

## Income in the American Community Survey: Comparisons to Census 2000

Charles Nelson, Edward Welniak, and Kirby G. Posey

U.S. Census Bureau  
Washington, DC 20233

*This paper reports the results of research and analysis undertaken by 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.*

### Income in the American Community Survey and Census 2000:

The long-term goal of the American Community Survey (ACS) is to produce more timely local area estimates. Currently, the only sources of Census Bureau local area income estimates are the decennial census (data only once every ten years) and the Small Area Income and Poverty Estimates (SAIPE) program (biennial county and state estimates of median household income). We designed the battery of income questions as well as the edit and allocation specifications for the ACS to be as similar as possible to that of the Census in order to evaluate the ACS's performance. There are, of course, some differences inherent with the ACS that forced us to make some changes. The most notable change between the Census and the ACS is the reference period, "past 12 months" versus "last calendar year." The purpose of this paper is to explain the major differences and provide a preliminary overview of the ACS results.

In Census 2000, a battery of eight detailed income questions and a total income question were posed to all persons 15 years old and over. The eight detailed income questions included (1) wages and salary, (2) self-employment income, (3) interest, dividend, and net royalty income, (4) Social Security, (5) Supplemental Security Income (SSI), (6) public assistance or welfare, (7) retirement, survivor, or disability income, and (8) any other sources of income received on a regular basis. These eight income questions were placed at the end of each set of person questions on the long form (sample) questionnaire so that response levels to the other less sensitive sample questions would not be adversely impacted. A "total income" question was first asked in the 1980 Census to aid in the resolution of income entry problems in the detailed questions. Test results indicated that a sizable number of all income recipients furnished responses to the total income question, but failed to provide answers to some or all parts of the detailed type of income questions, proving the total income item very important.

We used the same battery of income questions including the total income item in the ACS with one major exception, the reference period.

The biggest difference between collection methods in the ACS and the Census is the income reference period. The ACS collects data throughout the year on an on-going, monthly basis. The ACS asks for a respondent's income over the "past 12 months." The census, however, collected the income data for a fixed period of time - "during 1999" (the last calendar year.)

## The Split Panel Test:

To evaluate the impact of this reference period change, we conducted a "split panel" test of about 19,000 households (with a 49 percent mail response rate). We conducted this split panel test from October through December of 1997. This special survey collected income statistics using both terminologies, "past 12 months" and "income in calendar year 1996" for the two different randomly assigned treatment groups. The only statistical differences in median income estimates between the two reference periods occurred in the earnings categories, wages / salary and self-employment (see Table 1). Note: In 1997 we combined the question on Supplemental Security Income (SSI) and the question on public assistance or welfare in the same item.

Table 1: Comparison of Median Income of Individuals by Income Type and Reference Period (asked October - December, 1997)				
SPLIT PANEL TEST				
	PAST 12 MONTHS	CALENDAR YEAR 1996	DIFFERENCE	
Wages or Salary	\$25,000	\$26,000	\$1,000	*
Self-Employment	\$12,000	\$10,000	-\$2,000	*
Interest	\$1,500	\$1,400	-\$100	
Social Security	\$7,932	\$8,028	\$96	
Retirement	\$9,700	\$10,000	\$300	
Public Assistance	\$3,438	\$3,872	\$434	
Other Income	\$3,000	\$3,000	\$0	
Total Income	\$24,000	\$24,662	\$662	
Note: These results are from mail responses only. Source: U.S. Census Bureau, 1996 American Community Survey * Indicates statistical significant difference at 90-percent confidence level.				

We expected the 12-month reference period to yield slightly higher estimates because of the more recent reference period. As shown in Table 1, that is not always the case. For example, the median wage / salary income reported on forms asking about the PAST 12 MONTHS was \$25,000 for individuals. The median wage / salary income reported on forms asking about the PAST CALENDAR YEAR is \$26,000. For self-employment income, however, the "past 12 months" reference period yielded the higher figure. All other income items did not differ significantly between the two reference periods.

