

1996 AMERICAN COMMUNITY SURVEY vs. 1990 DECENNIAL CENSUS HOUSEHOLD SIZE AND CHARACTERISTICS BY RESPONSE MODE

Pamela W. Ferrari, U.S. Bureau of the Census

U. S. Bureau of the Census, Statistical Research Division, Rm. 3232-4, Washington, DC 20233

Key Words: Coverage, Household Roster, Data Collection/Response Mode

MOTIVATION FOR COVERAGE RESEARCH

It is essential to be able to accurately represent the target population in order to produce unbiased estimates. In both the Decennial Census and the Bureau's demographic surveys, there is evidence of significant coverage problems for some segments of the population. Under-coverage occurs when a person eligible for the Census (or survey) does not have a chance to be represented or included in the Census, either intentionally or unintentionally. Historically, there are coverage problems with low income and inner-city communities, with household surveys being worse than the Decennial Census. This is of major concern with the American Community Survey (ACS) data since the Bureau's long range plan is to possibly replace long form Census data with ACS data and therefore be able to provide local area long form data on an ongoing basis.

In an effort to identify any possible coverage problems associated with the ACS, a research project was initiated to compare coverage in the 1996 ACS to the 1990 Decennial Census by looking at the distribution of household size by various demographic characteristics and mode of response. From those results, we hope to identify possible causes, such as forms designs and field and processing procedures, that might contribute to the under-coverage and suggest further research and testing.

Other research projects will address within household coverage, look at residence rules, suggest alternative rostering and questionnaire design, assess the impact of nonresponse on coverage, evaluate the completeness of data for persons from large households, and experiment with methodologies to improve whole household coverage in frames for sampling.

ACS BACKGROUND

The Bureau of the Census conducted the 1996 ACS in four counties: Multnomah Oregon, Brevard Florida, Rockland New York, and Fulton Pennsylvania. The data was collected in monthly panels beginning in November 1995 and continuing through October 1996, although there was an interruption due to the Federal Government furlough in late December/early January. In the three

"urban" sites, prenotice letters, initial questionnaires, reminder cards, and replacement questionnaires for nonrespondents were mailed to sample households. However in Fulton County Pennsylvania, the "rural" site, Field Representatives delivered questionnaires to sampled addresses that had been prelisted since there is not a number-street address mailing address system. Questionnaires from all four sites were expected to be mailed back. Nonrespondents were followed up by Computer-Assisted Telephone Interviewing (CATI) in the three urban sites. CATI was not conducted in Fulton County Pennsylvania since telephone numbers could not be obtained without the number-street address. A one-in-three subsample of those nonrespondents who could not be contacted by telephone and of the mail nonrespondents in rural Fulton County Pennsylvania was followed up by Computer-Assisted Personal Interviewing (CAPI).

The ACS questionnaire obtains a roster of the household and collects labor force and economic data for the adult household residents and provides characteristics of children and housing units. To provide a roster of the household, the respondent lists the number of persons living in the household and then lists the names of the persons on subsequent pages. The respondent provides demographic characteristics for the first five persons listed. If six or more persons are listed, the characteristics for those persons are collected by telephone follow-up.

The mail return questionnaires undergo a clerical edit and telephone follow-up prior to data entry. Questionnaires from all three data collection modes are sent to a coding operation for several data items, proceed through the edit and allocation processing, weighting, and tabulations are then produced.

DATA USED FOR COVERAGE STUDY

The 1996 ACS data files that were produced from the edit and allocation procedure were used in this coverage study. The files at this point were unweighted because the population controls used in the weighting, which are derived from the 1990 Census counts, could mask any coverage problems. The CAPI cases, however, were given a weight of 3.27 in order to account for the one in three subsample and the lost month of data during the Federal Government furlough.

The 1990 Census files used were the one hundred percent data files, rather than the sample data files since

coverage was being evaluated. While the ACS questionnaires were more like the Census long form (sample data), the issue is whether the ACS adequately covered the entire population and not to compare responses or cooperation from similar questionnaires.

Distributions by household size, data collection mode and various demographic characteristics of the ACS and Census data were compared. Differences in ACS and Census household size by type of household could be evidence of coverage problems with the ACS and would point out the need for further investigation.

