COMPARISON OF ACS-CPS DATA ON INDUSTRY, OCCUPATION, AND CLASS OF WORKER: 2003

Barbara Downs
Marisa Tegler
Mai Weismantle

Housing and Household Economic Statistics Division
U.S. Census Bureau

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical issues are those of the authors and not necessarily those of the United States Census Bureau.

COMPARISON OF ACS-CPS DATA ON INDUSTRY, OCCUPATION, AND CLASS OF WORKER: 2003

INTRODUCTION

This report is one in a series that compares data from the American Community Survey (ACS) with data from the Current Population Survey (CPS). This report focuses on comparisons of national distributions of industry, occupation, and class of worker data between the 2003 ACS and the 2003 CPS.[1] In this analysis, we compare the 2003 ACS and 2003 CPS distributions, look for differences that are both statistically and substantively different, and for those found, offer possible explanations. The 2003 ACS did not sample the group quarters population, while the 2003 CPS includes information obtained from interviews with residents sampled from non-institutional group quarters, such as college dormitories, emergency and transitional shelters, and group homes. Therefore, the analysis universe contains both household and group quarters population, though group quarters housing units account for less than 1 percent of the CPS sample.[2] This variation in universes may be minimized because this report focuses on national percentage distributions.

[1] The 2003 data were chosen for this comparison because 2003 was the first year that both the ACS and CPS referenced the 2002 North American Industry Classification System (NAICS) to code industry responses and the first year that both the ACS and CPS referenced the 2000 Standard Occupational Classification (SOC) to code occupation responses in data processing (ACS referenced the 2000 SOC prior to 2003).

[2] On the ASEC 2003 file, there were 66 interviewed group quarters units out of 78,310 total interviewed units.

METHODOLOGY

The tables included in this report compare the most commonly tabulated data on industry, occupation, and class of worker from the ACS and CPS. Comparisons consist primarily of percentage-point differences between the two distributions. Tables display the ACS and CPS estimates, [3] the margins of error from which 90-percent confidence intervals of the estimates can be derived, and the difference between the two estimates. In the case of relative frequency distributions, the difference is calculated as the percentage-point difference between the two estimates. An asterisk (*) denotes statistically significant differences.

At the national level, the ACS and CPS variances were quite small, resulting in many statistically significant differences between the ACS and CPS distributions. In this report, we generally consider statistically significant differences of 0.5 percentage points or less as not important. This yardstick was developed to help focus the analysis, though it can vary based on the relative size of the category. For example, for population groups constituting a relatively large percentage of the population (i.e., office and administrative support occupations), a 0.5 percentage point difference in the estimates might be small, while for population groups constituting a smaller percentage of the population (i.e., mining occupations), a 0.5 percentage point difference could be quite large. This

[3] The most striking difference between the ACS and CPS labor force estimates is in the size of the civilian employed noninstitutional 16 years and over population, the universe on which the occupation data are based. The CPS and ACS national estimates of the employed population differ by more than five million with CPS being the larger of the two estimates. Because of the large difference in the size of the employed population, this analysis focuses on the percentage distributions of industry, occupation, and class of worker rather than the population estimates themselves.

decision is subjective, however, and users can apply their own standards to interpret the data presented in this report.

The remainder of this section examines differences in methodology between the two surveys.

Sample Frame

The ACS derives its sample frame from a national Master Address File (MAF) that the Census Bureau maintains. The MAF is continuously updated using the U.S. Postal Service Delivery Sequence File (DSF), ACS non-response follow up, updates from special census operations, and the Community Address Updating System (CAUS). The CPS sample uses address lists from the Decennial Census along with area sampling in places without address lists to produce its sample frame and updates it using the Building Permit Survey in order to account for new housing units.

The 2003 ACS surveyed a national sample of housing units, both occupied and vacant. Data were collected in a total of 1,235 counties out of the 3,141 counties in the United States. The sample is designed to provide estimates of housing and socio-economic characteristics for the nation, all states, most areas with a population of 250,000 or more, and selected areas of 65,000 or more.

The 2003 CPS surveyed a national sample of occupied housing units or households. Households in the CPS are interviewed for four consecutive months, not interviewed for

the next eight months, and then reactivated again for four more months. The sample is designed primarily to produce monthly estimates of the labor force characteristics of the civilian non-institutionalized population 16 years of age and older for the nation and all states. For comparison purposes, this report uses published CPS data in the form of annual averages.

