

Chapter 4

The Big Picture; Fundamentals of Differential Undercounts



Abstract The 2010 Census coverage error for the total population was very small by international and historic standards, but that masks some large coverage differences among groups. Basic differences in Census coverage by age, sex, race, Hispanic Origin, and tenure are explored in this Chapter. Key findings include, young children (age 0–4) have a higher net undercount than any other age group; males have a higher net undercount than females; Blacks, American Indians on reservations, and Hispanics have relatively high net undercounts; and renters have higher net undercount rates than homeowners. This Chapter is meant to provide an overview and foundation for much of the rest of the book.

4.1 Introduction

When the results of the 2010 Decennial Census were compared to the expected population based on Demographic Analysis (DA) and Dual-Systems Estimates (DSE), the difference was near zero (Velkoff 2011; U.S. Census Bureau 2012). The 2010 Census coverage rate is very good by historical and international standards.

However, the small difference between the expected population and the 2010 Census count for the total population has caused some people to conclude there was a high level of accuracy for the 2010 Census. For example, Rebecca Blank, then Deputy U.S. Commerce Secretary (2012, p. 1) said, “Today’s Census Bureau report shows that not only was the 2010 Census delivered on time and significantly under budget- but even more important, it was extremely accurate.” A similar sentiment was expressed by Census Bureau Director Groves (2012).

The high level of accuracy for the total population, however, masks crucial differences among demographic groups and that is the point of this book. It is the differential undercounts which are important for many uses of Census data discussed in Chap. 2. If all groups and all places were undercounted at the same rate, the undercount would hardly be a problem. For one thing, we could just inflate the number based on the net undercount rate.

The fundamental problem with differential Census coverage rates is that certain groups have higher net undercount rates than others. This Chapter provides

an overview of key data on Census coverage differentials by age, sex, race, Hispanic Origin, and tenure, to give readers a broad picture of Census undercounts and overcounts. The information in this Chapter will provide readers with data on key measures of Census coverage and establish a foundation for most of the remaining Chapters in this book. All the topics covered in this Chapter will be pursued in more detail in subsequent Chapters.

To keep this Chapter to a manageable size, I look primarily at data from the 2010 Census. Wherever it is feasible, I look at both DA and DSE results. In some cases, however, I draw only on data from DA and in other cases I draw only on data from DSE because one source is much better than the other for a particular population (for example, DA is much better for measuring the net undercount of young children) or data for a particular group are only available from one source (for example, DSE is the only source of information for many race groups). To provide a succinct portrait, only net undercounts and net overcounts are examined in this Chapter, while subsequent Chapters will look at omissions as well as net undercounts.

Data from the 2010 DSE analysis (the method was called Census Coverage Measurement or CCM in the 2010 Census) includes five demographic categories; age, sex, race, Hispanic Origin status, and tenure. Data from the 2010 DA analysis include detailed data by age and sex, but the data on race and Hispanic Origin are limited. Historically the only race groups in DA were Black and Non-Black. In the 2010 Census, the DA data are available for Black Alone and Black Alone or in Combination for the population under age 30. In 2010, DA figures for Hispanics are only available for the population under age 20. For almost all groups the DA results and the DSE results are similar. The major exception to that is for young children (see Table 3.1).

The DSE estimates are based on a sample and therefore include sampling error. So, differences from zero need to be tested to see if the observed differences from zero might be due to chance or random error. Statistical significance is based largely on the size of the difference from zero and the size of the sample on which the estimate is based. Larger differences and estimates from larger samples are more likely to be statistically significant. The statistical significance level used throughout this book is the same one used by the Census Bureau, namely 0.10. This means an observed result that is statistically significant would only happen by random chance alone one time out of ten.

4.2 Census Coverage Differentials by Age

The Census coverage estimates from DA are the best source for examining age differentials for at least four reasons. First, DA data on age are more detailed than that from Dual-Systems Estimates (DSE). Data are available by single year of age from DA but only for large age/sex groups from DSE. Second, DA estimates have been produced since 1950, so there is more historical data. Third, in the decade prior to the 2010 Census, staff at the Census Bureau investigated several issues related to the production of DA estimates (Robinson 2010; Bhaskar et al. 2010; Devine et al.

