A Three-Phase Model of Census Capture

Douglas Olson
U.S. Census Bureau, Washington, DC 20233[1]

Abstract

This paper examines census captures as measured by the Census 2000 Accuracy and Coverage Evaluation. The premise of the research is that an accurate census person enumeration usually requires three successful census collections: the housing unit, a household population roster, and individual listing; and that a census non-capture can be understood as coming from a failure at one of these steps. A logistic model is developed to measure the three phases of capture, and some population groups known to show high non-capture rates are examined to understand which collection phase(s) are most important in causing the high rate.

Keywords: Census coverage measurement; logistic regression modeling

Introduction

The 2010 Census Coverage Measurement (CCM) program is the present decade's continuation of similar programs dating back to at least 1980. Traditionally, the primary focus of such efforts has been on estimating the census undercount, both nationally and among sub-populations such as ethnic groups, although results from the 2000 coverage measurement program, called Accuracy and Coverage Evaluation or A.C. E., suggest that overcounting is also a census accuracy issue worthy of equal attention. Another purpose of coverage measurement is to improve future censuses by identifying subpopulations with high omission rates, in the hope that better enumeration techniques can be developed for them. This paper is a developmental, exploratory effort to measure the "why" of census omission, as well as the "who," by associating the characteristics of persons omitted from, or otherwise captured incompletely in, the census with the phases of census operations.

Background

The 2010 CCM program includes only housing units and their residents (i.e., it excludes group quarters) and excludes the region the Census Bureau refers to as remote Alaska. In this paper, that universe will be considered the entire nation. This universe is identical to the one that was used in the 2000 A.C.E[2]. Census coverage measurement uses a capture-recapture methodology to estimate undercount or overcount in the Census. Under a traditional capture-recapture formulation, two samples of a population are taken, and the fraction of the population captured in the second sample, but not in the first, is assumed to be the omission rate of the first sample. In the CCM, the first sample is the Census (it is not designed as a sample, but under the capture-recapture formulation, it is treated as one) and the second sample is the independent enumeration consisting of an independent housing unit listing (IL for "Independent Listing") and Person Interview (or PI).

---

[1] This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

[2] Coverage measurement is also performed in Puerto Rico, but those tabulations are treated completely separately from the mainland United States program and not included here.

The IL and PI actually have multiple phases with different names, but for purposes of this paper, we can subsume all the related operations into those two categories.

The independent enumerations of both housing units and persons are called the housing unit and person P samples. The fraction of persons or housing units included in the P sample but that can not be linked into the Census represents an estimate of the non-capture rate of persons or housing units in the Census. The process of comparing the P sample to the Census is called "matching" and the match rate and non-match rate represent estimates of census capture and non-capture probabilities, respectively. Because the mathematical theory of capture-recapture depends on independence of the capture probabilities in the two samples, the Census Bureau has traditionally divided each sample into post-strata defined on demographic or sometimes census operational characteristics, within which the census capture probabilities are treated as homogeneous for purposes of calculating "dual system estimates." The dual system estimate, or DSE, can be formulated:

$$DSE_i = Census_i \text{ x } Prob(CE_i) / Prob(Match_i)$$

where i is any person (or housing unit) in the Census, and CE_i and $Match_i$ are the individual's estimated correct enumeration and match probabilities, respectively. Estimation of correct enumeration probabilities are outside the scope of this paper, but P-sample persons can only match to correctly-enumerated census persons for the purposes used in this paper. In 2010, the Census Bureau is planning to use statistical modeling, in particular logistic regression, in place of the traditional post-stratification.

