RedFame

International Journal of Social Science Studies
Vol. 8, No. 6; November 2020
ISSN 2324-8033 E-ISSN 2324-8041
Published by Redfame Publishing
URL: http://ijsss.redfame.com

Are Self-Participation Rates Predictive of Accuracy in the U.S. Census?

William P. O'Hare[1]

[1] O'Hare Data and Demographic Services LLC, Cape Charles, VA 23310, United States

Correspondence: William P. O'Hare, O'Hare Data and Demographic Services LLC, Cape Charles, VA 23310, United States.

Received: August 6, 2020 Accepted: September 14, 2020 Available online: September 24, 2020

doi:10.11114/ijsss.v8i6.4967 URL: https://doi.org/10.11114/ijsss.v8i6.4967

Abstract

Metrics related to the U.S. Census have been widely available for several decades but there has been a dearth of studies examining the relationship among key metrics in the Census. This paper provides empirical evidence about the link between self-participation rates and census accuracy using data from the 1990, 2000, and 2010 U.S Censuses. The preponderance of the evidence shows lower self-participation rates are highly correlated with higher net undercounts and omissions rates for key socio-demographic groups and states. Nine out of 11 correlations examined in this paper are statistically significant and in the predicted direction. One key reason self-participation rates are associated with census accuracy is the fact that the population not captured in the self-participation operation goes into the households for the Nonresponse Followup (NRFU) operation. Census Bureau data show data collected in NRFU is not as accurate as that collected in self-response. The larger the share of data collected for a population that is collected in NFRU, the lower the quality of data for that group. The connection between self-participation rates and census accuracy mean the differential self-participation rates seen in the 2020 Census suggest patterns of net Census undercounts seen in the past are likely to be seen in the 2020 Census.

Keywords: census, self-response, accuracy, undercounts, omissions

1. Introduction

The U.S. Census Bureau tracks many different metrics of the Census to understand the quality of the census-taking process and the outcomes of the census-taking process (Fay et al. 1988, Robinson et al. 1993, Velkoff 2011; U.S. Census Bureau 2012a, b & c, Hogan et al. 2013, O'Hare 2019a). The metrics reported by the Census Bureau are the kinds of metrics the United Nations (2017) recommends collecting in a census.

A series of reports from the U.S. Census Bureau show variations in Census accuracy (net undercount and omissions rates) and substantial differences in self-response rates over the past 40 years. Despite the availability of statistical data on census self-response rates and the statements by the Census Bureau and others indicating a belief that self-response rates are linked to census accuracy, I can find no empirical studies which document that connection among demographic groups or across geographic areas. This study addresses that gap in the literature.

This study documents the systematic relationship between census self-response measures and Census accuracy based on empirical data from the 1990, 2000, and 2010 U.S. Censuses, using data from the U.S. Census Bureau. The study uses correlational analysis and data for key socio-demographic groups (race and tenure) and states. These are the only characteristics where both self-participation rates and census accuracy measures are available and consistent.

2. Review of the Literature

The Census Bureau and Census Bureau researchers have provided a stream of publications and data related to census accuracy and participation over the past 40 years. Following the 2010 Census, the Census Bureau (2012a, b, and d) issued a series of reports providing data on net undercounts and omissions for demographic groups and geographic areas based on the dual-systems estimation methodology. The Census Bureau also provided a report showing self-response rates (mail return rates and the mail response rates) in the 2010 Census by demographics (U.S. Census Bureau 2012d). State self-response rates were made available on the Census Bureau website (U.S. Census Bureau https://www2.census.gov/data/api-documentation/how-to-download-decennial-self-response-rates-for-all-counties-in-fl-from-the-census-api.pdf).

Following the 2000 Census, the Census Bureau reports on census accuracy were revised a couple of times because of

technical problems that were uncovered as the estimates were generated. The final report published in 2003, includes data from both the demographic analysis estimation methodology and the dual-systems estimation methodology operation (U.S. Census Bureau 2003a). The Census Bureau (2003b) also produced data on self-response rates in the 2000 Census.

Following the 1990 Census there was no report from the Census Bureau that included undercounts by race/ethnicity and tenure simultaneously. However, Census Bureau staffers (Hogan and Robinson 1993: Word 1997) produced reports which contained consistent data for these groups. State net undercount rates were made available by the Census Bureau on their website (U.S. Census Bureau, https://www.census.gov/dmd/www/pdf/understate.pdf).

The series of reports following the 1990, 2000, and 2010 Censuses provided extensive data showing variation on measures of census accuracy and self-response. However, I can find no empirical studies which document the connection between the self-response rates and the census accuracy rates in the reports from the 1990, 2000, and 2020 Census.

Despite the lack of strong empirical evidence, several researchers have suggested there is a relationship between self-response rates and census accuracy. The Census Bureau Task Force on the Undercount of Young Children (U.S. Census Bureau 2014a, page ii) concluded, "Research suggests that areas with lowest levels of cooperation have higher levels of coverage and nonresponse error." A Census Bureau Working Paper from the 1990s (Word 1997, page 1) notes that "response rates and net undercount rates may be causally linked..." With respect to lower self-response rates the Census Bureau (2018b, page 54) concluded, lower self-response rates would end up, "… resulting in more nonresponse followup (NRFU) field work, more proxy responses, and a lower-quality population count.". In the context of the potential failure of the 2020 U. S. Census, Swanson (2019, page 6) says, "A leading indicator of whether the 2020 census will be our last is the overall participation rate."

The connection between self-response rates and census accuracy is also underscored by the Census Bureau's decision to use a self-response related measure to identify Hard-to-Count areas in the 2020 Census. The Low-Response Score developed by Erdman and Bates (2017) is based on the mail return rates in the 2010 Census. In describing the Low-Response Score the Census Bureau (2014b, page 4) states, "This score identifies Block Groups and Tracts whose characteristics predict low Census Mail Return Rate and are highly correlated (negatively) with Census and survey participation." The implicit association here is that areas where self-response rates are low are more difficult to enumerate and likely to end up with a net undercount.