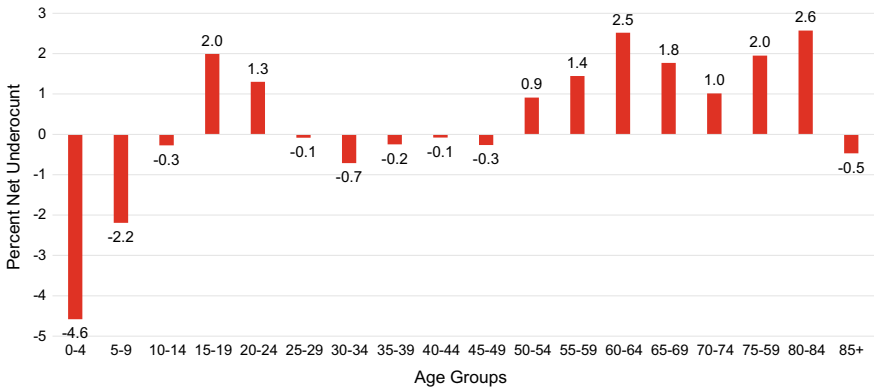


Fig. 4.1 Net undercount rates in 2010 census by five-year age groups. *Source* U.S. Census Bureau, May 2012 DA Release

2010). The increased input review and examination enhance the reliability of the 2010 DA estimates. Fourth, DSE estimates for the youngest ages greatly underestimate the net undercount (O’Hare et al. 2016a, b) so that data series cannot be used to examine the whole age spectrum.

Figure 4.1 shows the net coverage in the 2010 Census by five-year age groups based on DA. Figure 4.1 shows high net undercounts for the youngest population, those age 0–4, and somewhat smaller net undercounts for age 5–9. Young children had a higher net undercount than any other age group in the 2010 Census. The net undercount rate for age 0–4 is 4.6% and for age 5–9 it is 2.2%. No other age group has a net undercount rate of more than 0.7%. The high net undercount of young children has been noted in several recent publications (O’Hare 2014a, b, 2015; U.S. Census Bureau 2014b).

Figure 4.1 also shows relatively high net overcounts for two broad age-groups. The first group is those age 15–24. The second age group that has a net Census overcount is the population age 60 plus. Some of the potential explanations for undercounts and overcounts will be discussed in Chap. 5 along with a more detailed examination of differentials by age.

4.3 Census Coverage Differentials by Sex

There was a net overcount of 1.1% for females and a net undercount of 0.8% for males. As readers will see in Chap. 6, the overall coverage rates for males and females mask large differences across age and race groups. The gender differences in Census coverage are most pronounced in working-age adults.

The 2010 DSE analysis did not provide data on undercounts of males and females for the total population (all ages) but it provided information on males and females

Table 4.1 2010 census net undercount rates from by age and sex

Age	Males	Females
	Rate	Rate
18–29	–1.2	0.3
30–49	–3.6	0.4
50+	0.3	2.4

Source U.S. Census Bureau (2012, Table 12)

Negative sign reflects a net undercount. The signs here are reversed from the source report in order to keep directionality consistent within this publication

Figures in **BOLD** are statistically significantly different from zero

for three age groups over age 18. The results are shown in Table 4.1. The coverage differentials by sex from DSE are relatively consistent with the results of DA. Data from DSE show males age 18–29 and age 30–49 had net undercount rates that were statistically significantly different than zero while females experienced a net overcount in every age group examined here (although the rates for age 18–29 was not statistically significant). The net undercount estimate for males age 30–49 is particularly high at 3.6%. Both males and females age 50 and over have a net overcount.

4.4 Census Coverage Differentials by Race

I think it is fair to say the Census net undercount differentials for race and Hispanic Origin groups are among the most widely discussed, most well-known, and most contentious. The material in this section provides a brief overview of this issue.