In an ideal capture-recapture survey environment, a capture in the second phase that wasn't in the first phase should always represent an omission from the first phase. In the Census, it is entirely possible to enumerate someone in the first phase and still not be able to match to them if they are re-captured in the second phase. For example, the amount of information gathered for their census enumeration might be insufficient to be matchable. Also, many census persons for whom no individual information is available will be represented in the census count by an imputed person with characteristics that are as similar as can be estimated. Hence, a failure to match a P-sample person to a census enumeration does not imply a complete omission from the Census, but simply the absence of a potentially matchable person. Research into reasons for census persons being unmatchable are being developed separately for 2010 coverage measurement estimation (Mule 2008). This paper concentrates solely on estimating non-matches, some of which are complete omissions, but most of which probably represent a census person who was not matchable. Future research will hopefully develop a similar model for unmatchable census persons, to isolate the true omissions from the other causes.

Census enumeration procedures

Ignoring some background operations, technical details, and unusual cases, a census enumeration normally proceeds from the following steps:

- the housing unit is listed on the census address file
- the housing unit is delivered a census form, either by mail or by an enumerator
- the residents of the housing unit, if occupied, should return their form by mail
- if the household does not return its census form, it is contacted for nonresponse follow-up
- if repeated contact attempts fail, a household population and characteristics are imputed

There are a relatively small number of cases that deviate from the normal sequence:

- new housing units constructed after the census address list was completed can be added after the Census begins
- specialized enumeration techniques are used in some of the hardest to count areas (extremely sparse populations, Indian Reservations, areas with unusual mail distribution)
- some persons self-initiate their census inclusion (Be Counted forms)

From this outline, a "normal" census person enumeration requires three phases:

- the housing unit must be captured
- the household must complete a census roster
- the individual must be listed on that roster

In this paper, we will attempt to associate census non-captures (non-matches) with the failure of one of these phases.

The causes of census capture failure

A great deal of research is conducted during and after every Census, and every census testing phase, to determine possible reasons for capture failure. Among those commonly identified that can be associated with the three enumeration phases include:

The Census fails to enumerate the housing unit:

- addresses that are difficult to find (most common in rural and some urban areas)
- multiple entrances within a structure, causing failure to identify all units within a multi-unit dwelling
- newly constructed units, not on address lists in time for census inclusion

The household does not complete a (good quality) census roster:

- more common in households with young householders, renters, or ethnic minorities
- after repeated interview attempts, might result in a proxy interview with a non-household member, which can result in reported data of lower quality
- might result in a whole-household imputation

The household roster does not include all occupants, or gives incomplete information:

- most common among non-nuclear family members (grandchildren, boarders, roommates, roommate's children)
- in 2000, the length of the census form unintentionally encouraged complete reporting of only six household members

How the Census Bureau estimates non-captures

Census coverage measurement programs, since at least 1980 follow the same outline of steps:

- a sample of small geographic areas, called block clusters, is selected to represent the nation as a whole
- in those sample areas, enhanced census enumeration operations create an independent enumeration of persons and housing units called the P sample
- the results of the P sample are combined with the actual Census using a capture-recapture methodology, called "Dual System Estimation"

- traditionally, the Dual System Estimates have been calculated within post-strata defined by demographic characteristics
- in demographic groups that are sufficiently large, census operational characteristics, such as Type of Enumeration Area or Tract Return Rate, can also be used to create post-strata

Since the use of post-strata limits the number of characteristics that can be used, the Census Bureau is researching (and expecting to use) logistic modeling for 2010 coverage measurement.
Using modeling in place of post-stratification should allow the use of characteristics for population groups that had been too small to support them. For example, in the 2000 A.C.E. the characteristic "region" (four parts of the country) had been applied only to white homeowners because, being the largest population group, they had sufficient sample size to support being sub-divided into groups that small. A full-crossing of the post-stratification variables used in the 2000 A.C.E. would have produced 3,136 post-strata, but yielded only 416 final post-strata after collapsing for minimum cell size requirements (Census Bureau 2004).

