Predicting Tract-Level Net Undercount Risk for Young Children

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

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

Table of Contents

Executive Summary

In the 2010 U.S. Census, children under age 5 were more likely to be missed than any other age group. Census Bureau research shows the net undercount rate for young children (the percent of children who were missed minus the percent who were erroneously included) was nearly 5%. 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, based on the Census Bureau's Revised 2018 Experimental Demographic Analysis Estimates.

Our results suggest that two metrics currently being used to identify areas where young children are more likely to be missed by the census—the Low Response Score and tracts designated as hard-to-count (both based on 2010 Census mail return rates)—are not very good predictors of net undercount rates for young children in large counties.

A higher net undercount of young children in the largest counties (those with 250,000 people or more in 2010) is most closely associated with the following variables:

- Percent of adults ages 18 to 34 with less than a high school diploma, GED, or alternative.
- Percent of children under age 18 living in a female-headed household with no spouse present.
- Percent of children under age 6 living with a grandparent householder.
- Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well").
- Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born).
- Percent of persons living in renter-occupied households.

We applied the model coefficients from our county-level regression analysis to updated independent variables from the 2013-2017 American Community Survey for census tracts to produce current neighborhood-level predicted net undercount risk for children under age 5.

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, children under age 5 had higher net undercount and omissions rates than any other age group.[a] The net undercount rate for young children was 4.6%, and more than 10% 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 689 large U.S. counties—those with at least 5,000 children under age 5 in 2010. 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 689 counties accounted for about 93% of the national net undercount of children under age 5 in the 2010 Census. Census coverage of young children varies widely across these counties, ranging from a 19.3% net undercount to a net overcount of 12.0%.

A set of 40 potential explanatory variables is considered in our analysis, based on past Census Bureau research, conversations with Census Bureau staff, and our review of the literature on census accuracy. The 40 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. Then the results of the analysis based on 2010 county-level data are applied to tract-level data to produce neighborhood-level estimates of net undercount rates for young children.

For the largest counties (those with 250,000 people or more and at least 5,000 children under age 5), the most important explanatory variables are various measures related to living arrangements/family structure, recent immigration, and socioeconomic status. Interestingly, some of the variables thought to be closely related to census accuracy for the total population

---

[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.

(like poverty and racial/ethnic composition) are not statistically significant after controlling for other variables.

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.

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 Zelnik 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 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 Census 2000, 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 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 under age 5 in the 2010 Census. The net undercount rate for children under age 5 was 4.6%, 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%, which amounts to more than 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% in 1980 to 4.6% in 2010.[18] During the same period, the coverage rate for adults (ages 18 and older) changed from a net undercount of 1.4% to a net overcount of 0.7%. These diverging trends underscore the importance of examining undercounts of the population under age 5 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 Revised 2018 Experimental Demographic Analysis estimates.)

Our analysis responds to a stream of research calling for more subnational research on census undercount. Based on their analysis of Census 2000 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]

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 under age 5 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 Revised 2018 Experimental Demographic Analysis (DA) Estimates for young children—in combination with many potential explanatory variables—helps us determine which factors are most closely related to net undercount rates for young children.

A set of 40 potential explanatory variables is considered in our analysis, based on past Census Bureau research, conversations with Census Bureau staff, and our review of the literature on census accuracy. The 40 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.

Our analysis indicated that running two regression models—one for the largest counties and one for smaller counties—produced better results than a single model for all counties. Therefore, we divided the 689 counties with at least 5,000 children under age 5 in our data set into two groups: those with populations of 250,000 or more and those with fewer than 250,000 people, based on data from the 2010 Census.

Tables for counties with a total population of 250,000 or more in the 2010 Census are provided in the body of this report. Corresponding tables for counties with a total population of less than 250,000 are provided in Appendix A.

The model coefficients based on our analysis of the largest counties are then applied to updated (2013-2017) independent variables for census tracts in these counties to produce tract-level predicted net undercount rates for young children. The model coefficients based on our analysis of smaller counties are similarly applied to updated (2013-2017) independent variables for census tracts in these counties to produce predicted tract-level net undercount rates.

The tracts are then classified into three risk categories based on their predicted net undercount rates for young children. The percentage distribution of these predicted rates is used to select the cutoff value that puts the top 25% of tracts in the highest-risk category. Very high-risk tracts are those with a predicted net undercount rate of 8.3% or higher. High-risk tracts are those with predicted net undercount rates of 0 to 8.29%. Those tracts with a predicted net undercount rate less than 0% are labeled low-risk tracts. A net undercount rate less than 0% indicates a predicted net overcount of young children in that tract.

Demographic Analysis

The Census Bureau's Revised 2018 Experimental DA Estimates for young children provide an opportunity to assess subnational census results using a DA-like methodology for the population under age 5. Like the original national DA estimates, these estimates are based on a simple cohort-component demographic accounting equation that uses data on births, deaths, and net migration.

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. With respect to the results of the 2000 Census evaluation for the count of young children, the Census Bureau states:

"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]

The Census Bureau's Revised 2018 Experimental DA Estimates of net undercount rates for children under age 5, which we use in our analysis, have the same favorable qualities as the 2010 Census DA methodology. A more detailed description of the methodology for these experimental estimates is provided in Appendix B.

Data Sources

Our analysis is based on data for 689 large counties (those with at least 5,000 children under age 5 in the 2010 Census). The dependent variable in our analysis is the net undercount rate for the population under age 5 in these 689 counties, based on the Census Bureau's Revised 2018 Experimental DA Estimates. After an extensive, internal review process, the Census Bureau produced these estimates in the summer of 2019, and they kindly shared the data with us in September 2019.

Our analysis focuses on large counties—in terms of population size—rather than small ones because previous research indicates that population estimates are usually more accurate for large counties.[26] 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 Census counts and the Revised 2018 Experimental DA population estimates for many small counties might result from random error. In a county-level analysis, O'Hare found all the correlations between demographic variables and net undercounts for young children were more pronounced for large counties than all counties, which may reflect the impact of errors in smaller counties.[27]

There is another reason to focus on large counties: They include a disproportionate share of children who are missed in the census. The Census Bureau's Revised 2018 Experimental DA

Estimates indicate there were 21,015,226 young children in the United States on April 1, 2010, compared with a decennial census count of 20,201,362. This results in a revised net undercount of 813,864 young children and a revised net undercount rate of 3.9%, compared with a net undercount of 970,000 and a net undercount rate of 4.6% in the revised DA official estimates. The collective net undercount rate for the 689 counties in our analysis is 4.4%, and these counties account for approximately 93% of the total national net undercount of young children based on the Revised 2018 Experimental DA Estimates.

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 6.2%, compared with 4.5% for counties with populations between 500,000 and 999,999, 3.1% for counties with populations between 250,000 and 499,999, and 2.4% for counties with fewer than 250,000 people.

Data in Table 1 also show that the net undercount of young children is concentrated in the largest counties in the country. The counties with at least 500,000 people accounted for about 69% of the national net undercount of young children in the 2010 Census. The counties with 250,000 people or more in 2010 accounted for about 81% of the national net undercount of young children. So, understanding what drives net undercounts of young children in these large counties can help us understand net undercounts of children nationwide.

Table 1. Net Undercount of Young Children in 689 Large Counties, by County Population Size			
County Size	Number of children under age 5 (2010 Census)	Net undercount number	Net undercount rate
1 million or more	5,477,261	364,924	6.2
500,000 to 999,999	4,252,129	200,674	4.5
250,000 to 499,999	3,027,675	96,530	3.1
Less than 250,000	3,806,629	92,736	2.4
Total	16,563,694	754,864	4.4
Note: The number of children and net undercount rate are limited to counties with at least 5,000 children under age 5 in the 2010 Census.			
Source: PRB analysis of the U.S. Census Bureau's Revised 2018 Experimental Demographic Analysis Estimates.			

The Census Bureau provided net undercount rates but not the net numbers of young children who were undercounted. PRB calculated the undercount numbers based on the net undercount rates and the 2010 Census population counts of young children in the 689 large counties.

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. Readers are reminded that net undercounts are not the same as people missed in the census. Net undercounts are a product of people missed (omissions) on the one hand and erroneous enumerations (mostly double-counted) along with whole-person imputations on the other hand.[28] The number of people missed is reflected in omissions and the Census Bureau estimates there were about 2.2 million young children missed in the 2010 Census.[29]

Undercounts have sometimes been reported as negative numbers and sometimes as positive numbers.[30] In this report, net undercounts are consistently reported as positive numbers and rates and net overcounts as negative numbers and rates. Measuring net undercounts here as positive numbers and rates makes the correlation and regression 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 also included some additional variables recommended by Census Bureau staff. We identified 40 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.[31] Most of the social, economic, housing, and demographic measures for this analysis were derived from the 2006-2010 American Community Survey 5-year data set, while data on mail return rates are from the 2010 Census.