U.S. Decennial Census net undercounts are typically examined through a comparative lens. Historically the comparative paradigm used most often has been the comparison of Black and Non-Black populations (Fein 1989; West et al. 2014; Pre-witt 2003; Schwede et al. 2015). In large part, this perspective was based on data that has been available historically. Until recently, Blacks were the only race group identified consistently across states and over time in birth and death certificates data needed for DA estimates. But in many ways, the Blacks/Non-Black comparison also reflects the prevailing Black/White perspective on racial inequality that dominated the 1950s, 1960s, and 1970s. This publication expands that perspective to look at differential undercounts for other race groups as well.

For people not familiar with how the federal government collects data on race and Hispanic Origin, this arena can be confusing. Race and Hispanic Origin are measured in the U.S. Census as dictated by U.S. Office of Management and Budget (1997). Categorizing people by race has been a complicated part of the U.S. Census, ever since it was decided to count slaves as 3/5th of a person in the first Census in 1790.

Prewitt (2013) provides a detailed history of how the concepts and measurement of race (and Hispanic Origin) have changed since 1790 in the context of the Census.

Race and Hispanic Origin are two separate concepts in the definitions used by the federal government. In the Census, respondents are first asked if they are Hispanic or not, then they are asked about what race group(s) they identify with.

Starting in 1997, people have been allowed to mark as many race groups as they feel apply. For people who mark only one race, this is referred to as “race alone” for example Black Alone or Asian Alone. People who mark more than one race are added to the “race alone” group for form “Race Alone or in Combination”; for example, “Black Alone or in Combination.” In this construction, people can be in more than one group. Someone who marked both Black and White would be included in both Black Alone or in Combination and the White Alone or in Combination but not in the figures for Black Alone or White Alone.

Generally, the ‘race alone or in combination’ categories are used in this book, since that is the most inclusive way of identifying groups. It is also consistent with the recommendation the U.S. Office of Management and Budget which suggests that when data are used in a civil rights context, the race alone or in combination is the best grouping to use (U.S. Office of Management and Budget 2001). For many minority groups, the Census is largely seen in the context of civil rights enforcement (Leadership Conference Education Fund 2017).

The DSE net undercount rates for different races are shown in Table 4.2. Data for the Non-Hispanic White Alone population is provided because the rates for minority groups are typically contrasted to the rate for the Non-Hispanic White Alone population to develop differential net undercounts. The Non-Hispanic White Alone population is the numerical majority population in the nation.

In the 2010 Census, the Non-Hispanic White Alone population had a net overcount of 0.8%, and most of the other races had net undercounts. The only exception is American Indians Alone or in Combination living in an American Indian Area off reservations where there is a net overcount of 3.9%. In part, because this is such a small population, the net overcount rate for American Indians living off reservations is not statistically significant.

For Blacks Alone or in Combination the net undercount rate is 2.1% in the 2010 Census. There is a 2.9 percentage point difference between the net undercount rate for Blacks Alone or in Combination and the net overcount of Non-Hispanic Whites Alone noted above. This differential is not new. Census coverage data for Blacks are explored in more detail in Chap. 8.

For Asians Alone or in Combination the net undercount is zero. More data for Asians are presented in Chap. 9.

The largest net undercount in Table 4.2 is for American Indians Alone or in Combination on reservations at 4.9%. Overall, American Indian and Alaskan Natives Alone or in Combination had a net undercount of 0.2% which was not statistically significantly different than zero. More information on this population is presented in Chap. 10.

Table 4.2 2010 census estimates of percent net undercount by race (alone or in combination) and Hispanic origin

	Net undercount rate
U.S. total	0.0
Non-Hispanic White alone	0.8
<i>Race Alone or in Combination</i>	
Black	-2.1
Asian	0.0
American Indian and Alaskan Native	-0.2
On Reservation	-4.9
In American Indian Area Off Reservation	3.9
Balance of U.S.	0.1
Native Hawaiian and Pacific Islander	-1.0
Hispanic	-1.5

Source U.S. Census Bureau (2012, Table 8)

Except for the Non-Hispanic White Alone category, the race groups include Hispanics

A negative sign reflects an undercount. The signs here are reversed from the source report in order to keep directionality consistent within this publication

Figures in **BOLD** are statistically significantly different from zero

For Native Hawaiians or Pacific Islanders Alone or in Combination the net undercount rate is 1.0% which is not statistically different from zero. More information on this population are provided in Chap. 11.