New York City Demographer Joe Salvo (2020, page 1) states, "If an area has a low self-response rate, it means:

- More census enumerators will need to knock on doors to count residents in persons: and,
- It is more likely people in the area may be missed or counted inaccurately."

This statement by Salvo suggests a mechanism that would lead one to expect groups and areas with low response rates are likely to have lower census accuracy. The portion of the population not captured in the self-response operation in the Census fall into the Non-response Followup (NRFU) operation where data quality is lower than in self-response operation.

If the researchers cited above are correct there should be a strong statistical association between differences in self-response rates and measures of Census accuracy across socio-demographic groups and states. This study examines that relationship between self-response rates and census accuracy using empirical data from the 1990, 2000 and 2010 U.S. Censuses.

3. Methods

3.1 Background on Census Operations

Briefly, the U.S. Decennial Census can be thought of as having two major phases or operations: 1) a self-participation phase, and 2) a non-response followup (NRFU) phase (U.S. Census Bureau 2017a). A few minor Census operations like the Update/Leave operation, Update/Enumerate operation and group quarters count capture some people for the Census count but the vast majority of people included in the Census are counted in the self-response and NRFU phases.

The self-participation phase consists of households returning a completed questionnaire that was mailed to them by the Census Bureau (in 2020, for the first time in the Decennial Census, the self-response will include internet and telephone responses). Several weeks after Census day (April 1) the second phase of the Census begins, and households that did not return a completed Census questionnaire are visited by a Census enumerator to gather the information needed for the Census. This is referred to as NRFU operations by the Census Bureau.[1] While a few people belatedly self-respond

[1] In the 2020 Census the NRFU operational design will also use administrative records and third-party data to enumerate occupied housing units. See U.S. Census Bureau, 2020 Census Detailed Operational Plan: 18. Nonresponse Followup Operation (NRFU), April 2018 at 9, available at https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/NRFU-detailed-operational-plan.pdf

during the NRFU phase, and there are other Census operations (like update/ leave and update/enumerate) that are outside of these two operations, the bulk of Census respondents are captured in the self-response and NRFU operations of the Census.

3.2 Key Concepts and Terms

There are two key concepts used in this analysis. One is the idea of people self-responding to the Census. This concept is captured in mail return rates and mail response rates in 1990, 2000, and 2010. The second key concept is census accuracy, and that is captured by net undercount rates and omissions rates. All the data used in the analysis comes from the U.S. Census Bureau.

3.2.1 Self-Participation Rates

Self-participation rates in a census are one the most important early signals about how a census is unfolding. I use the term "self-participation rates" rather than self-response rates to help distinguish this concept from the mail response rates term, which is one ways of measuring self-participation, as explained below.

In this analysis, self-participation is measured by two closely related indicators: mail return rates and mail response rates. Detailed descriptions of mail return rates and mail response rates are provided by the Census Bureau (2014b, page 61). In simple terms, the mail return rate is the percentage of Census questionnaires that were returned from occupied households. Mail response rates are the percentage of Census questionnaires that were returned from all households whether they were occupied or not. Census mail-response rates are the only ones available as the census is being conducted.

For the groups and geographic areas (states) examined here there is typically a high correlation between mail return rate and mail response rates. The correlation across states in 2010 was +0.72. The correlation may be lower for smaller units of geography.

In the 1990, 2000, and 2010 Census, self-participation is reflected in the extent to which respondents mailed back completed census questionnaires. In the 2020 Census, self-participation includes responses by internet, phone, and mail. While the operationalization of self-response in the 2020 Census is somewhat different than previous Censuses, conceptually it is the same.

Mail return rates and the mail response rates used here are those calculated by the U.S. Census Bureau and they have been used by others in the analysis of Census data. The mail return rate has been used by U.S. Census Bureau (2018a & b) to measure self-participation in their analysis.

Where available, my analysis uses both mail return rates and mail response rates to measure self-participation rates, although generally they are highly correlated. The mail return rate was not available in the 1990 Census, so I use only the mail response rates to reflect the self-participation in 1990.

3.2.2 Census Accuracy: Net Undercounts and Omissions

Net undercount and omissions rates are both measures of Census accuracy, but they capture different parts of Census accuracy (O'Hare 2019b). I use the term "net undercounts" to include net undercounts and net overcounts (a net overcount can be seen as a positive net undercount). Demographers sometimes refer to net undercounts and net overcounts as census coverage. I use the term net undercount in this paper for two reasons. First, I believe the term net undercount is more widely understood than the term census coverage. Second, census coverage is sometimes used to include omissions. In this study, omissions are treated as a distinct dimension of census accuracy.

In addition, it should be noted that net undercounts and omissions are not the only kinds of census errors. For example, some people are counted in the Census, but counted in the wrong place. Other times, people are counted accurately, but one or more of their characteristics may be incorrect. For example, the Census Bureau estimates that a couple of million people marked the wrong sex in the 2010 Census because of a flaw in the way the questionnaire was designed. (U.S. Census Bureau 2011)

Omissions is the simplest and most straightforward concept of census accuracy. Omissions capture the number and share of a population that are missed in the Census and are defined by the Census Bureau (2012a, page 12) as, "omissions are people who should have been enumerated in the United States Census but were not."

The net undercount is a balance between people missed (omissions) and those included erroneously (those double counted and those inappropriately included in the census, like foreign tourists) and those imputed.[2] If the number of omissions is

[2] Imputations are people added to the Census count based on some evidence they exist. For example, if a housing unit looks occupied, but there is no self-response, and no one responds to an enumerator, the Census Bureau may impute people into the Census count.

higher than the number of erroneous inclusions and whole person imputations, there is a net undercount. If the number of erroneous inclusions and whole person imputations is larger than the number of omissions, there is a net overcount.