Potential explanatory variables were grouped into six domains or categories. The variables and domains are shown in Table 2 along with the dependent variable and the means and standard deviations for all variables for the counties with 250,000 people or more and at least 5,000 children under age 5. Table 2A in Appendix A shows the same information for the smaller counties used in this study—those with fewer than 250,000 people and at least 5,000 children under age 5.

The first rows in Tables 2 and 2A indicate that the net undercount rates for young children were higher in the largest counties than in the smaller ones. The mean undercount rate for the largest counties was 3.9% compared with 2.1% for the smaller counties included in the study. The standard errors for both groups are very similar. These results are consistent with past studies that found larger counties had higher net undercount rates for young children.[32]

Table 2. Means and Standard Deviations for All Potential Explanatory Variables for Counties with 250,000+ People, (2006-2010 ACS unless otherwise noted)

	Mean	Standard deviation
Revised 2018 Experimental Demographic Analysis net undercount rate for children under age 5, 2010 Census	**3.9**	**3.3**
Race and Hispanic Origin		
Percent minority (families with related children)	41.0	20.0
Socioeconomic Status		
Percent of families with children under age 5 in poverty	18.7	7.9
Percent of persons under age 5 in poverty	20.4	8.7
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	13.6	5.3
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	13.8	4.7
Percent of households that received public assistance or SNAP	9.2	4.6
Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	19.6	9.2
Percent of 16-24 year olds who are not in the labor force	39.2	5.5
Percent of 16-24 year olds who are unemployed	9.8	2.1
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	5.0	4.5
Family Structure and Living Arrangements		
Percent of households that are family households, female-headed, no spouse present	12.8	3.4
Percent of children under age 18 living in a female-headed household with no spouse present	25.2	8.5
Percent of children under age 6 living with a grandparent householder	8.8	3.3
Percent of households with a grandparent householder responsible for own grandchild	1.5	0.7
Percent of children who are not biological, adopted, or step-children of householder	10.1	3.3
Percent of all households that have 7+ people	1.4	0.9
Percent of children under age 5 who live with nonrelatives or in group quarters	1.4	0.6
Other Demographic Measures		
Percent of the population that are young children under age 5	6.7	1.1
Percent of the population that are children under age 18	24.5	3.1
Percent of the population ages 18-34	22.1	3.7
Percent of the population ages 35-64	39.8	2.8
Percent of the population ages 65+	12.2	3.3
Median age of the population	36.5	3.6
Percent of children ages 1-4 who have moved in the past year	21.2	5.0
Percent of children under age 18 who are foreign-born	4.3	2.3
Percent of adults ages 18+ who are foreign-born	16.2	11.0
Percent of persons who are foreign-born	13.2	8.8
Percent of children under age 18 who are not U.S. citizens	3.3	2
Percent of adults ages 18+ who are not U.S. citizens	8.9	6.1
Percent of persons who are not U.S. citizens	7.5	5
Percent of children under age 6 who are in immigrant families (child or 1+ parents are foreign-born)	25.8	14.6
Percent of all householders who are ages 15-34	21.3	5.1
Percent of children under age 6 who do not have health insurance (2008-12 ACS)	5.9	3.3
Percent of 3 and 4 year olds who are not enrolled in school	49.7	8.9
Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	21.7	16.3
Housing		
Percent of persons living in renter-occupied households	31.3	9.6
Percent of housing units that are vacant	8.1	3.1
Percent of housing units that are crowded (1.01 or more occupants per room)	2.9	2.5
Response/Return Rates		
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

Sources: U.S. Census Bureau: Revised 2018 Experimental Demographic Analysis Estimates, 2010 Census, and 2006-2010 and 2008-2012 American Community Surveys.

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 under age 5, the net undercount rate was 7.5% for Hispanics and 6.3% for children classified as black alone or in combination with one or more other races.[33] The original DA analysis showed the net undercount for young Hispanic children was 7.5%, but a recent update showed that the net undercount of Hispanic children under age 5 in the 2010 Census was 6.5%.[34] 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.[35] 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.[36]

A recent report by the U.S. Government Accountability Office also lists racial and ethnic minorities as a hard-to-count group.[37]

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 noted:

"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."[38]

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.[39] Fernandez et al. conclude that "children who are not found in the census are more disadvantaged than those who are in the census."[40]

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.[41] The percent of the population below the poverty line 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.[42]

Martin argued that residential ambiguity is a key factor in people being missed in surveys and the decennial census.[43] 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.[44] 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."[45]

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.[46] 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.[47] 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.[48]

The U.S. Census Bureau also reported that children in non-English or limited-English-speaking households are less likely to be counted.[49]

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.[50] 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.[51] Robinson et al. also listed renters as more difficult to enumerate than homeowners, and the U.S. Government Accountability Office listed renters as a hard-to-count population.[52]

The 2010 Census Coverage Measurement data show the population living in rental housing units had a net undercount while the population living in owner-occupied housing units had a net overcount.[53]

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.[54] 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.[55]

The Census Bureau's LRS is designed to predict self-response in the decennial census.[56] 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 in our analysis (those with populations of 250,000 or more), based on data from the Revised 2018 Experimental DA net undercount estimates for young children. Nearly all the counties (240 out of 261) exhibit a net undercount for young children, with a mean county-level net undercount rate of 3.9%. The county-level rates range from a 17.5% net undercount to a net overcount of 12.0%.

Table 3. Summary Statistics for 2010 Census Net Undercount Rates for Children Under Age 5 in Counties with 250,000+ People	
Number of counties	261
Number of counties with a net undercount	240
Percent of counties with a net undercount	92
Mean undercount rate	3.9
Standard deviation	3.3
Maximum net undercount rate	17.5
Maximum net overcount rate	-12.0
Note: In this report, net undercounts are reported as positive numbers and net overcounts as negative numbers.	
Source: PRB analysis of data from the U.S. Census Bureau Revised 2018 Experimental Demographic Analysis Estimates.	

Table 3A in Appendix A provides summary statistics for the 428 counties in our analysis with at least 5,000 children under age 5 and fewer than 250,000 people, based on data from the Revised 2018 Experimental DA Estimates of net undercount rates for young children. The net undercount of young children is widespread in these counties as well, with 317 of the 428 counties exhibiting a net undercount for young children and a mean county-level net undercount rate of 2.1%. The county-level rates range from a 19.3% net undercount to a net overcount of 10.2%.

Correlation Analysis

We started with a correlation analysis to identify the characteristics of counties that are most closely associated with the net undercount rate of young 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 appropriate for correlation analysis.

As described in the previous section, we began the analysis with the 40 potential explanatory variables listed in Table 2. The zero-order correlations between these variables and the net undercount rate of young children in the 261 largest counties are shown in Table 4. These correlations range from +0.61 for percent of families with related children in households where

the householder is a racial or Hispanic minority to -0.42 for final mail return rate in the 2010 Census. Most correlations are statistically significantly different than zero.

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 generally expected counties with higher percentages for the explanatory variables to also have higher net undercount rates.

Table 4. Correlations of Potential Explanatory Variables with Net Undercount Rate of Young Children, Counties with 250,000+ People