Modeling characteristics

To be usable for Dual System Estimation, a characteristic must be available for every person or housing unit in the Census and in the P sample. This limits the usable characteristics to two types: a) the questions the Census and independent Person Interview (PI) ask of each person (age, sex, race/ethnicity, tenure, relationship within household) or directly observe (number of dwellings in a multi-unit structure); or b) characteristics of the geographic area (Metropolitan Statistical Area, tract return rate, region). Characteristics that cannot be used include:

- those that exist only for the Census and not the P sample (e.g., census form type)
- those that exist only for the P sample and not the Census (e.g., moved since Census Day)
- those that will not generally be consistent in the Census and PI (e.g., proxy interview)

The last category is one reason that household relationship has not been used in post-stratification in the past. Since the independent PI interview is not necessarily with the same household member who completed the census form, the rostered persons could have a different reported relationship to those two reference persons. For 2010, the coverage measurement program is researching the possibility of using a household relationship covariate, that will attempt to select the most likely census reference person of all those rostered in the household in the independent PI, to construct relationships for all household members that are as nearly consistent as possible to those in the Census.

For purposes of the research included in this paper, a few obvious cases of inconsistent relationship reporting have been rearranged, such as when an older teen or very young adult was the independent interview respondent, and reported a household full of parents and siblings. In this case it is more likely that the Census reported a household with a householder or couple and their children. Also, categories of relationships for estimation, such as nuclear family member, have been designed to minimize category changes if each spouse is the reference person in the one of the interviews. For instance, stepchildren are categorized like natural-born children, so that their relationship category doesn't change depending on which spouse is the reference.

The characteristics used in this research include some used in past post-stratifications, some very similar to those but employed in a different way, and some that are completely new.

<u>Variables that have been commonly used in DSE post-stratification</u>

- race/Hispanic origin domain: seven categories
- Metropolitan Statistical Area (MSA) population size

- Type of Enumeration Area (TEA)
- region of the country: four areas
- sex (for adults 18 and older only)
- tenure (owner/renter)
- multi-unit structure size (single unit, 2-9, 10 plus)

Variables used in new way
- age (as a continuous variable called "splines;" previously was used in age-range groups)
- tract census return rate (as continuous rate; had been used in high/low groups)

New variables
- household relationship categories (nuclear family member, other relative, non-relative)
- age of householder/reference person
- seven-person household indicator (if more than six persons in the household)
- relationships present in the household (person lives alone, spouse present, other)
- tract rate of new construction (housing units added since census address canvassing)

Many of these characteristics have obvious and established correlation with census non-capture rates. Census omissions are more common, for instance, outside of MSAs than in them, and are more common in large MSAs than in medium or smaller ones (Kostanich 1999). Some of the characteristics above seem obviously associated with a particular one of these phases. The non-mailout type of enumeration area, for instance, has higher housing unit omission rates than mailout areas (Griffin 2006). Non-relatives are frequently not captured on a household population roster.

A word is in order about the age splines. The relationship of age to match rate is complicated with low match rates at ages near zero and 20, and high rates for school-age children and adults over 40. Under post-stratification, age range categories, 0-17, 18-29, 30-49 and 50-plus were used to partition the population. Since modeling allows the use of continuous variables, either linear or quadratic "splines" are fit into four ranges (which overlap at their endpoints for continuity): 0-17, 17-20, 20-50, and 50-80, the latter value being a topcode (Schellhamer 2009).

In this research, we will use the 2000 A.C.E. P-sample person research file as the primary input, along with the results of 2000 A.C.E. housing unit matching. This file includes all persons in the A.C.E. P sample, although some have been removed by zero-weighting, with their matching results and some supplemental adjustments resulting from the A.C.E. Revision II (Mule and Olson 2005). Since each person in the P sample resided in a housing unit that is also in the P sample, the results from housing unit matching can be merged onto the same file (Burcham 2002). In this way, each person in the P sample can have associated with them three match probabilities:

- the probability that their housing unit matched to one in the Census
- the probability that any person in the household matched to a census person (approximates whether a good household roster was captured)
- the probability that they individually matched