In some ways omissions are a better reflection than net undercount rates of who is missed in the Census. Former Census Bureau Director Robert Groves (2010), said "Most professionals would agree that the key issue is "gross" not "net" coverage, separately accounting for those enumerated more than once (double counted) and those not counted at all."

In the 2010 Census there were 10,042,000 erroneous enumerations, 5,993,000 whole person imputations and 15,999,000 omissions (U.S. Census Bureau 2012a, Table 3). Thus, the net undercount in the 2010 Census was near zero even though almost 16 million people were missed.

In another example, the net undercount of Asians in the 2010 Census was essentially zero, but there was an omissions rate of over 5 percent for Asians in the 2010 Census. The net undercount for young children in the 2010 Census was 4.6 percent, but the omissions rate was 10.3 percent (O'Hare 2019b). In the 2010 Census, the net undercount rate was near zero, while the omissions rate was 5.2 percent. Omissions provide different a different picture of census accuracy than net undercounts

Undercounts have sometimes been reported as a negative number by the U.S. Census Bureau (Velkoff 2011; King et al. 2018; Jensen et al. 2018) and sometimes as a positive number by the U.S. Census Bureau (U.S. Census Bureau 2012a &b). In this report, net undercounts are reported consistently as a positive number and net overcounts as a negative number. Measuring net undercounts here as a positive number makes the correlations easier to interpret.

4. Results: Statistical Relationships between Self-response and Census Accuracy

In this section, the empirical relationship between self-participation rates (mail return rates and mail response rates) and Census accuracy (net undercount and omissions rates) is examined.

4.1 Examination of Data from the 2010 Census

Table 1 shows the mail return rates, net undercount rates, and omissions rates for eight demographic groups defined by race, Hispanic origin, and tenure (i.e., owner or renter). These are the only demographic groups where I could find two concepts (self-participation, census accuracy rates) in consistently classified groups.

The correlation between the mail return rates and the net undercount rate for the eight groups shown in Table 1 is -0.78, and the correlation between the mail return rates and the omissions rates for the eight groups shown in Table 1 is -0.86. Both correlations are statistically significantly different than zero and indicate groups that have low mail return rates tend to have high net undercount and omissions rates.

Table 1 Mail Return Rates, Net Undercount Rates, and Omission Rates for Demographic Groups in the 2010 Census

	Mail Return Rates*	Net Undercount Rates**	Omission rates***
Total	75.8	-0.01	5.3
White Alone	82.5	-0.84	4.3
Black Alone	70.0	2.07	9.3
American Indian and Alaskan Native Alone	69.8	0.15	7.6
Asian Alone	75.4	-0.08	5.3
Native Hawaiian or Pacific Islander Alone	59.7	1.34	7.9
Hispanic	69.7	1.54	7.7
Population in Owner-Occupied Housing Units	85.8	-0.60	3.7
Population in Renter- Occupied Housing Units	66.9	1.09	8.5

* Source; U.S. Census Bureau (2012) 2010 Census Mail Response/Return Rates Assessment Report. 2010 Census Planning Memorandum Series, No. 198,

** Source: 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 (Net undercounts are shown as positive numbers)

*** Source: U.S. Census Bureau (2012) 2010 Components of Census Coverage for Race Groups and Hispanic Origin by Age, Sex and Tenure in the United States, DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-E-51

Table 2 shows mail return rates, mail response rates, net undercount rates, and omissions rates in the 2010 Census for

states along with the District of Columbia. Note that none of the state net undercount rates in Table 2 are statistically significantly different from zero (U.S. Census Bureau 2012b, Table 5). The fact that the 2020 Census state net undercount rates are highly clustered around zero is important because when there is little variation in one of the measures in a relationship, it is more difficult to determine the extent of any correlation or relationship.

Table 2. 2010 Census State Mail Response Rates, Mail Return Rates, Net Undercount Rates, and Omissions Rates

State	Mail Response Rate*	Mail Return Rate*	Net Undercount Rate**	Omissions Rate**		State	Mail Response Rate*	Mail Return Rate*	Net Undercount Rate**	Omissions Rate**
Alabama	62.5	78.4	0.13	7.7		Montana	64.6	80.4	-0.65	6.1
Alaska	55.6	74.8	-0.85	5.5		Nebraska	71.1	82.5	-0.54	3.1
Arizona	61.3	77.6	-0.42	7.3		Nevada	61.4	76.3	-0.04	6.9
Arkansas	62.3	77.0	-0.41	5.4		New Hampshire	64.4	79.4	0.6	5.0
California	68.2	76.9	0.26	5.1		New Jersey	67.6	78.1	-0.36	4.5
Colorado	67.2	79.1	-0.29	5.9		New Mexico	60.0	73.8	-0.16	7.7
Connecticut	69.5	79.1	-0.45	3.9		New York	64.6	75.8	-0.79	6.1
Delaware	64.1	80.0	0.55	6.2		North Carolina	64.8	80.7	0.52	7.6
District of Columbia	66.0	78.3	2.23	9.0		North Dakota	68.8	83.1	0.09	3.9
Florida	63.0	80.2	0.45	7.5		Ohio	69.0	80.8	-0.83	3.5
Georgia	62.5	77.2	0.91	7.3		Oklahoma	62.3	75.5	-1.08	6.4
Hawaii	64.1	76.8	-0.44	7.8		Oregon	66.9	79.8	0.02	4.0
Idaho	67.1	82.6	-0.03	5.8		Pennsylvania	70.2	82.3	0.14	4.5
Illinois	70.5	80.7	-0.48	4.6		Rhode Island	65.7	77.7	-0.81	5.9
Indiana	69.6	82.2	-0.67	3.6		South Carolina	64.7	81.4	0.41	5.2
Iowa	73.0	83.3	-0.28	2.6		South Dakota	67.1	82.7	0.1	4.9
Kansas	70.0	81.2	-0.67	3.7		Tennessee	67.1	80.3	0.12	5.8
Kentucky	65.7	81.0	-0.13	5.5		Texas	64.4	76.5	0.97	6.9
Louisiana	61.0	74.5	-0.38	6.8		Utah	68.6	80.4	-0.48	4.9
Maine	57.4	81.1	0.65	4.2		Vermont	60.3	79.7	1.29	5.4
Maryland	69.5	80.3	0.94	6.0		Virginia	69.0	80.8	0.57	5.8
Massachusetts	68.8	78.9	-0.52	5.7		Washington	67.2	79.9	-0.1	4.5
Michigan	67.7	83.7	-0.66	4.5		West Virginia	59.1	75.6	-1.43	7.7
Minnesota	74.1	85.6	-0.56	4.4		Wisconsin	73.5	85.1	-0.17	4.1
Mississippi	61.3	76.4	0.24	8.9		Wyoming	63.4	79.9	-0.51	6.4
Missouri	67.5	81.1	-0.66	4.5						