PRELIMINARY RESULTS

The most consistent finding across almost all characteristics examined and in three of the four sites is that the ACS has a higher percentage of smaller households and a lower percentage of larger households than the Census data. In Fulton County, Pennsylvania the four-person households was the only size category where there was a difference. This difference in larger/smaller household size distribution could be due to within household undercoverage in the ACS since that would tend to reduce household size, and hence produce a downshift in the household size distribution. Another possible explanation is that there have been changes in socio-demographic characteristics in the six years between the 1990 Census and the 1996 ACS. On the next page is a summary table of the distribution by household size for each of the four sites. Detailed tables by data collection mode and various demographic characteristics (tenure and the householder's race, Hispanic origin and sex) can be found in the appendix of the full version of this paper available from the author. Only Tables A1-A4 (Distribution of Household (HH) Size by Data Collection Mode for each of the four sites) are included at the end of this version of the paper. The results mentioned will have the Appendix table number referenced in parentheses.

The Fulton County Pennsylvania site has similar distributions for ACS and Census at the summary level, but tends to show the same pattern of differences as the other sites when further broken down. This is the rural site where the mail return rate for the ACS ran the lowest, but the final response rate was the highest. When these numbers are broken down by data collection mode (see Table A4 in the Appendix), we find that the large/small household size difference is quite apparent. In this rural site, coverage would not look very good at all if only mail response was attempted due to the high nonresponse to the mail. However, in total, the coverage is the best of the four sites.

When all four sites were broken down by data collection mode (Tables A1-A4 at the end), the difference in distributions was larger when comparing ACS and

Census mail returns than comparing CAPI to Census enumerator returns. It should be noted that only five persons could be listed on the ACS questionnaire (compared to seven on the 1990 Census form) and perhaps some respondents did not indicate that they had a larger household. This was probably not a major problem, or we would have seen larger differences in the largest size households. In Florida (Table A1), Oregon (Table A3), and New York (Table A2) the CATI distribution was generally quite different than the Census enumerator returns. In New York, the CATI distribution was fairly similar to the Census mail return; but not so in Florida and Oregon. The CAPI distribution was more similar to the Census enumerator returns, but still showed some difference in the large/small households. In Pennsylvania (Table A4), the larger household sizes were different when comparing mail returns and CAPI versus enumerator returns. A major caveat to looking at just the data collection mode is that there is too much weight for mail returns and too little for the enumerator returns in the Census numbers because the nonresponse adjustment did not account for mode.

The mail return comparisons showed that Florida, New York and Oregon have a higher percentage of one and two-person households in the ACS than Census and smaller percentage of three and larger person households. In Pennsylvania, ACS and Census had similar percentages for one and three-person households, while ACS had more two-person households and fewer four and larger person households. When CATI cases were compared to enumerator returns, there were fewer one-person households in ACS (opposite of mail returns) in all three sites with the CATI treatment. CAPI had a smaller percentage of one-person households in Oregon, a larger percentage of one-person households in Florida and similar percentages in New York and Pennsylvania.

The household size distributions were broken down by tenure, owner/renter (Tables A5-A8 in the Appendix of the full version). In Florida (Table A5), ACS had a higher percentage of owners than Census in all household size categories. In New York (Table A6), ACS had a higher percentage of owners than Census in one and two-person households, about the same for three-person households, and slightly lower percentage in four and larger person households. In Oregon (Table A7), ACS had a higher percentage of owners than Census in one-person households and about the same for two and larger person households. In Pennsylvania (Table A8), Census had a higher percentage of owners in one, four and six and larger person households, while the other size categories were similar. The same general pattern was observed when the tenure was broken down by mail returns. CATI generally had a smaller percentage of one