Modeling results

Table I
Wald Chi-Square Test Statistics of Modeling Characteristics

	Housing Unit			Any Person in Household			Individual		
	Parameter	std err	Wald	Parameter	std err	Wald	Parameter	std err	Wald
Intercept*	2.3980	0.2342	104.9	1.4849	0.1257	139.6	0.6676	0.0871	58.7
Renter	-0.0793	0.0383	4.3	-0.2354	0.0208	128.6	-0.1973	0.0145	184.3
Male 18+	0.0091	0.0226	0.2	0.0028	0.0122	0.1	-0.0162	0.0094	3.0
Tract Return Rate	1.6663	0.3740	19.8	1.5858	0.1827	75.3	1.3843	0.1258	121.1
Householder Age	0.0045	0.0037	1.4	0.0120	0.0020	37.2	0.0077	0.0013	36.5
New Constr Rate	-2.6560	0.3648	53.0	-1.7347	0.2521	47.3	-1.0864	0.2089	27.0
Non-Nucl Relative	0.0303	0.0502	0.4	0.1245	0.0267	21.7	-0.0308	0.0167	3.4
Non-Relative	-0.0329	0.0742	0.2	-0.0545	0.0356	2.3	-0.4447	0.0222	400.5
Small Multi-Unit	-0.4658	0.0657	50.3	-0.2242	0.0315	50.6	-0.1907	0.0239	63.4
Large Multi-Unit	0.2403	0.0921	6.8	0.0706	0.0358	3.9	0.0414	0.0274	2.3
Lives Alone	-0.1969	0.0823	5.7	-0.4629	0.0402	132.3	-0.3255	0.0333	95.7
No Spouse Present	-0.2169	0.0689	9.9	-0.3045	0.0385	62.6	-0.1950	0.0266	53.7
Adult in 7+ HH	-0.1016	0.2124	0.2	0.3749	0.1411	7.1	-0.5983	0.0508	138.6
Child 7th Oldest	0.1118	0.1833	0.4	0.2542	0.1238	4.2	-1.3548	0.0512	700.5
Race/Ethnicity									
Indian on Reserv	-0.0266	0.1904	0.0	0.0788	0.1395	0.3	0.0102	0.0858	0.0
Indian off Reserv	0.2461	0.2627	0.9	0.0539	0.1402	0.1	-0.0062	0.1051	0.0
Hispanic	0.1262	0.1273	1.0	0.0375	0.0738	0.3	0.0339	0.0461	0.5
Black	0.3878	0.1247	9.7	-0.0158	0.0710	0.0	-0.0761	0.0458	2.8
Pacific Islander	-0.6888	0.4444	2.4	-0.3525	0.2980	1.4	-0.2291	0.1782	1.7
Asian	-0.1242	0.1856	0.4	0.0640	0.0986	0.4	0.0508	0.0584	0.8
MSA/TEA									
Mailout, Medium	0.3651	0.0901	16.4	0.1535	0.0419	13.4	0.1061	0.0292	13.2
Mailout, Small	0.2834	0.1052	7.3	0.1671	0.0542	9.5	0.1073	0.0377	8.1
Mail, non-MSA	-0.1102	0.0983	1.3	0.0462	0.0582	0.6	0.0339	0.0416	0.7
MSA, Non-Mail	-0.1965	0.1043	3.5	-0.0884	0.0717	1.5	-0.0632	0.0509	1.5
Non-Mail/MSA	-0.6328	0.0888	50.8	-0.2986	0.0592	25.4	-0.2218	0.0440	25.5
Region									
Northeast	-0.1393	0.0755	3.4	-0.0073	0.0400	0.0	0.0146	0.0283	0.3
Midwest	0.2665	0.0700	14.5	0.1350	0.0366	13.6	0.1099	0.0251	19.2
South	-0.0929	0.0631	2.2	-0.1167	0.0337	12.0	-0.1013	0.0241	17.7
Age Splines									
0-17 Linear	0.0172	0.0181	0.9	-0.0029	0.0108	0.1	0.0246	0.0077	10.2
0-17 Quadratic	0.0163	0.0611	0.1	-0.1309	0.0348	14.2	0.0215	0.0242	0.8
17-20 Linear	-0.0178	0.0479	0.1	0.1263	0.0263	23.0	-0.0063	0.0185	0.1
20-50 Linear	0.0014	0.0116	0.0	0.0061	0.0064	0.9	-0.0052	0.0050	1.1
20-50 Quadratic	-0.0007	0.0012	0.3	0.0004	0.0007	0.4	-0.0012	0.0005	6.3
50-80 Linear	0.0001	0.0003	0.2	-0.0001	0.0002	0.1	0.0002	0.0001	2.7

