

Research Article

Assessing Net Coverage for Young Children in the 2010 U.S. Decennial Census

William P. O'Hare

The Annie E. Casey Foundation, 4228 Red Bankana Way, Ellicott City, MD 21042, USA

Correspondence should be addressed to William P. O'Hare; billohare1@gmail.com

Received 8 July 2014; Accepted 21 October 2014; Published 18 November 2014

Academic Editor: Kathryn Kost

Copyright © 2014 William P. O'Hare. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The U.S. Census Bureau's Demographic Analysis shows that the population aged 0 to 4 experienced a net undercount rate of 4.6 percent in the 2010 Decennial Census. This is more than twice as high as any other age group. Despite the fact that the relatively high net undercount of young children was uncovered more than fifty years ago, this problem has received little systematic attention from demographers. To help fill that gap in the literature, this study examines the accuracy of the count of children in the 2010 Decennial Census. The initial focus on all children shifts to a focus on young children (aged 0 to 4) where the net undercount rate is the highest. Discussion highlights some of the potential explanations for the findings.

1. Introduction

A high net undercount of young children has been documented historically in the U.S. Decennial Census and a high net undercount of children has been experienced in societies as varied as China, South Africa, Laos, and the former Soviet Union [1–7]. There is also evidence that young children are underreported in some of the U.S. Census Bureau's major surveys [8].

Previous research on the quality of the data from the 2010 U.S. Decennial Census [9–11] provides limited data on the net coverage rate for children but provides little detail and no ideas about why young children have such a high net undercount in the Census. This study extends the stream of research regarding the net undercount of children by providing a detailed examination of net undercounts and overcounts of children (population aged 0 to 17) in the 2010 U.S. Decennial Census. In particular, the net coverage of young children is contrasted to that of adults and older children in this study.

The analysis presented here rests largely on the results of the U.S. Census Bureau's Demographic Analysis (DA) method for assessing Census accuracy. The reasons for focusing on the results of Demographic Analysis rather than the U.S. Census Bureau's Dual-System Estimates results are explained later in the paper.

After providing background on the Demographic Analysis (DA) estimation methodology, DA estimates for children are compared to the 2010 Decennial Census counts to detect net undercounts and overcounts. Following a brief overview of historical trends, results are examined by single-year of age, sex, race, and Hispanic origin for the child population. That leads to a focus on the population aged 0 to 4 where the net undercount rate is the highest. Finally, the importance of focusing on age differentials and some ideas about potential explanations for the high net undercount of young children are discussed.

2. Demographic Analysis History and Methodology

Assessing the net undercounts in the U.S. Decennial Census is typically based on one of two methods: (1) Demographic Analysis (DA) and/or (2) Dual System Estimates (DSE). This study focuses on the results of DA for reasons that will be explained later in this section. Since there are already several detailed descriptions of the DA methodology available, I will only review the method briefly here [12–14].

The DA method has been used to assess the accuracy of decennial census figures for more than a half century and its origins are often traced back to an article by Price [15]. The

unexpectedly high number of young men who turned up at the first compulsory selective service registration on October 6, 1940, alerted scholars to the possibility of underenumeration in the 1940 Decennial Census. The selective service data also provided an independent population estimate for assessing the size of such underenumeration in the Decennial Census.

The relatively high net undercount among young children was uncovered early in the history of DA. In one of the first systematic efforts to use DA to examine decennial census results, Coale [16] found that children aged 0 to 4 had a relatively high net undercount rate in the censuses of 1940 and 1950. Siegel and Zelnik [17] also found a significant net undercount of children aged 0 to 4 in the 1950 and 1960 Decennial Censuses. Coale and Zelnick [18] found high net undercount rates for young children in the Decennial Censuses as far back as 1880. Coale and Rives [19] found very high undercount rates for young black children in every Decennial Census from 1880 to 1970. A recent reanalysis of historical U.S. Census data shows a significant net undercount of young children in the U.S. Censuses of the late 1800s and early 1900s [20]. Genealogical research shows a similar pattern of underreporting young children as far back as the 1850s [21].

The DA method compares the Census counts to independent estimates of the expected population based largely on vital events data. The DA method employed for the 2010 Decennial Census used one technique to estimate the population under age of 75 and another method to estimate the population aged 75 and older [22]. Since this study focuses on children, only the method used for people aged 0 to 74 is discussed here (people under age of 1 are classified as age of 0).

The 2010 DA estimates for the population aged 0 to 74 are based on the compilation of historical estimates of the components of population change: births (B), deaths (D), and net international migration (NIM). The data and methodology for each of these components are described in separate background documents prepared for the development and release of the Census Bureau's 2010 DA estimates [23–25].

As described by the U.S. Census Bureau the DA population estimates for age of 0 to 74 are derived from the basic demographic accounting equation (1) applied to each birth cohort:

$$P_{0-74} = B - D + NIM. \quad (1)$$

P_{0-74} is population for each single year of age from 0 to 74, B is number of births for each age cohort, D is number of deaths for each age cohort since birth, and NIM is net international migration for each age cohort.

For example, the estimate for the population aged 17 on the April 1, 2010 Decennial Census date is based on births from April 1992 through March 1993, reduced by the deaths to that cohort in each year between 1992 and 2010, and incremented by net international migration (NIM) of the cohort of each year over the 17-year period.

Births are by far the largest component of the population estimates in the DA estimates for children. Births account

for 97 percent of DA population estimate for age of 0 to 17 in 2010 and for 99.6 percent of the DA population estimate for the population aged 0 to 4 [26]. The Middle Series DA estimate of the population aged 0 to 4 released in December 2010 is comprised of 21,121,000 births, 154,000 deaths, and net international migration of 240,000. The Middle Series DA estimate for the population aged 0 to 17 released in December 2010 is comprised of 73 million births, 657,000 deaths, and a net international migration of 3.1 million.

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. After a thorough review of vital statistic prior to the 2010 Census, the U.S. Census Bureau [24, page 3] stated the following:

“The following assumptions are made regarding the use of vital statistics for DA:

- (i) Birth registration has been 100 percent complete since 1985.
- (ii) Infant deaths were underregistered at one-half the rate of the underregistration of births up to and including 1959.
- (iii) The registration of deaths for ages 1 and over has been 100 percent complete for the entire DA time series starting in 1935.”

The Census Bureau's conclusion rests heavily on the results of a test of birth registration completeness conducted by the National Center for Health Statistics (NCHS) that found over 99.2 percent of the births occurring between 1964 and 1968 were registered. Although some of the characteristics gathered on the birth certificate may be suspect, the registration of births is widely seen as complete.

In addition to regularly published totals, the Census Bureau receives microdata files from NCHS containing detailed monthly data on each birth and death that are used for DA estimates by race. Construction of DA estimates by race is discussed later in this paper.

The Census Bureau changed the way it calculated net international migration for the 2010 set of DA estimates. The current method relies heavily on data from the Census Bureau's American Community Survey (ACS) where the location of the residence one year ago (ROYA) is ascertained for everyone in the survey. The total number of yearly immigrants is derived from this question in each year of the ACS, and then that total number is distributed to demographic cells (sex, age, and race) based on an accumulation of the same data over the last five years of the ACS. Five years of ACS data are used to provide more stable and reliable estimates for small demographic groups. On the other hand, the five-year average may mask changes in trends over time. Given changing economic conditions, it would not be surprising if the pattern in 2008 to 2010 differed from the pattern before 2008; however, I suspect such errors would be small for the child population. In addition to estimates for the total population, NIM provides estimates for younger blacks (black alone and black alone or in combination) and for Hispanics under age of 20.