One difference between the two survey universes is that the CPS included a small number of individuals living at addresses that were housing units in 2000 but were later converted to noninstitutional group quarters (e.g. emergency and transitional shelters and group homes).

Sample Size and Mode of Data Collection

The 2003 ACS interviewed a total of 572,447 households. Data were collected continuously throughout the year using a combination of mail-out/mail-back questionnaires, Computer-Assisted Telephone Interviewing (CATI), and Computer-Assisted Personal Interviewing (CAPI). Each month a unique national sample of addresses received an ACS questionnaire. Addresses that did not respond were telephoned during the second month of collection if a phone number for the address was available, and personal visits were conducted during the third and last month of data collection for a subsample of the remaining nonresponding units. The 2003 ACS achieved an overall survey response rate, calculated as the initially weighted estimate of interviews divided by the initially weighted estimate of cases eligible to be interviewed, of 96.7 percent.

The 2003 CPS contained interviews from about 55,000 households each month of the year. All CPS data are collected via Computer-Assisted Telephone and Personal Interviews (CATI/CAPI). The response rate for the 2003 CPS was 92.8 percent.[4] The difference in response rates between the surveys may be influenced by the number of times each household was contacted for a survey. The ACS is administered once per year while the CPS contacts households up to four times per year.

Both the ACS and CPS employ experienced permanent interviewers for CATI and CAPI data collection.

Residence Rules

The ACS and the CPS employ different residence rules to determine which individuals in a household are eligible for interview; the ACS uses the concept of current residence, while the CPS uses a version of usual residence. This difference may contribute to variation in the universes on which social characteristics depend.

The ACS interviews everyone in the housing unit on the day of interview who is living or staying there for more than two months, regardless of whether or not they have a usual residence elsewhere. If a person who usually lives in the housing unit is away for more than two months at the time of the survey contact, he or she is not considered to be a current resident of that unit. This rule recognizes that people can have more than one place where they live or stay over the course of a year.

[4] The response rate is unweighted.

The CPS interviews everyone staying in the housing unit at the time of the interview who considers the housing unit as their usual residence or who have no usual residence elsewhere. In addition, the CPS also considers individuals who are temporarily absent at the time of interview but who consider the housing unit their usual residence.

The different residence rules result in one notable difference in the universe of the two surveys. Because the 2003 ACS excluded group quarters from the sample frame and interviewed individuals at their current residence, college students living in dormitories are not included in the ACS universe. ACS included those students if they were living off campus in a sampled housing unit for more than two months when that housing unit was contacted by the survey. In contrast, the CPS interviewers are instructed to include as household members any college students who are temporarily absent from the household, including those who are currently residing in college dormitories. The result being that the CPS sample universe should include more college students than the ACS sample universe.

Question Wording and Reference Periods

Differences between the ACS and the CPS in presentation and wording of questions may contribute to differences in estimates. For example, the CPS uses dependent questioning formats, where the respondent is asked to verify or update answers from previous interviews, rather than being asked the same question during each interview. The ACS respondents are interviewed only once, thus their responses to questionnaire items are

independent of prior responses. Also, the ACS asks only about a worker's primary job, whereas the CPS questionnaire repeats similar questions regarding a respondent's second job.[5] The 2003 questionnaire items are listed below.

ACS 2003 Industry, Occupation, and Class of Worker Questions.
The 2003 American Community Survey asks the industry, occupation, and class of worker items as follows:

Current or Most Recent Job Activity. Describe clearly this person's chief job activity or business last week. If this person had more than one job, describe the one at which this person worked the most hours. If this person had no job or business last week, give information for his/her last job or business.

35. Was this person—
-An employee of a PRIVATE FOR PROFIT company or business, or of an individual, for wages, salary, or commissions?
-An employee of a PRIVATE NOT FOR PROFIT, tax-exempt, or charitable organization?
-A local GOVERNMENT employee (city, county, etc.)?
-A state GOVERNMENT employee?
-A federal GOVERNMENT employee?
-SELF-EMPLOYED in own NOT INCORPORATED business, professional practice, or farm?
-SELF-EMPLOYED in own INCORPORATED business, professional practice, or farm?
-Working WITHOUT PAY in family business or farm?

36. For whom did this person work?
(*Name of company, business, or other employer*)

37. What kind of business or industry was this?
(*Describe the activity at the location where employed.* For example: hospital, newspaper publishing, mail order house, auto engine manufacturing, bank)

38. Is this mainly—
-Manufacturing?
-Wholesale Trade?
-Retail Trade?

[5] In both the ACS and the CPS, if the respondent has not worked in the five-year period prior to when they complete the survey questionnaire, they are instructed to skip over the class of worker, industry and occupation items.