*Intercept term absorbs one category of each categorical variable: Owner, Female 18+, White/Other race, Nuclear Family Relatives, Single Unit Structure, Spouse Present, Mailout/Large MSA, West region

The characteristics defined a model with 33 covariates: 24 categorical (0/1) variables, plus continuous variables for age of individual (6), age of householder, tract return rate and tract new construction rate. The models were fit for main effects using SAS Proc SurveyLogistic using a logit link, with the A.C.E. block cluster, the primary sampling unit, the only level of clustering used to estimate standard errors for the model. The housing unit, household, and individual match rates for each person were fit separately and independently. Each P-sample person was considered an observation in each model and weighted at their P-sample estimation weight from Census 2000 A.C.E. Revision II (Mule and Olson 2005), so a housing unit that included five residents appeared in the model five times.

Because the overall non-match rate for persons is only 8 percent, the most important modeling characteristics are usually negative predictors. The measure of significance of the intercept term is not of interest, because it is known that the match rate is not approximately 50 percent. The most important characteristics in each model, as measured by their Wald Chi-Square test statistics (with one degree of freedom) are ones that should be expected:

Housing unit non-match
- New Construction Rate (Wald=53.0)
- Non-Mailout, non-MSA area (50.8)
- Small Multi-unit dwelling of 2-9 units (50.3)
- Tract Return Rate (19.8)

Whole-household non-match
- Lives Alone (132.3)
- Renter (128.6)
- Tract Return Rate (75.3)
- Small Multi-unit dwelling (50.6)

Individual person non-match
- Child seventh (or more) oldest person in household (700.5)
- Non-relative (400.5)
- Renter (184.3)
- Adult in 7+ person household (138.6)
- Tract Return Rate (121.1)
-

Tract return rate shows up in all the lists, for reasons that could be either methodological or causal. The Census Bureau calculates tract return rate by dividing the actual number of census returns by the number of non-vacant housing units in the tract. Which implies that the denominator of the calculation includes housing units that were newly constructed and hence could not have returned a census form because they were never sent or given one. Hence, there is probably a significant correlation or confounding between the tract return rate characteristic and the new construction characteristic, even though census form return rates *per se* should not affect housing unit inclusion. For similar reasons, the tract return rate is highly correlated with many of the other characteristics in the models. Therefore, an interaction was added to the model between the tract return rate and all the other covariates.

Although the above model was useful for flagging significant characteristics, a residual analysis for model error suggests it has too great a propensity to estimate very low rates. An examination of the lowest modeled rates for persons revealed that the lowest 13,000 estimated individual person match rates averaged about 5 percentage points lower than the actual observed match rates for those persons. Experimentation suggested that the low rates could be appropriately modeled by changing the linking function to the complimentary log-log. That became the final model, along with the tract return rate interactions.

Some well-known populations with low match rates

Once the three models have been developed, a capture rate can be estimated for each individual. A comparison of observed and modeled rates can help to point out what phases of collection are most important in causing census non-captures of persons with various characteristics of interest. The purpose of the research in this paper is not to develop a better modeling method for calculating population estimates. It is designed as an exploratory technique to link census non-captures with the collection phases most likely to have produced them. When the 2010 CCM modeled results have been completed, an analysis similar to the one below, which uses 2000 A.C.E. results, will hopefully be useful in developing improved methods for future censuses. Since the investigations below are intended solely as examples in the use of a technique, no actual inferences about populations are implied by or should be inferred from these results.

