that for adults has been improving since the 1980 Census.

While several factors have been linked to the undercount of young children, most previous studies on this topic have been descriptive, rather than analytical in nature. Our analysis focuses on the factors that are most closely associated with the net undercount of children in the census. Data are based on the Census Bureau's new experimental Demographic Analysis (DA) estimates of the net undercount of young children and American Community Survey (ACS) estimates for the 261 largest counties.

Our results suggest that two data sets currently being used to identify areas where young children are more likely to be missed by the census—the percent of young children living in hard-to-count census tracts and the Low Response Score (based on mail return rates)—are not very good predictors of net undercount rates for young children in large counties.

Our preliminary results indicate that a higher net undercount of young children is most closely associated with the following variables:

- Percent of racial/ethnic minorities.
- Percent of households that are linguistically isolated.
- Percent of young children living with grandparent householders.
- Percent of young children living with nonrelatives or in group quarters.

The superior results of our model in predicting net undercount rates for young children indicate this line of research should be pursued further to better identify hard-to-count areas for young children in the 2020 Census.

We hope the results of this analysis will help advocates and others better target geographic areas and population subgroups for Get-Out-the-Count efforts to reduce the undercount of young children and help ensure an accurate 2020 Census.

1. Introduction

In the 2010 U.S. Census, young children (ages 0 to 4) had higher net undercount and omissions rates than any other age group.[a] The net undercount rate for young children was 4.6 percent, and more than 10 percent of young children were missed in the 2010 Census.

Given the high nationwide net undercount rate for young children, it would be useful to gain a better understanding of the geographic differences in census coverage rates for young children, and why children are more likely to be missed in certain areas than others. Yet, few studies have investigated the characteristics associated with children being missed in the census.[1] This is at least partly due to the lack of subnational undercount estimates that could be used as dependent variables in multivariate analysis.

This study takes advantage of a unique set of U.S. Census Bureau estimates to examine the factors associated with 2010 Census net undercount rates for young children in the 261-largest U.S. counties—those with populations of 250,000 or more. Only the most populous counties are used in this study because they are likely to have the most accurate estimates of net undercount of young children. Collectively, these 261 counties accounted for about 73 percent of the net undercount of young children in the 2010 Census. Census coverage of young children varies widely across these counties, ranging from a 14.3 percent net undercount to an overcount of 8.4 percent.

A set of 32 potential explanatory variables are considered in our analysis, based on past Census Bureau research and our review of the general literature on census accuracy. The 32 potential explanatory variables are sorted into six different domains including:

1. Race and Hispanic Origin.
2. Socioeconomic Status.
3. Family Structure and Living Arrangements.
4. Other Demographic Measures.
5. Housing.
6. Census Response/Return Rates.

Zero-order correlations between the net undercount rates for young children and the potential explanatory factors are examined first, followed by a multiple regression analysis.

Variables from two different domains—racial/ethnic diversity and children's living arrangements—have the strongest associations with young children's net undercount rates in our model, while some of the variables thought to be closely related to census accuracy for the total population (like percent of housing units that are rental units) are not statistically significant. Our results suggest that the variables most closely associated with the net undercount of young children are different from the variables linked to variation in self-response rates (mail return rates) in the 2010 Census. We hope the results of this analysis will help advocates and others better target geographic areas and population subgroups for Get-Out-the-Count efforts to reduce the undercount of young children and help ensure an accurate 2020 Census.

---

[a] Omissions are people who should have been counted in the Census but were not. Net undercounts represent a balance between two groups. One group is people omitted from the Census. The second group is erroneous enumerations (mostly people counted twice) and whole-person imputations.

2. Background on the Net Undercount of Young Children

The high net undercount of young children and the need for subnational data on census accuracy are not new issues. More than 100 years ago, Young stated, “Experience has shown that it is extremely difficult to ascertain the true number of children in any population by simple enumeration.”[2] A passage from a 1940 Decennial Census report underscores the point, “Underenumeration of children under 5 years old, particularly of infants under 1 year old, has been uniformly observed in the United States Census and in the Censuses of England and Wales and of various countries of continental Europe.”[3] With respect to the situation in the United States, this observation from more than 75 years ago is still largely true today. A recent report from an ad hoc Census Bureau Task Force on the Undercount of Young Children concluded, “The undercount of children under age five in the decennial census, and in surveys like the American Community Survey (ACS), is real and growing.”[4]

The high net undercount among young children was discovered early in the history of demographic analysis. Coale found children ages 0 to 4 had a high net undercount rate in the Decennial Censuses of 1940 and 1950.[5] Additional research by Siegel and Zelnik also found a significant net undercount of children ages 0 to 4 in the 1950 and 1960 Decennial Censuses.[6] Coale and Zelnick discovered high net undercount rates for young children in the decennial censuses as far back as 1880.[7] Coale and Rives found very high undercount rates for young black children in every decennial census from 1880 to 1970.[8] Genealogical research also shows a pattern of underreporting young children as far back as the 1850s.[9]

Research on subnational assessments of decennial census results is limited. However, following the 1970 Census, Siegel et al. examined census coverage for states and for various population groups by race and age.[10] They used several different approaches with varied results and did not focus on young children.

After the 1990 Decennial Census, Robinson et al. offered a set of undercount estimates for states for the total population (all ages), but the estimates are only evaluated at the regional level.[11] The authors also proposed alternatives for evaluating the 2000 Census at the state and sub-state levels and listed several reasons why such an evaluation is needed.

After the 2000 Census, Adlakha et al. used Census Bureau population estimates to assess decennial census counts for the population ages 0 to 9.[12] However, their results focused on regional-level differences in net undercount rates and did not show data separately for the population ages 0 to 4.

Cohn compared Census Bureau state population estimates to the 2010 Census counts for the total population (that is, all age groups) but did not analyze data for young children separately.[13] Cohn concludes that the decennial census counts and the population estimates are very similar for most states in terms of total population.

Mayol-Garcia and Robinson compared state population estimates for 0 to 4 and 0 to 9 age groups with corresponding counts from the 2010 Decennial Census, but only provided limited results.[14] Regarding the state-level data on net undercounts of the population ages 0 to 4,

Mayol-Garcia and Robinson report, “The relatively large differences noted nationally for 0-4 year olds are observed at the state level as well.”[15]

Data from the Census Bureau’s Demographic Analysis (DA) (a method for assessing census accuracy described in the next section) show an overall net undercount of 970,000 children ages 0 to 4 in the 2010 Decennial Census. The net undercount rate for the population ages 0 to 4 was 4.6 percent, which is more than twice as high as for any other age group.[16] Census Bureau research also shows that the omissions rate for young children in the 2010 Census was 10.3 percent, which amounts to 2.2. million young children being missed.[17]

The net undercount rate for the population ages 0 to 4 is not only higher than that of any other age group, but has been increasing over the past several decades. O’Hare shows that the net undercount rate for young children rose from 1.4 percent in 1980 to 4.6 percent in 2010.[18] During the same period, the coverage rate for adults (ages 18 and older) changed from a net undercount of 1.4 percent to a net overcount of 0.7 percent. These diverging trends underscore the importance of examining undercounts of the population ages 0 to 4 in more detail.[19]

O’Hare examined 2010 Census coverage rates at the state level for the population ages 0 to 4 and found substantial variation in coverage rates across the states.[20] The coverage rates used in O’Hare’s study were based on a comparison of the 2010 Census to the Census Bureau’s Vintage 2010 Population Estimates. He also found several demographic characteristics of states that were statistically significantly correlated with differences in coverage rates, including percent minority, percent linguistically isolated, and several socioeconomic and housing measures.

O’Hare also examined 2010 Census coverage rates for young children for counties based on comparisons of the Vintage 2010 Population Estimates to the 2010 Census count.[21] He found several county characteristics that were statistically significantly correlated with differences in coverage rates. (The net undercount estimates based on Vintage 2010 Population Estimates have been superseded by the Census Bureau’s DA research estimates released in April 2018.)