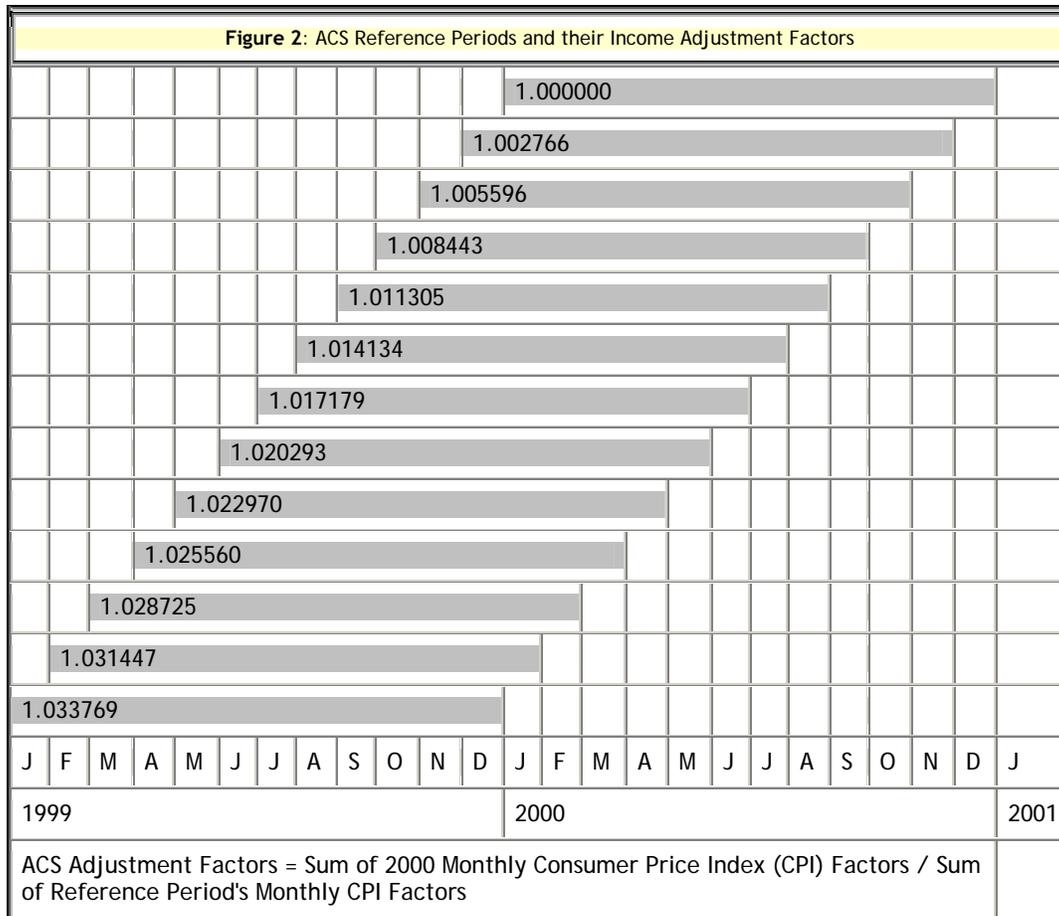
When looking at response rates (the percentage of respondents who answered "yes," "no," or "loss") by the different sources of income, the questionnaire with the "past 12 months" reference period produced slightly higher response rates for every income

source although only one income item, public assistance, was statistically significant over the two reference periods. (See Table 2.)

Table 2 - Response Rates for All Sources of Income - Split Panel Test				
Income	Period	Percent Reporting Income	Difference	
Wage/Salary	Past 12 months	73.93%		
	Calendar year 1996	72.84%	+1.09	
Self-Employment	Past 12 months	73.72%		
	Calendar year 1996	72.68%	+1.04	
Interest, etc.	Past 12 months	72.07%		
	Calendar year 1996	71.05%	+1.02	
Social Security	Past 12 months	75.78%		
	Calendar year 1996	74.65%	+1.13	
Retirement	Past 12 months	75.36%		
	Calendar year 1996	73.96%	+1.40	
Public Assistance	Past 12 months	77.12%		
	Calendar year 1996	75.38%	+1.74	*
Other Income	Past 12 months	76.40%		
	Calendar year 1996	75.23%	+1.17	
Total Income	Past 12 months	83.67%		
	Calendar year 1996	83.09%	+0.58	

Source: 1996 American Community Survey  
 \* Indicates statistical significant difference at 90-percent confidence level.