	Correlation	P-value
Race and Hispanic Origin		
Percent minority (families with related children)	0.61	<.0001
Socioeconomic Status		
Percent of families with children under age 5 in poverty	0.34	<.0001
Percent of persons under age 5 in poverty	0.35	<.0001
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	0.40	<.0001
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	0.22	0.0004
Percent of households that received public assistance or SNAP	0.32	<.0001
Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	0.36	<.0001
Percent of 16-24 year olds who are not in the labor force	0.29	<.0001
Percent of 16-24 year olds who are unemployed	0.10	0.0952
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	0.47	<.0001
Family Structure and Living Arrangements		
Percent of households that are family households, female-headed, no spouse present	0.53	<.0001
Percent of children under age 18 living in a female-headed household with no spouse present	0.50	<.0001
Percent of children under age 6 living with a grandparent householder	0.47	<.0001
Percent of households with a grandparent householder responsible for own grandchild	0.36	<.0001
Percent of children who are not biological, adopted, or step-children of householder	0.52	<.0001
Percent of all households that have 7+ people	0.12	0.0440
Percent of children under age 5 who live with nonrelatives or in group quarters	0.08	0.2214
Other Demographic Measures		
Percent of the population that are young children under age 5	0.01	0.8389
Percent of the population that are children under age 18	-0.14	0.0286
Percent of the population ages 18-34	0.23	0.0002
Percent of the population ages 35-64	-0.22	0.0002
Percent of the population ages 65+	0.03	0.6382
Median age of the population	-0.11	0.0718
Percent of children ages 1-4 who have moved in the past year	-0.05	0.4456
Percent of children under age 18 who are foreign-born	0.40	<.0001
Percent of adults ages 18+ who are foreign-born	0.41	<.0001
Percent of persons who are foreign-born	0.42	<.0001
Percent of children under age 18 who are not U.S. citizens	0.41	<.0001
Percent of adults ages 18+ who are not U.S. citizens	0.41	<.0001
Percent of persons who are not U.S. citizens	0.43	<.0001
Percent of children under age 6 who are in immigrant families (child or 1+ parents are foreign-born)	0.37	<.0001
Percent of all householders who are ages 15-34	0.08	0.1953
Percent of children under age 6 who do not have health insurance (2008-12 ACS)	0.10	0.1151
Percent of 3 and 4 year olds who are not enrolled in school	-0.17	0.0051
Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	0.45	<.0001
Housing		
Percent of persons living in renter-occupied households	0.47	<.0001
Percent of housing units that are vacant	0.30	<.0001
Percent of housing units that are crowded (1.01 or more occupants per room)	0.35	<.0001
Response/Return Rates		
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.42	<.0001
Final undeliverable as addressed rate, 2010 Census	0.11	0.0775

Sources: U.S. Census Bureau: Revised 2018 Experimental Demographic Analysis Estimates, 2010 Census, and 2006-2010 and 2008-2012 American Community Surveys.

However, a few of the zero-order correlations are negative. The statistically significant negative correlation (-0.42) between the 2010 mail return rate and the net undercount rate makes sense because a higher mail return rate should mean that fewer people, including young children, are undercounted.

For two of the variables—Percent of the population ages 35 to 64 and Median age—the negative correlations are likely explained by the higher risk of young children with young parents being missed in the census. Older parents ages 35 to 64 are more likely to self-respond in the census and less likely to leave a young child off the census questionnaire. Adults ages 35 to 64 are also less likely than younger adults to have children under age 5 in the household. Counties with older populations therefore have lower net undercount rates for young children.

The negative correlation (-0.17) between the percent of 3- and 4-year-olds not enrolled in school and the net undercount rate for young children may be explained by the relatively high rate of preschool enrollment among African American children, who are more likely than white children to be missed in the census.

There is also a statistically significant negative correlation (-0.14) between the percent of the population ages 0 to 17 and the net undercount rate for young children. In other words, counties with a higher share of children have lower net undercount rates for young children. It is not clear why this relationship exists.

The next step of our analysis was to reduce the list of 40 potential explanatory variables to a smaller set that could be used in a multiple regression analysis to estimate each independent relationship with the net undercount rate of young children. Variables that are highly correlated with each other cannot be used together in a multiple regression analysis without risking multicollinearity problems. We used three criteria to reduce the list of 40 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 21 independent variables that represented all six domains. Tables 5 and 5A (in Appendix A) show the interrelationships among these 21 independent variables.

Table 5. Intercorrelation Matrix for Counties with a Total Population of 250,000 or More

	Percent minority (families with related children)	Percent of persons under age 5 in poverty	Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	Percent of households that are female-headed family households, with no spouse present	Percent of children under age 18 living in a female-headed household with no spouse present	Percent of children under ages 6 living with a grandparent householder	Percent of children who are not biological, adopted, or step-children of householder	Percent of all households that have 7+ people	Percent of children ages 1-4 who have moved in the past year	Percent of children under age 18 who are foreign-born	Percent of children under age 18 who are not U.S. citizens	Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born)	Percent of all householders who are ages 15-34	Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	Percent of persons living in renter-occupied households	Percent of housing units that are vacant	Percent of housing units that are crowded (1.01 or more occupants per room)	Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	Final undeliverable as addressed rate, 2010 Census
Percent minority (families with related children)	1																				
Percent of persons under age 5 in poverty	0.47	1																			
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	0.57	0.55	1																		
Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	0.47	0.92	0.45	1																	
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	0.71	0.25	0.50	0.24	1																
Percent of households that are female-headed family households with no spouse present	0.72	0.71	0.46	0.75	0.39	1															
Percent of children under age 18 living in a female-headed household with no spouse present	0.52	0.77	0.28	0.81	0.07	0.83	1														
Percent of children under age 6 living with a grandparent householder	0.67	0.68	0.55	0.69	0.39	0.71	0.63	1													
Percent of children who are not biological, adopted, or step-children of householder	0.74	0.69	0.57	0.71	0.38	0.74	0.70	0.93	1												
Percent of all households that have 7+ people	0.48	0.20	0.51	0.19	0.60	0.22	-0.14	0.38	0.35	1											
Percent of children ages 1-4 who have moved in the past year	0.09	0.40	0.22	0.34	-0.19	0.16	0.31	0.19	0.28	-0.06	1										
Percent of children under age 18 who are foreign-born	0.56	-0.08	0.29	-0.09	0.75	0.07	-0.09	0.02	0.10	0.33	-0.23	1									
Percent of children under age 18 who are not U.S. citizens	0.62	0.01	0.41	-0.03	0.77	0.11	-0.05	0.10	0.18	0.39	-0.13	0.98	1								
Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born)	0.62	-0.09	0.41	-0.12	0.81	0.09	-0.16	0.09	0.15	0.50	-0.26	0.93	0.92	1							
Percent of all householders who are ages 15-34	0.24	0.31	-0.06	0.35	0.04	0.21	0.30	0.05	0.18	0.11	0.49	0.02	0.07	-0.06	1						
Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	0.69	0.47	0.42	0.48	0.52	0.68	0.52	0.50	0.55	0.36	0.07	0.34	0.36	0.38	0.19	1					
Percent of persons living in renter-occupied households	0.64	0.49	0.30	0.54	0.46	0.56	0.62	0.38	0.55	0.18	0.21	0.41	0.44	0.39	0.50	0.57	1				
Percent of housing units that are vacant	0.39	0.65	0.30	0.61	-0.06	0.57	0.75	0.53	0.60	-0.14	0.41	-0.16	-0.09	-0.21	0.24	0.40	0.34	1			
Percent of housing units that are crowded (1.01 or more occupants per room)	0.69	0.34	0.61	0.32	0.84	0.39	0.07	0.51	0.52	0.80	-0.06	0.57	0.62	0.70	0.11	0.51	0.52	-0.03	1		
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.73	-0.44	-0.42	-0.45	-0.54	-0.66	-0.47	-0.53	-0.59	-0.39	-0.12	-0.35	-0.40	-0.43	-0.30	-0.89	-0.65	-0.37	-0.57	1	
Final undeliverable as addressed rate, 2010 Census	0.08	0.48	0.32	0.38	-0.15	0.19	0.39	0.34	0.37	-0.13	0.42	-0.19	-0.12	-0.25	0.01	0.06	0.06	0.67	-0.08	-0.03	1

Sources: PRB analysis of data from the U.S. Census Bureau's Revised 2018 Experimental DA Estimates, the 2006-2010 American Community Survey, and the 2010 Census.

Regression Analysis

The 21 independent variables that were selected from the group of 40 potential variables 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 measures currently used to predict geographic differences in the 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 (MRR) for the 2010 Census, which is the basis for the Census Bureau's Low-Response-Score Model, and the net undercount rate for young children.[57] This model shows a statistically significant relationship with the final mail return rate, but explains only 17% of the variation in the net undercount rate of young children among the largest counties (adjusted r-square = .172).

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 rate of young children. HTC tracts are defined here as those with mail return rates of 73% or less in 2010. The regression coefficient indicates that large counties with higher shares of children living in hard-to-count census tracts are also likely to have higher rates of net undercount. Although it is statistically significant, the HTC measure explains only about 20% of the variation in the net undercount rate of young children (adjusted r-square = .200). While the HTC measure is slightly more predictive than the MRR, the adjusted r-square for this measure is much lower than that for our model.

These results indicate that census accuracy for young children is likely driven by different factors than those associated with accuracy for other age groups or the population in general. The lack of a strong relationship 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 Bureau's tract-level model for the Low Response Score, even though it was significant in their block group-level model.[58]

Table 6A (Appendix A) shows results from the same regression models for the smaller counties in our study. We believe the results for the largest counties may be more reliable because the dependent variable is measured more reliably for larger counties. Therefore, we provide the results for smaller counties in Table 6A, but we do not discuss the results in this section.