* Source: Downloaded from CUNY Mapping system https://www.censushardtocountmaps2020.us/

**Source: 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. U.S. Census Bureau, Washington, DC. G-01 Table 14 (Net undercounts are shown as a positive number in this table)

Given the low level of measurable variation in the state-level net undercount rates in the 2010 Census, correlation with self-response rates is expected to be low. Indeed, the correlation between mail response rates and net undercount rates across the states is -0.09, and the correlation between mail return rates and net undercount rates is 0.06. Neither correlation is statistically significant.

The correlation between state mail response rates and omissions rates is -0.61 and the correlation between 2010 state mail return rates and omissions rates is -0.63. Both correlations are statistically significantly different from zero and indicate that states with low mail response rates tend to have higher net undercount and omissions rates.

4.2 Examination of Data from the 2000 Census

Table 3 shows self-participation rates and net undercount rates from the 2000 Census for eight demographic groups. Note that the racial groups are not defined exactly the same way in the two Census Bureau reports from which the data were taken but they are very similar. This is a minor point and unlikely to significantly impact the correlation. The correlation coefficient between mail return rates and net undercount rates in Table 3 is -0.97, which is remarkably high and statistically significantly different from zero. Demographic groups that have relatively low mail return rates tend to have relatively high net undercount rates.

Table 3. Mail Return Rates and Net Undercount Rates in the 2000 Census for Eight Demographic Groups				
Groups	Mail Return Rate*		Groups	**Net Undercount Rates (A.C.E. Revision II)
White Alone	81.8		Non-Hispanic White	-1.13
Black Alone	64.3		Non-Hispanic Black	1.84
Asian Alone	74.6		non-Hispanic Asian	-0.75
Pacific Islander Alone	59.4		Hawaiian or Pacific Islander	2.12
Hispanic	69.2		Hispanic	0.71
American Indian Alone	70.7		AIAN Off Reservations	0.62
Owner-Occupied	84.8		Homeowner	-1.25
Renter-Occupied	65.9		Renter	1.14
*Source: U.S. Census Bureau : DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-54, Table 1				
**Source: U.S. Census Bureau (2003) Census 2000 Mail Return Rates, Census 2000 Evaluation A.7.b, Herbert Stackhouse and Sarah Brady, January 30, Table 10 (Net undercounts are shown as positive numbers)				

Table 4 shows the self-participation rates and net undercount rates for states along with the District of Columbia in the 2000 Census. State omissions rates were not available in the 2000 Census.

Table 4. Census 2000 Mail Return Rates, Mail Response Rates, and Net Undercount Rates for States								
State	Mail Return Rates*	Mail Response Rates**	Net Undercount Rates***		State	Mail Return Rates*	Mail Response Rates**	Net Undercount Rates***
Alabama	76.9	61	-0.34		Montana	83.2	68	0.49
Alaska	76.1	56	0.29		Nebraska	84.6	75	-0.81
Arizona	76.9	63	-0.32		Nevada	75.0	66	0.50
Arkansas	79.0	64	-0.09		New Hampshire	80.5	67	-1.15
California	78.6	70	0.13		New Jersey	78.7	68	-0.52
Colorado	80.8	70	-0.01		New Mexico	78.1	62	0.02
Connecticut	80.2	70	-0.75		New York	75.2	63	-0.25
Delaware	78.0	63	-0.32		North Carolina	78.3	64	-0.15
District of Columbia	72.1	60	1.54		North Dakota	85.4	72	-1.43
Florida	77.8	63	-0.64		Ohio	82.1	72	-1.27
Georgia	79.1	65	0.27		Oklahoma	77.9	64	-0.20
Hawaii	75.7	60	0.22		Oregon	81.1	68	-0.35
Idaho	83.5	67	-0.41		Pennsylvania	82.4	70	-0.91
Illinois	80.2	69	-1.42		Rhode Island	76.9	67	-1.14
Indiana	81.6	69	-1.66		South Carolina	76.5	58	-0.36
Iowa	85.6	76	-1.44		South Dakota	86.8	74	-1.28
Kansas	81.8	71	-1.28		Tennessee	77.1	65	-0.41
Kentucky	79.9	66	-0.48		Texas	75.3	64	0.05
Louisiana	75.2	60	-0.09		Utah	79.6	68	-0.10
Maine	80.1	61	-1.20		Vermont	81.1	60	-1.12
Maryland	79.3	69	0.25		Virginia	81.3	72	0.27
Massachusetts	79.0	69	-1.00		Washington	78.6	66	-0.21
Michigan	83.7	71	-0.95		West Virginia	80.7	64	-0.73
Minnesota	86.1	75	-1.70		Wisconsin	87.3	75	-1.50
Mississippi	78.3	63	-0.41		Wyoming	83.6	66	-0.39
Missouri	82.2	69	-1.35		United States	67	67	0.48
* Source: U.S. Census Bureau, State Mail Return Rates 2010 Census downloaded on August 21 at https://www2.census.gov/dssd/2010_census_public_rates/excel/								
** Source; Census 2000 Final Response Rates October 6, 2011.								
*** Net undercount as a percent of the total population. Source: U.S. Census Bureau (2003) A.C.E. Revision II- Adjusted Data for States, Counties, and Places, DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-60, Table 1 (Net Undercounts Shown as a positive number)								

The correlation between mail response rates and net undercount rates across the states is -0.52 and the correlation between mail return rates and net undercounts is -0.66. Both correlations are statistically significantly different from zero and indicate states with low mail response rates and low mail return rates tend to have high net undercount rates.