3. Methodology and Data Sources

The present analysis extends the work of previous researchers by examining 2010 Census county-level coverage rates for the population ages 0 to 4 based on updated U.S. Census Bureau estimates. Previous studies have largely been descriptive, while this study is more analytical in nature. Using the Census Bureau's new experimental DA net undercount estimates for young children—in combination with a large number of potential explanatory variables—will help us determine which factors are most closely related to net undercounts of young children.b

This analysis responds to a stream of research calling for more subnational research on census undercount. Based on their analysis of 2000 Decennial Census data compared to subnational population estimates, Adlakha et al. recommend researchers "…expand the current demographic analysis to include subnational benchmarks in the 2010 Census evaluation."22 Mayol-Garcia and Robinson conclude: "More studies are needed on the patterns of this population age group compared to the results of the previous censuses."23 The final report of the Census Bureau's Task Force on the Undercount of Young Children states: "This work must look below the national level to determine if certain areas, populations, or census operations were more likely to have these errors."24

The Census Bureau's experimental DA net undercount estimates for young children provide an opportunity to assess subnational census results using a DA-like methodology for the population ages 0 to 4. Like the original national DA estimates (described below), these estimates are based on a simple cohort-component demographic accounting equation that uses number of births, deaths, and net migration.

Demographic Analysis

The DA methodology used to assess census accuracy at the national level compares the decennial census results to independent population estimates to ascertain undercounts and overcounts for the total population and for selected age, sex, and racial/Hispanic groups. One of the major limitations of the DA technique for measuring the census undercounts for most demographic groups is that it can only be applied at the national level because subnational population estimates for those ages 10 and older include census errors (people who may be missed or counted twice) from the previous census. However, population estimates for those under age 10 are not based on the previous census. Consequently, young children are one of the few demographic groups for which this method can be used. The methodology employed here to develop net undercount rates compares Vintage 2010 population estimates to the 2010 Census counts.

The DA methodological approach is well-suited for analyzing census coverage of young children. With respect to the results of the 2000 Census evaluation for the count of young children, the Census Bureau states:

---

[b] See King et al. (2018) for more information on the methodology used to produce the Census Bureau's experimental DA net undercount estimates.

"The Demographic Analysis estimate for this age group is more accurate than those for other age groups because the estimate for young children depends primarily on recent birth registration data which are believed to be highly accurate."[25]

In comparing the DA results to Dual Systems Estimates results in the 2000 Decennial Census, Zeller also concluded:

"Since the Demographic Analysis estimate for young children depended on highly accurate recent birth registration data, the Demographic Analysis estimate is believed to be more accurate."[26]

The favorable view of the DA methodology is related to the simplicity of the method and the quality of the key data, that is, data on births and deaths. Nearly all (99.6 percent) of the estimated population from the national DA estimates for those ages 0 to 4 in 2010 is derived from birth data.[27] Heavy dependence on birth certificate data and the high quality of those data provide a strong foundation for county population estimates for the population ages 0 to 4. The birth and death data used in the Census Bureau's DA estimates come from the U.S. National Center on Health Statistics (NCHS), and these records are widely viewed as being accurate and complete.[28]

The Census Bureau's 2018 experimental DA net undercount estimates for young children, which we use in our analysis, have the same favorable qualities as 2010 Census DA methodology. The county population estimates are derived using the formula in Equation 1, which is the cohort-component technique used by the U.S. Census Bureau to estimate the household population across time based on its exposure to mortality, fertility, and migration.

$$P1 = P0 + B - D + NDM + NIM \quad (1)$$

Where:

P1 = Population at the end of the year.
P0 = Population at the beginning of the year.
B = Births during the year.
D = Deaths during the year.
NDM = Net domestic migration during the year.
NIM = Net international migration during the year.

Births and deaths for these estimates are taken from the Federal-State Cooperative for Population Estimates (FSCPE). The authors felt these birth and death figures were slightly better than those from the National Center for Health Statistics. The difference between the two series are quite small.[29]

Migration between counties is captured in the Census Bureau administrative records technique, which uses federal tax records to estimate such migration.[30]

Net international migration estimates have been updated by King et al. based on the results of the 2010 Mexican Survey of Occupation and Employment (ENOE), which asked about migration.[31] These updates resulted in substantial changes in the net undercount of young children in many counties, particularly counties with large numbers of Hispanics. The number of

counties with net undercount rates of more than 10 percent decreased from 279 using the Vintage 2010 estimates to 203 with the 2018 DA estimates.[32]

County-level census coverage rates for young children are derived by comparing the Census Bureau's experimental DA net undercount estimates for the population ages 0 to 4 to the 2010 Decennial Census counts for this age group. This methodology for examining census coverage at the state and local level has been used by several analysts in the past, including several demographers at the Census Bureau.[33]

Data Sources

Our analysis is based on data for the 261 most populous counties in the country (those with populations of 250,000 or more). We focused on large counties because previous research indicates that population estimates are usually more accurate for larger counties than for smaller ones.[34] Obtaining accurate estimates is more difficult for relatively small population subgroups (such as the population ages 0 to 4). Consequently, it is reasonable to expect that the differences between the 2010 Decennial Census counts and the Census Bureau's experimental DA population estimates for many small counties would be fraught with random error. In county-level analysis, O'Hare found all the correlations between demographic variables and net undercounts for young children were more pronounced for larger counties than all counties, which may reflect the impact of errors in smaller counties.[35]

There is another reason to focus on large counties: They make up a disproportionate share of children who are missed in the census. The revised 2018 DA estimates indicate there were 21,015,226 young children in the United States on April 1, 2010, compared to a decennial census count of 20,201,362.[36] This means there was a revised net undercount of 813,864 and a revised net undercount rate of 3.9 percent, compared to a net undercount of 970,000 and a net undercount rate of 4.6 percent in the analysis based on the original DA estimates. The net undercount for the 261 counties is 590,860 and these counties account for 73 percent of total net undercount of young children.

Table 1 shows that even among large counties, there are important distinctions by county size. The counties with populations of one million or more have a collective net undercount rate of 5.0 percent, compared to 4.2 percent for counties with populations between 500,000 and 999,999, and 3.7 percent for counties with populations between 250,000 and 499,999. More than a third (36 percent) of the net undercount for young children in the 2010 Census is accounted for by the largest 39 counties. Half of the net undercount of young children is accounted for by the 80 largest counties in the country.

Table 1. Net Undercount of Young Children in Large Counties, by County Population Size

County Size	Number of children ages 0-4 (Census 2010)	Net undercount number	Net undercount rate
1 million or more	5,477,261	290,078	5.0
500,000 to 999,999	4,252,129	184,568	4.2
250,000 to 499,999	3,027,675	116,214	3.7
All large counties	12,757,065	590,860	4.4

Source: PRB analysis of the U.S. Census Bureau's experimental net undercount rates for the population ages 0 to 4.

The dependent variable in our analysis is the Census Bureau's experimental DA net undercount rate for the population ages 0 to 4 in the 261 largest U.S. counties. These estimates were first presented at the 2018 Population Association of America Conference.[37] The net undercount rates for children ages 0 to 4 for the largest 261 counties were kindly provided to us by the Census Bureau.

Net undercount is probably the most widely used metric for measuring census accuracy. It is the measure used by both the DA method and the Dual-Systems Estimates method of the Census Bureau.

Undercounts have sometimes been reported as negative numbers and sometimes as positive numbers.[38] In this report, net undercounts are consistently reported as positive numbers and net overcounts as negative numbers. Measuring net undercounts here as positive numbers makes the correlations and multivariate results easier to interpret and explain.