Statistics on emigration of the foreign-born population from the U.S. are based on a residual method comparing data from the 2000 Decennial Census to later ACS estimates to develop rates and then applying those rates to observed populations [27].

Emigration of U.S. citizens (net native migration) is derived by examining census data from several other countries [28]. This method of estimating out migration of native-born is problematic for a couple of reasons. Data are not available for every country and the quality of some foreign censuses is suspect. However, with few exceptions [29], it is widely felt that emigration has little impact on population estimates for young children.

There are four major limitations to DA. First, it is only routinely available for the nation as a whole. The fact that many people move after birth is a barrier to employing this method at the subnational level. While attempts have been made to produce subnational DA estimates [31–33], they have not been widely used. Subnational analysis can only be done for the population under age of ten, because the Census Bureau's population estimates for age of 10 and higher are linked to the previous Decennial Census.

Second, DA estimates are only available for a few race/ethnic groups. Historically the estimates have only been available for black and non-black groups. This restriction is due the lack of race specificity and consistency for data collected on the birth and death certificates historically. The only group that has been identified relatively consistently over time is blacks (African-Americans). Recent change in how the Census Bureau collects data on race raises questions about the comparability of the data for blacks in the 2010 Census relative to earlier census, but racial trends over time are not an issue in this analysis.

The 2010 DA estimates include data for Hispanics for the first time, but only for the population under age of 20. Hispanics under age of 20 were included in the DA estimates in 2010 because Hispanics have been consistently identified in birth and death certificates since 1990.

The third limitation of the DA estimates is that they only supply net undercount/overcount figures. A zero net undercount could be the result of no one being missed (omissions) or double counted (erroneous enumerations) or it could be the result of ten percent of the population being missed and ten percent counted twice.

The fourth limitation of the DA methodology is the lack of any measures of uncertainty for the estimates. However, it should be noted that, in the December 2010 DA release, the Census Bureau released five different estimate series based on five sets of assumptions about births, deaths, and net international migration to reflect some of the uncertainty regarding the DA estimates. For the population aged 0 to 17, the estimates ranged from 75,042,000 to 76,222,000, and for the population aged 0 to 4 the estimates ranged from 21,181,000 to 21,265,000. This is relatively small band of uncertainty compared to the estimated net undercount.

Despite these limitations, DA has been used for many decades, the underlying data and methodology are strong, and it has provided useful information for those trying to understand the strengths and weaknesses of the U.S.

Decennial Census. According to Robinson [34, page 1], *"The national DA estimates have become the accepted benchmark for tracking historical trends in net Census undercounts and for assessing coverage differences by age, sex, and race (Black, all other)."*

DA is particularly useful for assessing the accuracy of the Decennial Census count of young children for two reasons. First, one of the major uncertainties in using DA to assess the accuracy of total population counts is the assumptions about net international migration that must be made. For most age groups, net international migration is subject to more error because of the greater uncertainty of some specific elements such as undocumented immigrants and emigration. According to Bhaskar et al. [35, page 1], *"The largest uncertainty in the Demographic Analysis (DA) estimates comes from the international migration component."*

Given the highly reliable vital statistics data, the net internal migration component is the weakest link in the DA methodology. The DA estimates released in May 2012 assume a net international migration of only 244,000 out of a population of 21,172,000 for age of 0 to 4 (the 244,000 figure was obtained from Census Bureau staff). Net international migration accounts for only 1.1 percent of the DA estimate for the population aged 0 to 4. Therefore, errors in this component of population change would not have a big impact on the final DA population estimate for the 0 to 4 age group.

In preparing for the December 2010 DA release the Census Bureau developed five estimation series with differing assumptions to reflect uncertainty. In those five series the net international migration assumptions for the population aged 0 to 4 varied from 214,000 to 297,000 [36]. This provides some guidance about the size of potential errors in immigration estimates used in DA. The net international migration estimate for the population aged 0 to 4 in the December 2010 middle series DA estimates was 240,000 which is 30,000 more than the lowest estimate in the five series. If the net international migration component for children aged 0 to 4 was overestimated by 30,000, the net undercount of children aged 0 to 4 would be 4.4 percent instead of the observed value of 4.6 percent and the net undercount of young children would still be much higher than any other age group.

For older children, net international migration plays a bigger role. For the population aged 14 to 17, the net international migration assumptions for the five DA series released in December 2010 range from 1.023 million for the low series to 1.424 million in the high series and compose 6.1 percent and 8.3 percent of the DA estimate, respectively. The net international migration figures presented in the December 2010 DA release are not broken out by race or Hispanic origin status.

The point here is that errors in the estimate of net international migration are not likely to have a significant impact on the DA estimates of children, particularly, young children.

The second reason DA is the preferred method for assessing the net undercount of young children is improved quality of vital events data. Improvement in the birth certificate data over time is a major reason the Census Bureau is now producing DA estimates for Hispanic under age 20. In the five

DA scenarios provided in the 2010 DA estimates released in December 2010, the birth and death assumptions are identical for people under age of 18 in all five series, which reflects the high level of reliability and credibility given to the vital events data.

The other major source of data on undercounts and overcounts is the Census Bureau's Dual Systems Estimates (DSE) methodology. The DSE uses the results of from a Post-Enumeration Survey (PES) to develop an estimate of the "true" population which is then compared to the Census counts. The DSE approach for 2010 is called Census Coverage Measurement but DSE has been given other names in previous censuses. The 2010 Decennial Census is the first one where DSE has produced data for the population aged 0 to 4.

In the context of comparing the results of DSE and DA in the 2000 Census, and noting the generally consistent results, the U.S. Census Bureau [37, page v] concludes,

"The primary exception to the consistency of results occurs for children aged 0–9. While the A.C.E. Revision II estimates a small net overcount for children 0–9 (the estimate was not statistically significantly different from zero), Demographic Analysis estimated a net undercount of 2.56 percent. 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."

A National Research Council report [38, page 254] made the same observation about the inconsistency of DA and DSE estimates for young children and the authors note, "No explanation for this discrepancy has been advanced."

The issue of correlation bias in the DSE approach has been discussed by other researchers [39, 40]. Correlation bias refers to the situation where groups that are likely to be missed in the Decennial Census count are also likely to be missed in the Post-Enumeration Survey used in the DSE methodology. The existence of correlation bias in the DSE method is already recognized for the adult black male population. Currently adjustments in the DSE estimates for black males are made to correct for correlation bias based on sex ratios obtained from DA, but no adjustments are made for young children, in part, because there is not a widely accepted method for doing so. O'Hare et al. [41] document the inconsistency between DSE and DA estimates for young children and suggest that uncorrected correlation bias may result in an underestimation of the undercount for young children in the DSE methodology. The existence of correlation bias violates the assumption of independence that is required for the DSE methodology to work properly and correlation bias results in underestimating net undercounts.

In the absence of any other explanation for the large difference in net undercount estimates for young children between the DA method and the DSE method, uncorrected correlation bias in the DSE is the leading explanation for the observed differences.

In the analysis shown here, I rely exclusively on DA estimates. I believe the strengths of DA methodology make

it a particularly good technique for estimating the number of children. The data and method for DA estimates of children are very strong. Moreover, in the decade prior to the 2010 Census, staff at the Census Bureau investigated a number of issues related to the production of DA estimates [42]. The increased input, review, and examination enhance the likelihood that the 2010 DA estimates are accurate and credible. The inconsistency between DSE and DA estimates for young children and the potential correlation bias in the DSE raises questions about the accuracy of DSE estimates for young children.