4.3 Examination of Data from the 1990 Census

The only self-response rates available for states in the 1990 Census were mail response rates. Table 5 shows mail response

rates and net undercount rates in the 1990 Census for seven demographic groups. The correlation between the mail response rates and the net undercount rates is -0.60, which is statistically significantly different from zero and indicates that groups with lower self-response rates tend to have higher net undercount rates.

Table 5. 1990 Census Mail Response Rates and Net Undercount Rates for Seven Demographic Groups				
	Mail Response Rates*			Net Undercount Rate **
Non-Hispanic White	78.0		Non-Hispanic White	0.7
Black	56.6		Black	4.4
American Indians, Eskimo and Aleut	63.1		American Indian, Eskimo and Aleut	12.2
Asian and Pacific Islanders	66.1		Asians and Pacific Islander	2.4
Hispanic Origin	63.4		Hispanic Origin	5
Owners	82.3		Owners	0.1
Renters	61.3		Renters	4.3
*Source: Derived from Word, D.L., (1997) "Who Responds ? Who Doesn't?: Analyzing Variation in Mail L Response Rates During the 1990 Census, Population Division Working Paper No . 19, Table 2.0				
**Source: Hogan, H and Robinson G (1993) What the Census Bureau's Coverage Evaluation Programs tell Us About Differential Undercounts ; Paper Delivered at the 1993 Research Conference on Undercounted Ethnic Populations, May 5-7, Richmond VA. , Table 3 (net undercounts shown as a positive number)				

Table 6 shows 1990 Census mail response rates and net undercount rates for states and the District of Columbia. Omissions rates were not available in 1990. The correlation coefficient between mail response rates and net undercount rates, shown in Table 6 is -0.56, which is statistically significantly different from zero and indicates that states with lower self-response rates tend to have higher net undercount rates.

Table 6 1990 Census Mail Response Rates and Net Undercount Rates for States						
state	1990 Mail Response Rates*	1990 Net Undercount of Persons**		state	1990 Mail Response Rates*	1990 Net Undercount of Persons**
Alabama	62.0	1.8		Montana	67.0	2.4
Alaska	52.0	2.0		Nebraska	74.0	0.7
Arizona	62.0	2.4		Nevada	61.0	2.3
Arkansas	65.0	1.8		New Hampshire	63.0	0.8
California	65.0	2.7		New Jersey	65.0	0.6
Colorado	67.0	2.1		New Mexico	62.0	3.1
Connecticut	66.0	0.6		New York	62.0	1.5
Delaware	68.0	1.8		North Carolina	63.0	1.9
District of Columbia	56.0	3.4		North Dakota	72.0	0.7
Florida	61.0	2.0		Ohio	75.0	0.7
Georgia	63.0	2.2		Oklahoma	63.0	1.8
Hawaii	62.0	1.9		Oregon	67.0	1.9
Idaho	70.0	2.2		Pennsylvania	73.0	0.3
Illinois	68.0	1.0		Rhode Island	62.0	0.1
Indiana	72.0	0.5		South Carolina	58.0	2.0
Iowa	76.0	0.4		South Dakota	74.0	1.0
Kansas	72.0	0.7		Tennessee	65.0	1.8
Kentucky	69.0	1.6		Texas	61.0	2.8
Louisiana	58.0	2.2		Utah	67.0	1.7
Maine	58.0	0.7		Vermont	64.0	1.1
Maryland	70.0	2.1		Virginia	70.0	2.0
Massachusetts	64.0	0.5		Washington	67.0	1.8
Michigan	72.0	0.7		West Virginia	65.0	1.4
Minnesota	76.0	0.4		Wisconsin	77.0	0.6
Mississippi	62.0	2.1		Wyoming	61.0	2.2
Missouri	69.0	0.6		U.S. Total		1.6
* Source: U.S. Census Bureau 1990 Census Page (/main/www.cen1990.html) 1990 Mail Response rates by 1990 Geography Boundaries.						
** Source: U.S. Census Bureau https://www.census.gov/dmd/www/pdf/understate.pdf						

4.4 Summary of Relationship between Self-response Rates and Census Accuracy.

Table 7 summarizes the correlations between self-participation rates and census accuracy for the 1990, 2000 and 2010 Decennial Censuses. Based on empirical statistical relationships seen in the past three decennial censuses, groups and states that have lower self-response rates are likely to have higher net undercounts and omissions rates.

Table 7. Summary of Statistical Relationships between Census Self-Response Rates and Census Accuracy

	Correlation Coefficient	Statistical Significance (One-Tailed Test)
2010 Mail Return Rates and Net Undercount Rates for Eight Demographic Groups	**-0.78**	0.011
2010 Mail Return Rates and Omissions Rates for Eight Demographic Groups	**-0.86**	0.003
2010 Mail Response Rates and Net Undercount Rates for States	-0.09	0.265
2010 Mail Response Rates and Omissions Rates for States	**-0.61**	< .0001
2010 Mail Return Rates and Net Undercount Rates for States	0.06	0.338
2010 Mail Return Rates and Omissions Rates for States	**-0.63**	< .0001
2000 Mail Return Rates and Net Undercount Rates for Eight Demographic Groups	**-0.97**	< .0001
2000 Mail Response Rates and Net Undercount Rates for States	**-0.52**	< .0001
2000 Mail Return Rates and Net Undercount Rates for States	**-0.66**	< .0001
1990 Mail Response Rates and Net Undercount Rates for Seven Demographic Groups	**-0.60**	0.077
1990 Mail Response Rates and Net Undercount Rates for States	**-0.56**	< .0001

Note correlations in **BOLD** are statistically significant at the .90 level or higher.