Respondents who were interviewed in February 2000 and asked about their income from February 1999 through January 2000 had their income adjusted by a factor of 1.031447. On the other hand, respondents who were interviewed in October and whose income reference period was from October 1999 through September 2000 had their incomes adjusted by a factor of 1.008443. These factors were computed as the sum of the 2000 CPI monthly adjustment factors divided by the sum of the CPI monthly adjustment factors for the reference period. The result is that all ACS respondents in a given year end up with incomes in terms of consistent 2000 dollars. Note that on the Public Use files, for disclosure avoidance purposes, there is only one adjustment factor for all respondents. This is to keep from disclosing the month of interview to help protect the individual confidentiality of our respondents. That factor will be the average of the 12 monthly factors.

Also note that the ACS is not designed to produce income statistics for use on a monthly basis. We collect these data over a 12-month period then aggregate it to an annual amount.

### Edits and Allocations

For the most part, we used the same editing scheme for the Census 2000 Supplemental Survey (C2SS) and ACS as we did for Census 2000. For example, in the consistency edits for both surveys, we checked for things like obvious monthly amounts reported

for income sources like social security, public assistance or retirement income. We also checked for common mistakes respondents sometimes make with decimal points. For example, we checked for someone who might have reported \$100,000 in public assistance income where they should have reported \$1,000. We checked for respondents who might have intended to report a "loss" in self-employment income or in interest, dividends, or net rental income but forgot to mark the "loss" box based on their total reported income. We checked identically reported amounts in the wage/salary question and the total income question. Many times, respondents will confuse the first income item for the total income field and report their total income twice. We used the total income field to resolve differences in the reporting of the components when we could.

One difference between the two surveys was in the method used for data capture. In Census 2000, most of the enumerator-filled and mail-return questionnaires were processed using Optical Character Recognition, or OCR, an image scanning system. The ACS did not employ this system of data capture. The ACS employed actual "keyers" during the data capture operation. In the Census, OCR interpreted numeric handwritten income entries then performed a data quality check to help ensure that the number read was accurate. If the entry failed to meet an accuracy threshold, an image of the item in question was displayed to a clerk who then edited the response. Because of the differences in data capture methodology we modified the census edits slightly.

For missing income data, we designed the allocation and imputation specifications for the C2SS / ACS basically the same as those used in Census 2000. In both surveys, we sorted the file first by sex. That is, we stored usable data (hot deck values) separately for males and females so that missing economic characteristics are allocated to a male only from another male; never from a female and vice versa. Then, for each sex, we further stored usable data according to race and ethnicity, educational attainment, living arrangements (that is, husband or wife, other family reference person, other family members and unrelated individuals), and finally by residence (or whether a person did or did not live in a metropolitan area.) For example, if a respondent failed to report earnings data but did provide answers to other items such as occupation, class of worker, weeks worked last week, and age, we matched their reported data to that of another respondent who fully reported all items. When we located a match, we simply substituted the fully reported earnings data for the missing information.

Because the ACS data are collected on a monthly basis, at the time of imputation we will have access to a much smaller sample than we did during the census (about 3 million housing units for a full year of data for the ACS compared to around 17 million during the Census). Therefore, we found it necessary to borrow some matching schemes from the Current Population Survey processing system that does not rely so much on the nearest geographic neighbor as a donor as we did during the 1990 Census. That is, in the ACS, we used different "keys" when attempting to match to an information donor. Because we made this change in the ACS we also employed the approach in Census 2000. We started the process by looking for donors in fairly detailed matrices hoping to match as closely as possible to another person's characteristics. If we fail to locate a match, we dropped back to a more general and less detailed matrix or a second "key" in order to find a match.