Table 6. Multivariate Regressions Predicting the County-Level Net Undercount Rate for Children Under Age 5, Counties with 250,000+ People												
	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	35.51	4.27	<.0001	0.00	1.86	0.31	<.0001	0.00	-2.84	0.68	<.0001	0.00
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.40	0.05	<.0001	-0.42								
Percent of children under age 5 living in hard-to-count census tracts					0.09	0.01	<.0001	0.45				
Percent of persons under age 5 in poverty									-0.01	0.04	0.75	-0.03
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative									-0.22	0.05	<.0001	-0.31
Percent of children under age 18 living in a female-headed household with no spouse present									0.27	0.04	<.0001	0.70
Percent of children under age 6 living with a grandparent householder									0.19	0.07	0.01	0.18
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")									0.14	0.07	0.04	0.19
Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born)									0.12	0.02	<.0001	0.52
Percent of persons living in renter-occupied households									-0.08	0.03	0.00	-0.22
Adjusted r-square	0.172				0.200				0.523			

Sources: PRB analysis of data from the U.S. Census Bureau's Revised 2018 Experimental DA Estimates, the 2006-2010 American Community Survey, and the 2010 Census.

Model 3

The final step of our county-level analysis was to identify which of the 21 explanatory variables are independently associated with the net undercount of young children, and to determine if this set of variables could be used to improve geographic targeting efforts to reduce the undercount of young children in the 2020 Census.

Our final model—Model 3—includes all the explanatory variables that were statistically significant for either set of counties (for example, those with 250,000 people or more and those with fewer than 250,000 people). The 21 potential explanatory variables were reduced using a stepwise process. This method identifies the best one-variable model—the one variable that by itself explains the most variation in the dependent variable—and compares it to the best two-variable model. Models were run and variables were added and deleted to generate the best-fitting model. The best two-variable model was compared to the best three-variable model and so on until no improvements in model fit occurred when additional variables were added to the model. Several alternative measures for "best fit" were used and they all produced similar results. This procedure was conducted separately for larger counties (250,000 people or more) and smaller counties (fewer than 250,000 people). At every step, results were checked to make sure that multicollinearity was not an issue.

In Table 6, the multiple adjusted r-square, presented at the bottom of panel 3, shows that this model explains 52% of the variation in the net undercount rates of young children in the largest counties (adjusted r-square = .523). This represents a substantial improvement over the results from Models 1 and 2 using just the percent of children living in hard-to-count tracts or the 2010 mail return rate. These results are similar to those the Census Bureau obtained from the model they used to develop their tract-level Low Response Score (r-squared was 0.55).[59] Note that the Percent of young children living in hard-to-count tracts was not a statistically significant variable in the stepwise regression models. This again suggests that census accuracy for young children is driven by different factors than census accuracy for the population in general.

Our interpretation of the factors most highly predictive of differences in net undercount rates for young children focuses on the results of the largest counties as shown in Table 6, since these are likely to be the most reliable results given the more precise estimates of dependent and independent variables. Several variables are statistically significant and in the predicted direction. Grouping the individual variables into domains or categories may provide a better conceptual understanding of the findings.

Results show that household structure and living arrangements are important for understanding differences in the net undercount rates of young children. Two of the statistically significant variables representing living arrangements or family structure are: Percent of children under age 18 living in a female-headed household with no spouse present and Percent of children under age 6 living with a grandparent householder.

The Percent of children under age 18 living in a female-headed household with no spouse present is an important predictor of county differences in net undercount rates for young children, and the coefficient is positive. This means that counties with higher shares of children living in female-headed households with no spouse present have higher net undercount rates for young children after controlling for other factors. This relationship has also been found in other research; an analysis by Fernandez and her Census Bureau colleagues found children living in female-headed households had higher likelihood of being missed in the census

compared with children in married-couple families.[60] In addition, female-headed households are more likely to be poor, which is another characteristic often associated with being hard-to-count.[61]

The Percent of children under age 6 living with a grandparent householder is statistically significant and the coefficient is positive, meaning counties with larger shares of young children living in grandparent-headed households tend to have higher rates of undercount for young children. Other research has also identified young children living with grandparents as being vulnerable to being missed in the census.[62]

These two measures of family structure or living arrangements that were statistically significant in the model likely reflect unstable housing situations and produce what Martin calls "residential ambiguity" and what Brown and Manning refer to as "family boundary ambiguity."[63]

Two other variables that are in the predicted direction and statistically significant are related to recent immigration. These are: Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well") and Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born). The coefficients for both variables are positive. Counties with larger shares of households that are linguistically isolated, and counties with a large share of children in immigrant families tend to have higher rates of net undercount of young children. This relationship probably reflects recent immigration; for example, Jensen et al. show that recent immigrants have lower coverage rates in the American Community Survey.[64] Other analysts have found recent immigrants to be a hard-to-count population. Fernandez and colleagues found that young children were more likely to be missed in the 2010 Census when half or more of the household members were foreign-born.[65]

With respect to the socioeconomic domain, the Percent of adults ages 18 to 34 with less than a high school diploma, GED, or alternative, is statistically significant in a negative direction, meaning counties with a higher share of younger adults with low educational attainment tend to have lower net undercount rates for young children. It is not clear why this relationship exists after other factors are controlled.

The Percent of young children living in families with incomes below the poverty line is not statistically significant for the largest counties. This suggests that when other variables are controlled, poverty itself is not associated with the net undercount of young children. However, the Percent of young children in poverty is statistically significant for smaller counties. This is a surprising finding for large counties 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.[66] Also, the Low Response Score produced by the Census Bureau found poverty to be statistically significant after other measures had been controlled.[67] And the General Accountability Office lists low-income persons as a hard-to-count characteristic.[68]

More broadly, education and income are key dimensions of socioeconomic status and are widely thought to be related to census accuracy. Yet our analysis shows that for young children, other factors are more powerful explanatory variables than education and family income.

Finally, the housing variable shows that the Percent of persons living in renter-occupied households is statistically significant but is negatively associated with the net undercount rate for young children. This means the more renters in a county, the lower the net undercount rate

for young children after other factors have been controlled. This finding for the share living in renter-occupied households is surprising. The zero-order correlation was positive (+0.47). It could be that counties with higher shares of persons living in renter-occupied households also have higher shares of younger householders. Or, counties with large shares of renters, like those with college towns, may be areas with relatively few children.

5. Tract-Level Analysis

Those interested in targeting outreach efforts to reduce the high net undercount of young children in the census typically focus on areas smaller than counties—such as census tracts. However, the Census Bureau does not publish net undercount rates for young children at the census-tract level. In addition, our results suggest that two measures currently being used to identify areas where young children are more likely to be missed—the 2010 Census mail return rate and the Low Response Score (also based on mail return rates)—are not very good predictors of net undercount rates for young children in large counties. To address this gap, we apply coefficients from our county-level regression models to more recent population data at the census-tract level to produce tract-level estimates of the net undercount rate for young children in the 689 counties included in our study.

There are 57,932 census tracts located in the 689 counties included in this analysis. This is about 80% of all census tracts in the country. We exclude tracts with fewer than 25 children under age 5 because predicted net undercount rates in such tracts would not be reliable. This leaves a total of 56,638 census tracts in our database, with 43,654 located in the largest counties (250,000 people or more) and 12,984 located in smaller counties (fewer than 250,000 people). The concentration of tracts in the largest counties is helpful because these are the counties where our regression model explained the most variance in the net undercount rate of young children.

Although the Census Bureau only provided experimental DA estimates of net undercount rates for young children in 2010 for the 689 counties that had at least 5,000 children under age 5, these counties include 82% of the total number of children under age 5 enumerated in the 2010 Census. The 56,638 tracts in these counties with at least 25 children under age 5 include 99.9% of all children under age 5 enumerated in these counties in the 2010 Census.

We first estimate 2010 tract-level net undercount rates for young children by applying the coefficients from our separate models for larger counties (250,000+ people) and smaller counties (fewer than 250,000 people) to tract-level values for the independent variables from the 2006-2010 ACS. Because there are no observed 2010 tract-level net undercount rates for young children, we are not able to evaluate our tract-level estimates with typical measures such as Mean Absolute Percent Error (MAPE) or Mean Algebraic Percent Error (MALPE). Instead, we aggregate our tract-level estimates to counties and compare the results to the Census Bureau's Revised Experimental DA estimates for these 689 large counties.

The results in Table 7 show that our tract-level estimates of the net undercount rate for young children are slightly higher than the Census Bureau's estimates when aggregated to the county level. However, our predicted average rates at the county level are reasonable and follow the same pattern by county size. There is a larger difference between our county-level estimates (aggregated from tracts) and the Census Bureau's estimates in the range of undercount rates at the county level. Our county-level estimates have an overcount rate of only 1.6% compared with 12% in the Census Bureau's estimates. This is not surprising given that our model includes

those independent variables most strongly associated with net undercount rather than overcount. Similarly, our maximum county-level undercount rate is also lower than the Census Bureau's—12.8% versus 19.3%. This suggests that our tract-level model produces somewhat conservative estimates of the net undercount rate for children under age 5 when aggregated to the county level.