The Race groups in DA are much more limited than in DSE, but the findings for Blacks and Non-Blacks are very similar to those of DSE. Based on DA analysis, Velkoff (2011) shows a Black Alone or in Combination undercount of 2.5% in the 2010 Census compared to a net overcount of 0.5% for the Non-Black population. The DA method provides data comparing Blacks and Non-Blacks for each Census from 1950 onward and these are presented in Chap. 8.

4.4.1 Hispanics

Table 4.2 shows the net undercount rate for Hispanics was 1.5% in the 2010 Census compared to a net overcount of 0.8% for Non-Hispanic White alone. More data on Census coverage for Hispanics are provided in Chap. 7.

The Hispanic population is treated as homogeneous group with respect to undercount estimates from the Census Bureau. However, it is important to recognize that subgroups within the Hispanic Populations may well have Census coverages rates that are different than the group as a whole (O'Hare 2017). The same is true for Asians. This idea is addressed in more detail in Chaps. 7 and 9.

Table 4.3 2010 census coverage rates by tenure

	Net undercount rate
Population living in owner-occupied housing	0.6
Population living in renter-occupied housing	-1.1

Source U.S. Census Bureau (2012, Table A)

A negative sign reflects an undercount. The signs here are reversed from the source report in order to keep directionality consistent within this publication

Figures in **BOLD** are statistically significantly different from zero

4.5 Census Coverage Differentials by Tenure

Table 4.3 shows the 2010 Census coverage rates for the population living in owner-occupied housing units and the population living in renter-occupied housing units. There is a net overcount for owners (0.6%) and a net undercount for renters (1.1%). Both figures are statistically significant.

This pattern of Census coverage for homeowners and renters is long standing. To some extent, homeownership may reflect socio-economic status. More information about Census coverage by tenure are presented in Chap. 12.

4.6 Other Groups Missed in the Census

Although Demographic Analysis and Dual-Systems Estimates are the best estimates available for coverage in the U.S. Census, there are some important limitations. For example, the groups for which these two programs produce data are limited. The Census Bureau research only measures Census coverage for five demographic characteristics including;

- Age
- Gender
- Race
- Hispanic Origin Status
- Tenure.

There are other groups which are widely believed to be undercounted in the Census but for which there are no direct undercount estimates from the Census Bureau. Some of these groups are identified by former Under-Secretary of Commerce following the release of data from the 2010 Census. In response to the 2010 Census results being release, then Under-Secretary of Commerce Blank (2012) stated,

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.

Table 4.4 Lists of hard-to-count populations from three sources

<ul style="list-style-type: none"> • Racial and ethnic minorities 	<ul style="list-style-type: none"> • Irregular housing units 	<ul style="list-style-type: none"> • Single and divorced males
<ul style="list-style-type: none"> • Persons who do not speak English fluently 	<ul style="list-style-type: none"> • Lack of cooperation/trust 	<ul style="list-style-type: none"> • Recent migrants
<ul style="list-style-type: none"> • Lower income persons 	<ul style="list-style-type: none"> • Communication/Language 	<ul style="list-style-type: none"> • Unemployed
<ul style="list-style-type: none"> • Homeless persons 	<ul style="list-style-type: none"> • Renters 	<ul style="list-style-type: none"> • Minority ethnic groups
<ul style="list-style-type: none"> • Undocumented immigrants 	<ul style="list-style-type: none"> • Socio-Economic Status 	<ul style="list-style-type: none"> • Private renters
<ul style="list-style-type: none"> • Young mobile persons 	<ul style="list-style-type: none"> • Residential mobility 	<ul style="list-style-type: none"> • Those who share a dwelling with other households or with a business
<ul style="list-style-type: none"> • Children 	<ul style="list-style-type: none"> • Non-City style/non-traditional addresses” 	
<ul style="list-style-type: none"> • Persons who are angry at and/or distrust the government 		
<ul style="list-style-type: none"> • LGBT persons 		
<i>Source</i> U.S. Census Bureau’s National Advisory Committee on Racial, Ethnic, and Other Population, presentation by Julie Dowling May 27, (2016, page 2)	<i>Source</i> Bruce and Robinson (2007)	<i>Source</i> Simpson and Middleton (1997)

The factors mentioned by Dr. Blank are commonly believed to be related to Census undercounts, but the Census Bureau does not produce undercount estimates for any of these characteristics (poverty, mobility, language, education, or awareness of the Census).