3. Using DA to Estimate the Black Population

As stated earlier, historically black is the only race assessed using DA because black is the only racial category where data have been collected relatively consistently in the birth and death certificates over time, although recent changes in how the Census Bureau collects data on race make comparisons between the results of the 2010 Census and earlier Censuses precarious.

The Census Bureau faces a set of challenges in producing DA estimates for the black population. In discussing the use of vital statistics for DA estimates by race the Census Bureau [43, page 4] concludes, "...developing the estimates for DA race categories comes with a more complex, and substantial set of challenges."

The "some other race" category is a response category for the race question in the 2010 Decennial Census but not in birth or death certificates. Because the birth certificate data does not have a "some other race" category, the Census Bureau constructed a set of modified race categories from the 2010 Decennial Census in which respondents in the "some other race" category are distributed to black and non-black categories. Thus for making comparisons between DA estimates and the 2010 Decennial Census counts for blacks and non-blacks, one must use the 2010 Decennial Census modified race tabulations available on the Census Bureau's website. Assigning people from the "some other race" category to non-black, black alone, or the black alone or in combination categories provides an opportunity for error.

For some groups, the modified race tabulations are substantially different from the unmodified tabulations. For the population aged 0 to 4, the unmodified 2010 Decennial Census count for black alone was 2,903,000 but the figure based on modified race concept was 3,055,000 which amounts to a 5 percent difference. For age of 0 to 4, the unmodified count of black alone or in combination was 3,538,000 but it was 3,905,000 on the modified file, which amounts to a 10 percent difference.

A second issue is the fact that Decennial Census respondents in 2000 and 2010 could mark more than one race. Prior to the 2000 Census, respondents were only allowed to mark one race in the Decennial Census, which meant the race data from the Decennial Census and from vital events records were consistent in this regard. In 1997, the U.S. Office of Management and Budget [44] updated Statistical Policy

Directive 15 requiring federal data collection efforts to allow respondents to mark more than one race.

This issue is further complicated by the fact that it was not until 2003 that the federal government issued new standard birth certificate and death certificate forms allowing respondents to mark more than one race. Moreover, birth and death certificate data are collected by states and states only changed to the new form slowly over time. As of 2010, there were only 33 states that were using the new birth certificates, which allows respondents to mark more than one race.

To employ the DA methodology, the “more than one race” data from the birth (and death) certificates had to be put into black and non-black categories, based on both single-race and multiple-race reported by mother and fathers. NCHS provided the Census Bureau with both the multiple races that are reported and the multiple-race response “bridged” to the pre-1997 OMB single-race categories. Details about the bridging method are provided by NCHS on their website (http://www.cdc.gov/nchs/nvss/bridged_race.htm).

In addition, for the DA release of May 2012, DA estimates were provided for “black alone” as well as “black alone or in combination” so birth certificate data had to be put into these two different racial categories. The DA release in December 2010 only provided data for black alone.

A third issue is that birth certificate forms only record the race of the mother and father. Thus, the race of the child must be inferred from the race of the parent(s) in contrast to the Census where a child is put in a race category by the persons filling out the census questionnaire, which is typically the child's parent. This is further complicated by a significant level of missing data. While data on the race of mother is relatively complete, many birth certificates are missing the race of the father. In 2009, 19 percent of birth certificate forms did not contain the race of the father [45].

When both parents report the same race, that is the race assigned to the child. When the two parents report different races on the birth certificate, newborns are assigned to a race category based on the reported race of their mother and father and on parent-child race relationships seen in the Decennial Census data [46].

This is also an issue for Hispanic newborns who have a Hispanic and non-Hispanic parent and a similar approach is used. Assignment of race on death certificates is also a potential problem but deaths contribute very little to the DA estimates for children [47].

Mixed race parentage is a bigger statistical issue for young children than older people because increased rates of intermarriage over time mean more children today are likely to have parents with different races and Hispanic status. One study found that about 15 percent of marriages in 2010 involved spouses of different race or ethnicity compared to 7 percent in 1980 [48].

It is not difficult to imagine that parents of mixed racial background might report the race of a child differently on a Census questionnaire than estimated by the Census Bureau from the race of mother and father on the birth certificate form.

Given the issues described above, one should view DA estimates for blacks (alone or alone or in combination)

cautiously. Small differences in estimates could be due to methodological issues rather than real differences.

4. Data Sources

For the analysis presented in this paper, I used the figures from the revised DA estimates issued by the U.S. Census Bureau in May 2012 for all groups except Hispanics. In May 2012 the Census Bureau issued revised Demographic Analysis estimates, for the total population, the black alone population, the black alone or in combination population, the not black alone population, and the not black alone or in combination population [49]. Because no Hispanic DA estimates were provided in the May 2012 release, data for Hispanics used in this analysis are taken from the Middle Series of Census Bureau's DA estimates issue in December 2010.

As in the past, the DA program for 2010 produced estimates by age, sex, and race. However, the 2010 DA analysis included three new facets. For the first time, the Census Bureau provided DA estimates for the Hispanic population under age of 20. Secondly, in order to reflect some of the uncertainty in the DA estimates, the Census Bureau produced five sets of estimates based on different assumptions about vital events and net international migration for the DA estimates released in December 2010 [50]. It should be noted, however, that the revised DA estimates issued in May 2012 were only for the middle series and did not include Hispanics. Third, for the first time the Census Bureau published DA estimates of the black alone and the black alone or in combination populations for those under age of 30.

In the remainder of this study, the differences between the Decennial Census counts and DA estimates are shown as the Decennial Census count *minus* the DA estimate. This calculation is often labeled “net census coverage error” in other research. A negative number implies a net undercount and a positive number implies a net overcount which is consistent with the presentation of 2010 DA analysis by Velkoff [9]. This may be a point of confusion because in some past studies a similar measure called net undercount rate which subtracts the Decennial Census counts from the DA (or DSE) estimates has been used. In that construction, a negative figure implies an overcount. I chose to use the net decennial census coverage error because I feel having an undercount reflected by a negative number is more intuitive.

In converting the differences between Decennial Census counts and DA estimates to percentages, the difference is divided by the DA estimate. Estimates are shown rounded to the nearest thousand for readability.

5. Brief Historical Overview

Before examining the 2010 Census results it might be useful to look at historical trends in the net undercount of children and adults in the U.S. Decennial Census. Figure 1 shows net undercount rates for the adult population (aged 18 and over), the total child population (aged 0 to 17), and the young child population (aged 0 to 4) for each U.S. Decennial

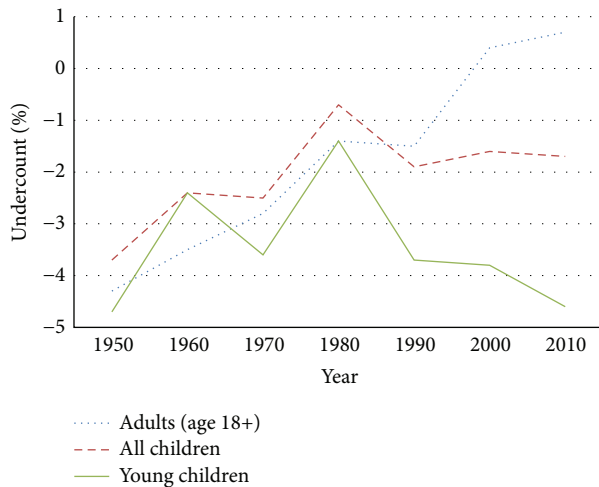


FIGURE 1: Census coverage rates for young children, all children, and adults: 1950 to 2010.

