

## Findings from the Census 2000 Accuracy and Coverage Evaluation

Howard Hogan and David C. Whitford, U.S. Bureau of the Census  
Decennial Statistical Studies Division, Washington, D.C. 20233<sup>1</sup>

**Key words:** Census 2000, Accuracy and Coverage Evaluation

### What does this article discuss?

The Census Bureau designed the Accuracy and Coverage Evaluation (A.C.E.) for Census 2000 to permit correction of the initial census results to account for systematic patterns of net undercount. This paper presents findings of the A.C.E. and documents the quality of its operations.

The A.C.E. was an independent survey based on the concepts of dual system estimation (See Hogan, 2000). It was designed to produce reliable estimates of net under- and/or over-count for the entire census, for geographic areas of interest in the census, and for any of a host of demographic characteristics—age, race, sex—for which one might desire census coverage statistics. We will present some of the coverage statistics produced by the A.C.E. survey and will compare these results with those of the 1990 Post-Enumeration Survey. This paper also reviews key A.C.E. design feature operations in order to begin assessing the overall success of the design. A more detailed assessment of A.C.E. accuracy, especially in comparison to census accuracy, is presented in a companion article by Kostanich, Singh and Hogan (2001).

In February of 2000, the senior Census Bureau professional staff reviewed the results of the A.C.E. and concluded that, in light of the information then available, they could not recommend adjusting the census files then to be released. Although they concluded that the survey was of high quality, the inconsistency between the demographic analysis estimates and the A.C.E. estimates raised the possibility of an as-yet undiscovered problem in the A.C.E. The results presented in this paper should be considered in light of this possibility.

### What is the Census net undercount as measured by the A.C.E.?

The A.C.E. results indicate that Census 2000 reduced both net and differential undercoverage compared to the levels measured by the 1990 Post-Enumeration Survey (PES). These percentages are for the total U.S. population:

- The net national undercount is estimated to have been reduced from the 1990 rate of 1.6 percent (0.2 percent standard error) to 1.2 percent (0.1 percent standard error).
- The estimated percent net undercount for the Black population dropped from 4.4 percent (0.6 percent standard error) to 2.1 percent (0.3 percent standard error), and
- The estimated percent net undercount for the Hispanic population dropped similarly from 5.0 percent (0.8 percent standard error) to 2.8 percent (0.4 percent standard error).
- The estimated percent net undercount for children dropped from 3.2 percent (0.3 percent standard error) to 1.5 percent (0.2 percent standard error).

Table 1 gives the 1990 and the 2000 results by race and Hispanic origin. Because of differences in the way race data were collected, the comparison between censuses is not exact.

The improvements demonstrated in Census 2000 do not mean that complete coverage has been achieved or that differential coverage has been eliminated. On the contrary, the A.C.E. indicates that Census 2000 perpetuated longstanding patterns of differential coverage, with minority groups exhibiting lower coverage rates.

---

<sup>1</sup> Howard Hogan is Chief, and David C. Whitford is Assistant Division Chief for Statistical Programs Management, Decennial Statistical Studies Division U.S. Census Bureau. This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

**Table 1: Total Population Percent Net Undercount by Race and Hispanic: 2000 and 1990**

Characteristic	2000		1990	
	Rate (%)	Standard Error (%)	Rate (%)	Standard Error (%)
<b>Race</b>				
Total	1.2	0.1	1.6	0.2
One Race	1.1	0.1		
White	0.8	0.1	0.9	0.2
Black or African American	2.1	0.3	4.4	0.5
American Indian or Alaska Native	3.4	0.8	4.5	1.2
Asian, Native Hawaiian or Other Pacific Islander	1.1	0.6	2.3	1.4
Asian	1.0	0.6		
Native Hawaiian or Other Pacific Islander	4.4	2.4		
Some Other Race	3.0	0.4		
Two or More Races	1.9	0.2		
<b>Race Alone or In Combination With One or More Other Races</b>				
White	0.8	0.1		
Black or African American	2.1	0.3		
American Indian or Alaska Native	2.6	0.6		
Asian	1.1	0.6		
Native Hawaiian or Other Pacific Islander	3.7	1.9		
Some Other Race	2.9	0.4		
<b>Hispanic and Race</b>				
Hispanic	2.8	0.4	5.0	0.8
Non-Hispanic	0.9	0.1	1.2	0.2
White alone	0.7	0.1	0.7	0.2

The 2000 data are for 2000 census race categories for the total population. As a result, these numbers may differ from the numbers derived directly from the A.C.E. 448 dual system estimates which are for household estimates and estimation race/ethnicity domains. The 1990 data are for 1990 census race categories for the total population. As a result, these numbers may differ from the numbers derived directly from the PES 357 dual system estimates which included most noninstitutional Group Quarters in addition to the household population.