DISTRIBUTION BY HOUSEHOLD SIZE

Number of Persons in Household	BREVARD FLORIDA		ROCKLAND NEW YORK		MULTNOMAH OREGON		FULTON PENNSYLVANIA	
	ACS	CENSUS	ACS	CENSUS	ACS	CENSUS	ACS	CENSUS
1	7,065 <u>26.90%</u> 0.27%	38,216 23.69%	2,673 <u>19.76%</u> 0.34%	15,058 17.75%	12,391 <u>32.81%</u> 0.24%	77,249 31.91%	315 21.23% 1.06%	1,064 20.71%
2	10,483 39.91% 0.30%	64,245 39.82%	4,142 <u>30.64%</u> 0.40%	23,481 27.68%	12,822 <u>33.95%</u> 0.24%	80,527 33.26%	507 34.16% 1.23%	1,683 32.76%
3	3,953 <u>15.05%</u> 0.22%	27,048 16.77%	2,444 <u>18.08%</u> 0.33%	16,093 18.97%	5,415 <u>14.34%</u> 0.18%	36,379 15.03%	286 19.27% 1.02%	978 19.04%
4	2,981 <u>11.35%</u> 0.20%	20,346 12.61%	2,314 <u>17.12%</u> 0.32%	16,367 19.30%	4,350 <u>11.52%</u> 0.16%	29,071 12.01%	220 <u>14.82%</u> 0.92%	885 17.23%
5	1,246 4.74% 0.13%	7,786 4.83%	1,102 <u>8.15%</u> 0.24%	7,816 9.22%	1,708 <u>4.52%</u> 0.11%	11,747 4.85%	112 7.55% 0.69%	367 7.14%
6-7	483 <u>1.84%</u> 0.08%	3,244 2.01%	597 <u>4.42%</u> 0.18%	4,583 5.40%	858 <u>2.27%</u> 0.08%	5,952 2.46%	43 2.90% 0.44%	151 2.94%
>=8	54 <u>0.21%</u> 0.03%	443 0.27%	244 1.81% 0.11%	1,417 1.67%	217 <u>0.57%</u> 0.04%	1,168 0.48%	1 0.07% 0.07%	9 0.18%
Total	26,265	161,328	13,517	84,815	37,762	242,093	1,484	5,137

The shaded numbers are standard errors and the underlined ACS proportions indicate they are significantly different at the .05 level from the corresponding census proportions.

and two-person households for both owners and renters than the Census enumerator returns. In owner occupied households, CAPI had a higher percentage of one-person households than the Census enumerator returns in all but Pennsylvania and fewer or similar percentages in two-person households. In renter occupied units the percentages varied by site and household size.

We next examined the race/Hispanic origin of the person who was listed on line one of the questionnaire (Tables A13-A20 in the Appendix). This in some sense made a household variable from a person characteristic, as did race and sex of the person on line one, so that similar comparisons could be made. Looking at the mail

returns first, we found that in Oregon (Tables A17 and A18) most of the race categories had ACS with larger percentages (and in some cases similar percentages) in one and two-person households and similar or less in the larger household sizes. New York (Tables A15 and A16) and Florida (Tables A13 and A14) showed the same pattern. In Pennsylvania (Tables A19 and A20), where there is little racial diversity (The majority of people are White/Not Hispanic and in fact the tables are essentially household size by data collection mode.), there were similar percentages for one-person households, ACS had a larger percentage in two-person households, and Census had larger percentages in three and more person

households. When we looked at CATI/CAPI and enumerator returns, the CATI treatment behaved very differently. CAPI was similar to the enumerator returns in Pennsylvania. However, the overall pattern of the ACS having higher percentages in the smaller household sizes and lower percentages in the larger household sizes continued.

This same general pattern was also apparent when sex (Tables A9-A12 in the Appendix) and age (no tables included) were looked at across the data collection modes. An interesting result not directly related to coverage, however, was that there is a difference between ACS and Census as to who is listed on line one in multi-person households. ACS more often lists a female on line one, while Census more often lists a male. When we looked at the total population, rather than only the line one person, the differences in gender disappeared. The instructions for the roster list were generally the same for both ACS and Census, but perhaps there is a difference in who fills out the questionnaire or who is interviewed.

WHERE DO WE GO FROM HERE?

At this point in time, we see that the ACS finds more one-person households than the Census and that ACS may have an undercoverage problem with larger households. Since the undercoverage in larger households could be due to nonresponse, we should look at the data after the nonresponse weighting. We plan to continue to look at additional characteristics, such as income and education to see if there are particular segments that have coverage problems. There is an additional year of data collected during 1997 in the same

four sites, as well as in four new sites, that can be examined. In addition, we want to research alternative rostering techniques and instructions to respondents.