Census from 1950 to 2010. There have been two very distinct periods between 1950 and 2010 in terms of the net undercount trends of adults and children. Between 1950 and 1980, the net undercount rates for all groups generally declined steadily and the differences between the net undercount rate of children and adults were not large. Particularly, the net undercount of the adult population went from 4.3 percent in 1950 to 1.0 percent in 1980. For all children the net undercount rate fell from 3.5 percent to 0.7 percent over the same period and for young children the net undercount rate went from 4.7 percent in 1950 to 1.4 percent in 1980.

However, following 1980, the trends diverged, specially between adults and young children. Between 1980 and 2010 the net undercount rate for adults continued to fall to the point that there were small net overcounts in 2000 and 2010. The net undercount rates for young children fell from 1.4 percent in 1980 to 4.6 percent in 2010. The net undercount rate for all children remained relatively steady between 1980 and 2010.

Another perspective regarding Figure 1 is that the net undercount rate for young children in 1980 is an outlier [51]. If one disregards the net undercount rate for young children in 1980, it appears that net undercount rate for young children started high and remained high over the 1950 to 2010 period, and the change between 1980 and 2010 is minor.

In any case, the net undercount rate for young children has been higher than that for adults in nearly every U.S. Decennial Census since 1950 and the gap between young children and adults is very large in 2010.

6. 2010 Demographic Analysis Results by Age

In the 2010 Decennial Census there was a net overcount of 0.1 percent of the total population based on DA, which translates into 400,000 people. However, the small net overcount for the total population masks important differences among some age groups. Velkoff [9] shows the 0.1 percent net overcount for the entire population masks a 0.7 percent net overcount for

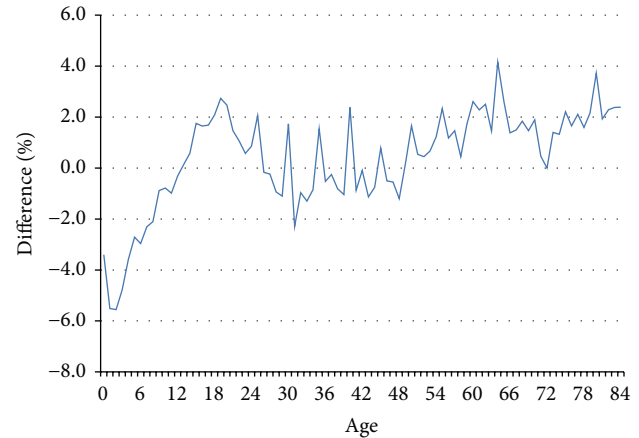


FIGURE 2: Percent difference between 2010 Census counts and DA estimates by single year of age: 0 to 84.

adults (aged 18 and older) and a 1.7 percent net undercount for children (aged 0 to 17). In population numbers, these reflect a net undercount of 1.3 million children and a net overcount of 1.7 million adults.

This underscores the extent to which the small difference between the 2010 Decennial Census and the DA estimates for the total population conceals important differences by age. This theme is repeated when examining data for children by age.

Figure 2 shows the net undercount and overcount figures from the 2010 Decennial Census by single year of age for age of 0 to 84. The age-specific estimates from DA closely match the Decennial Census counts with the exception of three age groups. There is a large net undercount for people under age of 10, particularly aged 0 to 4, a large net overcount for young adults (roughly aged 18 to 24), and a large net overcount for the population aged 60 to 80. Figure 2 also shows the effect of “age heaping” where people prefer reporting their age in figures ending with “0” or “5.”

Figure 2 shows that young children not only have a higher net undercount than any other age group, but also have a higher level of coverage error than any other age group regardless of the direction of the error.

The net overcount of 18- to 24-year-olds is widely believed to be due to the fact that many people in this age group are counted in the home of their parents as well as where they reside most of the time, for example, in a college dormitory. The net overcount of 60- to 80-year-olds may be due to the large portion of this population who have second homes and are counted in both places. On the other hand, there is no similar commonly accepted explanation for the high net undercount of young children.

As Figure 2 suggests, there are big differences in net undercounts and overcounts for children based on their age. Figure 3 shows the net undercount rates for children by single year of age and underscores the extent to which undercount rates for children vary by age.

There are three key points that can be derived from Figure 3. First, the highest net undercount rates are found

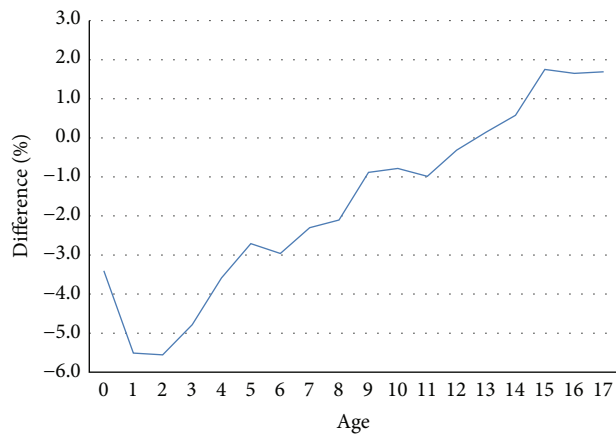


FIGURE 3: Percent difference between 2010 Census counts and DA estimates by single year of age: 0–17.

among the youngest children, particularly aged 0 to 4. More than three-quarters of the 1.3 million person net undercount for the population aged 0 to 17 can be accounted for by those aged 0 to 4, where the net undercount is about one million people.

Second, there is a net overcount rate for people aged 14 to 17, and all of the net overcount of children aged 14 to 17 is accounted for by a high net overcount of Hispanics and blacks in this age range (see Figure 5). The DA figures released in December 2010 show an estimated net overcount of 183,000 Hispanics aged 14 to 17, or 5.4 percent difference. It is easy to imagine that at least some of this difference is due to errors in estimates of net international migration. Unfortunately, neither the May 2010 nor the December 2010 DA release provide migration assumptions by race or Hispanic origin. However, it seems unlikely that net international migration would have a big impact on the DA results released May 2012 which showed that black alone or in combination aged 14 to 17 was 84,000 more than the Census count which amounts to a 2.9 percent difference.

Third, there is a very clear age gradient along the age range from age of 1 to 17. The net undercount rate declines steadily from age of 1 to age of 13 and there is a net overcount in the 14- to 17-year-old age group. The correlation coefficient between age and net undercount rate for the population aged 0 to 17 is -0.98 . I am not aware of any theories about why this relationship exists.

Interestingly, the net undercount rate of those aged 0 is much lower than the rate for those in the age of 1 or 2. There was a new instruction added to the 2010 Census questionnaire reminding respondents to include newborns, which may explain this anomaly.

7. Single Year of Age by Sex, Race, and Hispanic Origin

Several major demographic groups show the same age gradient seen for the total child population. Data in Figure 4, which shows net undercount estimates by single year of

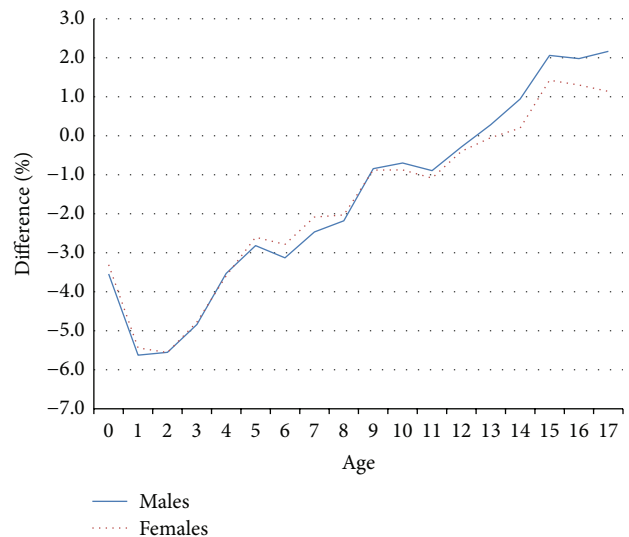


FIGURE 4: Percent difference between 2010 Census counts and DA estimates by sex by single year of age: 0–17.

age for males and females, indicates there are virtually no differences in net undercount rates between males and females at the youngest ages. When children enter the middle teens, however, the net overcount rate of males becomes noticeably higher than that of females which is similar to the pattern among young adults. I suspect the higher overcount of males in the 14- to 17-year-old age group may also be related to undetected net immigration from abroad, where males typically outnumber females, but this deserves further research.