The Census 2000 percent net undercount for the Black and the Hispanic populations, 2.1 percent (0.3 percent standard error) and 2.8 percent (0.4 percent standard error) respectively, remain significantly higher than the national total. The same is true for the American Indian/ Alaska Native population with an undercount of 3.4 (0.8 percent standard error). The undercount for the Asian population (1.0 with a standard error of 0.6) is now indistinguishable from that of the Non-Hispanic white population (0.7 with a standard error of 0.1).

The dual system estimate from the A.C.E. was computed for each of 448 post-strata.<sup>2</sup> These post-strata comprise the entire housing unit/household universe (except remote Alaska). One of the most important findings of the A.C.E. is the reduction of the undercount for children. The A.C.E. measured an undercount of the total 0-17 year old population of 1.5 percent (0.2 percent standard error) compared with 3.2 percent (0.3

<sup>2</sup> Because of collapsing of cells due to small sample size, separate estimates were computed for only 416 post-strata, with the same undercount then assigned to each of the collapsed post-strata.

percent standard error) in 1990.

Summing over age and sex allows one to analyse the results for 64 categories. Since these post-strata are mutually exclusive and designed to be approximately homogeneous with respect to the undercount, they provide a useful approach to understanding the undercount of the household population as directly measured by the A.C.E. Appendix tables 1 and 2 give the results of the A.C.E. for these 64 post-stratum groups, together with their standard errors.

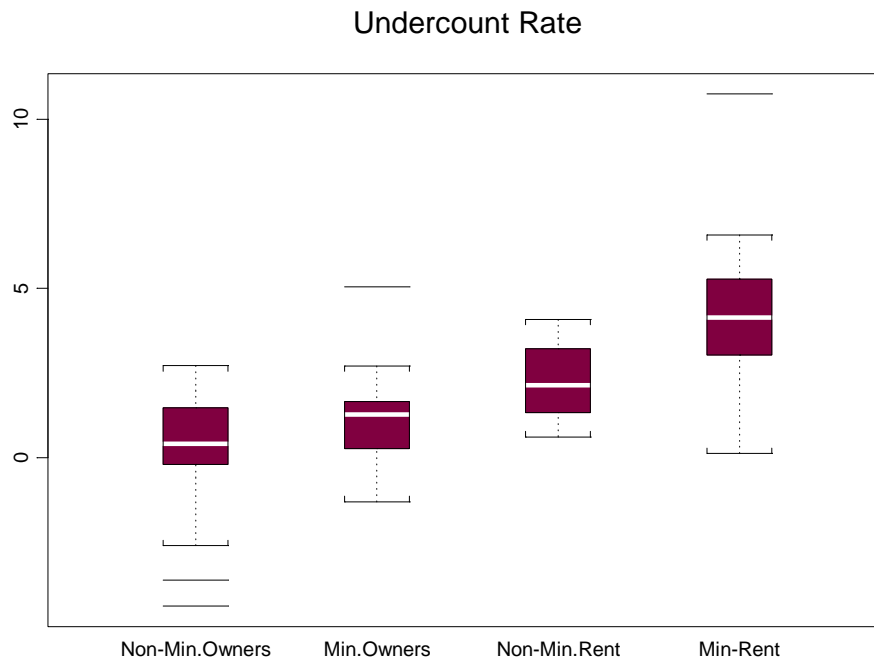
Tenure continues to be an important characteristic related to coverage. The A.C.E. indicates that the pattern of differential coverage continues despite improvements in Census 2000. The percent net undercount for non-owners was 2.7 percent (0.3 standard error) as compared with an estimated net undercount for owners of 0.4 percent (0.1 standard

error). This is a distinct improvement over the percent net undercount for non-owners in the 1990 census, which is estimated at 4.5 percent (0.4 standard error). Still, the A.C.E. indicates that the estimated undercount for this population is significant as well.