Table 7. Comparison of Tract-Level Estimates of Net Undercount Rates for Children Under Age 5 for 2006-2010 with the Revised 2018 Experimental DA Estimates for Counties in 2010

	Number of counties	Average predicted undercount rate	Standard deviation	Median	Minimum	Maximum
Tract Estimates Aggregated to County Level						
All tracts rolled to counties	689	3.2	2.3	2.9	-1.6	12.8
Tracts in smaller counties (<250,000 people)	428	2.6	2.0	2.3	-1.5	12.5
Tracts in largest counties (250,000+ people)	261	4.2	2.4	4.0	-1.6	12.8
Revised 2018 Experimental DA Estimates						
All counties	689	2.8	3.5	2.6	-12.0	19.3
Smaller counties (<250,000 people)	428	2.1	3.4	1.9	-10.2	19.3
Largest counties (250,000+ people)	261	3.9	3.3	3.4	-12.0	17.5

Note: Estimates are limited to counties with at least 5,000 children under age 5 in the 2010 Census and to tracts within those counties with at least 25 children under age 5 in the 2010 Census.
Sources: PRB analysis of U.S. Census Bureau Revised 2018 Experimental DA Estimates, 2010 Census, and 2006-2010 American Community Survey.

To provide updated estimates of net undercount rates for young children, we apply coefficients from our separate models for the larger and smaller counties to more recent tract-level data for our independent variables from the 2013-2017 ACS. Table 8 compares our tract-level results for 2006-2010 with those from 2013-2017. The average predicted undercount rate for all tracts increased from 4.2% to 4.8%, but rose from 4.7% to 5.4% among the largest counties. Since the regression coefficients are the same for both periods, the higher predicted undercount rates in 2013-2017 suggest that tract-level characteristics have shifted since 2006-2010 to increase the share of children who are at risk of being undercounted. The minimum and maximum predicted undercount rates are higher for tracts than for counties, ranging from an overcount of nearly 19% to an undercount of 39% in 2013-2017.

Table 8. Tract-Level Estimates of Net Undercount Rates for Children Under Age 5, 2006-2010 and 2013-2017
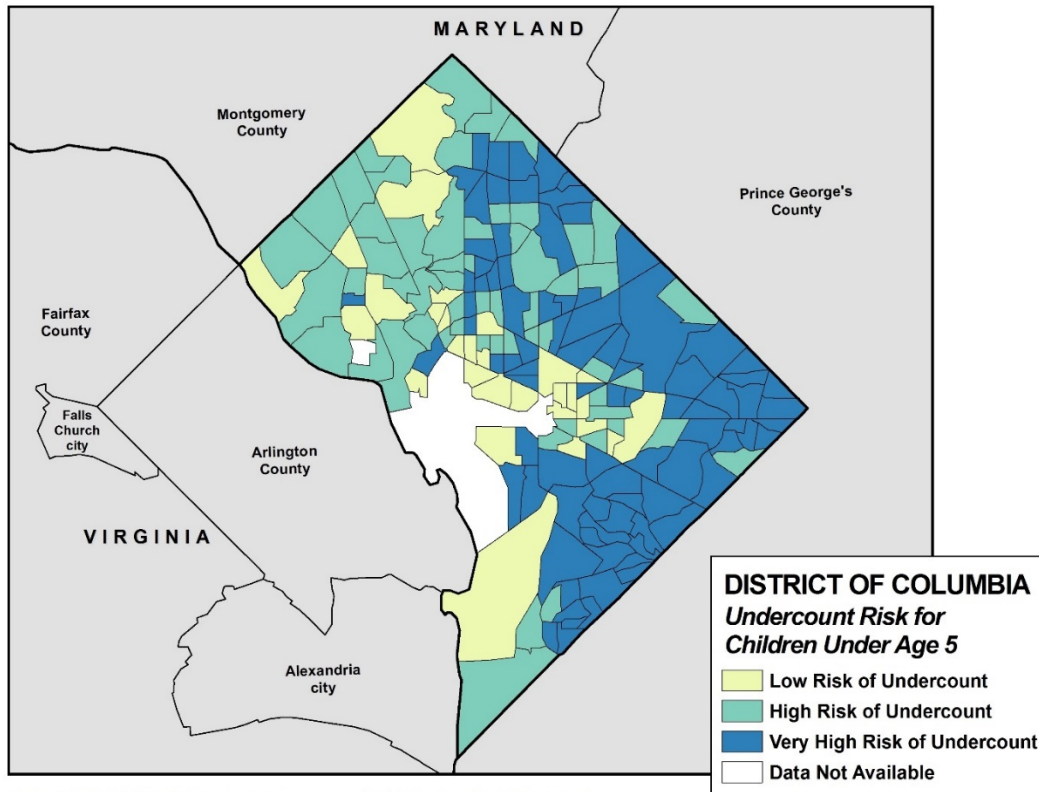
	Number of tracts	Average predicted undercount rate	Standard deviation	Median	Minimum	Maximum
Tract-level Estimates, 2013-2017						
All tracts	56,638	4.8	5.5	4.1	-18.6	39.3
Tracts in smaller counties (< 250,000 people)	12,984	2.9	4.3	2.0	-7.8	27.3
Tracts in largest counties (250,000+ people)	43,654	5.4	5.7	4.9	-18.6	39.3
Tract-level Estimates, 2006-2010						
All tracts	56,638	4.2	5.5	3.4	-21.4	39.2
Tracts in smaller counties (<250,000 people)	12,984	2.5	4.3	1.6	-9.6	28.5
Tracts in largest counties (250,000+ people)	43,654	4.7	5.7	4.1	-21.4	39.2

Note: Estimates are limited to counties with at least 5,000 children under age 5 in the 2010 Census and to tracts within those counties with at least 25 children under age 5 in the 2010 Census.
Sources: PRB analysis of U.S. Census Bureau Revised 2018 Experimental DA Estimates, 2010 Census, and 2006-2010 and 2013-2017 American Community Surveys.

Finally, tracts are classified into three risk categories based on their predicted net undercount rates for children under age 5. The percentage distribution of these predicted rates is used to select the cutoff value that puts the top 25% of tracts in the highest-risk category. Very high-risk tracts are those with a predicted net undercount rate of 8.3% or higher. High-risk tracts are those with a predicted net undercount rate of 0 to 8.29%. Those tracts with a predicted net undercount rate less than 0% are labeled low-risk tracts. A net undercount rate less than 0% indicates a predicted net overcount of young children in that tract. Figure 1 shows how the predicted risk of undercounting young children varies across census tracts in Washington, DC.

Figure 1. Undercount Risk for Children Under Age 5 by Census Tract, District of Columbia

Source: Population Reference Bureau analysis of U.S. Census Bureau data.

Our results suggest that about 25% of children under age 5 (4.1 million) live in tracts classified as having a "very high risk of undercount," while nearly four-fifths (81%) of young children live in a neighborhood with a high or very high risk of a net undercount (see Table 9).

Table 9. Distribution of Children Under Age 5 by Tract-Level Risk Status, 2014-2018

	Number of Young Children	Percent of Young Children
Low risk of undercount or potential overcount	3,095,045	19
High risk of undercount	9,290,040	56
Very high risk of undercount	4,065,149	25
Total	16,450,234	100

Note: The number and percent of children are limited to counties with at least 5,000 children under age 5 in the 2010 Census and tracts within those counties with at least 25 children under age 5 in the 2010 Census.

Sources: PRB analysis of U.S. Census Bureau Revised 2018 Experimental Demographic Analysis Estimates, 2010 Census, and 2014-2018 American Community Survey.

The figures in Table 9 for all children in the 689 counties mask large differences among racial and ethnic groups. The percent of children under age 5 living in tracts with a very high risk of undercounting young children ranges from a high of 48% among African American children to a low of 9% among non-Hispanic white children (see Table 10). A large share of young Hispanic/Latino children also live in neighborhoods with a very high risk of undercount (38%). Data from the 2010 Census indicate these two groups of young children had exceptionally high net undercounts in the 2010 Census. Special attention should be devoted to neighborhoods where young black and Hispanic children are concentrated.