Sometimes respondents will indicate they received a particular type of income but will not report how much they received. The ACS and Census 2000 treated respondents who reported income "reciency" with no amount differently than those respondents who left an income question completely blank. That is, we allocated missing amounts from different matrices depending on whether the respondent marked the "yes" box for that item. In other words, we first try to establish a pattern of reciency and then go from there. The allocation process in both surveys was basically the same.

The biggest difference between the allocation processes between the two surveys, at least at this time, is the sample size. We believe that once the ACS is fully implemented and sample size increases, the results of the ACS allocation procedure will more closely mirror that of the Census because we are more likely to locate a better match for non-respondents.

**Results:**

Median household income estimates were generally lower in the C2SS / ACS than Census 2000 after adjusting the 1999 dollar values from Census 2000 for inflation. (See Table 3 for the median household income figures).

Table 3. Comparisons of Median Household Income in 2000 Dollars: C2SS/ACS, Census 2000, and the CPS

	C2SS Estimate (in dollars)	Census 2000 (Adjusted for inflation - in dollars)	Percent Difference	Statistically Significant Difference (denoted by "**")	Current Population Survey (CPS) 2000 Dollars	Percent Difference	Statistically Significant Difference (denoted by "**")
United States	41486	43396	4.4		41990	1.2	
Alabama	33433	35274	5.2		35424	5.6	
Alaska	52906	53292	0.7		52847	-0.1	
Arizona	38547	41912	8.0		39783	3.1	
Arkansas	33188	33256	0.2		29697	-11.8	
California	46617	49078	5.0		46816	0.4	
Colorado	46391	48778	4.9		48240	3.8	
Connecticut	54129	55735	2.9		50172	-7.9	
Delaware	47695	48962	2.6		50365	5.3	
District of Columbia	40926	41466	1.3		41222	0.7	
Florida	38054	40115	5.1		38856	2.1	
Georgia	41295	43849	5.8		41901	1.4	

	C2SS Estimate (in dollars)	Census 2000 (Adjusted for inflation - in dollars)	Percent Difference	Statistically Significant Difference (denoted by "**")	Current Population Survey (CPS) 2000 Dollars	Percent Difference	Statistically Significant Difference (denoted by "**")
Hawaii	51587	51483	-0.2		51546	-0.1	
Idaho	37569	38826	3.2		37611	0.1	
Illinois	45708	48145	5.1		46064	0.8	
Indiana	40794	42954	5.0		40865	0.2	
Iowa	38179	40786	6.4		40991	6.9	
Kansas	40264	41980	4.1		41059	1.9	
Kentucky	32862	34796	5.6		36265	9.4	
Louisiana	30985	33653	7.9		30718	-0.9	
Maine	36608	38483	4.9		37266	1.8	
Maryland	52447	54632	4.0		54535	3.8	
Massachusetts	49801	52188	4.6		46753	-6.5	
Michigan	43387	46158	6.0		45512	4.7	
Minnesota	47753	48683	1.9		54251	12.0	
Mississippi	32728	32376	-1.1		34299	4.6	
Missouri	37145	39200	5.2		45097	17.6	
Montana	33103	34126	3.0		32777	-1.0	
Nebraska	37379	40560	7.8		41750	10.5	
Nevada	42401	46069	8.0		45758	7.3	
New Hampshire	50969	51118	0.3		50926	-0.1	
New Jersey	54276	56986	4.8		50405	-7.7	
New Mexico	33428	35272	5.2		35093	4.7	
New York	43735	44841	2.5		40744	-7.3	
North Carolina	37784	40492	6.7		38317	1.4	
North Dakota	34376	35759	3.9		35996	4.5	
Ohio	39745	42323	6.1		42962	7.5	
Oklahoma	34135	34515	1.1		32432	-5.3	
Oregon	39090	42282	7.5		42499	8.0	
Pennsylvania	39661	41445	4.3		42176	6.0	
Rhode Island	43778	43495	-0.7		42197	-3.7	