In addition, the undercount for minority renters also remains high. The non-owner undercount for non-Hispanic Blacks was estimated to be 3.6 (0.5 standard error), for Hispanics 4.3 (0.6 standard error), for Asians 1.6 (1.0 standard error), for Hawaiians and Pacific Islanders 6.6 (4.1 standard error), and for American Indians not on reservations 5.6 (2.0 standard error).

Figure 1 displays the distribution of the percent net undercounts by tenure for the 64 post-stratum groups. It is clear from this graph that tenure is becoming equally important with minority status in predicting the undercount.

Figure 1. Percent Net Undercount for 64 Post-strata Groups Classified by Tenure and Minority Status.



One interesting aspect of the A.C.E. results is the amount of net over-count measured in many groups. Of the 64 post-stratum groups, eleven had measured net over-counts, and 53 measured net undercounts, although a 90 percent confidence interval would contain zero for 37 of these 64 groups. At the detailed post-stratum level, 141 had measured net over-counts and 307 measured net

undercounts. Although the net national undercount was 1.18 percent or 3.26 million, in fact any adjustment would have resulted in adding 4.26 million people to account for measured net undercounts and subtracting 1 million to account for measured net over-counts, for a gross change (sum of the absolute net errors) of 5.26 million.

## How well did the A.C.E. operations work?

### Sampling:

The A.C.E. sample was nearly twice as large as that used in 1990, which, together with several modifications in the sample design resulted in a much smaller sample variance.

Comparisons by post-strata between 1990 and 2000 are necessarily inexact as the universe differs (2000 includes only the household population) and the exact post-stratum definitions are different. Still, some comparisons are instructive. The standard error for owners was reduced from 0.2 percent to 0.1 percent, and the standard error for non-owners fell from 0.4 percent to 0.3 percent. The measured standard error fell for all comparable race/origin groups and for each age/sex group. The estimated standard error was comparatively high for the two groups estimated separately for the first time: Hawaiian and Pacific Islanders (2.8 percent) and American Indians and Alaskan Natives living off reservation (1.3 percent). The estimated standard error for American Indians living on reservations fell dramatically from 5.3 percent to 1.2 percent. The standard error for non-Hispanic Asians was 0.6 percent. For Hispanics, it was 0.4 percent, and for non-Hispanic Blacks it was 0.3 percent.

Appendix Tables 1 and 2 give the estimated percent net undercount and standard errors for the 64 major post-stratum groups. The standard errors for several groups are above 1 percent and for a few small groups are up to about 4 percent. Because the populations of these groups are small, their high variances will have only limited impact on geographic variance.

At the state level, the median coefficient of variation (CV) for state population totals dropped from 0.4 percent in 1990 to 0.2 percent in 2000. More important, the median CV for the congressional districts dropped from 0.5 percent to 0.3 percent. Similar drops in the CV of 40 percent to 50 percent were estimated for counties and places larger than 100,000.

This decrease in sampling variance is due to the much larger sample size of the A.C.E. relative to the PES: 300,913 housing units in 11,303 clusters for the A.C.E., versus 165,000 housing units in approximately 5,000 clusters for the 1990 PES. Better measures of population size in the sample selection of block clusters, better subsampling methods, better methods of treating "small blocks," and a reduction in the variability of sampling weights all contributed to this reduction.

### Interviewing

The A.C.E. operations began around the Nation in late 1999 when the Census Bureau listed all the addresses of

housing units in the A.C.E. sample clusters. The list was then compared with the census list and a final workload of the addresses to be interviewed was produced.

One change from 1990 was the introduction of computer assisted personal interviewing on laptop computers for the A.C.E. interview and, additionally, doing some of these interviews by telephone. The Census Bureau implemented a telephone program to enhance the efficiency and quality of the A.C.E. interview. The Census Bureau believed that shortening the elapsed time from Census Day to the A.C.E. enumeration would improve data quality and that beginning interviewing early in a more easily controlled environment would allow the A.C.E. supervisors to gain valuable experience in conducting interviews and in operating their laptop computers before training the enumerators. (Figure 2 in the Appendix displays Gantt charts depicting how the A.C.E. survey fit into the overall census schedule and how the timing of the A.C.E. interview and census follow up operations fit together.) The Census Bureau designed this process to maintain the independence between the A.C.E. and the other Census 2000 operations.