Table 10. Children Under Age 5 Living in Tracts With Very High Risk of Undercounting Young Children, by Race and Hispanic Origin, 2014-2018

	Percent in census tracts with a very high risk of undercounting young children
All children	25
Hispanic/Latino	38
Black alone	48
Asian alone	28
American Indian/Alaska Native alone	31
Native Hawaiian and Other Pacific Islander alone	31
Non-Hispanic White alone	9
Two or more races	22

Notes: The percentage shown includes all children living in counties with at least 5,000 children under age 5 in the 2010 Census and in tracts within those counties with at least 25 children under age 5 in the 2010 Census. Regarding "Hispanic/Latino," Hispanics can be of any race, so they are also included in the other racial categories with the exception of White alone.

Sources: PRB analysis of U.S. Census Bureau Revised 2018 Experimental Demographic Analysis Estimates, 2010 Census, and 2013-2017 and 2014-2018 American Community Surveys.

Our results show that the risk of undercounting young children also varies widely across the 689 large counties included in the analysis. While 53% of children under age 5 in Washington, DC live in a census tract with a very high risk of undercounting young children, the corresponding share in King County, Washington is only 12%. The characteristics of neighborhoods in these two counties help explain the difference in risk status. For example, while 14% of children under age 6 in DC live with a grandparent householder, only 6% of young children in King County fall in that category. Table 11 shows how the characteristics of tracts with a very high risk of undercounting young children differ from those for all tracts in the 689 counties and from those for the United States as a whole.

Table 11. Characteristics Associated with the Risk of Undercounting Children Under Age 5, 2013-2017			
	United States	All census tracts in large counties	Census tracts with a very high risk of undercounting young children
Children under age 5 living in families with incomes below 100% of the poverty level (%)	22	22	34
Adults ages 18 to 34 with less than a high school diploma (%)	12	11	14
Children under age 18 year living in a female-headed household with no spouse present (%)	25	26	43
Children under age 6 living with a grandparent householder (%)	11	11	19
Households that are linguistically isolated (%)	4	5	11
Children under age 6 living in immigrant families (child is foreign-born or at least one parent is foreign-born) (%)	25	28	43
Population living in renter-occupied housing units (%)	35	36	50

Note: Percents below the national level are limited to counties with at least 5,000 children under age 5 in the 2010 Census and to tracts within those counties with at least 25 children under age 5 in the 2010 Census.

Sources: PRB analysis of U.S. Census Bureau Revised 2018 Experimental DA Estimates, 2010 Census, and 2013-2017 American Community Survey.

6. Limitations

Note 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. Moreover, smaller counties typically have higher levels of random errors in both the dependent and independent variables.

In addition, most of the independent variables used in this study come from the Census Bureau's American Community Survey. These estimates are derived from a sample and are therefore subject to sampling error, as well as non-sampling error. Many of the independent variables represent the population under age 5, which is a relatively small population with larger standard errors.

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

7. 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 under age 5. 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.

For the most part, 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 12 shows the six variables that were statistically significant in our model for counties with 250,000 people or more, along with the six most powerful variables from the Census Bureau's Low Response Score (LRS) Model.

Table 12. Comparison of the Six Statistically Significant Variables in Our Model to the Six Most Powerful Estimators in the Erdman-Bates Model for the Low Response Score	
Our Undercount Model (County-Level, Counties with 250,000+ People)	**Erdman-Bates Low Response Score Model (Census-Tract Level)**
Percent of children under age 18 living in a female-headed household with no spouse present	Number of persons per household
Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born)	Percent of population ages 65+
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	Percent of housing units that are renter occupied
Percent of persons living in renter-occupied households	Percent of housing units that are vacant
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	Percent of population that is non-Hispanic white
Percent of children under age 6 living with a grandparent householder	Median home value
Note: Variables are shown in descending order by the strength of their associations with the dependent measures (net undercount rate and low response, respectively).	
Source: Erdman and Bates (2017), *Public Opinion Quarterly.*	

There is only one common variable in the two lists (renters). But even here there is an important difference. In the LRS model, the higher the share of renters in a census tract, the lower the self-response rate after other factors have been controlled. For young children, the higher the percentage of renters in a county, the lower the net undercount rate of young children after other factors have been controlled.

The lack of similarity in the two lists 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). In discussing the undercount of young children, the Census Bureau says "Most often, people who do return the forms just forget to count everyone under their roof."[69] One recent study found 18% of low-income parents of young children were not sure they would count their young children in the

census.[70] Other studies have found a lot of confusion among adults about whether young children are supposed to be included in the census or not.[71]

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 the 689 large counties in this study. The percent of young children living in hard-to-count census tracts (defined here as those with a mail return rate of 73% or less in the 2010 Census) has a correlation of +0.45 with net undercount rates of children under age 5. The mail return rate (what the LRS tries to estimate) has a correlation of -0.42. In our multiple regression models, the percent of young children living in hard-to-count tracts explains 20% of the variance in net undercount rates for young children among the 261 largest counties, while the mail return rate explains 17%. In comparison, our final model explains 52% of the variance in net undercount rates for young children across the 261 largest counties.

8. Summary and Conclusions

The data examined here indicate that the net undercount rate for the population under age 5 varies substantially across large counties, but almost all counties with 250,000 people or more 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 under age 5. In the 261 largest counties with 250,000 people or more, there was a net undercount of 662,128 children under age 5, which accounts for 81% of the nationwide net undercount for this age group. So, understanding what drives net undercounts of young children in these large counties can help us understand net undercounts of children nationwide. In addition, information about where the net undercount rates for young children are the highest should help child advocates and others pinpoint the places that deserve special attention in the 2020 Census.

Our analysis finds that two metrics 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.

Instead, we find the best predictor variables for net undercount of young children are:

- Family structure and living arrangements
  - Percent of children under age 18 living in female-headed households with no spouse present.
  - Percent of young children under age 6 living with a grandparent householder.
- Recent immigration
  - Percent of households that are linguistically isolated.
  - Percent of children under age 6 living in immigrant families.

- Socioeconomic status
  - Percent of persons living in renter-occupied households

    o   Percent of persons ages 18 to 34 with less than a high school diploma, GED, or alternative.

The superior results of our model in terms of predicting net undercount rates for children under age 5 indicate this line of research should be pursued to better identify areas with a greater risk of undercounting young children in the future. In the meantime, with the 2020 Census rapidly approaching, we hope this study and our new undercount risk measure 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.

9. Appendix A. Tables for Counties with Fewer than 250,000 People

Table 2A. Means and Standard Deviations for All Potential Explanatory Variables for Counties with Fewer than 250,000 People, (2006-2010 ACS unless otherwise noted)

	Mean	Standard deviation
Revised 2018 Experimental Demographic Analysis net undercount rate for children under age 5, 2010 Census	**2.1**	**3.4**
Race and Hispanic Origin		
Percent minority (families with related children)	25.4	18.7
Socioeconomic Status		
Percent of families with children under age 5 in poverty	20.8	8.1
Percent of persons under age 5 in poverty	22.6	9.0
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	14.3	5.7
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	14.4	6
Percent of households that received public assistance or SNAP	10.4	4.2
Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	21.7	8.4
Percent of 16-24 year olds who are not in the labor force	37.4	6.6
Percent of 16-24 year olds who are unemployed	9.9	2.6
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	2.3	2.9
Family Structure and Living Arrangements		
Percent of households that are family households, female-headed with no spouse present	12.0	3.2
Percent of children under age 18 living in a female-headed household with no spouse present	24.1	7.4
Percent of children under age 6 living with a grandparent householder	9.5	4.4
Percent of households with a grandparent household responsible for own grandchild	1.7	0.9
Percent of children who are not biological, adopted, or step-children of householder	10.4	3.7
Percent of all households that have 7+ people	1.2	0.8
Percent of children under age 5 who live with nonrelatives or in group quarters	1.6	1.0
Other Demographic Measures		
Percent of the population that are young children under age 5	6.6	1.1
Percent of the population that are children under age 18	24.6	3.1
Percent of the population ages 18-34	21.8	5.2
Percent of the population ages 35-64	39.6	3.3
Percent of the population ages 65+	12.9	3.3
Median age of the population	37.0	4.1
Percent of children ages 1-4 who have moved in the past year	22.5	5.8
Percent of children under age 18 who are foreign-born	2.1	1.6
Percent of adults ages 18+ who are foreign-born	6.7	5.6
Percent of persons who are foreign-born	5.5	4.4
Percent of children under age 18 who are not U.S. citizens	1.6	1.4
Percent of adults ages 18+ who are not U.S. citizens	4.1	3.9
Percent of persons who are not U.S. citizens	3.4	3.1
Percent of children under age 6 who are in immigrant families (child or 1+ parents are foreign-born)	12.5	9.4
Percent of all householders who are ages 15-34	20.9	5.7
Percent of children under age 6 who do not have health insurance (2008-12 ACS)	6.2	3.5
Percent of 3 and 4 year olds who are not enrolled in school	55.1	9
Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	11.6	15.6
Housing		
Percent of persons living in renter-occupied households	27.4	7.9
Percent of housing units that are vacant	8.3	2.8
Percent of housing units that are crowded (1.01 or more occupants per room)	2.2	1.8
Response/Return Rates		
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	80.7	3.7
Final undeliverable as addressed rate, 2010 Census	12.7	4.9
Sources: U.S. Census Bureau: Revised 2018 Experimental Demographic Analysis Estimates, 2010 Census, and 2006-2010 and 2008-2012 American Community Surveys.		