Of the 11 correlations shown in Table 7, all but two were in the predicted direction and statistically significant. There was a reasonable explanation for the correlations that were low and not statistically significant, namely, little variation in one of the variables.

The preponderance of empirical evidence from examining relationships seen in the recent U.S. Censuses indicates lower self-participation rates for groups or states will likely lead to an increase in net undercounts and omissions. The magnitude of the correlations varies from one Census to the next, but the correlations are consistent in showing a strong negative correlation between self-participation rates and census accuracy.

5. Discussion

The likely explanation for the relationship between self-response rate and census accuracy is the fact that households that do not self-respond end up in the non-response followup (NFRU) universe where the Census Bureau must send out an enumerator to get information from a nonresponding household. With respect to the NRFU process relative to the self-reporting operation in the 2020 Census, former Census Bureau Director Kenneth Prewitt (2020) stated, "The census self-reporting phase successfully reached 62 percent of the population. An achievement to applaud. But the next phase is orders of magnitude more difficult. We're in the NRFU/Hard-to-count territory."

Data collected in the NRFU phase of the Census is generally less accurate than the data gathered in the self-response phase. In the 2010 Census, only 88 percent of the NRFU responses were correct compared to 97 percent of self-responses (authors calculation from U.S. Census Bureau 2012b, Table 14). The Census Bureau's (2012e, Table 8) 2010 Census Reinterview Survey Evaluation Report found that self-responses were more accurate than interview enumerated responses in five of six variables. For Tenure, Relationship, Age, Hispanic or Latino, and Race self-responses were more accurate but for interviewer/enumerator collected data sex was more accurate.

For one thing, the portion of the population that is most willing to respond to the Census typically does so in the self-response phase. The population less willing to respond to the census is concentrated in the NRFU population. In addition, some people are reluctant to provide information to a person at their front door. Given the COVID-19 pandemic, many respondents will be even less likely to open their door to a stranger.

In addition, since the NRFU operation occurs months after the April 1 Census date, some respondents may have difficulty accurately recalling exactly who was in the household on that date. As Martin (2007, page 429) states, "Respondents interviewed months after April 1 may find it difficult to recall accurately when a move occurred."

For some portion of the NRFU universe the Census Bureau may have to rely on a proxy response. A proxy response occurs when someone who may know a little about the population living in a housing unit (perhaps a neighbor or a landlord) provides Census data for the household. A paper by the Census Bureau shows that the quality of response from

the self-response portion of the Census is much more accurate than responses from the NRFU proxy response portion (U.S. Census Bureau, 2018b, Table 12). The data showed that 97.3 percent of the responses from the Mailout/Mail back portion of the Census were correct, as compared to just 70.2 percent of those from the NRFU proxy responses.

This analysis makes it clear that NRFU and post-processing operations (such as imputation and de-duplication) in the Census do not fully account for differences in self-response rates. In addition to the lower quality of the data collected in NRFU, many people are missed in the Census because they are left off a census questionnaire for a housing unit that self-responded (Hogan 1993). When people are left off Census questionnaires that are returned, there is no census operation designed to correct this mistake.

The errors highlighted in this article have occurred despite the Census Bureau's attempts to eliminate or reduce such errors. Over the past several years the Census Bureau has undertaken numerous attempts to get a more accurate census (O'Hare, 2019a, Chapter 14) including a paid advertising program and a robust partnership program introduced in the 2000 Census. In the 2020 Census it was thought that providing more ways to respond to the Census (online, by phone, or by mail) would help encourage respondents.

There are several reasons why one might expect the 2020 Census to be more difficult than the one in the 2010 Census. As this article is being written, we are in the middle of the 2020 Census data collection period. The 2020 U.S. Census has been subject to a number of challenges including a nationwide pandemic in the middle of the data collection period, underfunding throughout the decade, and unprecedented political interference (O'Hare 2020; Sullivan 2020).

The Census Bureau's (2019) Census Barriers and Motivators Study (CBAMS) survey conducted early in 2018 found 53 percent of respondents mistakenly thought census data was used," To locate people living in the country without documentation." And 63 percent thought the Census was used "to help police and FBI keep track of people who break the law." Moreover, 28 percent of the CBAMS respondents mistakenly thought the Census Bureau would not keep answers confidential and 24 percent mistekenly thought the Census Bureau would share answers with other government agencies. People with this level of distrust in the Census, were unlikely to self-respond and resist responding to an enumerator in the NFU process.

In addition, the climate of fear among immigrants has greatly escalated recently. Even before the administration attempted to add a citizenship question to the 2020 Census, Census Bureau researchers (U.S. Census, 2017b; Meyer and Goerman 2018; U.S. Census Bureau 2017b and 2017c) found respondents less willing to cooperate given the growing climate of fear and mistrust. After careful review, U.S. Census Bureau researchers Meyers and Goerman (2018, Slide 24) conclude, "During multilingual pretesting studies conducted in 2017 and 2018, respondents expressed concerns about participating in the Census Bureau surveys because of fears about their confidentiality.

In a series of focus groups among Latino adults regarding the 2020 Census conducted by National Association of Latino Elected Officials (NALEO) , they concluded, "Hesitation, fear and cynicism arose among focus group participants when they saw a version of the questionnaire. The citizenship question raised the most concerns, anxiety increased as participants considered the reality of providing their information to the current administration." (National Associations of Latino Elected Officials, 2018, slide 3). Data from a recent NALEO (2020) survey indicate nearly half of all Hispanics think the citizenship question is still on the 2020 Census questionnaire

Based on a series of interactions with interviewees and Census Bureau field staff, the Census Bureau (2017c, page 7) concluded, "Overall, these findings, in various languages from respondents, Field Representatives, and Field Supervisors across the country who have participated in recent projects are raising concerns with CSM regarding potential barriers to respondents participation in the 2020 Census, as well as other Census Bureau surveys."