Figure 5 shows net undercount rates for black alone or in combination, Hispanic, and a group labeled “not black alone or in combination and not Hispanic” children by single year of age based on the 2010 DA estimates series. The not black alone or in combination and not Hispanic group shown in Figure 5 is a derived group and not one that is used by the Census Bureau. The not black alone or in combination and not Hispanic figures are derived by subtracting the number in the black alone or in combination category and the number in the Hispanic category from the total number for each age group. Since there is no figure for the non-Hispanic white population to compare with the minority populations to gauge the racial/Hispanic differentials, the not black alone or in combination and not Hispanic category is used as a proxy population for the non-Hispanic white population.

Data from the 2010 Decennial Census shows that non-Hispanic whites are 90 percent of the not black alone or in combination and not Hispanic population aged 0 to 17 and 92 percent of the not black alone or in combination and not Hispanic population aged 0 to 4.

There is a minor problem with double counting black Hispanics in this methodology. The 2010 Decennial Census shows that there were 285,000 children aged 0 to 4 who were both black alone or in combination and Hispanic, which translates into 1.4 percent of the total population aged 0 to 4 in 2010. There were 838,000 children aged 0 to 17 who were

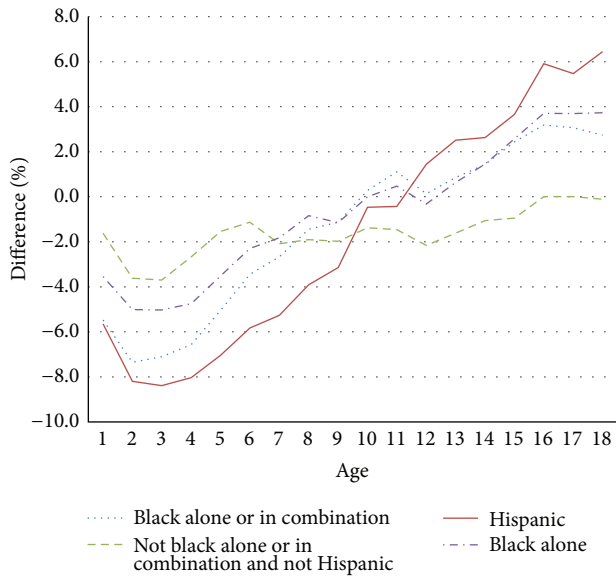


FIGURE 5: Difference between 2010 Census count and DA estimates for race and Hispanic groups by single year of age: 0–17.

both black alone or in combination and Hispanics, which translates into 1.1 percent of the total population aged 0 to 17. While the black alone or in combination figures and the Hispanic figures include a small number of people who are included in both groups, the not black and not Hispanic population provides a much better comparison group than the alternatives of total, non-black, or non-Hispanic.

The age gradient reflected in Figure 3 for all children is seen consistently among all the race/Hispanic groups examined here. While the exact levels and age breaks differ slightly, the same pattern is seen for black alone, black alone or in combination, Hispanic, and not black alone or in combination and not Hispanic children. As one moves up in age, the net undercount decreases and turns into a net overcount in the early teen years; this is followed by increasing overcounts in later years of adolescence.

The age gradient is steeper for blacks and Hispanics than for others. While racial minorities account for a disproportionately large share of the net undercount of children aged 0 to 4, they experienced higher net overcounts in the 14- to 17-year-old age group. In fact, blacks and Hispanics account for all of the net overcount for the population aged 14 to 17.

The net undercount rates for “black alone” and “black alone or in combination” are similar for ages of 10 to 17 but differ substantially for the population below age of 10 and particularly for ages of 0 to 4. The net undercount rate for population black alone aged 0 to 4 is 4.4 percent but it is 6.3 percent for black alone or in combination in this age group (see Table 1). Of course, the only difference between these two series is the “black in combination” population. The number of blacks in combination estimated from the birth certificate data used in the DA estimates is much higher than the number reported in the census for younger children. For the population aged 0 to 4, the net undercount rate for blacks in combination is about 15 percent.

The high net undercount rate of young black and Hispanic children relative to others is consistent with much of the past research on Decennial Census undercount differentials which shows racial minorities often have higher net undercount rates than others. Historically the comparative paradigm used most often has been the comparison of black and non-black populations [52, 53]. But the non-black population has changed over the past few decades in ways that make the comparisons between black and non-black less meaningful than in the past. Fifty years ago, the vast majority of the non-black population was non-Hispanic whites, but a growing share of the non-black population is now Hispanic.

8. The Population Aged 0 to 4

Since children aged 0 to 4 have the highest net undercount rate of any age group in the 2010 Decennial Census, the remainder of this section will focus on that age group.

Table 1 shows the 2010 Decennial Census net undercount rates and numbers for population aged 0 to 4 by various demographic characteristics. Overall, there was a net undercount rate of 4.6 percent and a net undercount of about 970,000 people aged 0 to 4. All groups had a relatively high net undercount rate, but there are substantial differences among groups.

The net undercount rates for males and females are nearly identical at 4.6 percent and 4.5 percent, respectively. There are significant differences for undercount and overcount rates of young adults by sex, which is typically attributed to differences in living arrangements. Given the similarity of living arrangements for young children, it is not surprising that there is little difference by sex for young children.

The net undercount rate for all people aged 0 to 4 was 4.6 percent, but the rate for black alone or in combination children in this age range was 6.3 percent, the rate for black alone was 4.4 percent, and the rate for Hispanic children in this age range was 7.5 percent. For the not black alone or in combination and not Hispanic category, which is a proxy group for non-Hispanic whites, the net undercount rate for age of 0 to 4 was only 2.7 percent.

The fact that the racial/Hispanic minorities have higher net undercount rates than the proxy population for non-Hispanic whites (not black alone or in combination and not Hispanic) is consistent with past research. The fact that young Hispanics have a higher net undercount rate than any other group is not surprising when one examines the extent to which Hispanic households have many of the hard-to-count characteristics identified by the Census Bureau such as being highly mobile, disproportionately in rental housing, and more likely to experience language barriers.

The difference between the net undercount rate for black alone (4.4 percent) and black alone or in combination (6.3 percent) raises several questions. For example, it poses problems for those wondering which of the two figures to use. It may be worth noting that, in a civil rights context, guidance from the U.S. Office of Management and Budget [54] suggests the figure for black alone or in combination is the more appropriate measure to use. It is also worth noting

TABLE 1: Difference between 2010 Decennial Census counts and Middle Series DA estimates for the population aged 0 to 4, by sex, race, and Hispanic origin.