Potential Explanatory Variables

This section discusses the rationale for selecting potential independent variables for our models. We considered explanatory variables related to the general accuracy of decennial census counts as well as those specifically focused on young children. We identified 32 potential explanatory variables for our analysis. We selected these variables based on the indicators used in the Census Bureau's Hard-to-Count (HTC) scores and Low Response Scores (LRS), research reports on the undercount of young children by the Census Bureau, and our review of the research literature in this area.[39] Most of the social, economic, housing, and demographic measures for this analysis were derived from the 2008-2012 American Community Survey 5-year estimates, while data on mail return rates are from the decennial census.

Variables were grouped into six domains. The variables and the domains are shown in Table 2, along with the dependent variable and the means and standard deviations for all variables.

Table 2. Means and Standard Deviations for All Potential Explanatory Variables		
	Mean	Standard deviation
Demographic Analysis (DA) net undercount percentage, 2010 Census (ages 0-4)	**3.9**	**2.6**
Race and Hispanic Origin		
Percent non-Hispanic black alone (ages 0-4)	15.3	14.3
Percent Hispanic or Latino (ages 0-4)	24.5	18.4
Percent minority (ages 0-4)	50.9	18.6
Percent non-Hispanic black alone (ages 5-17)	16.3	15.6
Percent Hispanic or Latino (ages 5-17)	21.6	18.5
Percent minority (ages 5-17)	47.5	20.1
Percent black alone (families with related children)	16.4	15.1
Percent Hispanic or Latino (families with related children)	18.6	17.2
Percent minority (families with related children)	42.3	20.0
Socioeconomic Status		
Percent of families with children ages 0-4 in poverty	20.4	8.1
Percent of persons ages 0-4 in poverty	22.2	8.9
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	13.0	5.2
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	4.9	4.4
Percent of population ages 18+ with limited English proficiency (do not speak English "very well")	9.7	7.8
Family Structure and Living Arrangements		
Percent of children ages 0-5 living with a grandparent householder	9.5	3.2
Percent of children who are not biological, adopted, or step-children of householder	10.3	3.2
Percent of all households that have 7+ people	1.5	1.0
Percent of children ages 0-4 who live with nonrelatives or in group quarters	1.8	0.6
Other Demographic Measures		
Percent of children ages 1-4 who have moved in the past year	20.6	4.8
Percent of adults ages 18+ who are foreign-born	16.4	11.0
Percent of adults ages 18+ who are not U.S. citizens	8.8	6.0
Percent of all householders who are ages 15-34	20.5	5.1
Percent of children ages 0-4 living in Hard-to-Count (HTC) census tracts	21.7	16.2
Housing		
Percent of persons living in renter-occupied households	32.9	9.6
Percent of housing units that are vacant	8.1	3.2
Percent of household population living in multi-unit buildings	21.4	12.4
Percent of housing units that are crowded (1.01 or more occupants per room)	3.1	2.6
Response/Return Rates		
Final initial questionnaire mail return rate, 2010 Census	76.0	4.9
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	79.3	3.5
Final undeliverable as addressed rate, 2010 Census	10.0	4.2
Final Census 2000 short form mail return rate	80.9	4.3
Final Census 2000 mail return rate (short form and long form)	79.4	4.4

Race and Hispanic Origin Status

Certain racial/ethnic groups face higher risks of being missed in the decennial census. The DA results for 2010 show that among children ages 0 to 4, the net undercount rate was 7.5 percent for Hispanics and 6.3 percent for children classified as black alone or in combination with one or more other races.[40] A recent update showed that the net undercount of Hispanic children ages 0 to 4 in the 2010 Census was 6.5 percent, rather than the 7.5 percent shown in the 2010 DA results.[41] Consequently, one would expect counties with relatively large numbers of young children who are Hispanic and/or black alone or in combination to have higher net undercount rates for the population ages 0 to 4.

Several measures of race and Hispanic Origin status are statistically significant predictors of the Census Bureau's LRS, controlling for other factors.[42] The Census Bureau's HTC score also includes measures of race and Hispanic Origin. O'Hare found racial composition to be related to net undercounts for young children in states and counties.[43]

A recent report by the U.S. Government Accountability Office also lists racial and ethnic minorities as a hard-to-count group.[44]

Socioeconomic Status

It is widely believed that socioeconomic status, and poverty in particular, is associated with census coverage. For example, in response to the release of 2010 Census results, former Undersecretary of Commerce Rebecca Blank said,

"However, as has been the case for some time, today's release shows that certain populations were undercounted. More work remains to address persistent causes of undercounting, such as poverty, mobility, language isolation, low levels of education, and general awareness of the survey."[45]

Fernandez et al. found the ratio of household income to the poverty threshold was related to the likelihood of young children being missed in the 2010 Census.[46] Fernandez et al., conclude that "children who are not found in the Census were more disadvantaged than those who are in the census."[47]

Research by Robinson et al. and the U.S. Government Accountability Office also identified socioeconomic status as a barrier to being counted in the census.[48] The percent of the population below poverty remained a statistically significant predictor of the Census Bureau's LRS, after controlling for other factors.

Family Structure and Living Arrangements

Family structure and children's living arrangements may also increase the risk of being missed in the census. A recent report by the U.S. Government Accountability Office lists "Complex households including those with blended families, multi-generational, or non-relatives" as a hard-to-count group.[49]

Martin argued that residential ambiguity is a key factor in people being missed in surveys and the decennial census.[50] People who are not clearly attached to one specific household—and

particularly those who are not closely related to the person completing the questionnaire—are more likely to be missed. Young children living in complex families meet these criteria.

The Census Bureau also found that the living arrangements of young children are related to the likelihood of being missed in the decennial census.[51] Young children are more likely to be missed if they are living in households where they are not closely related to the householder, living in a single-parent household rather than a married-couple household, and living in a complex household. A child is also more likely to be missed if he/she "lives in a household that is large, multigenerational, or includes extended or multiple families."[52]

Other Demographic Measures

Several other demographic measures are also related to census coverage. Robinson et al. and the Census Bureau cited residential mobility as a factor that can increase the risk of being missed in the census.[53] Other demographic measures linked to census coverage include the percent of the population that is foreign-born and the percent who are not U.S. citizens. Jensen et al. reported recent immigrants are under-reported in the ACS.[54] Age of the householder is another variable included in this domain; younger adult householders had a lower self-response rate in the 2010 Census than householders in other age groups.[55]

The U.S. Census Bureau also reported that children in non-English or limited-English-speaking households are less likely to be counted.[56]

Housing

The 12-factor HTC score developed by the Census Bureau in the 1990s and used in Census 2000 and the 2010 Census was based on six housing characteristics and six population characteristics.[57] Robinson et al. listed irregular housing as a barrier to accurate enumerations in the census because irregular housing may not be included in the Master Address File.[58] Robinson et al. also listed renters as more difficult to enumerate than homeowners, and the U.S. General Accountability Office listed renters as a hard-to-count population.[59]

The 2010 Census Coverage Measurement data show the population living in rental housing units had a net undercount while the population on living in owner-occupied housing units had a net overcount.[60]

Self-Response Rates

Data collected through the Census Bureau's self-response operation are more accurate than the data collected during the Nonresponse Followup (NRFU) phase.[61] The Census Bureau found that the nonmatch rate (similar to the omissions rate) for young children was much lower for respondents during the self-response operation than in the NRFU operation.[62]

The Census Bureau's LRS is designed to predict self-response in the decennial census.[63] Implicit in the model is the idea that people living in geographic areas with low levels of self-response will be harder to count.

4. Results

Table 3 provides summary statistics for the 261 largest counties, based on data from the Census Bureau's experimental DA net undercount estimates for young children. The net undercount of young children is widespread in these counties, with 250 of the 261 counties exhibiting a net undercount for young children and a mean county-level net undercount rate of 3.9 percent. The county-level rates range from a 14.3 percent net undercount in Webb County, Texas to a net overcount of 8.4 percent in Pinal County, Arizona.