A.C.E. interviewing was an operational success. The A.C.E. interviewing finished on schedule by September 1, 2000, in every local census office except the Hialeah office, where census nonresponse follow up census interviewing finished late (September 11, 2000) due to local difficulties. The utilization of the Computer Assisted Personal Interviewing (CAPI) instrument was a major accomplishment. The timely interviews allowed the Census Bureau to conduct a quality assurance program more efficiently than when using paper and pencil and to have an orderly completion of interviewing. There were no major disruptions or delays introduced.

Twenty-nine percent of the total A.C.E. workload was completed during the telephone phase (April 24 through June 13). These A.C.E. interviews were conducted much closer to census day (April 1) than had been possible in 1990, thereby reducing recall bias (the phenomenon of a respondent not remembering the actual situation several months earlier). By design, the telephone phase was restricted to a limited universe of households which were both unlikely to have any exposure to continuing census operations and where we were most confident that the census form corresponded to the ACE sample address. These were primarily households that had mailed back their questionnaires, that had included a telephone number on the questionnaire, and that did not live in certain multi-unit or rural structures. The Census Bureau's conservative use of this interview mode meant that more than 99 percent of the telephone cases were classified as complete or partial interviews and were conducted with a household member.

The automated Computer Assisted Personal Interviewing (CAPI) increased the quality of the data captured in the A.C.E. interviews, as the instrument included data edits to ensure a predetermined quality of data before the interview was considered complete. This was not possible with the paper and pencil 1990 instrument. CAPI insured that the interviewer followed the correct path through the interview. It also allowed quick feedback to the interviewers. The Census Bureau's observations and debriefings indicated that CAPI instilled the interviewers with a sense of professionalism and purpose. Observations also indicated that the use of laptop computers enhanced the respect and cooperation exhibited toward the interviewers by the respondent households thereby leading to improved A.C.E. data quality. However, there were a couple of small problems with the CAPI instrument that had minor impacts on quality.

The Nonresponse Conversion Operation (NRCO) was designed to "convert" nonresponse cases, that is, to obtain A.C.E. information for nonresponding households. On a national basis, the NRCO operation successfully converted 70.8 percent of its cases to complete interviews and 14.1 percent to partial interviews. Only 2.2 percent of the NRCO cases finished as refusals.

A.C.E. interview rates were very high. The A.C.E. asked questions about both the household living at the address on census day and the current household. Because of this, there are two measures of household nonresponse. The interview rate for occupied housing units on census day was 97.1 percent; on the date of the A.C.E. interview, the interview rate for occupied housing units was 98.8 percent. These rates compare favorably to the approximately 98.4 percent (unweighted) in the 1990 Post-Enumeration Survey.

The unweighted rates for 2000 were 97.0 and 98.9, respectively. Due to the high rate of response, most of the noninterview adjustment factors were very close to one. Consequently, this operation did not change the final weights very much. This helps to keep down the variance of the survey weights.

Missing data rates for characteristic data were very low, ranging from 1.4 percent to 2.4 percent. Compared to the 1990 PES, the rates of characteristic missing data are slightly higher for the age and sex characteristics and slightly lower for tenure and race. Again, this is indicative of good quality interviewing.

The goal of A.C.E. interviewing quality assurance was to ensure that the interviewers did, in fact, visit the designated households, and to prevent systematic errors caused by of lack of knowledge or understanding. The evidence

indicates that the A.C.E. interviewing quality assurance operation was properly implemented and successful.

Of the QA interviews, 4.9 percent were randomly selected. The 95.1 percent of cases not in randomly selected QA can be assumed to have a remaining error rate similar to that of the randomly selected QA cases (0.13 percent). However, 171 of the remaining errors were corrected in the targeted QA sample.

### Matching and Follow-up

Matching refers to the process of determining whether an individual enumerated in the A.C.E. was the same person as an individual enumerated in the census. The matching and follow-up process also determines whether a census record in the E-Sample<sup>3</sup> was complete and correct. Errors in matching can significantly affect undercount estimates; highly accurate matching and processing are an important component of A.C.E. methodology.

Although matching error was not identified as a significant problem with the 1990 PES, the Census Bureau made significant improvements to the matching process in the 2000 A.C.E. design. The A.C.E. computer matched the P-sample to the census using the Census Bureau's Statistical Research Division Record Linkage System, a system that the Census Bureau has been developing, testing and using for nearly two decades. Clerical personnel at a centralized location reviewed records that were not matched by the computer matcher. The Census Bureau utilized an ample staff of over 200 clerks, 46 technicians, and 16 analysts so that each successive level of review could perform quality assurance on the previous level. Higher level staff independently reviewed a sample of each employee's work, a process designed to identify random matching errors. Each of the matching levels improved on the previous level. The clerks matched what the computer could not. The technicians worked on any cases the clerks could not resolve and performed the quality assurance on the clerks' cases. Then the analysts finished any cases the technicians could not resolve and performed quality assurance on the technicians' cases.