We will take a closer look at the data collection mode results, especially CATI and CAPI, to see if the mode is affecting coverage, particularly within household coverage. It should be noted that the CATI group is really a different population than the CAPI group since CAPI cases are nonrespondents to CATI or those cases where telephone numbers could not be obtained. In a sense, the CATI cases are "easier" to enumerate. The CATI effect, presumably due to its centralized nature, has been seen in other surveys.

Perhaps the most important additional investigations can be performed using the 1998 Census Dress Rehearsal results in Columbia, South Carolina since the ACS will also have that location as a sample site at exactly the same time. The six year time lag between the 1990 Census and the 1996 ACS will not be an issue here. The intense publicity program for responding to the census may spill over to the ACS since they are being conducted at the same time, so that the lack of the Decennial promotion effect will not be as apparent as with the 1996 and 1997 ACS where there was no Decennial advertising. We should be able to see more clearly real differences in coverage instead of having to guess whether any differences are due to true demographic change over time or coverage problems. There will still, however, be the difference of the Census which counts persons where they are on April 1 versus the ACS which counts persons where they are the majority of the year.

This paper reports results of research and analysis undertaken by the Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.

Table A1

**DISTRIBUTION OF HH SIZE BY DATA COLLECTION MODE
BREVARD COUNTY FLORIDA**

Number of Persons in Household	1996 ACS				1990 CENSUS		
	Mail	CATI	CAPI	Total	Mail Return	Enumerator Return	Total
1	4,963	565	1,537	7,065	28,266	9,950	38,216
	<u>26.57%</u>	<u>23.23%</u>	<u>29.82%</u>		22.53%	27.73%	
	0.32%	0.86%	0.64%				
2	8,104	796	1,583	10,483	52,886	11,359	64,245
	<u>43.39%</u>	32.73%	30.71%		42.16%	31.66%	
	0.36%	0.95%	0.64%				
3	2,574	418	961	3,953	20,631	6,417	27,048
	<u>13.78%</u>	17.19%	18.65%		16.45%	17.88%	
	0.25%	0.77%	0.54%				
4	2,003	383	595	2,981	15,314	5,032	20,346
	<u>10.72%</u>	<u>15.75%</u>	<u>11.55%</u>		12.21%	14.02%	
	0.23%	0.74%	0.45%				
5	767	178	301	1,246	5,747	2,039	7,786
	<u>4.11%</u>	<u>7.32%</u>	5.84%		4.58%	5.68%	
	0.15%	0.53%	0.33%				
6-7	243	83	157	483	2,327	917	3,244
	<u>1.30%</u>	<u>3.41%</u>	<u>3.05%</u>		1.85%	2.56%	
	0.08%	0.37%	0.24%				
>=8	25	9	20	54	276	167	443
	<u>0.13%</u>	0.37%	0.38%		0.22%	0.47%	
	0.03%	0.12%	0.09%				
Total	18,679	2,432	5,154	26,265	125,447	35,881	161,328

37 cases missing for 1990 Census.

Table A2

**DISTRIBUTION OF HH SIZE BY DATA COLLECTION MODE
ROCKLAND COUNTY NEW YORK**

Number of Persons in Household	1996 ACS				1990 CENSUS		
	Mail	CATI	CAPI	Total	Mail Return	Enumerator Return	Total
1	1,602	211	860	2,673	10,224	4,834	15,058
	<u>19.32%</u>	<u>14.43%</u>	22.87		16.19%	22.32%	
	0.43%	0.92%	0.68%				
2	2,731	361	1,050	4,142	18,134	5,347	23,481
	<u>32.93%</u>	24.69%	27.91		28.71%	24.68%	
	0.52%	1.13%	0.73%				
3	1,484	267	693	2,444	12,235	3,858	16,093
	<u>17.89%</u>	18.26%	18.43		19.37%	17.81%	
	0.42%	1.01%	0.63%				
4	1,453	318	543	2,314	12,663	3,704	16,367
	<u>17.52%</u>	<u>21.75%</u>	14.43		20.05%	17.10%	
	0.42%	1.08%	0.57%				
5	618	157	327	1,102	5,815	2,001	7,816
	<u>7.45%</u>	10.74%	8.70%		9.21%	9.24%	
	0.29%	0.81%	0.46%				
6-7	284	107	206	597	3,245	1,338	4,583
	<u>3.42%</u>	7.32%	5.48%		5.14%	6.18%	
	0.20%	0.68%	0.37%				
>=8	122	41	82	245	837	580	1,417
	1.47%	2.81%	2.17%		1.32%	2.68%	
	0.13%	0.43%	0.24%				
Total	8,294	1,462	3,761	13,517	63,153	21,662	84,815

59 cases missing for 1990 Census.