Table 3. Summary Statistics for 2010 Census Net Undercount Rates for Children Ages 0 to 4 in 261 Large Counties

Number of counties	261
Number of counties with a net undercount	250
Percent of counties with a net undercount	96
Mean undercount	3.9
Standard deviation	2.6
Maximum net undercount	14.3%
Maximum net overcount	-8.4%

Note: In this report, net undercounts are reported as positive numbers and net overcounts as negative numbers.

Correlational Analysis

We started with a correlational analysis to identify the characteristics of counties that are most closely associated with the undercount of children. We examined factors across six domains, ranging from indicators typically used to identify hard-to-count areas for the entire population (for example, mail return rates and living in hard-to-count census tracts), along with race and ethnicity and other demographic characteristics, socioeconomic status, family structure/living arrangements, and housing-related factors. All the indicators examined in this section are continuous variables and therefore amenable to correlational analysis.

As stated earlier, we began the analysis with 32 potential explanatory variables. The zero-order correlation between these variables and the net undercount of young children in the 261 large counties are shown in Table 4. These correlations range from +0.44 for percent of families with related children in households where the householder is a racial or Hispanic minority to -0.31 for final initial questionnaire mail return rate in the 2010 census. Most correlations are statistically significantly different than zero. (Correlation coefficients over.10 are statistically significantly different than zero at a .10 level for a two-tailed test.)

Recall that the net undercount rate is measured as a positive number in this report. Most correlations in Table 4 are in the predicted direction—that is, we expected counties with higher rates on the explanatory variables to also have a higher net undercount. Also, the negative

correlations between the mail return rate indicators and the net undercount rate make sense because a higher mail return rate should mean that fewer people are undercounted.

Table 4. Correlations of Potential Explanatory Variables with Net Undercount of Young Children		
	Correlation	P-value
Race and Hispanic Origin		
Percent non-Hispanic black alone (ages 0-4)	0.29	<.0001
Percent Hispanic or Latino (ages 0-4)	0.21	0.0
Percent minority (ages 0-4)	0.41	<.0001
Percent non-Hispanic black alone (ages 5-17)	0.29	<.0001
Percent Hispanic or Latino (ages 5-17)	0.21	0.0
Percent minority (ages 5-17)	0.40	<.0001
Percent black alone (families with related children)	0.30	<.0001
Percent Hispanic or Latino (families with related children)	0.26	<.0001
Percent minority (families with related children)	0.44	<.0001
Socioeconomic Status		
Percent of families with children ages 0-4 in poverty	0.16	0.0
Percent of persons ages 0-4 in poverty	0.18	0.0
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	0.30	<.0001
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	0.38	<.0001
Percent of population ages 18+ with limited English proficiency (do not speak English "very well")	0.35	<.0001
Family Structure and Living Arrangements		
Percent of children ages 0-5 living with a grandparent householder	0.41	<.0001
Percent of children who are not biological, adopted, or step-children of householder	0.36	<.0001
Percent of all households that have 7+ people	0.08	0.2
Percent of children ages 0-4 who live with nonrelatives or in group quarters	0.12	0.0
Demographics		
Percent of children ages 1-4 who have moved in the past year	-0.13	0.0
Percent of adults ages 18+ who are foreign-born	0.30	<.0001
Percent of adults ages 18+ who are not U.S. citizens	0.30	<.0001
Percent of all householders who are ages 15-34	-0.11	0.1
Percent of children ages 0-4 living in Hard-to-Count (HTC) census tracts	0.28	<.0001
Housing		
Percent of persons living in renter-occupied households	0.11	0.1
Percent of housing units that are vacant	0.18	0.0
Percent of household population living in multi-unit buildings	0.15	0.0
Percent of housing units that are crowded (1.01 or more occupants per room)	0.23	0.0
Response/Return Rates		
Final initial questionnaire mail return rate, 2010 census	-0.31	<.0001
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 census	-0.26	<.0001
Final undeliverable as addressed rate, 2010 census	0.09	0.1
Final Census 2000 short form mail return rate	-0.30	<.0001
Final Census 2000 mail return rate (short form and long form)	-0.30	<.0001

However, two of the zero-order correlations are not in the predicted direction. There is a negative correlation (-0.13) between percent of children ages 1 to 4 who moved in the past year and the net undercount rate for young children. In this set of counties, the higher the mobility of children ages 1 to 4, the lower the net undercount rate. There is also a negative correlation (-0.11) between percent of householders ages 15 to 34 and the net undercount rate for young

children. It is not clear why these relationships are negative, but the fact that the correlation coefficients are barely statistically significant may be related to the unpredicted outcome.

The next step of our analysis was to reduce the list of 32 potential explanatory variables down to a smaller set that could be used in a multiple regression analysis to estimate each independent relationship with the net undercount of young children. We used three criteria to reduce the list of 32 variables. First, we identified the variables within each domain that had relatively high zero-order correlations with the dependent variable. Second, we examined those variables to determine which variables had relatively low zero-order correlations with remaining variables in the domain. Third, we selected those variables that had relatively low zero-order correlations with variables in other domains. This left us with a set of 14 independent variables that represented all six domains. Table 5 shows the interrelationships among these 14 independent variables.

Table 5. Intercorrelation Matrix

	Demographic Analysis (DA) net undercount percentage, 2010 Census,(ages 0-4)	Percent minority (families with related children)	Percent of persons ages 0-4 in poverty	Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	Percent of households that are linguistically isolated (no one ages 14+ speaks English very well)	Percent of children ages 0-5 living with a grandparent householder	Percent of all households that have 7+ people	Percent of children ages 0-4 who live with nonrelatives or in group quarters	Percent of children ages 1-4 who have moved in the past year	Percent of all householders who are ages 15-34	Percent of children ages 0-4 living in hard-to-count census tracts	Percent of persons living in renter-occupied households	Percent of housing units that are vacant (and not intended for seasonal, recreational, or occasional use)	Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	Final undeliverable as addressed rate, 2010 Census
Percent minority (families with related children)	0.44	1.00													
Percent of persons ages 0-4 in poverty	0.18	0.47	1.00												
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	0.30	0.72	0.68	1.00											
Percent of households that are linguistically isolated (no one ages 14+ speaks English very well)	0.38	0.70	0.24	0.70	1.00										
Percent of children ages 0-5 living with a grandparent householder	0.41	0.65	0.64	0.71	0.40	1.00									
Percent of all households that have 7+ people	0.08	0.48	0.20	0.62	0.59	0.39	1.00								
Percent of children ages 0-4 who live with nonrelatives or in group quarters	0.12	0.26	0.19	0.30	0.17	0.31	0.22	1.00							
Percent of children ages 1-4 who have moved in the past year	-0.13	0.12	0.45	0.15	-0.16	0.17	-0.04	0.15	1.00						
Percent of all householders who are ages 15-34	-0.11	0.25	0.31	0.07	0.03	0.03	0.10	-0.07	0.45	1.00					
Percent of children ages 0-4 living in hard-to-count census tracts	0.28	0.69	0.45	0.56	0.52	0.48	0.36	0.13	0.07	0.21	1.00				
Percent of persons living in renter-occupied households	0.11	0.65	0.49	0.49	0.47	0.34	0.20	0.44	0.25	0.52	0.56	1.00			
Percent of housing units that are vacant (and not intended for seasonal, recreational, or occasional use)	0.18	0.36	0.67	0.30	-0.09	0.47	-0.17	0.10	0.42	0.21	0.37	0.33	1.00		
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.26	-0.73	-0.42	-0.59	-0.54	-0.52	-0.40	-0.17	-0.10	-0.31	-0.89	-0.64	-0.34	1.00	
Final undeliverable as addressed rate, 2010 Census	0.09	0.08	0.53	0.20	-0.16	0.31	-0.14	0.18	0.44	0.01	0.06	0.08	0.67	-0.03	1.00

Regression Analysis

The 14 independent variables that were selected from the group of 32 were used in multiple regression analyses to estimate their independent relationships with net undercount rates for the population ages 0 to 4.