	C2SS Estimate (in dollars)	Census 2000 (Adjusted for inflation - in dollars)	Percent Difference	Statistically Significant Difference (denoted by "**")	Current Population Survey (CPS) 2000 Dollars	Percent Difference	Statistically Significant Difference (denoted by "**")
South Carolina	36439	38320	4.9		37570	3.0	
South Dakota	34840	36460	4.4		36475	4.5	
Tennessee	36559	37574	2.7		34096	-7.2	
Texas	39398	41260	4.5		38609	-2.0	
Utah	45536	47252	3.6		47550	4.2	
Vermont	40505	42220	4.1		39594	-2.3	
Virginia	47125	48235	2.3		47163	0.1	
Washington	45246	47304	4.4		42525	-6.4	
West Virginia	29089	30687	5.2		29411	1.1	
Wisconsin	42209	45253	6.7		45088	6.4	
Wyoming	38614	39157	1.4		39629	2.6	
Pima County, Arizona	35223	37985	7.3		NE		
Jefferson County, Arkansas	30838	32373	4.7		NE		
San Francisco County, California	57608	57064	-1.0		NE		
Tulare County, California	32508	35117	7.4		NE		
Broward County, Florida	40569	43082	5.8		NE		
Lake County, Illinois	67907	69208	1.9		NE		
Black Hawk County, Iowa	36686	38510	4.7		NE		
Calvert County, Maryland	65497	68146	3.9		NE		
Hampden County, Massachusetts	37963	41044	7.5		NE		
Madison County, Mississippi	44036	48538	9.3		NE		

	C2SS Estimate (in dollars)	Census 2000 (Adjusted for inflation - in dollars)	Percent Difference	Statistically Significant Difference (denoted by "**")	Current Population Survey (CPS) 2000 Dollars	Percent Difference	Statistically Significant Difference (denoted by "**")
Douglas County, Nebraska	41418	44651	7.2		NE		
Bronx County, New York	27743	28533	2.8		NE		
Rockland County, New York	67268	70240	4.2		NE		
Franklin County, Ohio	42056	44160	4.8		NE		
Multnomah County, Oregon	40321	42656	5.5		NE		
Schuylkill County, Pennsylvania	32955	33790	2.5		NE		
Sevier County, Tennessee	34640	35878	3.4		NE		
Fort Bend County, Texas	62660	65961	5.0		NE		
Harris County, Texas	41177	44020	6.5		NE		
Yakima County, Washington	33523	35990	6.9		NE		
NE = No Estimate							

Median household income for the nation was more than four percent higher in Census 2000 than in the C2SS. Median household income in Arizona and Nevada were eight percent higher in the Census than in the C2SS. At the site level, Madison County, Mississippi's median household income was more than nine percent higher in the Census than in the C2SS. Only three states in the C2SS had median household incomes slightly higher than that of the Census: Hawaii, Mississippi, and Rhode Island. The only ACS site that came in slightly higher than the census was San Francisco County, California.

We looked into several possible explanations for the higher income amounts in the Census. The difference could be the result of certain census / ACS differences such as survey design, mode of interviewing, different data processing methods, as well as the different reference periods.

As far as the reference periods are concerned there is one possible explanation for the lower figures for the "PAST 12 MONTHS." A respondent might read the instructions

and "key in" on the word "Month" and report "monthly" amounts in the income fields. The income edits are designed to catch monthly amounts and would have caught many of these types of mistakes but some of the higher monthly amounts indicated by respondents could have remained unchanged thus lowering the median income. As a precautionary measure and to help prevent this situation, we changed the instructions slightly on the questionnaire in the late 1990s to emphasize annual amounts. The instructions were changed to read "Mark (X) the "Yes" box for each type of income this person received, and give your best estimate of the TOTAL AMOUNT during the PAST 12 MONTHS. (NOTE: The "past 12 months" is the period from today's date one year ago up through today.)" The C2SS / 2000 ACS questionnaire reflected this change although monthly income amounts still may be at least partially to blame for the lower median household incomes.

Another possible reason for the higher income amounts in Census 2000 is the data capture methodology. As previously stated, we used the Optical Character Recognition (OCR) technology to read handwritten data entry by machine in the census. We know that OCR produces higher income amounts than having actual "keyers" record the data. Still, that would not account for the results of a previous study of 1996 ACS income data in comparison to results from the 1990 Census. The 1990 Census did NOT employ the OCR technology. (For results of the 1996 ACS comparisons of household income to the results of the 1990 Census, see Table 4).