The shaded numbers are standard errors and the underlined ACS proportions indicate they are significantly different at the .05 level from the corresponding Census proportions (mail vs. mail, CATI vs. enumerator return, and CAPI vs. enumerator return).

Table A3

**DISTRIBUTION OF HH SIZE BY DATA COLLECTION MODE
MULTNOMAH COUNTY OREGON**

Number of Persons in Household	1996 ACS				1990 CENSUS		
	Mail	CATI	CAPI	Total	Mail Return	Enumerator Return	Total
1	8,875	776	2,740	12,391	54,532	22,717	77,249
	<u>33.93%</u>	<u>25.59%</u>	<u>31.96%</u>		30.92%	34.57%	
	0.29%	0.79%	0.50%				
2	9,490	889	2,443	12,822	62,070	18,457	80,527
	<u>36.28%</u>	29.31%	28.49%		35.19%	28.09%	
	0.30%	0.83%	0.49%				
3	3,407	566	1,442	5,415	25,996	10,383	36,379
	<u>13.03%</u>	<u>18.66%</u>	<u>16.82%</u>		14.74%	15.80%	
	0.21%	0.71%	0.40%				
4	2,849	458	1,043	4,350	21,133	7,938	29,071
	<u>10.89%</u>	<u>15.10%</u>	12.17%		11.98%	12.08%	
	0.19%	0.65%	0.35%				
5	1,000	195	513	1,708	8,124	3,623	11,747
	<u>3.82%</u>	<u>6.43%</u>	5.99%		4.61%	5.51%	
	0.12%	0.43%	0.26%				
6-7	445	122	291	858	3,873	2,079	5,952
	<u>1.70%</u>	<u>4.02%</u>	3.39%		2.20%	3.16%	
	0.08%	0.36%	0.20%				
>=8	89	27	101	217	648	520	1,168
	0.34%	0.89%	<u>1.18%</u>		0.37%	0.79%	
	0.04%	0.17%	0.12%				
Total	26,155	3,033	8,574	37,762	176,376	65,717	242,093

47 cases missing for 1990 Census.

Table A4

**DISTRIBUTION OF HH SIZE BY DATA COLLECTION MODE
FULTON COUNTY PENNSYLVANIA**

Number of Persons in Household	1996 ACS			1990 CENSUS		
	Mail/Delivery	CAPI	Total	Mail Return	Enumerator Return	Total
1	201	114	315	833	231	1,064
	20.79%	22.15%		20.06%	23.45%	
	1.31%	1.83%				
2	373	134	507	1,408	275	1,683
	<u>38.57%</u>	25.95%		33.91%	27.92%	
	1.57%	1.93%				
3	191	95	286	769	209	978
	19.75%	18.35%		18.52%	21.22%	
	1.28%	1.70%				
4	138	82	220	729	156	885
	<u>14.27%</u>	15.82%		17.56%	15.84%	
	1.12%	1.61%				
5	53	59	112	296	71	367
	<u>5.48%</u>	11.39%		7.13%	7.21%	
	0.73%	1.40%				
6-7	10	33	43	111	40	151
	<u>1.03%</u>	<u>6.33%</u>		2.67%	4.06%	
	0.32%	1.07%				
>=8	1	0	1	6	3	9
	0.10%	0.00%		0.14%	0.30%	
	0.10%	0.00%				
Total	967	517	1,484	4,152	985	5,137

2 cases missing for 1990 census.

The shaded numbers are standard errors and the underlined ACS proportions indicate they are significantly different at the .05 level from the corresponding Census proportions (mail vs. mail, CATI vs. enumerator return, and CAPI vs. enumerator return).