Models 1 and 2

We first evaluated how well two key factors currently used to predict geographic differences in net undercount of the total population explain geographic variation in the net undercount of young children. The first panel in Table 6—Model 1—shows the bivariate relationship between the final mail return rate for the 2010 Census, which is the basis for the Census Bureau's Low-Response-Score Model, and the net undercount of young children.[64] This model shows a weak, but statistically significant relationship, with the final mail return rate explaining only 6 percent of the variation in the net undercount rate of children (adjusted r-square=.062).

Table 6. Multivariate Regressions Predicting the County-Level Percent Undercount of Children Ages 0-4 (n=261)

	Model 1				Model 2				Model 3			
	Parameter Estimate	Standard Error	P-value	Standardized Parameter Estimate	Parameter Estimate	Standard Error	P-value	Standardized Parameter Estimate	Parameter Estimate	Standard Error	P-value	Standardized Parameter Estimate
Intercept	19.22	3.60	<.0001	0.00	2.93	0.26	<.0001	0.00	3.18	0.88	0.00	0.00
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.19	0.05	<.0001	-0.26								
Percent of children ages 0-4 living in hard-to-count census tracts					0.05	0.01	<.0001	0.28	0.01	0.01	0.45	0.06
Final undeliverable as addressed rate, 2010 Census									0.04	0.05	0.41	0.06
Percent minority (families with related children)									0.06	0.01	<.0001	0.48
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative									-0.12	0.07	0.07	-0.24
Percent of persons ages 0-4 in poverty									0.03	0.03	0.29	0.12
Percent of households that are linguistically isolated (no one ages 14+ speaks English very well)									0.28	0.06	<.0001	0.47
Percent of children ages 0-5 living with a grandparent householder									0.20	0.07	0.01	0.24
Percent of all households that have 7+ people									-0.98	0.22	<.0001	-0.37
Percent of children ages 0-4 who live with nonrelatives or in group quarters									1.03	0.33	0.00	0.22
Percent of all householders who are ages 15-34									0.06	0.04	0.13	0.12
Percent of children ages 1-4 who have moved in the past year									-0.10	0.04	0.01	-0.17
Percent of persons living in renter-occupied households									-0.14	0.03	<.0001	-0.52
Percent of housing units that are vacant (and not intended for seasonal, recreational, or occasional use)									0.00	0.09	0.99	0.00
Adjusted r-square	0.062				0.077				0.379			

The second panel in Table 6—Model 2—shows the results for the bivariate relationship between the percent of children living in hard-to-count (HTC) census tracts and the net undercount of young children. HTC tracts are defined here as those with mail return rates are 73 percent of less. Although the regression coefficient indicates that counties with higher shares of children living in hard-to-count census tracts are also likely to have higher rates of net undercount, our results indicate a weak relationship between these two variables. Although it is statistically significant, the HTC measure explains less than 8 percent of the variation in the net undercount rate of children (adjusted r-square=.077). These results indicate that census accuracy for young children is likely driven by different factors than census accuracy for other age groups or the population in general. Census Bureau research found that roughly 80 percent of young children who were missed in the 2010 Census resided in households that returned a census questionnaire (or at least were on the Master Address List).[65] The lack of a strong connection between mail return rates and the net undercount of young children may explain why the percent of young children (under age 6) was not a statistically significant variable in the census-tract-level model for the Low Response Score, even though it was in the block-group-level model.[66]

Model 3

The final step of our analysis was to identify which of the 12 explanatory variables—other than the 2010 mail return rate and the percent of young children in hard-to-count census tracts—are independently associated with the net undercount of children, and to determine if this set of variables could be used to improve geographic targeting efforts to reduce the undercount of children. Our final model—Model 3—includes all 12 explanatory variables as well as the percent of children in hard-to-count census tracts. The final model could not include both the 2010 mail return rates and the percent living in hard-to-count tracts because those two variables are highly correlated with each other (r= -.89). We kept percent of children living in hard-to-count tracts because it performed slightly better than the mail return rates in the bivariate models.

The multiple adjusted r-square, presented at the bottom of panel 3, shows that this expanded model explains 38 percent of the variation in the net undercount of children (adjusted r-square=.379). This represents a big improvement over the results using just the percent of children living in hard-to-count tracts or the mail return rate. Additionally, the percent of children living in hard-to-count tracts is no longer statistically significant from zero (β=.01, p=.45) once the additional 12 explanatory variables are added to the model. This again suggests that census accuracy for young children is driven by different factors than census accuracy for the population in general.

The percent of families with related children that are in racial or Hispanic minority families has a significant, positive association with net undercount rates for young children. The relatively high net undercounts and omissions rates for racial and Hispanic minorities seen in numerous other analyses would lead one to expect this result. The fact that these variables remain statistically significant even after controlling for other factors is also not surprising.

For example, O'Hare also found the percent of children ages 0 to 4 who are black or Hispanic was statistically significantly related to net undercounts, even after controlling for population size and population growth from 2000 to 2010.[67] Fernandez et al. found young children in every race/Hispanic group other than non-Hispanic white alone had higher odds of being missed in the 2010 Census based on comparing young children in Administrative Records to 2010 Census records.[68]

Some researchers think race and Hispanic origin status are only associated with census accuracy because they reflect a cluster of other attributes like socioeconomic status. For example, Schwede et al. state,

"Though there is no reason to believe that race or ethnicity in and of itself leads to coverage error, it seems that some underlying variables associated in past studies with undercounting may also be correlated with race (e.g., mobility, complex living situations, and language isolation)."[69]

In other words, race is widely seen as a proxy for a combination of factors related to the likelihood of being missed in the census.

But our analysis shows that race is statistically significant even after many socioeconomic factors have been controlled. The LRS model developed by Erdman and Bates also found several measures of race were statistically significant after measures of socioeconomic status and other variables were controlled.[70] Fernandez et al. also found measures of race and ethnicity were statistically significant even after controlling for many other factors.[71]

With respect to the socioeconomic domain, only one of the three variables continue to be statistically related to the undercount of children once all explanatory variables are included in the model. Counties with larger shares of households that are linguistically isolated tend to have higher rates of net undercount of children. This relationship probably reflects recent immigration; for example, Jensen et al. shows that recent immigrants have lower coverage rates in the American Community Survey.[72] Furthermore, a large share of children at risk of being missed in the census live in households with noncitizen immigrants. PRB shows that 20 percent of the population ages 0 to 4 live in households with at least one noncitizen, which is a higher share than in any other age group.[73]

Low education levels of adults (percent of adults ages 18 or older with less than a high school diploma, GED, or alternative) is not statistically significant once other factors have been controlled. Likewise, the poverty rate for children ages 0 to 4 is not statistically significant. This is a surprising finding given the results of Fernandez et al., which show young children in low-income households have a higher risk of being missed in the census.[74] Also, education and income are key dimensions of socioeconomic status and are widely thought to be related to census accuracy. Our analysis shows that for young children, other factors are more powerful explanatory variables than income and education.