	2010 Decennial Census count (in 1000s)	2010 DA estimate (in 1000s)	Numeric difference between Decennial Census count and DA estimate	Percent difference between Decennial Census count and DA estimate
Total	20,201	21,171	-970	-4.6
Female	9,882	10,353	-471	-4.5
Male	10,320	10,821	-501	-4.6
Black alone	3,055	3,195	-140	-4.4
Black alone or in combination	3,658	3,905	-247	-6.3
Not black alone	17,146	17,976	-830	-4.6
Not black alone or in combination	16,544	17,268	-724	-4.2
Hispanic	5,114	5,528	-414	-7.5
Not Hispanic	15,087	15,643	-556	-3.6
Not black alone and not Hispanic	12,032	12,448	-416	-3.3
Not black alone or in combination and not Hispanic	11,429	11,738	-309	-2.7

Note. The “not black alone or in combination and not Hispanic” category and the “not black alone and not Hispanic” category are not categories used by the Census Bureau. Please see the text for an explanation of how the figures for these categories were derived. DA estimates by Hispanic origin are only available for the December 2010 DA release. Responses of “some other race” from the 2010 Census are modified. This results in differences between the populations for specific race categories shown for the 2010 Census population in this table and those in the original 2010 Census data. For more information, see <<http://www.census.gov/popest/data/historical/files/MRSF-01-US1.pdf>>.

Source: [30].

that both black alone and black alone or in combination net undercount rates are considerably higher than the rate for the proxy non-Hispanic white population.

The difference between black alone and black alone or in combination demonstrates the difficulty of accurately estimating detailed race categories from birth certificate information about race of parents that match the way race is reported for children in the Decennial Census. Recall that new birth certificates were introduced in 2003 which allowed parents to mark more than one race for the first time. The switch to the more than one race option for parents on the birth certificates starting in 2003 corresponds to the increase in the number of newborns classified as “black in combination” which is most heavily reflected in the population of 0 to 4 in 2010. Given current demographic trends in the U.S. which indicate increasing levels of intermarriage and multiracial births, the ability to code birth certificate data to match Census responses will be important in the future use of the DA methodology.

There was a net undercount of slightly less than one million people aged 0 to 4 in the 2010 Decennial Census including a net undercount of 247,000 blacks alone or in combination and a net undercount of 414,000 Hispanics. The combination of young black alone or in combination and young Hispanic children accounts for about two-thirds of the total net undercount in this age group even though they only account for about 40 percent of the population in this age range. Moreover, the net undercount of black alone or in combination and Hispanic children aged 0 to

4 (about 671,000) accounts for more than half of the total net undercount of all people under age of 18 (1.3 million) even though this group comprises only 11 percent of that population aged 0 to 17.

It should be noted that only DA data available for Hispanic is from the December 2010 release, but this estimate of the undercount for Hispanic aged 0 to 4 may be a little high. In the 2010 DA release, the Census Bureau had to rely on estimates of births and deaths for 2008, 2009, and the first quarter of 2010. The revised estimates for the total and the black groups that were released in May 2012 benefited from the availability of actual birth and death records for 2008, 2009, and the first quarter of 2010, and the actual vital event data indicate the Census Bureau’s estimates for the number of births in 2008–2010 period used in the 2010 DA release were a little too high and thus the DA estimates for the youngest ages were a little high. However, this is likely to have only a very small impact. For the overall population the addition of corrected birth data changed the net undercount from 4.7 percent in the December 2010 DA release to 4.6 percent in the May 2012 DA release.

9. Discussion

The comparative perspective used in this study focuses primarily on differences by age and I would argue that this is an important perspective to keep in mind as we move toward the 2020 Decennial Census. While this perspective has been employed in the past in limited cases, it has not

received the same amount of attention as racial differentials. The gap between young children and adults (5.3 percentage points in 2010) is now much larger than the gap between the blacks alone population and the non-black alone population of all ages (0.5 percentage points in 2010) [9]. I believe the field should focus on differences by age as well as differences among racial/ethnic groups in future investigations of Decennial Census undercount differentials. The evidence presented here indicates that among children aged 0 to 4 there are large differentials by race and Hispanic origin status.

DA figures from the May 2012 release indicate there was a net undercount of about 146,000 black alone or in combination males aged 25 to 29 which amounts to a net undercount rate of 9.7 percent. The 25 to 29 age groups were selected here because there is no data for the black alone or in combination over age of 30. Table 1 shows the total net undercount for all children aged 0 to 4 was 970,000 and for blacks alone or in combination aged 0 to 4, the net undercount was 247,000. While the net undercount rate for young black males is higher than the net undercount rates for the population aged 0 to 4, the number of people affected is much higher for the population aged 0 to 4.

Despite the high net undercount of young children in the past U.S. Censuses the topic has received little attention in the professional literature. When census coverage data have been made available in the past for children, they are often a small part of a larger volume related to assessing the quality of the decennial census count [55–58].

It should also be noted that the population with the highest net undercount among young children, Hispanics, is growing more rapidly than any other group. In 2000, Hispanics made up 19.4 percent of the population aged 0 to 4 and by 2010 they were 25.3 percent of this age group.

It is clear that people under age of 18 should not be treated as a homogeneous group with respect to age, as has sometimes been done in the past. For example, prior to the 2010 Census, there were no estimates for the population aged 0 to 4 from the DSE method. Analyses that fail to make a distinction among age groups of children are likely to find interpretation of findings difficult. The explanation for why young children experience a high net undercount is likely to be quite different than the explanation for why teens (aged 14 to 17) have a net overcount. Moreover, combining the aged 0 to 4 population and the aged 14 to 17 population into one group masks the differences between the Decennial Census and the DA estimates in both groups.

While the high net undercount rate for young children is clear, the reasons for such a high rate are not. In fact, the conventional wisdom in survey research suggests the presence of children in a household increases response rates. Groves and Couper [59, page 138] offer this succinct summary of the relationship between children in the household and cooperation in survey research, *“Without exception, every study that has examined response or cooperation finds positive effects of the presence of children in the household.”* It should be noted that there is no distinctions made in this review about the age of the children in the household and that household response is not exactly the same as getting data on a child. A household could respond to the Census, for

example, but fail to include a young child on the Census questionnaire. The heterogeneity of net census undercount rates for children of different ages suggests that age of child should be examined more closely in terms of the impact of children of different ages on surveys response rates.

There is no commonly accepted explanation for the high net undercount of young children in the census, but a few possibilities are discussed below (for more information on this topic see [60]). It is possible that young children were concentrated in households where the whole household was not included in the Census count but it is also possible that young children lived in households that were included in the Census, but young children were left off the census questionnaire. Unfortunately, there is no good evidence about how many children are missed because they lived in a household where the entire household was missed and how many were missed in households where a census questionnaire was returned but the child was not included. Developing data which sheds light on relative importance of whole household misses compared to within household misses for young children would be very helpful in devising strategies to mitigate this problem.

It is possible that the way the Decennial Census data are collected and/or processed contributes to a high net undercount of young children. For example, the continuation form used by the Census Bureau to capture information on persons in large households may have a differential impact on young children. The 2010 Decennial Census mail-out questionnaire only contains room for complete information for six people in the household. If there are more than six people in the household, the Census Bureau must follow up to get complete information for the members of the household. The 2010 American Community Survey shows that 10.1 percent of young children live in such large households, compared to 3.5 percent of adults. Therefore, any problem following up with these types of households would affect young children disproportionately.

It is possible that the way the Census Bureau imputes age to cases where age is not provided or the data provided is implausible may result in an underestimation of young children. Perhaps, too many people had their age imputed as age of 14 to 17, and too few had their age imputed as age of 0 to 4. However, the percent of children who had their age imputed is relatively small (between one and two percent) so it seems unlikely that this is a major factor.