Groups with high net undercount rates are often referred to as hard-to-count (HTC) populations. According to the U.S. Census Bureau (2016, p. 2);

Hard-to-count populations face physical, economic, social, and cultural barriers to participation in the Census, and require careful consideration as part of a successful communications strategy.

It is important to recognize that just because the Census Bureau does not produce net undercount estimates for a group that does not mean they are not undercounted. While there are no Census undercount rates for many HTC groups other data such as Mail Return Rates (Leadership Conference Education Fund 2017) or the percent living in hard-to-count areas are sometimes available to shed light on potential Census coverage issues (U.S. Census Bureau 2014a).

Three different lists of hard-to-count populations and/or factors associated with being hard-to-count are presented in Table 4.4. These lists are meant to provide a flavor for the kinds of groups that are often thought to be missed at high rates in the Census.

Table 4.5 Hard-to-count groups compiled by U.S. General Accountability Office based on Perusal of Census Bureau documents

Complex households including those with blended families, multi-generational, or non-relatives
Cultural and linguistic minorities
Displaced persons affected by a disaster
Lesbian gay bisexual transgender queer/questioning persons
Low income persons
Persons experiencing homelessness
Persons less likely to use the Internet and others without Internet access
Persons residing in places difficult for enumerators to access, such as buildings with strict doormen, gated communities, and basement apartments
Persons residing in rural or geographically isolated areas
Persons who do not live in traditional housing
Persons who do not speak English fluently (or have limited English proficiently)
Persons who have distrust of the government
Persons with mental and/or physical disabilities
Persons without a high school diploma
Racial and ethnic minorities
Renters
Undocumented immigrants (or recent immigrants)
Young children
Young, mobile persons

Source U.S. General Accountability Office (2018, Table 1)

Table 4.5 shows a list of hard-to-count populations compiled by the U.S. General Accountability Office based on looking over many Census Bureau documents. Clearly there are many groups that are thought be hard-to-count where the net undercount and omissions are not calculated by the Census Bureau.

4.7 Summary

Several groups have been identified as having relatively high net Census undercounts rates. Groups with the highest net undercount rates include:

- The net undercount for young children (age 0–4) in the 2010 Census was 4.6% which is higher than any other age group.
- Males have a net undercount (0.8%) and females have a net overcount (1.1%).
- Renters have a net undercount (1.1%) while homeowners have a net overcount (0.6%).
- Race and Hispanic Origin Groups with the largest net undercounts are;

- Black Alone or in Combination (2.1%),
- Hispanics (1.5%),
- American Indian Alone or in Combination on reservations (4.9%).

Each of these differential undercounts will be explored in more detail in subsequent Chapters.