Household structure and living arrangements are important for understanding the net undercount of children. The percent of children ages 0 to 5 living with a grandparent householder is statistically significant: Counties with larger shares of young children living in grandparent-headed households tend to have higher rates of undercounted children. Other research has also identified young children living with grandparents as being vulnerable to being missed in the census.[75]

The percent of children living with nonrelatives or in group quarters is also statistically significant. Note that the share of young children living with nonrelatives or in group quarters is very low in most large counties. The mean is 1.8 percent among the 261 largest counties. To the extent this measure captures some dimension of complex living arrangements, our results fit previous research findings and theory. Young children living in complex families and households

have been seen as more vulnerable to being missed in the census. Fernandez et al. found children other than the son or daughter—including grandchildren, other relatives, foster child, and other non-relatives—had higher odds of being missed in the 2010 Census.[76] Also, Census Bureau analysts found higher miss rates (technically called nonmatch rates) for stepchildren, grandchildren, other relative children, and foster children or other unrelated child when they compared children found in the post-enumeration survey to those in the census records.[77]

The percent of households with 7 or more people is also statistically related to the undercount rate for children. This variable has a non-significant zero-order correlation with net undercount of +0.08, but in the multivariate context it has a coefficient of -0.98. This could indicate a problem with multicollinearity—counties with a higher percentage of large households also have larger shares of minority families, linguistically isolated households, and less-educated adults. But it could also be that there is a confounding relationship. Because the percent of large households is very low across the large counties (with a mean of 1.5 percent), and these counties overlap with the counties with higher rates among factors associated with net undercount (such as, minority family, linguistically isolated households, and low levels of education), the zero-order correlation appears to be zero (or even positive). But, after controlling for the share of minority families, linguistically isolated households, and less educated adults, counties with a higher percentage of large households may have a lower rate of net undercount.

In the other demographic domain, the percent of householders who are ages 15 to 34 is non-significant. This is surprising because young householders have lower self-response rates.[78] The final mail return rate for householders ages 18 to 24 was 55.4 percent in the 2010 Census compared to 83.1 percent for householders ages 45 to 64.[79] It could be that the net undercount of young children is driven by young families living in households with older adult householders, such as the grandparent of the young children. Census Bureau research shows that a large share of young mothers were missed in the 2010 Census along with their young children.[80] Alternatively, it may be that counties with a large percentage of young householders are also counties with a large share of renters, and may have relatively few young children.

The percent of children ages 1 to 4 who have moved in the past year is significant and negatively associated with the net undercount of children. Although this is an unexpected finding as residential mobility is considered a factor that can increase the risk of being missed in the census, it is consistent with the zero-order correlations.[81]

Finally, the housing variables show that the percent of persons living in renter-occupied households is negatively associated with the net undercount of children, and the percent of vacant housing in a county is unrelated to the undercount of children. The finding for renter-occupied households is surprising. The zero-order correlation was positive (r=.11), but as stated earlier, it could be counties with higher rates of renter-occupied households are also counties with higher rates of younger householders and may be counties with relatively few children.

Although the share of vacant housing units was positively associated with the net undercount of children in zero-order correlations, in the multivariate context it is non-significant. This indicates that other county characteristics are correlated with the share of vacant housing and do a better job of explaining the net undercount of children.

5. Limitations

It should be noted that there is a degree of uncertainty in both the county-level net undercount rates for young children and in the estimates for potential explanatory variables. While the Census Bureau does not provide margins of error for the experimental DA net undercount rates, we know these estimates are subject to error based on the uncertainty in the official DA estimates that were released in December 2010.

In addition, most of the independent variables used in this study come from the Census Bureau's American Community Survey. Since these estimates are derived from a sample, there is sampling error associated with these estimates. These estimates are also subject to non-sampling error.

While the estimates presented here involve some estimation error, these figures are the best data available to understand the geographic distribution of young children missed in the 2010 Census.

6. Discussion and Implications

As we approach the 2020 Census, there is a need to develop better information about the geographic areas to target for a complete count of young children. This paper reviews the data currently available and describes an alternative to existing measures of census accuracy. Our analysis shows that young children are different from most other groups in terms of census coverage, so general methods designed to locate hard-to-count areas may not work as well for young children.

By and large, the factors most closely associated with the net undercount of young children are not the same factors that are used to predict census participation for the total population. Table 7 shows the seven variables that were statistically significant in our model along with the seven most powerful variables from the Census Bureau's LRS Model.

Table 7. Comparison of the Seven Statistically Significant Variables in Our Model to the Seven Most Powerful Estimators in the Erdman-Bates Model for the Low Response Scores

Our Undercount Model (County level)	**Erdman-Bates Low Response Model (Census-Tract Level)**
Percent of children ages 0-4 who live with nonrelatives or in group quarters	Number of persons per household
Percent of all households that have 7+ people	Percent of population ages 65+
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	Percent of housing units that are renter occupied
Percent of children ages 0-5 living with a grandparent householder	Percent of housing units that are vacant
Percent of persons living in renter-occupied households	Percent of population that is non-Hispanic white
Percent of children ages 1-4 who have moved in the past year	Median home value
Percent minority (families with related children)	Percent of housing units that are single-unit structures

Note: Variables are shown in descending order by the strength of their associations with the dependent measures (undercount and low response, respectively).

Source: Erdman and Bates (2017), *Public Opinion Quarterly.*

There are only two common variables in the two lists (renters and racial/ethnic composition). But even here there is an important difference. In the LRS model, the higher the percent of renters in a census tract, the lower the self-response rate after other factors have been controlled. For young children, the higher the percent of renters in a county, the lower the net undercount rate of young children after other factors have been controlled.

This is not too surprising since the LRS focuses on mail return rates, and evidence indicates most young children are missed in households that probably returned a census questionnaire (or at least were on the Master Address file).

The factors that increase the propensity to self-respond in the census are probably not the same factors that determine whether young children are left off the returned census questionnaire.

Another key finding is the fact that two of the main metrics currently available to identify hard-to-count places are not very good at predicting net undercount rates for young children across large counties. The percent of young children living in hard-to-count census tracts (defined here

as those with a mail return rate of 73 percent or less in the 2010 Census) has a correlation of 0.28 with net undercount rates. The mail return rate (what the LRS tries to estimate) has a correlation of -0.26. In our multiple regression models, the percent of young children living in hard-to-count tracts explains 8 percent of the variance in net undercount rates for young children among the 261 counties, while the mail return rate explains 6 percent. In comparison, our final model explains 38 percent of the variance in net undercount rates for young children across the 261 counties.

The superior results of our model in terms of predicting net undercount rates for young children indicate this line of research should be pursued to better identify hard-to-count areas for young children in the 2020 Census.

7. Summary and Conclusions

The data examined here indicate that the national net undercount rate for the population ages 0 to 4 varies substantially across counties, but almost all large counties had a net undercount of young children in the 2010 Census. Moreover, the data show that larger counties account for the vast majority of the national net undercount for the population ages 0 to 4. In the 261 largest counties based on total population, there was a net undercount of 591,000 persons ages 0 to 4, which accounts for almost three-quarters of the nationwide net undercount for this age group. This information about where the net undercount rates for young children are the highest should help the Census Bureau and child advocates pinpoint the places that deserve special attention in the 2020 Census.

Two data sets currently being used to highlight hard to count areas for young children—percent of young children living in hard-to-count census tracts and the LRS (based on mail return rates)—are not very good predictors of net undercount rates for young children across the 261 largest counties.

Based on this preliminary analysis, the best predictor variables for net undercount of young children are:

- Percent of racial/ethnic minorities.
- Percent of households that are linguistically isolated.
- Percent of young children living with grandparent householders.
- Percent of young children living with nonrelatives or in group quarters.

We hope the results of this analysis will help advocates and others better target geographic areas and population subgroups for Get-Out-the-Count efforts to reduce the undercount of young children and help ensure an accurate 2020 Census.

References

[1] O'Hare, W.P. "Estimating the Net Undercount of Young Children in the 2010 U.S. Decennial Census at the County Level," poster presented at the conference of the Population Association of America, Boston, Mass., April 2014; William P. O'Hare, "State-Level 2010 Census Coverage Rates for Young Children," Population Research and Policy Review 33, no. 6 (2014): 797-816; and William P. O'Hare, "Geographic Variation in the 2010 U.S. Census Coverage Rates for Young Children a Look at Counties," International Journal of Social Science Studies 5, no. 9 (2017): 11-24.

[2] Allyn A. Young, "The Enumeration of Children," Publications of the American Statistical Association 7, no. 53 (1901): 21.

[3] U.S. Census Bureau, Population Differential Fertility: 1940 and 1910 Standardized Fertility Rates and Reproduction Rates (Washington, D.C.: Government Printing Office, 1944): 32-3, appendix A, "Completeness of Enumeration of Children Under 5 Years Old, U.S. Census of 1940 and 1910, 1940 Census."