-Other (agriculture, construction, service, government, etc.)?

39. What kind of work was this person doing?
(For example: registered nurse, personnel manager, supervisor of order department, secretary, accountant)

40. What were this person's most important activities or duties?
(For example: patient care, directing hiring policies, supervising order clerks, typing and filing, reconciling financial records)

CPS 2003 Industry, Occupation, and Class of Worker Questions.
The 2003 Current Popualtion Survey asks the industry, occupation, and class of worker items as follows:

IODP1. Last month, it was reported that you worked for (input company name). Do you still work for (input company name) at your MAIN job?
- Yes
- No

IODP2. Have the usual activities and duties of your job changed since last month?
- Yes
- No

IODP3. Last month you were reported as (a/an) (input occupation) and your usual activities were (input duties). Is this an accurate description of your current job?
- Yes
- No

IO1INT. Now I have a few questions about the (job/main job) (at which you (last worked/worked) (last week/the week before last) /from which you were absent/from which you are on layoff). Were you employed by government, by a private company, a nonprofit organization, or were you self employed or working in the family business?
- Government
- Private for profit company
- Nonprofit organization including tax exempt and charitable organizations
- Self employed
- Working in the family business

IO1GVT. Would that be the federal, state, or local government?
- Federal
- State
- Local (county, city, township)

IO1INC. Was this business incorporated?
- Yes
- No

IO1WP. Were you working for pay?
- Yes

- No

IO1NMP. What is the name of the (company/nonprofit organization) for which you worked at your MAIN job? (DO NOT READ TO RESPONDENT: name of company, business, organization or other employer)

IO1NMG. What is the name of the government agency for which you worked at your MAIN job?

IO1NMB. What is the name of (your business / the name of the business for which you worked)?

IO1IND. What kind of business or industry is this? (READ IF NECESSARY: What do they make or do where you worked?)
(PLEASE PROVIDE MORE DETAILED INFORMATION FOR THE CODERS.)

IO1MFG. (ASK IF NECESSARY:) Is this business or organization mainly manufacturing, retail trade, wholesale trade, or something else?
- Manufacturing
- Retail trade
- Wholesale trade
- Something else

IO1OCC. What kind of work (do/did) you do, that is, what (is/was) your occupation? (For example: plumber, typist, farmer.)
(PLEASE PROVIDE MORE DETAILED INFORMATION FOR THE CODERS.)

IO1DT. What (are/were) your usual activities or duties at this job? (For example: typing, keeping account books, filing, selling cars, operating printing press, laying brick.)
(PLEASE PROVIDE MORE DETAILED INFORMATION FOR THE CODERS.)

IOPRO1. Now I have a few questions about your second job. (READ IF NECESSARY: By second job I mean the one that you work the second most hours.)

From this point, the CPS instrument guides the interviewer through questions IO1INT through IO1DT for the respondent's second job, if they have one.

Item Nonresponse

Item nonresponse is the failure of an individual to provide complete and usable information for a data item. Item allocation rates are often used as a measure of the level of item nonresponse. These rates are computed as the ratio of the number of eligible people or households for which a value was allocated during the editing process for a

specific item to the number of people or households eligible to have responded to that item.

For the 2003 ACS, weighted allocation rates for the industry, occupation, and class of worker items were 6.6 percent, 6.9 percent, and 5.5 percent, compared with 10.3 percent, 10.3 percent, and 10.2 percent unweighted allocation rates for the 2003 CPS annual data.[6]

Data Editing and Imputation Procedures

ACS and CPS edit and imputation rules are designed to ensure that the final edited data are as consistent and complete as possible. These rules are used to identify and account for missing, incomplete, and contradictory responses. In each case where a problem is detected, pre-established edit rules govern its resolution.

The ACS and the CPS employ two principal imputation methods: relational imputation and hot deck allocation. Relational imputation assigns values for blank or inconsistent responses on the basis of other characteristics on the person's record or within the household. Hot deck allocation supplies responses for missing or inconsistent data from similar responding housing units or people in the sample.

Both the ACS and CPS editing procedures employ logical checking routines to produce consistency within an individual's responses. Respondents must have occupations

[6] The universe for these items in ACS is the civilian household population age 16 years and older. In CPS, it is the civilian non-institutionalized population age 16 years and over. However, the ACS allocation rates are weighted estimates while the CPS allocation rates are unweighted.

consistent with their industries, class of worker, income per week, and educational attainment. For example, if a respondent has less education than a Master's degree, but reports that he/she is a dentist, the occupation is changed to dental hygienist. Additionally, if one of the items is missing, the information is imputed from a donor with similar responses to other items. Allocation routines using hot decks generally stratify the donors and recipients of the hot deck by their age, sex, race, and other characteristics.