The results indicate computer matching of 69.6 percent of the P-sample and 64.4 percent of the E-sample. The computer matcher assigned matches very conservatively. Numerous studies over the years have shown that this operation produces insignificant numbers of false matches. Therefore, only questionable matches, possible matches, and nonmatches are required to be clerically reviewed.

---

<sup>3</sup> The E-sample refers to the sample of census data defined person records selected for inclusion in the A.C.E. The P-sample refers to the independent sample of people included in the initial A.C.E. interview.

Still, some review of computer matches is undertaken when appropriate.

We have quality assurance results only on the quality of the clerical matching in the before follow-up stage and the first three stages of after follow-up. The Census Bureau measures matching quality relative to the results that would be produced by the Census Bureau's most experienced and best trained matchers, the 16 analysts permanently employed by the Census Bureau. The quality of the matching process is further measured in terms of changes made by the next level of review; this process tends to overstate the matching error, as not all changes are the result of erroneous matching. However, given these caveats, the outgoing quality rate (the final match rate) for before follow-up was well more than 99 percent. For after follow-up, the outgoing quality rate was also well more than 99 percent. These rates are calculated based on the before follow-up and the after follow-up workload and not on the total number of sample cases, that is, they do not include the cases matched by computer. These rates exceed expectations and are indicative of high quality matching.

Person follow-up is also an important A.C.E. process. The follow-up resolves possible matches and, most importantly, determines which E-sample nonmatches are, nonetheless, correctly enumerated in the census. The person follow-up interviews were conducted either by permanent census field staff or by experienced decennial interviewers and the quality assurance operation was targeted at ensuring that the interview was conducted. Of the randomly selected person follow-up quality assurance

cases, 0.45 percent resulted in a discrepancy, that is, only 0.45 percent determined that the person follow-up interview may not have been conducted. We can assume that the remaining 84,843 cases not randomly selected for quality assurance have the same rate of failure or roughly 400 cases total that may have not been conducted. In addition, we corrected 84 of those cases in the targeted samples.

## References

Singh, R., Kostanich, D. and Hogan, H., "Census 2000: Assessing A.C.E. Quality," Proceeding of the Survey Methods Research Section, American Statistical Association Joint Statistical Meetings, Atlanta, Georgia, August, 2001.

Hogan, H., "Accuracy and Coverage Evaluation: Theory and Application." Proceeding of the Survey Methods Research Section, American Statistical Association Joint Statistical Meetings, Indianapolis, Indiana, August, 2000.

Childers, D. and Fenstermaker, D., "Accuracy and Coverage Evaluation: Overview of the Design," DSSD Census 2000 Procedures and Operations Memorandum Series S-DT-02, U.S. Census Bureau, Washington, D.C., 2000.

Davis, P., "Accuracy and Coverage Evaluation: Dual System Estimation Results," DSSD Census 2000 Procedures and Operations Memorandum Series B-9\*, U.S. Census Bureau, Washington, D.C., 2001