Anecdotal evidence suggests that some respondents deliberately leave young children off the Census questionnaire. They may leave a young child off the form because they think the Census Bureau is not interested in including young children, or they may leave a young child off the form because they do not want the government to know there is a child living in the household unit. In a small test sample, Nichols and colleagues [61] found respondents thought the Census Bureau was less interested in collecting the names of children than the names of adults. In checking the consistency of Census reports, ethnographic research related to the 2010 Decennial Census [62, page viii] found: *“the 0–4 age cohort had the greatest proportion of inconsistency among all age cohorts.”*

It is also plausible that the living arrangements and/or locations of young children may be the driving force behind their high net undercount rates. The Census Bureau identified several factors that are related to hard-to-count populations and children are overrepresented in most of those categories [63].

Gaining a better understanding of why young children have such a high net undercount will be important in terms of reducing such an undercount in the 2020 Decennial Census.

10. Conclusions

The net undercount of people aged 0 to 4 in the 2010 Decennial Census is higher than any other age group, but this is not a new problem. A passage from the 1940 U.S. Decennial Census [64, page 32] reads, “Underenumeration of children under 5 year old, particularly of infants under one year old, has been uniformly observed in the United States Decennial Census and in the Censuses of England and Wales and of various countries of continental Europe.” This observation from almost 70 years ago is still largely true today. The net undercount rate for the population aged 0 to 4 in the 2010 Decennial Census is almost identical to the net undercount rate experienced by this age group in the 1950 Census.

The high net undercount rate for young children is driven primarily by the high net undercount of young minority (blacks alone or in combination and Hispanic) children. The net undercount rates for young black alone or in combination and Hispanic children are more than twice as like as those for not black and not Hispanic children aged 0 to 4.

Given the changing demographics of the nation, racial and Hispanic minorities are destined to be an increasing share of the total population. Increasing our collective ability to count such groups accurately will become bigger factor in the overall accuracy of future Decennial Census counts. Moreover, as such minority groups grow in number and political power, demands for developing accurate assessments of overcount and undercounts will grow.

The results of this study suggest that efforts to minimize overall net undercounts in future censuses should focus on families, households, and neighborhoods with high concentrations of black or Hispanic children aged 0 to 4.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] W. P. O'Hare, *The Overlooked Undercount: Children Missed in the Decennial Census*, The Annie. E. Casey Foundation, Baltimore, Md, USA, 1999.
- [2] W. P. O'Hare, *Why Are Young Children Missed So Often in the Census*, KIDS COUNT Working Paper, The Annie E. Casey Foundation, Baltimore, Md, USA, 2009.
- [3] W. P. O'Hare, “The net undercount of children in the 2010 decennial census,” in *Proceedings of the Applied Demography Conference*, University of Texas, San Antonio, Tex, USA, January 2012.
- [4] K. K. West and J. G. Robinson, “What do we know about the undercount or children?” Population Division Working Paper, U.S. Decennial Census Bureau, 1999.
- [5] D. Goodkind, “Child underreporting, fertility, and sex ratio imbalance in China,” *Demography*, vol. 48, no. 1, pp. 291–316, 2011.
- [6] B. A. Anderson and B. D. Silver, “Estimating census undercount from school enrollment data: an application to the Soviet censuses of 1959 and 1970,” *Demography*, vol. 22, no. 2, pp. 289–308, 1985.
- [7] B. A. Anderson, “Undercount in China's 2000 decennial census in comparative perspective,” PSC Research Report, No. 04-565, Population Studies Center, University of Michigan, Ann Arbor, Mich, USA, 2004.
- [8] W. P. O'Hare, E. Jensen, and B. C. O'Hare, “Assessing the undercount of young children in the U.S. decennial census: implications for survey research and potential explanations,” in *Proceedings of the American Association of Public Opinion Research Conference*, Boston, Mass, USA, May 2013.
- [9] V. Velkoff, “Demographic evaluation of the 2010 census,” in *Proceedings of the PAA Annual Conference*, Washington, DC, USA, March 2011.
- [10] H. Hogan, P. J. Cantwell, J. Devine, V. T. Mule Jr., and V. Velkoff, “Quality and the 2010 Census,” *Population Research and Policy Review*, vol. 32, no. 5, pp. 637–662, 2013.
- [11] V. T. Mule, *2010 Census Coverage Measurement Estimates Report: Summary Estimates of Coverage of Persons in the United States*, DSSD 2010 Census Coverage Measurement Memorandum Series 2010-G-01, U.S. Census Bureau, Washington, DC, USA, 2012.
- [12] J. G. Robinson, “Coverage of population in decennial census 2000 based on demographic analysis: the history behind the numbers,” Working Paper No. 91, U.S. Census Bureau, 2010.
- [13] C. L. Himes and C. C. Clogg, “An overview of demographic analysis as a method for evaluating census coverage in the United States,” *Population index*, vol. 58, no. 4, pp. 587–607, 1992.
- [14] U.S. Decennial Census Bureau, *The Development and Sensitivity Analysis of the 2010 Demographic Analysis Estimates*, Population Division, 2010.
- [15] D. O. Price, “A check on the underenumeration in the 1940 census,” *The American Sociological Review*, vol. 12, pp. 44–49, 1947.
- [16] A. J. Coale, “The population of the United States in 1950 classified by age, sex and color—a revision of decennial census figures,” *Journal of the American Statistical Association*, vol. 50, pp. 16–54, 1955.
- [17] J. S. Siegel and M. Zelnik, “An evaluation of coverage in the 1960 decennial census of population by techniques of demographic analysis and by composite methods,” in *Proceedings of the Social Statistics Section of the American Statistical Association*, pp. 71–85, American Statistical Association, Washington, DC, USA, 1966.
- [18] A. J. Coale and M. Zelnick, *New Estimates of Fertility and Population in the United States*, Princeton University Press, Princeton, NJ, USA, 1963.
- [19] A. J. Coale and W. Rives, “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*, vol. 39, no. 1, pp. 3–36, 1973.