C2SS median household income at the national level matched up more closely with that of the Current Population Survey (CPS) estimates. We compared the 2000 ACS results with the March 2001 CPS (in 2000 dollars with 2000 population controls). The CPS national median household income estimate was \$41,990, only 1.2 percent higher than the ACS estimate. The C2SS estimates of median household incomes were lower than those of the CPS in 34 states and the District of Columbia. The C2SS estimates came in higher than those of the CPS in 16 states.

It is interesting (and somewhat puzzling) to note that the Census 2000 estimate of median household income at the national level in 2000 dollars (\$43,396) was 5.9 percent higher than the calendar-year 2000 estimate from the CPS. This is puzzling because conventional wisdom tells us that the CPS, which has a much more detailed set of income questions than Census 2000/C2SS/ACS, would result in higher income levels. Thus, the fact that the C2SS median estimate at the national level is slightly lower median than the CPS estimate is not surprising. The fact that the Census 2000 median is quite a bit higher than the CPS is, however, quite surprising. Of the 3 major Census Bureau household survey-based estimates of median incomes at the national level, the outlier is the Census 2000 estimate, not the C2SS or CPS estimate. This relationship is generally true that the state level as well. For example, the Census 2000 estimate of median household income in California was \$49,078, about 5 percent higher than both the CPS and C2SS estimates.

Table 4: ACS Median Household Income vs. 1990 Census estimates: (1996 Dollars)			
	1989	1996	1996
	Census Adjusted Medians	True ACS Medians	ACS Medians (Grouped Data)
Site:			
Brevard Co., Florida	\$38,635	<b>\$32,728</b>	\$32,735
Multnomah Co., Oregon	\$34,073	<b>\$32,732</b>	\$32,777
Rockland Co., New York	\$66,722	<b>\$60,163</b>	\$60,155
Fulton Co., Pennsylvania	\$30,034	<b>\$28,058</b>	\$27,996

When making comparisons between the Census income figures and those of the ACS, it is also important to keep in mind the different modes of data collection. That is, when gathering information from households who failed to respond via mail, the ACS used computer-assisted telephone interviewing (CATI) and computer-assisted personal interviewing (CAPI) where the 1990 Census did not. This is yet another possible explanation for the differences in median household income.

Also, as previously discussed, the 1997 split-panel test showed that respondents report wage or salary income significantly lower when asked about the “past 12 months” than they do when asked about the last calendar year. Wage or salary income makes up most (about 75 percent) of aggregate household income. Further analysis is needed in order to determine why respondents report lower wage / salary amounts for the “past 12 months” versus the last calendar year. (For components of aggregate household income see Figure 3.)

## Conclusions

This paper presented income comparisons between the Census 2000 Supplementary Survey (C2SS) and Census 2000 at the national and state level. In doing so, it also provided a summary of the major differences between the two income data sources, in terms of data collection, capture, and processing, and provided very preliminary assessments of the possible role these differences may have played. CPS estimates of income at the national and state level were also provided in order to put the C2SS/Census 2000 differences in perspective.

The most obvious limitation to this study is that we don’t know the “truth,” meaning that we don’t know the true income of the C2SS and Census 2000 survey respondents. Thus, it is impossible to make firm judgments about the relative data quality of these quite different sources of sub-national survey-based income estimates.

The work presented here is very preliminary; there is much that needs to be done as we make the transition from the Census long form to the ACS as the Nation's preeminent source of sub-national income and poverty estimates. At this point we really don't come close to understanding the reasons why income estimates from Census 2000 were quite a bit higher than the C2SS, though we have thoughts on some of the factors behind these differences. More work has to be done on examining which of these factors played major roles in these differences. One of the Census 2000 evaluation projects, the CPS/Census 2000 Exact Match Project, may shed light on Census 2000 income reporting/imputation patterns that may in turn shed light on why the C2SS and Census 2000 figures differed. It is clear that we are just at the beginning stages of understanding why Census 2000 and C2SS income figures differ.

Figure 3: Aggregate Household Income and its Components (United States – C2SS)