[4] Deborah H. Griffin to Frank A. Vitrano, U.S. Census Bureau, Washington, D.C., memorandum, "Final Task Force Report: Task Force on the Undercount of Young Children," February 2, 2014.

[5] Coale, A.J., "The Population of the United States in 1950 Classified by Age, Sex and Color-a Revision of Census Figures," Journal of the American Statistical Association 50, (1955):16-54.

[6] Siegel, J. S. and Zelnik, M. "An Evaluation of Coverage in the 1960 U.S. Census of Population by Techniques of Demographic Analysis and by Composite Methods," Proceedings of the Social Statistics Section of the American Statistical Association (1966): 71-85.

[7] Coale, A.J. and Zelnick, M. New Estimates of Fertility and Population in the United States (Princeton, NJ: Princeton University Press, 1963).

[8] Coale, A.J. and Rives, N.W. "A Statistical Reconstruction of Black Population of the United States: 1880 to 1970: Estimates of True Numbers by Age and Sex, Birth Rates, and Total Fertility," Population Index 39, no. 1 (1973): 3-36.

[9] Adams, J.W. and Kasakoff, A.B. "Estimates of U.S. Decennial Census Underenumeration Based on Genealogies," Social Science History 15, no. 4 (1991): 527-43.

[10] Siegel, J. S., Passel, J S., Rives, N.W., and Robinson, J. G. "Developmental Estimates of the Coverage of the Population of States in the 1970 Census: Demographic Analysis," Current Population Reports, P23-65 (December 1977).

[11] Robinson, G.J., Bashir, A., Das Dupta, P., and Woodward, K.A. "Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis," Journal of the American Statistical Association 88, no. 423 (1993): 1061-71.

[12] Adlakha, A.L. et al., "Assessment of Consistency of Census Data with Demographic Benchmarks at the Subnational Level," Census 2000 Evaluation 0.20 U.S. Census Bureau (August 18, 2003).

[13] Cohn, D. "State Population Estimates and Census 2010 Counts: Did they match?" Pew Social and Demographic Trends (Jan. 12, 2011).

[14] Mayol-Garcia, Y. and Robinson, J.G. "Census 2010 Counts Compared to the 2010 Population Estimates by Demographic Characteristics," poster presented at the conference of the Southern Demographic Association, Tallahassee, FL, Oct. 2011.

[15] Mayol-Garcia “Census 2010 Counts Compared to the 2010 Population Estimates by Demographic Characteristics.”

[16] O’Hare, W.P. The Undercount of Young Children in the U.S. Decennial Census (Springer Publishers, 2015).

[17] U.S. Census Bureau, “Investigating the 2010 Undercount of Young Children – A New Look at 2010 Census Omissions by Age” (July 26, 2016).

[18] O’Hare, W.P. "Historical Examination of Net Coverage Error for Children in the U.S. Decennial Census: 1950 to 2010," Center for Survey Measurement Study Series, no. 2014-03 (2014), accessed at www.census.gov/srd/papers/pdf/ssm2014-03.pdf, on Feb. 11, 2019.-

[19] O’Hare, "Historical Examination of Net Coverage Error for Children in the U.S. Decennial Census: 1950 to 2010.”

[20] O’Hare, “State-Level 2010 Census Coverage Rates for Young Children.”

[21] O’Hare, W.P. “Geographic Variation in the 2010 U.S. Census Coverage Rates for Young Children a Look at Counties,” International Journal of Social Science Studies 5, no. 9 (Redframe Publishing, Sept. 2017): 11-24.

[22] Adlakha, A.L. et al., “Assessment of Consistency of Census Data with Demographic Benchmarks at the Subnational Level,” Census 2000 Evaluation 0.20 U.S. Census Bureau (August 18, 2003).

[23] Mayol-Garcia “Census 2010 Counts Compared to the 2010 Population Estimates by Demographic Characteristics.”

[24] U.S. Census Bureau, “Final Task Force Report: Task Force on the Undercount of Young Children,” Memorandum for Frank A. Vitrano (February 2, 2014): ii.

[25] U.S. Census Bureau, Technical Assessment of A.C.E. Revision II, (Washington, DC: Government Printing Office, 2003): v.

[26] Zeller, A., “Inconsistency Between Accuracy and Coverage Evaluation Revision II and Demographic Analysis Estimates for Children 0 to 9 Years of Age,” paper presented at the annual conference of the American Statistical Association, 2006: 320.

[27] U.S. Census Bureau, “National Demographic Analysis Tables: 2010” (December 2010), accessed at www.census.gov/coverage_measurement/demographic_analysis/.

[28] Devine, J. et al., “The Use of Vital Statistics in the 2010 Demographic Analysis Estimates,” (2010) Census Bureau Working Paper no. 88, accessed at www.census.gov/population/www/documentation/twps0088/twps0088.pdf.

[29] King, H., Ihrke, D., and Jensen, E., “Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census,” paper presented at the annual conference of the Population Association of America, Denver, CO, April 2018.

[30] U.S. Census Bureau, “Documentation for the Revised 2010 Demographic Analysis Middle Series Estimates” (2012).

[31] King, “Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census.”

[32] King, “Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census,” table 4.

[33] O’Hare, W.P. “State-Level 2010 Census Coverage Rates for Young Children,” Population Research and Policy Review 33, no. 6 (2014): 797-816; Siegel, J. S., Passel, J S., Rives, N.W., and Robinson, J. G. “Developmental Estimates of the Coverage of the Population of States in the 1970 Census: Demographic Analysis,” Current Population Reports, P23-65 (December 1977); Robinson, G.J., Bashir, A., Das Dupta, P., and Woodward, K.A. “Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis,” Journal of the American Statistical Association 88, no. 423 (1993): 1061-71; Adlakha, A.L. et al., “Assessment of Consistency of Census Data with Demographic Benchmarks at the Subnational Level,” Census 2000 Evaluation 0.20 U.S. Census Bureau (August 18, 2003); Mayol-Garcia, Y. and Robinson, J.G. “Census 2010 Counts Compared to the 2010 Population Estimates by Demographic Characteristics,” poster presented at the conference of the Southern Demographic Association, Tallahassee, FL, Oct. 2011; U.S. U.S. Census Bureau, “Final Task Force Report: Task Force on the Undercount of Young Children,” Memorandum for Frank A. Vitrano (February 2, 2014); and Cohn, D. “State Population Estimates and Census 2010 Counts: Did they match?” Pew Social and Demographic Trends (Jan. 12, 2011).

[34] Felton, G.T. “Evaluation of Population Estimation Procedures for Counties: 1980,” Current Population Reports P25-984 (1986); Davis, S.T. “Evaluation of Postcensal County Estimates for the 1980s,” Current Population Reports P25-963 (1994), O’Hare, W.P. "How to Evaluate Population Estimates," American Demographics 10, no. 1 (January 1988); and Yowell, T. and Devine, J., (2013). “Evaluating Current and Alternative Methods to Produce 2010 County Population Estimates” Population Division Working Paper 100, U.S. Census Bureau (July 2013).

[35] O’Hare, W.P. “Geographic Variation in the 2010 U.S. Census Coverage Rates for Young Children a Look at Counties,” International Journal of Social Science Studies 5, no. 9 (Redframe Publishing, Sept. 2017): 11-24.

[36] King, H., Ihrke, D., and Jensen, E., “Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census,” paper presented at the annual conference of the Population Association of America, Denver, CO, April 2018: 13.

[37] King, “Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census.”