Table 3A. Summary Statistics for 2010 Census Net Undercount Rates for Children Under Age 5 in Counties with Fewer than 250,000 People

Number of counties	428
Number of counties with a net undercount	317
Percent of counties with a net undercount	74
Mean undercount rate	2.1
Standard deviation	3.4
Maximum net undercount rate	19.3
Maximum net overcount rate	-10.2

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

Source: PRB analysis of data from the U.S. Census Bureau's Revised 2018 Experimental Demographic Analysis Estimates.

Table 4A. Correlations of Potential Explanatory Variables with Net Undercount Rate of Young Children, Counties with Fewer than 250,000 People

	Correlation	P-value
Race and Hispanic Origin		
Percent minority (families with related children)	0.45	<.0001
Socioeconomic Status		
Percent of families with children under age 5 in poverty	0.35	<.0001
Percent of persons under age 5 in poverty	0.37	<.0001
Percent of adults ages 18+ with less than a high school diploma, GED, or alternative	0.34	<.0001
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	0.20	<.0001
Percent of households that received public assistance or SNAP	0.33	<.0001
Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	0.38	<.0001
Percent of 16-24 year olds who are not in the labor force	0.12	0.0124
Percent of 16-24 year olds who are unemployed	0.00	0.9371
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	0.29	<.0001
Family Structure and Living Arrangements		
Percent of households that are family households, female-headed, no spouse present	0.47	<.0001
Percent of children under age 18 living in a female-headed household with no spouse present	0.48	<.0001
Percent of children under age 6 living with a grandparent householder	0.35	<.0001
Percent of households with a grandparent household responsible for own grandchild	0.39	<.0001
Percent of children who are not biological, adopted, or step-children of householder	0.42	<.0001
Percent of all households that have 7+ people	0.13	0.0095
Percent of children under age 5 who live with nonrelatives or in group quarters	0.10	0.0367
Other Demographic Measures		
Percent of the population that are young children under age 5	0.17	0.0004
Percent of the population that are children under age 18	0.02	0.6385
Percent of the population ages 18-34	0.16	0.0007
Percent of the population ages 35-64	-0.26	<.0001
Percent of the population ages 65+	-0.03	0.5910
Median age of the population	-0.17	0.0004
Percent of children ages 1-4 who have moved in the past year	0.20	<.0001
Percent of children under age 18 who are foreign-born	0.13	0.0082
Percent of adults ages 18+ who are foreign-born	0.21	<.0001
Percent of persons who are foreign-born	0.20	<.0001
Percent of children under age 18 who are not U.S. citizens	0.16	0.0008
Percent of adults ages 18+ who are not U.S. citizens	0.22	<.0001
Percent of persons who are not U.S. citizens	0.21	<.0001
Percent of children under age 6 who are in immigrant families (child or 1+ parents are foreign-born)	0.14	0.0048
Percent of all householders who are ages 15-34	0.12	0.0105
Percent of children under age 6 who do not have health insurance (2008-12 ACS)	0.07	0.1608
Percent of 3 and 4 year olds who are not enrolled in school	-0.09	0.0622
Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	0.32	<.0001
Housing		
Percent of persons living in renter-occupied households	0.38	<.0001
Percent of housing units that are vacant	0.33	<.0001
Percent of housing units that are crowded (1.01 or more occupants per room)	0.24	<.0001
Response/Return Rates		
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.36	<.0001
Final undeliverable as addressed rate, 2010 Census	0.06	0.2050

Sources: U.S. Census Bureau: Revised 2018 Experimental Demographic Analysis Estimates, 2010 Census, and 2006-2010 and 2008-2012 American Community Surveys.

Table 5A. Intercorrelation Matrix for Counties with a Total Population Less than 250,000

	Percent minority (families with related children)	Percent of persons under age 5 in poverty	Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	Percent of households that are female-headed family households, with no spouse present	Percent of children under age 18 living in a female-headed household with no spouse present	Percent of children under ages 6 living with a grandparent householder	Percent of children who are not biological, adopted, or step-children of householder	Percent of all households that have 7+ people	Percent of children ages 1-4 who have moved in the past year	Percent of children under age 18 who are foreign-born	Percent of children under age 18 who are not U.S. citizens	Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born)	Percent of all householders who are ages 15-34	Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	Percent of persons living in renter-occupied households	Percent of housing units that are vacant	Percent of housing units that are crowded (1.01 or more occupants per room)	Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	Final undeliverable as addressed rate, 2010 Census
Percent minority (families with related children)	1																				
Percent of persons under age 5 in poverty	0.48	1																			
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative	0.48	0.56	1																		
Percent of children under age 18 living in households that received public assistance, SNAP, or SSI	0.49	0.89	0.53	1																	
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")	0.60	0.29	0.49	0.28	1																
Percent of households that are female-headed family households with no spouse present	0.75	0.69	0.45	0.72	0.25	1															
Percent of children under age 18 living in a female-headed household with no spouse present	0.62	0.78	0.32	0.78	0.09	0.87	1														
Percent of children under age 6 living with a grandparent householder	0.61	0.56	0.53	0.60	0.44	0.65	0.52	1													
Percent of children who are not biological, adopted, or step-children of householder	0.68	0.66	0.61	0.72	0.37	0.73	0.66	0.89	1												
Percent of all households that have 7+ people	0.49	0.21	0.48	0.25	0.64	0.27	0.01	0.49	0.44	1											
Percent of children ages 1-4 who have moved in the past year	0.20	0.33	0.14	0.33	0.02	0.27	0.32	0.04	0.21	-0.01	1										
Percent of children under age 18 who are foreign-born	0.37	-0.05	0.22	-0.10	0.59	-0.11	-0.11	-0.06	-0.01	0.24	0.02	1									
Percent of children under age 18 who are not U.S. citizens	0.44	0.06	0.33	0.01	0.64	-0.01	-0.03	0.02	0.10	0.30	0.09	0.98	1								
Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born)	0.47	0.01	0.40	-0.01	0.71	-0.02	-0.10	0.07	0.12	0.40	0.03	0.87	0.88	1							
Percent of all householders who are ages 15-34	0.20	0.05	-0.29	0.04	0.09	0.08	0.15	-0.20	-0.10	0.00	0.32	0.27	0.28	0.13	1						
Percent of children under age 5 living in Hard-to-Count (HTC) census tracts, 2010 Census	0.66	0.42	0.31	0.44	0.51	0.52	0.42	0.57	0.58	0.49	0.14	0.16	0.21	0.23	0.13	1					
Percent of persons living in renter-occupied households	0.51	0.50	0.09	0.46	0.23	0.46	0.61	0.13	0.33	0.07	0.49	0.34	0.39	0.28	0.63	0.35	1				
Percent of housing units that are vacant	0.47	0.62	0.41	0.58	0.21	0.59	0.62	0.55	0.63	0.18	0.30	-0.04	0.03	0.00	0.02	0.44	0.38	1			
Percent of housing units that are crowded (1.01 or more occupants per room)	0.64	0.38	0.55	0.40	0.75	0.37	0.18	0.62	0.62	0.83	0.09	0.38	0.45	0.52	0.06	0.65	0.30	0.33	1		
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.69	-0.46	-0.37	-0.50	-0.48	-0.60	-0.47	-0.60	-0.67	-0.44	-0.25	-0.16	-0.24	-0.26	-0.24	-0.82	-0.47	-0.47	-0.63	1	
Final undeliverable as addressed rate, 2010 Census	0.07	0.29	0.28	0.28	0.08	0.12	0.17	0.25	0.30	0.10	0.14	-0.04	0.01	0.03	-0.20	0.14	0.10	0.39	0.21	-0.15	1

Sources: PRB analysis of data from the U.S. Census Bureau's Revised 2018 Experimental DA Estimates, the 2006-2010 American Community Survey, and the 2010 Census.