- [20] J. D. Hacker, "New estimates of census coverage in the United States, 1850–1930," *Social Science History*, vol. 37, no. 1, pp. 71–101, 2012.
- [21] J. W. Adams and A. B. Kasakoff, "Estimates of decennial census underenumeration based on genealogies," *Social Science History*, vol. 15, no. 4, pp. 527–543, 1991.
- [22] K. K. West, "Using medicare enrollment file for the DA 2010 estimates," in *Proceedings of the Applied Demography Conference*, University of Texas at San Antonio, San Antonio, Tex, USA, January 2010.
- [23] J. G. Robinson, "Coverage of population in decennial census 2000 based on demographic analysis: the history behind the numbers," Working Paper no. 91, U.S. Census Bureau, 2010.
- [24] J. Devine, L. Sink, B. DeSalvo, and R. Corte, "The use of vital statistics in the 2010 demographic analysis estimates," Census Bureau Working Paper No. 88, 2010.
- [25] R. Bhaskar, M. Scopilliti, F. Hollman, and D. Armstrong, "Plans for producing estimates of net international migration for the 2010 demographic analysis estimates," Census Bureau Working Paper 90, 2010.
- [26] U.S. Census Bureau, Release of Demographic Analysis, 2010.
- [27] Demographic Analysis Research Team, "Estimates of net international migration in demographic analysis," in *Proceedings of the Demographic Analysis Conference*, Population Division, U.S. Census Bureau, Washington, DC, USA, December 2010.
- [28] J. Schachter, "Estimating native emigration from the United States," Memorandum, U.S. Census Bureau, 2008.
- [29] J. Pitkin and J. Park, "The gap between births and decennial census counts of children born in California: undercount or transnational movement?" in *Proceedings of the Population Association of America Conference*, Philadelphia, Pa, USA, March 2005.
- [30] "U.S. Census Bureau, 2010 Demographic Analysis," December 2010–May 2012.
- [31] Y. Mayol-Garcia and J. G. Robinson, "Census 2010 counts compared to the 2010 population estimates by demographic characteristics," in *Proceedings of the Southern Demographic Association Conference*, Tallahassee, Fla, USA, October 2011.
- [32] J. G. Robinson, A. Bashir, and E. W. Fernandez, "Demographic analysis as an expanded program for early coverage evaluation of the 2000 census," in *Proceedings of the Census Bureau Annual Research Conference*, Arlington, Va, USA, March 1993.
- [33] A. L. Adlakha, J. G. Robinson, K. K. West, and A. Bruce, "Assessment of consistency of decennial census data with demographic benchmarks at the subnational level," Decennial Census 2000 Evaluation 0.20 Final Report, 2003.
- [34] J. G. Robinson, *Accuracy and Coverage Evaluation: Demographic Analysis Results*, DSSD Decennial Census 2000 Procedures for operations Memorandum Series B-4, U.S. Decennial Census Bureau, 2000.
- [35] R. Bhaskar, M. Scopilliti, F. Hollman, and D. Armstrong, "Estimates of net international migration for the 2010 demographic analysis estimates," in *Paper Prepared for U.S. Census Bureau Workshop: 2010 Demographic Analysis Technical Review*, p. 1, Suitland Md, USA, January 2009.
- [36] U. S. Census Bureau, Tables released at December 2010 Conference Table 8, 2010.
- [37] U.S. Census Bureau, *Technical Assessment of A.C.E. Revision II*, U.S. Decennial Census Bureau, 2003.
- [38] National Research Council, "The 2000 census: counting under adversity," in *Panel to Review the 2000 Census*, C. F. Citro, D. L. Cork, and J. L. Norwood, Eds., p. 254, Committee on National Statistics, Division of Behavioral and Social Science and Education, The National Academy Press, Washington, DC, USA, 2004.
- [39] K. W. Wachter and D. A. Freedman, "The fifth cell: correlation bias in U.S. census adjustment," Tech. Rep. no. 570, Department of Statistics, University of California, Berkeley, Calif, USA, 1999.
- [40] R. Shores, "Accuracy and coverage evaluation revision II adjustment for correlation bias," DSSD A.C.E. Revision II Memorandum Series PP-53, U.S. Bureau of the Census, 2002.
- [41] W. P. O'Hare, J. G. Robinson, K. K. West, and V. T. Mule, "Comparing demographic analysis and dual-systems estimates results for children," in *Proceedings of the Southern Demographic Association Conference*, Williamsburg, Va, USA, October 2012.
- [42] J. G. Robinson, "Coverage of population in decennial census 2000 based on demographic analysis: the history behind the numbers," Working Paper no. 91, US Census Bureau, The History Behind the Numbers, 2010.
- [43] J. Devine, L. Sink, B. DeSalvo, and R. Cortes, "The use of vital statistics in the 2010 demographic analysis estimates," *Census Bureau Working Paper no.*, vol. 88, 2010.
- [44] U.S. Office of Management and Budget, "Revisions to the standards for the classification of federal data on race and ethnicity," Statistical Policy Directive 15, Federal Register Notice, October 1997.
- [45] J. A. Martin, B. E. Hamilton, and S. J. Ventura, *Births: Final Data for 2009, National Vital Statistics System*, vol. 60, no. 1, National Center for Health Statistics, Hyattsville, Md, USA, 2009.
- [46] J. Ortman, F. W. Hollmann, and C. E. Guarneri, "Estimating multiple-race births for the 2010 demographic analysis estimates," in *Proceedings of the Applied Demography Conference*, San Antonio, Tex, USA, January 2012.
- [47] E. Aries, W. S. Schauman, K. Eschbach, P. D. Sorlie, and E. Backlund, "The Validity of race and Hispanic Origin reporting on death certificates in the United States," *National Center for Health Statistics, Vital and Health Statistics*, vol. 2, no. 148, 2008.
- [48] W. Wang, *The Rise of Intermarriage*, Pew Research Center's Social & Demographic Trends, 2012.
- [49] U. S. Census Bureau, *Documentation for the Revised 2010 Demographic Analysis Middle Series Estimates*, U.S. Census Bureau, Washington, DC, USA, 2012.
- [50] U.S. Census Bureau, Tables released at December 2010 Conference Table 8, December 2010.
- [51] H. Hogan, "Comments in the undercount of young children," in *Proceedings of the Joint Statistical Meetings*, Montreal, Canada, 2013.
- [52] D. J. Fein, *The social sources of decennial census omission: racial and ethnic differences in omission rates in recent U.S. censuses [Dissertation]*, Department of Sociology, Princeton University, Princeton, NJ, USA, 1989.
- [53] M. Anderson, C. Citro, and J. Salvo, Eds., *Encyclopedia of the U.S. Census: From the Constitution to the American Community Survey (ACS)*, Sage, New York, NY, USA, 2012.
- [54] U.S. Office of Management and Budget, "Guidance on aggregation and allocation of data on race for use in civil rights monitoring and enforcement," May 2001.
- [55] U.S. Census Bureau, *Estimates of Coverage of Population by Sex, Race and Age: Demographic Analysis 1970*, Decennial Census of Population and Housing, Evaluation and Research Program, 1974.

- [56] R. E. Fay, J. S. Passel, J. G. Robinson, and C. D. Cowan, "The coverage of the population in the 1980 census," Decennial Census of Population and Housing, Evaluation and Research Reports PHC80-E4, U.S. Decennial Census Bureau, Washington, DC, USA, 1988.
- [57] J. G. Robinson and A. Adlaka, *Comparison of A.C.E. Revision II Results with Demographic Analysis*, DSSD A.C.E. Revision II Estimates Memorandum Series #PP-41, US Bureau of the Census, 2002.
- [58] G. Robinson and A. Adlaka, *Comparison of A.C.E. Revision II Results with Demographic Analysis*, DSSD A.C.E. Revision II Estimates Memorandum Series #PP-41, U.S. Bureau of the Census, 2002.
- [59] R. M. Groves and M. P. Couper, *Nonresponse in Household Interview Surveys*, John Wiley & Sons, New York, NY, USA, 1998.
- [60] D. H. Griffin, *Final Task Force Report: Task Force on the Undercount of Young Children*, Memorandum for Frank A. Vitrano, U.S. Census Bureau, Washington, DC, USA, 2014.
- [61] E. Nichols, R. King, and J. Childs, "Small-Scale Testing Pilot Test Results: Testing email subject lines and address collection screens and Census opinion questions using a nonprobability panel," Internal memorandum to Burton Reist. U.S. Census Bureau, 2014.
- [62] L. Schwede and R. Terry, *Comparative Ethnographic Studies of Enumeration Methods and Coverage Across Race and Ethnic Groups*, Census Planning Memorandum Series, No. 255, U.S. Census Bureau, 2013.
- [63] W. P. O'Hare, "Potential explanations for the high net undercount of young children in the U.S. census," in *Proceedings of the Applied Demography Conference*, University of Texas, San Antonio, Tex, USA, January 2014.
- [64] U.S. Census Bureau, "Population differential fertility: 1940 and 1910; standardized fertility rates and reproduction rates, appendix A," in *Completeness of Enumeration of Children under 5 Years Old in the Decennial Census of 1940 and 1910*, 1940, p. 32, U.S. Decennial Census, Washington, DC, USA, 1944.