Zero-year olds

Although few people would describe an infant as being "zero years old," the Census Bureau classifies persons by age at last birthday, which places an infant less than one year old into this category. Children in large households of seven or more persons have been excluded due to the extreme nature of the reporting pattern of very young children in such households.

Capture Probabilities

	Zero Year Olds				One/Two Year Olds			
	Obs	s.e.	Mod	s.e.	Obs	s.e.	Mod	s.e.
Structure	.974	.003	.975	.002	.977	.002	.975	.002
Household	.948	.004	.947	.003	.950	.003	.948	.003
Individual	.901	.005	.911	.003	.920	.003	.913	.003

The overall match rate for zero year olds in A.C.E. 2000 was .901, considerably lower than the .920 rate for one- and two-year olds; the standard error (s.e.) of the difference is .005. Overall, both the observed and modeled capture rates for the two age groups are nearly the same for their structure and household capture probabilities. While the model predicts almost the same individual capture rate for the two age groups, the observed rate for zero-year-olds is distinctly below their modeled rate. A breakdown by household relationship largely explains the difference:

Capture Probability of Zero-Year Olds

	Child of Householder				Other Household Relationship			
	Obs	s.e.	Mod	s.e.	Obs	s.e.	Mod	s.e..
Structure	.976	.003	.975	.002	.966	.012	.974	.002
Household	.951	.004	.947	.003	.931	.012	.949	.003
Individual	.915	.005	.918	.003	.804	.016	.865	.003

The observed capture rates of zero-year-olds who are the child of their householder track the modeled rates very well. But those who have a different household relationship (for example grandchildren, or children of non-relatives such as roommates) have a particularly high within-household non-match rate compared to their modeled value. The observed within-household
non-match rate for "other relationship" infants was .127 (.931 - .804) while the modeled rate was .084, a difference of .043 (s.e.=.005).

The modeled structure and household capture rates for the two groups above were very similar, and the observed rates for the householder's children mirror those closely, while the observed rates for the other relationship children were somewhat lower in all phases, suggesting that the presence of a zero-year-old who is not the householder's child occurs disproportionately among hard-to-count areas and populations. The high within-household non-match rate for other relationship zero-year-olds suggests that householders completing census rosters have a high tendency to omit them.

Hispanic Male Renters 18-29

Under the post-stratification used in the 2000 A.C.E., two of the four highest non-match rates estimated for any of the 416 final post-strata were for groups consisting of Hispanic male renters age 18-29 (18-29 was an age grouping used to partition every demographic group, except those too small to support it). The model used in this research includes the characteristics Hispanic, male and age, as well as other characteristics that are common among this group such as non-related household status and living in a low census return rate area. Since the number of Blacks in Census 2000 was very similar to the number of Hispanics, it is convenient to compare them side-by-side:

Capture Probability of Hispanic and Black Male Renters18-29

	Hispanic				Black			
	Obs	s.e.	Mod	s.e.	Obs	s.e.	Mod	s.e.
Structure	.969	.005	.972	.003	.976	.006	.976	.002
Household	.882	.008	.899	.006	.881	.006	.882	.004
Individual	.761	.011	.796	.007	.787	.011	.778	.004

The model reflects the phases of capture for Blacks very well, showing only minor variations from the observed probabilities at each phase. For Hispanics, the structure capture rate is modeled a bit high (difference = .003, s.e. = .005), the modeled household capture rate is high by .017 (s.e.=.004) and the individual rate is modeled high by .035 (.s.e=.008), suggesting that a model used to publish estimated undercount rates for these population groups (which the model in this paper is not) would need to include interactions for this group, whose capture rates are not adequately described by their individual characteristics.