Controls and Weighting

There are notable differences in the selection of controls and the calculation of weights between the two surveys that may lead to differences in estimates. The ACS and CPS are both weighted to account for the probability of selection and housing unit nonresponse.

After the initial weighting, data from the ACS and CPS are both controlled to be consistent with independent population estimates. Data from the 2003 ACS were controlled, at the county level, to the July 2003 independent estimates of the population of individuals and housing units. The 2003 CPS was controlled to independent national estimates of the population of individuals. Because the ACS controlled to both the total population and the total number of housing units, the ACS files contained both person weights and housing unit weights. The CPS does not measure housing unit characteristics and therefore does not control to an independent housing unit estimate. CPS files do not contain an independent housing unit weight but instead use the weight of the householder to produce household estimates if required.

After the application of the population controls to the CPS data, a second set of weights is created for the Bureau of Labor Statistics to use in tabulating their labor force data. In this weighting adjustment, "composited" estimates of the current month's labor force status (employed, unemployed, and not in the labor force) are used as controls for these "composited" weights. The ACS employs no similar procedure.

RESULTS

Starting at the higher aggregations of industry and occupation categories, this report compares each category's percentage distribution value and tests differences between the surveys for the total employed population[7] and by sex (Tables 1 and 3). Based on the results found in those analyses, detailed estimates were created and tested to discover the source of the higher-order category differences (Table 2). Table 4 compares the major and detailed Class of Worker classifications in the ACS and CPS, for the total employed population and by sex. The outcomes from these comparisons may lead to additional research and testing of these items.

Occupation

The results of an analysis of six major occupation categories are shown in Table 1. The analysis reveals noteworthy differences in two categories: "Management, professional, and related occupations" (-0.7%) and "Sales and office occupations" (0.5%). Both sexes report "Management, professional, and related occupations" less often in ACS than in CPS (-0.7% and -0.6%, respectively, for male and female). Conversely, both males and females also report "Sales and office occupations" more often in ACS than in CPS (0.5%

[7] "Total employed population" refers to the employed civilian population age 16 years and over.

each). More ACS occupation groups are statistically different from CPS among males than among female groups.

Separating these major occupation groups into more detailed categories sheds some light on which subgroups may contribute to the differences found in the major aggregated categories. In Table 2, the "Management, professional and related occupations" category is divided into two separate subcategories and then further divided into ten minor classification subgroups. The minor classification subgroup entitled "Management occupations (including farmers and farm managers)" has an ACS-CPS difference of -1.0%. This negative value means that this minor classification subgroup is reported more often in the CPS than the ACS. Because this ACS-CPS difference is the largest of all the minor classification subgroups, it appears, then, that the key to the ACS-CPS difference in "Management, professional and related" is the measurement of the subgroup "Management occupations (including farmers and farm managers)."

The significant differences between estimates for the "Sales and office occupations" major category (0.5%) appears to be driven by the ACS-CPS difference in measurement of "Office and administrative support occupations" (0.7%), which is offset by a difference in the opposite direction of "Sales and related occupations" (-0.2%). In other words, respondents to the ACS are more likely to report "Office and administrative support occupations," while respondents to the CPS are more likely to report working in "Sales and related occupations."

Industry

Focusing only on those differences greater than 0.5%, none of the industry classifications differ between surveys. However, two of the ACS and CPS categories' percentages have statistically significant differences of 0.5%, the threshold for discussion. The differences are found in "Wholesale trade" (0.5%) and "Public administration" (0.5%), with respondents to the ACS more often reporting these industries.

Similar to the occupation findings, statistically significant differences vary by sex. The two major industry group differences ("Wholesale Trade" and "Public Administration") appear to be due more to the ACS-CPS differences measured for males (0.6% and 0.7%, respectively), than for females (all statistically significant differences are smaller than 0.5%).

Because the significant differences in Table 3 are on the cusp of the threshold designated in this report, the industry subcategories were broken out to explore the source of each significant difference. No patterns emerged from this exploration. We do not present a separate table of these distributions.