**Appendix Table 1: Census 2000 A.C.E. 64 Post-Stratum Groups - Percent Net Undercount**

Race/Hispanic Origin Domain Number*	Tenure	MSA/TEA	High Return Rate				Low Return Rate			
			NE	MW	S	W	NE	MW	S	W
<b>Domain 7</b> (Non-Hispanic White or "Some other race")	Owner	Large MSA MO/MB	0.81	0.01	0.36	-0.38	-3.62	-2.61	2.19	1.14
		Medium MSA MO/MB	0.30	-0.12	0.46	-0.28	-4.39	-0.33	0.66	1.81
		Small MSA & Non-MSA MO/MB	-0.25	0.14	0.44	0.30	2.29	2.61	2.09	2.71
		All Other TEAs	1.84	-1.11	1.34	0.85	0.56	-0.16	0.15	1.59
	Non-Owner	Large MSA MO/MB	1.82				1.02			
		Medium MSA MO/MB	0.61				2.83			
		Small MSA & Non-MSA MO/MB	2.45				3.61			
		All Other TEAs	1.64				4.08			
<b>Domain 4</b> (Non-Hispanic Black)	Owner	Large MSA MO/MB	1.63				-1.31			
		Medium MSA MO/MB	0.07				0.46			
		Small MSA & Non-MSA MO/MB	4.18				3.42			
		All Other TEAs	2.64				0.12			
	Non-Owner	Large MSA MO/MB	1.46				0.04			
		Medium MSA MO/MB	1.66				1.08			
		Small MSA & Non-MSA MO/MB	3.52				4.98			
		All Other TEAs	4.88				10.74			
<b>Domain 5</b> (Native Hawaiian or Pacific Islander)	Owner	2.71								
	Non-Owner	6.58								
<b>Domain 6</b> (Non-Hispanic Asian)	Owner	0.55								
	Non-Owner	1.58								
<b>American Indian or Alaska Native</b>	<b>Domain 1</b> (On Reservation)	Owner	5.04							
		Non-Owner	4.10							
	<b>Domain 2</b> (Off Reservation)	Owner	1.60							
		Non-Owner	5.57							

- For Census 2000, persons can self-identify with more than one race group. For post-stratification purposes, persons are included in a single Race/Hispanic Origin Domain. This classification does not change a person's actual response. Further, all official tabulations are based on actual responses to the census.
- A negative net undercount denotes a net overcount.

**Table 2: Census 2000 A.C.E. 64 Post-Stratum Groups - Standard Error of Percent Net Undercount**

Race/Hispanic Origin Domain Number*		Tenure	MSA/TEA	High Return Rate				Low Return Rate			
				NE	MW	S	W	NE	MW	S	W
<b>Domain 7</b> (Non-Hispanic White or "Some other race")	Owner	Large MSA MO/MB	0.43	0.36	0.87	0.45	1.05	1.43	1.54	2.09	
		Medium MSA MO/MB	0.85	0.28	0.42	0.38	1.52	0.84	1.10	2.79	
		Small MSA & Non-MSA MO/MB	1.33	0.40	0.43	0.57	3.60	2.12	1.08	1.49	
		All Other TEAs	1.06	0.39	0.97	1.66	2.17	1.21	0.65	1.89	
	Non-Owner	Large MSA MO/MB	0.63				1.01				
		Medium MSA MO/MB	0.71				1.24				
		Small MSA & Non-MSA MO/MB	0.51				1.24				
		All Other TEAs	0.94				1.67				
<b>Domain 4</b> (Non-Hispanic Black)	Owner	Large MSA MO/MB	0.56				1.24				
		Medium MSA MO/MB									
		Small MSA & Non-MSA MO/MB	1.07				1.86				
		All Other TEAs									
	Non-Owner	Large MSA MO/MB	0.66				1.05				
		Medium MSA MO/MB									
		Small MSA & Non-MSA MO/MB	0.96				2.08				
		All Other TEAs									
<b>Domain 3</b> (Hispanic)	Owner	Large MSA MO/MB	0.52				1.26				
		Medium MSA MO/MB									
		Small MSA & Non-MSA MO/MB	1.01				2.09				
		All Other TEAs									
	Non-Owner	Large MSA MO/MB	0.67				1.12				
		Medium MSA MO/MB									
		Small MSA & Non-MSA MO/MB	1.55				4.12				
		All Other TEAs									
<b>Domain 5</b> (Native Hawaiian or Pacific Islander)		Owner	3.83								
		Non-Owner	4.07								
<b>Domain 6</b> (Non-Hispanic Asian)		Owner	0.87								
		Non-Owner	0.98								
<b>American Indian or Alaska Native</b>	<b>Domain 1</b> (On Reservation)	Owner	1.45								
		Non-Owner	1.42								
	<b>Domain 2</b> (Off Reservation)	Owner	1.95								
		Non-Owner	2.02								

- For Census 2000, persons can self-identify with more than one race group. For post-stratification purposes, persons are included in a single Race/Hispanic Origin Domain. This classification does not change a person's actual response. Further, all official tabulations are based on actual responses to the census.
- For Census 2000, persons can self-identify with more than one race group. For post-stratification purposes, persons are included in a single Race/Hispanic Origin Domain. This classification does not change a person's actual response. Further, all official tabulations are based on actual responses to the census.
- A negative net undercount denotes a net over-count.