Table 6A. Multivariate Regressions Predicting the County-Level Net Undercount Rate for Children Under Age 5, Counties with Fewer than 250,000 People												
	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	28.83	3.41	<.0001	0.00	1.28	0.20	<.0001	0.00	-4.02	0.65	<.0001	0.00
Final mail return rate (initial questionnaire and replacement questionnaire), 2010 Census	-0.33	0.04	<.0001	-0.36								
Percent of children under age 5 living in hard-to-count census tracts					0.07	0.01	<.0001	0.32				
Percent of persons under age 5 in poverty									-0.06	0.03	0.07	-0.15
Percent of adults ages 18-34 with less than a high school diploma, GED, or alternative									-0.03	0.04	0.40	-0.06
Percent of children under age 18 living in a female-headed household with no spouse present									0.23	0.04	<.0001	0.50
Percent of children under age 6 living with a grandparent householder									0.07	0.05	0.18	0.08
Percent of households that are linguistically isolated (no one ages 14+ speaks English "very well")									0.31	0.09	0.00	0.26
Percent of children under age 6 who are in immigrant families (child is foreign-born or at least one parent is foreign-born)									0.00	0.03	0.94	-0.01
Percent of persons living in renter-occupied households									0.04	0.03	0.22	0.08
Adjusted r-square	0.125				0.102				0.298			

Sources: PRB analysis of data from the U.S. Census Bureau's Revised 2018 Experimental DA Estimates, the 2006-2010 American Community Survey, and the 2010 Census.

10. Appendix B. Census Bureau Methodology for the Revised 2018 Experimental Demographic Analysis Estimates for Young Children

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%) of the estimated population from the national DA estimates for those ages 0 to 4 in 2010 is derived from birth data.[72] 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.[73]

In March 2018, the Census Bureau developed a first set of experimental state- and county-level estimates of net coverage error for young children under age 5 in the 2010 Census. These experimental estimates were not part of the Census Bureau's official 2010 Demographic Analysis estimates. The subnational estimates were produced using a cohort-component method of population change for births, deaths, and domestic migration, and a stock method for the estimates of net international migration, as shown below:

P1 = P0 +B-D+NDM + NIM

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.

A more detailed description of the data sources and methods used to produce the March 2018 experimental estimates is included in a conference paper presented at the annual meeting of the Population Association of America.[74]

In July 2019, the Census Bureau made a revision to the method used to estimate the domestic migration component. In the original series, the method intended to capture the migration of 0-year-olds during their first year of life. To this end, the domestic migration rates were applied at the end of the period over which they were born. For the other ages, the rates were applied at the beginning of the period. This approach treated migration inconsistently by age, produced an extra migration period for all ages, and overstated domestic migration overall.

For the revised series, the Census Bureau applied domestic migration rates at the beginning of the period to ensure domestic migration estimates only exist for ages 1-4 at the end of the period. This also ensured that there are an appropriate number of migration periods for all ages. In addition, the revised method aligns more with the approach used to produce the official

county-level population estimates. The overall impact of the new domestic migration method is that there is less domestic migration during the period as illustrated below:

National Sum of County-to-County Domestic Migrants for Ages 1 to 4 by Source			
	Initial DA Estimates	Revised DA Estimates	American Community Survey
2006-2010	6,108,838	4,121,362	4,247,744
Note: The universe for the migration question in the American Community Survey question is the population aged 1 and older.			
Source: U.S. Census Bureau, Demographic Analysis Program, 2006-2010 American Community Survey.			

County-level census coverage rates for young children are derived by comparing the Census Bureau's Revised Experimental DA net undercount estimates for children under age 5 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.[75]

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 Young, A.A. "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 Griffin, D.H. "Final Task Force Report: Task Force on the Undercount of Young Children," Memorandum for Frank A. Vitrano, U.S. Census Bureau, Washington, D.C., 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 Zelnik, 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, Y. and Robinson, J.G. "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), https://www2.census.gov/programs-surveys/decennial/2020/program-management/memo-series/2020-report-2010-undercount-children-ommissions.pdf.

[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, W.P. "State-Level 2010 Census Coverage Rates for Young Children," Population Research and Policy Review 33, no. 6 (2014): 797-816"

[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 and Robinson. "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] 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).

[27] 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.

[28] O'Hare, W.P. "Understanding Who Was Missed in the 2010 Census," Population Reference Bureau, Washington, DC., 2019, https://www.prb.org/understanding-who-was-missed-in-the-2010-census/

[29] Hogan, H. and Griffin, D. "Investigating the 2010 Undercount of Young Children—A New Look at 2010 Census Omissions by Age," issued July 26, 2016., U.S. Census Bureau, Washington DC., https://www2.census.gov/programs-surveys/decennial/2020/program-management/memo-series/2020-report-2010-undercount-children-ommissions.pdf

[30] 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).

[31] 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), https://www.census.gov/content/dam/Census/newsroom/press-kits/2018/counting-young-children-in-2020-census.pdf.

[32] O'Hare, W.P. "Geographic Variation in 2010 U.S. Census Coverage Rates for Young Children: A Look at Counties," *International Journal of Social Science Studies,* Vol. 5, No. 9 (Redframe Publishing, Sept. 2017): 11-24.

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

[34] 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.

[35] 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.

[36] 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.

[37] 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).

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

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

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

[41] Robinson, et. al. "The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis"; and U.S. Government Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups." GAO-18-599, July 2018, https://www.gao.gov/assets/700/693450.pdf.

[42] U.S. Government Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups." GAO-18-599, July 2018, https://www.gao.gov/assets/700/693450.pdf.

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

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

[45] U.S. Census Bureau, "Counting Young Children in the 2020 Census" (2018), https://www.census.gov/content/dam/Census/newsroom/press-kits/2018/counting-young-children-in-2020-census.pdf.

[46] 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."

[47] 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.

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

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

[50] Robinson et al. "The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis."

[51] Robinson et al. "The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis."

[52] Robinson et al. "The Planning Database: Decennial Census Data for Historical, Real-time, and Prospective Analysis," and U.S. Government Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups." GAO-18-599, July 2018, https://www.gao.gov/assets/700/693450.pdf.

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

[54] 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).

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

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

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

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

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

[60] Fernandez, "Using Administrative Records and the American Community Survey to Study the Characteristics of Undercounted Young Children in the 2010 Census."

[61] U.S. Government Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerating Hard-to-Count Groups," GAO-18-599, July 2018, https://www.gao.gov/assets/700/693450.pdf.

[62] U.S. Census Bureau, "Investigating the 2010 Undercount of Young Children –Analysis of Census Coverage Measurement Results" (July 2016); U.S. Census Bureau, "Investigating the 2010 Undercount of Young Children -- Analysis of Census Coverage Measurement Results," January 2017, https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-2017_04-undercount-children-analysis-coverage.pdf; U.S. Census Bureau, "Memorandum 2017.04: Investigating the 2020 2010 Undercount of Young Children –Analysis of Census Coverage Measurement Results," https://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/memo-series/2020-memo-2017_04.html.

[63] Martin, E. "Strength of Attachment: Survey coverage of people with tenuous ties to residences, *Demography,* 44, no. 2 (2007): 427-44; Brown S. L. and Manning, W.D. "Family Boundary Ambiguity and the Measurement of Family Structure: The Significance of Cohabitation," *Demography, V*ol. 46, no. 1 (2009): 85-101.

[64] 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).

[65] Fernandez, "Using Administrative Records and the American Community Survey to Study the Characteristics of Undercounted Young Children in the 2010 Census."

[66] Fernandez, "Using Administrative Records and the American Community Survey to Study the Characteristics of Undercounted Young Children in the 2010 Census."

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

[68] U.S. Government Accountability Office, "2020 Census: Actions Needed to Address Challenges to Enumerate Hard-to-Count Groups" GAO-18-599, July 2018, https://www.gao.gov/assets/700/693450.pdf.

[69] U.S. Census Bureau, Children Under 5 Most Undercounted in Last Census, November 13, 2019, page 2.

[70] Lake Research Partners, Webinar for the Partnership for America's Children: "Census 2020 Messaging Testing Results" September 10, 2019.

[71] Article 1 and Quadrant Strategies, "Recent polling on perceptions of the 2020 US Census, with a particular focus on counting children." November 2019.

[72] U.S. Census Bureau, "National Demographic Analysis Tables: 2010" (December 2010), https://www.census.gov/data/tables/2012/demo/popest/revised-2010-demographic-analysis-etimates.html.

[73] Devine, J. et al., "The Use of Vital Statistics in the 2010 Demographic Analysis Estimates," (2010) Census Bureau Working Paper no. 88, https://www.census.gov/population/www/documentation/twps0088/twps0088.pdf.

[74] 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, https://paa.confex.com/paa/2018/mediafile/ExtendedAbstract/Paper21374/King%20Ihrke%20Jenson%20PAA%20website.pdf.

[75] 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. 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).