References

- Bhaskar, R., Scopilliti, M., Hollman, F., & Armstrong, D. (2010). Plans for producing estimates of net international migration for the 2010 demographic analysis estimates. Census Bureau Working Paper No. 90.
- Blank, R. (2012, May 22). Statement by Deputy U.S. Commerce Secretary Rebecca Blank on release of the data measuring census accuracy.
- Bruce, A., & Robinson, J. G. (2007). *Tract-level planning database with census 2000 census data*. Washington, DC: U.S. Census Bureau.
- Devine, J., Sink, L., DeSalvo, B., & Cortes, R. (2010). The use of vital statistics in the 2010 demographic analysis estimates. Census Bureau Working Paper No. 88.
- Fein, D. J. (1989). *The social sources of census omission: Racial and ethnic differences in omission rates in recent U.S. censuses*. Dissertation in Department of Sociology, Princeton University, Princeton, NJ.
- Groves, R. (2012, May 30). How good was the 2010 census? A view from the post-enumeration survey, Directors Blog.
- Leadership Conference Education Fund. (2017). Fact sheet: The census and civil rights. Downloaded on June 13, 2017 from <http://civilrightsdocs.info/pdf/Census/Fact-Sheet-Census-and-Civil-Rights.pdf>.
- O'Hare, W. P. (2014a). Assessing net coverage error for young children in the 2010 U.S. decennial census. *Center for Survey Measurement Study Series (Survey Methodology #2014-02)*. U.S. Census Bureau. Available online at <http://www.Census.gov/srd/papers/pdf/ssm2014-02.pdf>.
- O'Hare, W. P. (2014b). Historical examination of net coverage error for children in the U.S. decennial census: 1950 to 2010. *Center for Survey Measurement Study Series (Survey Methodology #2014-03)*. U.S. Census Bureau. Available online at <http://www.Census.gov/srd/papers/pdf/ssm2014-03.pdf>.
- O'Hare, W. P. (2015) *The undercount of young children in the U.S. Decennial Census*. Springer Publishers.
- O'Hare, W. P., Robinson, J. G., West, K., & Mule, T. (2016a). Comparing the U.S. decennial census coverage estimates for children from the demographic analysis and coverage measurement surveys. *Population Research and Policy Review*, 35(5), 685–704.
- O'Hare, W. P., Mayol-Garcia, Y., Wildsmith, E., & Torres, A. (2016b). The invisible ones: How Latino children are left out of our nation's census count. Child Trends Hispanic Institute & National Association of Latino Elected Officials, Child Trends, Washington DC.
- O'Hare, W. P. (2017). *Counting all Californians in the 2020 census*. The Census Project. https://Censusproject.files.wordpress.com/2017/10/calif-report-10-16-2017_format-final.pdf.
- Prewitt, K. (2003). *Politics and science in census taking*, in series The American People: Census 2000, Russell Sage Foundation and Population Reference Bureau, Population Reference Bureau, Washington, DC.
- Prewitt, K. (2013). *What is your race: The census and our flawed efforts to classify Americans*. Princeton, NJ: Princeton University Press.
- Robinson, J. G. (2010). Coverage of population in census 2000 based on demographic analysis: The history behind the numbers. U.S. Census Bureau, Working Paper No. 91.

- Schwede, L., Terry, R., & Hunter, J. (2015). Ethnographic evaluations on coverage of hard-to-count minority in the US decennial censuses. In R. Tourangeau, B. Edwards, T. P. Johnson, K. M. Wolter, & N. Bates (Eds.), *Hard-to-survey populations* (pp. 293–315). Cambridge, England: Cambridge University Press.
- U.S. Census Bureau. (2012). *2010 census coverage measurement estimation report: Summary of estimates of coverage for persons in the United States*. DSSD 2010 Census Coverage Measurement Memorandum Series #2010-G-01. Washington, DC: U.S. Census Bureau.
- U.S. Census Bureau. (2014a). Planning Data Base. Available at https://www.census.gov/research/data/planning_database/2014/.
- U.S. Census Bureau. (2014b, February 2). *Final task force report: Task force on the undercount of young children*. Memorandum for Frank A. Vitrano. Washington, DC: U.S. Census Bureau.
- U.S. Census Bureau. (2016). *Developing an integrated communication strategy: Select topics in international census*. Washington, DC: U.S. Census Bureau.
- U.S. General Accountability Office. (2018, July). 2020 census: Actions needed to address challenges to enumerating hard-to-count groups, GAO-18-599.
- U.S. Office of Management and Budget. (1997, October 30). *Revisions to the standards for the classification of federal data on race and ethnicity*. Statistical Policy Directive 15, Federal Register Notice.
- U.S. Office of Management and Budget. (2001). *Guidance on aggregation and allocation of data on race for use in civil rights monitoring and enforcement*.
- Velkoff, V. (2011). *Demographic evaluation of the 2010 census*. Presentation at the Population Association of America Annual Conference.
- West, K., Devine, J., & Robinson, J. G. (2014). *An assessment of historical demographic analysis estimates for the Black male cohorts of 1935–39*. Paper presented at the Annual Meeting of the American Statistical Association, Boston, MA.
- Simpson, L., & Middleton, E. (1997). *Who is missed by a National Census? A review of empirical results from Australia, Britain, Canada, and the USA* (p. 1). The Cathie Marsh Centre for Census and Survey Research, University of Manchester, UK.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