6. Conclusion

The preponderance of empirical evidence from the 2010, 2000 and 1990 Census shows lower self-participation rates lead to higher net undercount rates and omissions rates. This connection between self-participation rates and census accuracy underscores the importance of monitoring self-participation rates in the Census. Differences in self-participation rates in the 2020 Census suggests that differential undercount patterns of the past are likely to be repeated in the 2020 Census.

The empirical relationship between self-participation rates and census accuracy has been recognized by the Courts. After listening to several days of expert testimony and reading many expert witness reports in the New York citizenship question case Judge Furman (*State of New York v United States Department of Commerce*, page 133) concluded, "The Court concludes just that: Dr. O'Hare's testimony provides affirmative evidence that self-response declines among specific subpopulations directly cause net undercounts of those subpopulations. For the purposes of this litigation, a preponderance of the evidence supports that conclusion."

In the Maryland citizenship question case (*Kravitz v. United States Department of Commerce:* 2019, page 50) Judge Hazel concluded, "The court next finds that demographic groups with lower self-response rates are more likely to be

undercounted. This straight line between lower participation ultimate undercount is supported by common sense and the preponderance of the evidence." In *Kravitz v. United States Department of Commerce:* 2019, page 51 Judge Hazel also concluded " Thus, the Court is comfortable finding that Plaintiffs have demonstrated a causal relationship between decreased Census participation and an increased likelihood of net undercounting by a preponderance of the evidence.

Acknowledgments

Author wishes to thank numerous colleagues who provided advice on this article**.**

References

Erdman, C., & Bates, N. (2017). THE LOW RESPONSE SCORE (LRS) A METRIC TO LOCATE, PREDICT, AND MANAGE HARD-TO-SURVEY POPULATIONS. *Public Opinion Quarterly, 81*(1), 144-156. https://doi.org/10.1093/poq/nfw040

Fay, R. E., Passel, J. S., Robinson, J. G., & Cowan, C. D. (1988). *The Coverage of the Population in the 1980 Census. U.S. Census of Population and Housing, Evaluation and Research Reports,* PHC80-E4. U.S Census Bureau, Washington, DC.

Furman, (2019). *FINDINGS OF FACT AND CONCLUSIONS OF LAW, Case 1-28-cv-02921-JMF Document 574 Filed 01/15/19, STATE OF NEW YORK et al.,* Plaintiffs -v- United States Department of Commerce et al. defendants.

Groves, R. (2010). *Quality in a Census, Some Overview Thoughts.* Census Bureau's Directors Blog, September 9, 2010. Retrieved from https://www.census.gov/newsroom/blogs/director/2010/09/quality-in-a-census-some-overview-thoughts.html

Hazel. (2019). Kravitz v. United States Department of Commerce IN THE UNITED STATES DISTRICT COURT FOR THE DISTRICT OF MARYLAND, Southern Division, Case No. GJH -18-1041, Document 125, filed 4/5/2019.

Hogan, H. (1993). The 1990 Post-Enumeration Survey: Operations and Results. *Journal of the American Statistical Association, 88*(423), 1047-1060. https://doi.org/10.1080/01621459.1993.10476374

Hogan, H., & Robinson, G. (1993). *What the Census Bureau's Coverage Evaluation Programs Tell Us About Differential Undercounts.* Paper delivered at the 1993 Research Conference on Undercounted Ethnic Populations, May 5-7, Richmond VA.

Hogan, H., Cantwell, P. J., Devine, J., Mule, V. T., & Velkoff, V. (2013). Quality and the 2010 Census. *Population Research and Policy Review, 23,* 673-662. https://doi.org/10.1007/s11113-013-9278-5

Jensen, E., Benetsky, M., & Knapp, A., (2018). *A Sensitivity Analysis of the Net Undercount for Young Hispanic Children in the 2010 Census.* Poster presented at the Population Association of America annual Conference, Denver, CO. April

King, H., Ihrke, D., & Jensen, E. (2018). *Subnational Estimates of Net Coverage Error for the Population Aged 0 to 4 in the 2010 Census.* Paper presented at the Population Association of America annual Conference, Denver, CO. April

Martin, E. (2007). STRENGTH OF ATTACHMENT: SURVEY COVERAGE OF PEOPLE WITH TENUOUS TIES TO RESIDENCES. *Demography, 44*(2), 427-440. https://doi.org/10.1353/dem.2007.0014

Meyers, M., & Goerman, P. (2018). *Respondents Confidentiality Concerns in Multilingual Pretesting Studies and Possible Effects on Response Rates and Data Quality for the 2020 Census.* Paper delivered at the annual Conference of the American Association of Public Opinion Research, Denver CO., May 16-19, 2018

National Association of Latino Elected Officials (2018). "Census 2020: Research and Messaging," Power points from presentation on Sept 12.

National Association of Latino Elected Officials (2020). *Countdown to Census 2020: New Data on Latino Perception.* Retrieved February 10, 2020, from https://naleo.org/COMMS/PRA/NEF%20Census%202020%20_2_10_FINAL.pdf

O'Hare W. P. (2019a). *Differential Undercounts in the U.S. Census: Who is Missing?* Springer publishers. Retrieved from https://link.springer.com/book/10.1007/978-3-030-10973-8

O'Hare, W. P. (2019b). *Understanding Who Was Missed in the 2010 Census.* Population Reference Bureau, Washington DC. Retrieved February 10, 2020, from https://www.prb.org/understanding-who-was-missed-in-the-2010-census/

O'Hare, W. P. (2020). *The Politicization of the 2020 Census.* PAA Affairs, Fall 2020, The Population Association of America, Washington DC. Retrieved from https://higherlogicdownload.s3.amazonaws.com/POPULATIONASSOCIATION/3e04a602-09fe-49d8-93e4-1dd0069a7f14/UploadedImages/Documents/PAA_Affairs/PAA-Fall_20_.pdf