[38] Velkoff, V. “Demographic Evaluation of the 2010 Census,” paper presented at the annual conference of the Population Association of America, Washington, DC, March 2011; King, “Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census.”; Jensen, E., Benetsky, M., and Knapp, A. “A Sensitivity Analysis of the Net Undercount for Young Hispanic Children in the 2010 Census,” poster presented at the annual conference of the Population Association of America, Denver, CO. April 2018; U.S. Census Bureau, “Summary of Estimates of Coverage for Persons in the United States” DSSD 2010 Census Coverage Measurement Memorandum Series #2010-G-01, (2012); and U.S. Census Bureau, “Components of Census Coverage for the Household Population in the United States” DSSD 2010 Census Coverage Measurement Memorandum Series #2010-G-04, (2012).

[39] Robinson, J. G., Johanson, C., and Bruce, A. “The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis,” paper presented at the Joint Statistical Meetings, Salt Lake City, UT, 2007; Erdman, C. and Bates, N. “The Low Response Score (LRS) A Metric to Locate, Predict, and Manage Hard-To-

Survey Population,” Public Opinion Quarterly 81, no. 1 (2017): 144-156; and U.S. Census Bureau, “Counting Young Children in the 2020 Census” (2018), accessed at www.census.gov/content/dam/Census/newsroom/press-kits/2018/counting-young-children-in-2020-census.pdf, on Feb. 13, 2019.

[40] O'Hare, W.P. The Undercount of Young Children in the U.S. Decennial Census (Springer Publishers, 2015).

[41] Jensen, E., Benetsky, M., and Knapp, A. "A Sensitivity Analysis of the Net Undercount for Young Hispanic Children in the 2010 Census."

[42] Erdman, C. and Bates, N. "The Low Response Score (LRS) A Metric to Locate, Predict, and Manage Hard-To-Survey Population," Public Opinion Quarterly 81, no. 1 (2017): 144-156.

[43] O'Hare, W.P. "State-Level 2010 Census Coverage Rates for Young Children," Population Research and Policy Review 33, no. 6 (2014): 797-816 and O'Hare, W.P. "Geographic Variation in the 2010 U.S. Census Coverage Rates for Young Children a Look at Counties," International Journal of Social Science Studies 5, no. 9 (Redframe Publishing, Sept. 2017): 11-24.

[44] U.S. General Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups" Report to Congressional Requesters GAO -18-599 (2018).

[45] Blank, R. "Statement by Deputy U.S. Commerce Secretary Rebecca Blank on Release of the Data Measuring Census Accuracy," (May 22, 2012): 1.

[46] Fernandez, L., Shattuck, R., and Noon, J. "Using Administrative Records and the American Community Survey to Study the Characteristics of Undercount Young Children in the 2010 Census," CARRA Working Paper Series #2018 - 05 (May 25, 2018): tables 6 and 7.

[47] Fernandez, "Using Administrative Records and the American Community Survey to Study the Characteristics of Undercount Young Children in the 2010 Census,": 13.

[48] Robinson, "The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis." and U.S. General Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups."

[49] U.S. General Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups."

[50] Martin, E. "Strength of Attachment: Survey Coverage of People with Tenuous Ties to Residences," Demography 44, no. 2 (2007): 427-440.

[51] U.S. Census Bureau, "Investigating the 2010 Undercount of Young Children –Analysis of Census Coverage Measurement Results" (July 2016).

[52] U.S. Census Bureau, "Counting Young Children in the 2020 Census" (2018), accessed at www.census.gov/content/dam/Census/newsroom/press-kits/2018/counting-young-children-in-2020-census.pdf, on Feb. 13, 2019.

[53] Robinson, J. G., Johanson, C., and Bruce, A. "The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis," paper presented at the Joint Statistical Meetings, Salt Lake City, UT, 2007 and U.S. Census Bureau, "Counting Young Children in the 2020 Census."

[54] Jensen, E., Benetsky, M., and Knapp, A. "A Sensitivity Analysis of the Net Undercount for Young Hispanic Children in the 2010 Census."

[55] Letourneau, E. "Mail Response/Return Rates Assessment," 2010 Census Planning Memoranda Series, no. 198 (May 30, 2012).

[56] U.S. Census Bureau, "Counting Young Children in the 2020 Census."

[57] Robinson, “The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis.”

[58] Robinson, “The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis.”

[59] Robinson, “The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis,” and U.S. General Accountability Office, “2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups.”

[60] U.S. Census Bureau, “2010 Census Coverage Measurement Estimation Report” DSSD 2010 Census Coverage Measurement Memorandum Series #2010-E-51, 2012.

[61] Brown, J.D. et al., “Understanding the Quality of Alternative Citizenship Data Sources for the 2020 Census,” U.S. Census Bureau White Paper (August 6, 2018).

[62] U.S. Census Bureau, Investigating the 2010 Undercount of Young Children –Analysis of Census Coverage Measurement Results” (July 2016).

[63] Erdman, “The Low Response Score (LRS) A Metric to Locate, Predict, and Manage Hard-To-Survey Population.”

[64] Erdman, “The Low Response Score (LRS) A Metric to Locate, Predict, and Manage Hard-To-Survey Population.”

[65] U.S. Census Bureau, Investigating the 2010 Undercount of Young Children –Analysis of Census Coverage Measurement Results” (July 2016); and Fernandez, L., Shattuck, R., and Noon, J. “Using Administrative Records and the American Community Survey to Study the Characteristics of Undercount Young Children in the 2010 Census,” CARRA Working Paper Series #2018 - 05 (May 25, 2018).

[66] Erdman, “The Low Response Score (LRS) A Metric to Locate, Predict, and Manage Hard-To-Survey Population,”: table 1.

[67] O’Hare, “State-Level 2010 Census Coverage Rates for Young Children.”

[68] Fernandez, “Using Administrative Records and the American Community Survey to Study the Characteristics of Undercount Young Children in the 2010 Census.”

[69] Schwede, L., Terry, R., & Hunter, J., “Ethnographic evaluations on coverage of HTC minority in the US Decennial Censuses,” in Hard-to-Survey Populations, Edited by Tourangeau, R., Edwards, B., Johnson, T.P., Wolter, K.M., and Bates, N. (Cambridge University Press, 2015): 293.

[70] Erdman, “The Low Response Score (LRS) A Metric to Locate, Predict, and Manage Hard-To-Survey Population.”

[71] Fernandez, “Using Administrative Records and the American Community Survey to Study the Characteristics of Undercount Young Children in the 2010 Census.”

[72] Jensen, E., Renuta, B., and Scopilliti, M., Demographic Analysis 2010: Estimates of Coverage of the Foreign-born Population in the American Community Survey. U.S. Census Bureau, Population Division, Working Paper No. 163 (June 2015).

[73] Jarosz, B. “Citizenship Question Risks a 2020 Census Undercount in Every State, Especially Among Children,” accessed at www.prb.org/citizenship-question-risks-a-2020-census-undercount-in-every-state-especially-among-children/, on Feb. 19, 2018.

[74] Fernandez, “Using Administrative Records and the American Community Survey to Study the Characteristics of Undercount Young Children in the 2010 Census.”

[75] U.S. Census Bureau, “Investigating the 2010 Undercount of Young Children –Analysis of Census Coverage Measurement Results” (July 2016).

[76] Fernandez, “Using Administrative Records and the American Community Survey to Study the Characteristics of Undercount Young Children in the 2010 Census,” table 7.

[77] U.S. Census Bureau, “Investigating the 2010 Undercount of Young Children –Analysis of Census Coverage Measurement Results” (July 2016), Table 10.

[78] Letourneau, E. “Mail Response/Return Rates Assessment.”

[79] Letourneau, E. “Mail Response/Return Rates Assessment.”

[80] U.S. Census Bureau, “2020 Census Research and Testing Investigating the 2010 Undercount of Young Children – Examining the Coverage of Young Mothers” (June 7, 2016).

[81] U.S. Census Bureau, “Counting Young Children in the 2020 Census” (2018), accessed at www.census.gov/content/dam/Census/newsroom/press-kits/2018/counting-young-children-in-2020-census.pdf, on Feb. 13, 2019.