Class of Worker

The major class of worker categories are: private wage and salary workers; government workers; self-employed workers in own not incorporated business; and unpaid family workers. All four categories are statistically different between ACS and CPS, but only

two of those differences are greater than 0.5%. The greatest differences are found in the "Private wage and salary workers" (-0.7%) and "Government workers" (0.9%) categories.

Subdivision of the major class of worker categories shows the ACS-CPS difference in "Private wage and salary workers" to be the result of off-setting statistically significant differences in the "Private for-profit" (-2.2%) and "Private not-for-profit" (1.5%) classifications. These same differences appear when the data are cross-classified by sex, indicating a systematic difference in the measurement of these classes of workers, with the ACS respondents reporting fewer "Private for-profit" workers and the CPS reporting fewer "Private not-for-profit" workers. During the coding process, the ACS coders can alter the class of worker entries if the alteration provides a more accurate match with the industry, occupation, and job duties information given by the respondent. For example, if the respondent claims to be an employee of a private for profit company, but later provides information on working for the town government, the coder could change the class of worker code to indicate that the respondent is a local government employee. CPS protocols do not allow this practice. This difference in protocols may explain some of the measurement difference.

The difference in government work appears to be the cumulative effect of small measurement differences in the detailed government categories.

SUMMARY

Statistical testing between the 2003 ACS and CPS annual average percentage distributions for occupation, industry, and class of worker confirms the overall comparability of the ACS data for these items. Across all three variables, there are only three differences exceeding the threshold for comparability. However, several findings from our analyses are worth noting. Measurement of occupation groups does not differ for four of the six groups, but closer examination may be needed to understand why ACS underestimates the "Management, professional and related occupations" and overestimates "Sales and office occupations" relative to the CPS data. The ACS distribution of persons by industry is, in most cases, similar to the CPS, with no differences exceeding the threshold for discussion.

Comparing ACS and CPS measurement of class of worker does yield significant differences. First, the two main types of "Private wage and salary workers" differ between surveys. "Private for-profit" is measured less often in ACS than CPS, and the converse occurs for "Private not-for-profit", with a net undercount of the overall category in ACS relative to CPS. The other main categories of class of worker also differ between the surveys. The reasons for the differences in these estimates cannot be explained with the data at hand, but these results suggest that further investigation of the origin of the differences is warranted. Perhaps slight question differences or coding protocols yield these differences.

Overall, the ACS and CPS estimates of industry, occupation, and class of worker are comparable. Further investigation may shed light on the determinants of the few noted differences.

Appendix 1: Methodological Differences between ACS and CPS for survey year 2003

	ACS 2003	CPS 2003
Universe differences:	Civilian household population age 16 years and over	Civilian noninstitutional population age 16 years and over
Data Collection Mode:	Paper questionnaires, Computer-assisted telephone (CATI) and personal (CAPI) laptop interviewing	Computer-assisted personal (CAPI) laptop interviewing and Computer-assisted telephone (CATI) interviewing [1]
Interviewing frequency:	Interviewed once	Interviewed for 4 months, out of rotation for 8 months, back in for 4 more months [2]
Sample Frame:	Master Address File (MAF)	Decennial Census, Building Permit Survey, and area samples of new construction
Sample Size:	829,000 housing unit addresses	72,000 addresses a month (Of those, about 60,000 are found to be eligible for interview.)
Actual interviews:	Approx. 572,447 households interviewed in 2003	Approx. 55,000 households interviewed a month [3]
Response Rate:	96.7%	92.8%
Coverage Rate:	94.1%	89.0%
Number of Items in Questionnaire:	Mail / CATI / CAPI	CAPI
Class of Worker	1 / 4 / 1	4
Industry	3	5
Occupation	2	3
Item Allocation Rates:		
Class of Worker	5.5%	10.2%
Industry	6.6%	10.3%
Occupation	6.9%	10.3%
Industry and Occupation Coding:	Used the 2002 Census classification systems (based on the 2002 NAICS and 2000 SOC)	Same as ACS

(1) CAPI is used for the 1st and 5th month in sample. All other months are CATI. All months have CAPI non-response follow-up.

(2) These data cover 2003, but the actual months during which these data were collected are unknown.

(3) This translates into approximately 666,000 interviewed households per year but households are rotated in and out of the survey (see Sample Size and Data Collection Mode).

Source: U.S. Census Bureau

List of Basic Tables

Table 1. Comparison of Major Occupation Group Percentages by Sex: 2003

Table 2. Comparison of Detailed Occupation Groups: 2003

Table 3. Comparison of Major Industry Groups by Sex: 2003

Table 4. Comparison of Major Class of Worker Groups by Sex: 2003