Prewitt, K. (2020). Oral Presentation before the House Committee on Oversight and Reform, U.S. House of Representatives. Retrieved July 29, 2020, from https://oversight.house.gov/sites/democrats.oversight.house.gov/files/documents/Testimony%20Prewitt.pdf

Robinson, G. J., Bashir, A., Das Dupta, P., & Woodward, K. A. (1993). Estimates of Population Coverage in the 1990 United States U.S. Decennial Census Based on Demographic Analysis. *Journal of the American Statistical Association*, *88*(423), 1061-1071. https://doi.org/10.2307/2290801

Salvo, J. (2020). *Weekly, Report on 2020 Census Self-Response Rates issued on May 14," New York city Department of City Planning – Population Division.* Retrieved from file:///C:/Users/Owner/Documents/Bills%20Folder/RESPONSE%20RATES/Self-Response%20May%2014%20(1).pdf

Sullivan T. A. (2020). *Census 2020: Understanding the Issues.* Springer Press. https://doi.org/10.1007/978-3-030-40578-6

Swanson, D. A. (2019). *THE END OF THE CENSUS.* PAA Affairs, Population Association of America, Fall 2019. Retrieved from http://www.populationassociation.org/wp-content/uploads/PAA-Fall19.pdf

U.S Census Bureau (2017c). *Respondent Confidentiality Concerns*. MEMORANDUM FOR Associate Directorate for Research and Methodology (ADRM) Sept 20.

U.S. Census Bureau (2003a). *A.C.E. Revision II- Adjusted Data for States, Counties, and Places*. DSSD REVISION II MEMORANDUM SERIES #PP-60, U.S. Census Bureau, Washington, DC.

U.S. Census Bureau (2003b). "Census 2000 Mail Return Rates," Census2000 Evaluation A.7.b, Herbert Stackhouse and Sarah Brady, January 30, U.S. Census Bureau, Washington, DC.

U.S. Census Bureau (2011). *Same-Sex Couple Household Statistics from the 2010 Census.* WORKING PAPER NUMBER SEHSD-WP2011-26, MARTIN O'CONNELL AND SARAH FELIZ September 2011 https://www.census.gov/library/working-papers/2011/demo/SEHSD-WP2011-26.html

U.S. Census Bureau (2012a). *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. U.S. Census Bureau, Washington, DC.

U.S. Census Bureau (2012b). *2010 Census Coverage Measurement Estimation Report: Components of Census Coverage for the Household Population in the United States.* DSSD 2010 Census Coverage Measurement Memorandum Series #2010-G-04. U.S. Census Bureau, Washington, DC.

U.S. Census Bureau (2012c). 2010 Components of Census Coverage for Race Groups and Hispanic Origin by Age, Sex and Tenure in the United States, DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-E-51, U.S. Census Burau, Washington, DC.

U.S. Census Bureau (2012d). *2010 Census Mail Response/Return Rates Assessment Report.* 2010 Census Planning Memorandum Series, No. 198, U.S. Census Bureau, Washington, DC.

U.S. Census Bureau (2012e). *2010 Census Content Reinterview, Survey Evaluation Report.* REPORT NUMBER CPEX-206, June 13, GIANNA DUSCH and FRED MEIER, U.S. Census Bureau, Washington DC. https://www2.census.gov/programs-surveys/decennial/2010/program-management/5-review/cpex/2010-cpex-206.pdf

U.S. Census Bureau. (2014a). *Final Task Force Report: Task Force on the Undercount of Young Children.* Memorandum for Frank A. Vitrano, U.S. Census Bureau, Washington, DC. February 2.

U.S. Census Bureau (2014b). *2014 Planning Data Base, U.S. Census Bureau, Washington, DC.* Retrieved from https://www.census.gov/research/data/planning_database/2014/docs/PDB_Block_2014-11-20a.pdf

U.S. Census Bureau (2017a) 2020 Census Operational Plan: A Design for the 21st Century, Version 3.0, September, U.S. Census Bureau, Washington, DC. Retrieved from https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/2020-oper-plan3.pdf

U.S. Census Bureau. (2017b). *Respondent Confidentially Concerns and Possible Effects on Response Rates and Data Quality for the 2020 Census.* presentation by Mikelyn Meyers at the Census Bureau National Advisory Committee Meeting, November 2, 2017.

U.S. Census Bureau (2018a). *Technical Review of the Department of Justice Request to Add citizenship status Question to the 2020 Census*. Abowd, J. M, Memorandum for Wilbur L. Ross Jr., January 19, 2018.

U.S. Census Bureau (2018b). Understanding the Quality of Alternative Citizenship Data Sources for the 2020 Census U.S. Census Bureau White Paper, August 6, Brown, J.D., Heggeness, M. L, Dorinskii, S.M, Warren, L. and Yi, M CES 18-38, August 6. Retrieved from https://www2.census.gov/ces/wp/2018/CES-WP-18-38.p

U.S. Census Bureau (2019). 2020 Census Barriers, Attitudes, and Motivators Study Survey Report, Version 2.0, January 24, McGreeny, K., Kriz, B., Mullenax, S. Kail, L, Walejko, G., Vines, M., Bates, N. and Trejo Y. U.S. Census Bureau, Washington DC.

United Nations (2017). Principles and Recommendations for Population and Housing Censuses, Revision 3." United Nations, 2017. Retrieved from https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf

Velkoff, V. (2011). *Demographic Evaluation of the 2010 Census*, Paper presented at the 2011 Population Association of America Annual Conference, Washington, DC, March.

Word, D. L. (1997). *Who Responds/ Who Doesn't? Analyzing Variation in Mail Response Rates During the 1990 Census*, Population Division Working Paper No. 19, U.S. Census Bureau.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.