Adults 18-24

The youngest adults, those ages 18-24, have traditionally been a high undercount group. (They are also a high overcount group, due to being counted in multiple residences, although that is beyond the scope of this paper.) They are frequently counted in the Census as the adult children at their parent's house, and other times at households alone or with similarly-aged persons, usually as renters. Since renters and adult children both have a tendency to be individually omitted within captured households, persons in this age group should be expected to have high non-match rates regardless of their living situation, and they do:

Capture Probability of Adults Age 18-24
Oldest Person in Household is Aged:

	Thirty-three or more				Thirty-two or less			
	Obs	s.e.	Mod	s.e.	Obs	s.e.	Mod	s.e.
Structure	.979	.002	.980	.002	.970	.002	.970	.002
Household	.955	.003	.953	.002	.892	.003	.897	.002
Individual	.891	.004	.885	.003	.829	.004	.833	.003

Adults 18-24 who live with someone 33 years or older have a much higher capture rate than those who don't, .891 to .829 (s.e. of difference = .005). The within-household non-match rates are nearly the same between the two groups, at .063 and .064 (s.e. of each = .003). But those with a 33-plus year-old housemate (usually a parent, who's distinctly older than that) belong to households with a capture probability of .955, while those whose oldest housemate is 32 or less, belong to households whose capture probability is less than .900. So the most important reporting distinction for this group is household-level reporting, although the difference in the structure-level capture rate is also significant (.979 vs. .970, s.e. of difference = .003).

Conclusions

No extraordinary results have been found here, and none should have been expected. The research describes an effort to statistically explain differences in capture probability due to causes that are well-known and that have already been thoroughly researched. The research in this paper suggests that modeling could be used to detect effects that had not previously been capturable under the old post-stratification methodolgy, and that the rates so estimated usually accurately reflect the probabilities under study.

The research further suggests that census capture problems associated with populations known to have high non-capture rates can be traced to particular phases of the census process, and that such knowledge could hopefully be used to design improvements into future censuses.

References

Burcham, Joseph (2002). "Input Files for Housing Unit Dual System Estimates," DSSD Census 2000 Procedures and Operations Memorandum Series #Q-59.

de le Puente, Manuel (1994). "A Multivariate Analysis of the Census Omission of Hispanics and Non-Hispanic Whites, Blacks, Asians and American Indians: Evidence From Small Area Ethnographic Studies."

Griffin, Richard (2005). "Net Error Estimation for the 2010 Census," DSSD 2010 Census Coverage Measurement Memorandum Series #2010-E-01.

Griffin, Richard (2006). "Accuracy and Coverage Evaluation: Housing Unit Dual System Estimation Results," DSSD Census 2000 Procedures and Operations Memorandum Series Q-91.

Kostanich, Donna (1999). "Accuracy and Coverage Evaluation Survey: Logistic Regression Modeling for Poststratification Variable Selection," DSSD Census 2000 Procedures and Operations Memorandum Series #Q-6.

Mule, Tom and Richard Griffin (2008). "2010 Census Coverage Measurement Estimation Methodology," DSSD 2010 Census Coverage Measurement Memorandum Series #2010-E-18.

Mule, Tom and Doug Olson (2005). "A.C.E. Revision II - Computer Specifications for Research Files of A.C.E. Revision II Person Data," DSSD 2010 Census Coverage Measurement Memorandum Series#2010-E-02.

Schellhamer, Teresa, Thomas Mule and Andrew Keller (2009). "Using Partial Residual Plots in Modeling Covariates for Net Coverage Estimation," DSSD 2010 Census Coverage Measurement Memorandum Series #2010-E-15.

U.S. Census Bureau (2004). "Accuracy and Coverage Evaluation of Census 2000: Design and Methodology."

U.S. Census Bureau (2006). "Housing Vacancies and Homeownership Annual Statistics: 2005."