September 3, 2009

Investigation of User Preferences for Measures of Sampling Error to be Displayed on American Community Survey Data Products and Modification of Definitions of these Measures

FINAL REPORT

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

LYNN WEIDMAN
KATHLEEN T. ASHENFELTER
STATISTICAL RESEARCH DIVISION

STUDY SERIES
(Statistics #2009-01)

Project Report: Investigation of User Preferences for Measures of Sampling Error to be Displayed on American Community Survey Data Products and Modification of Definitions of these Measures

Lynn Weidman
Kathleen T. Ashenfelter

Statistical Research Division
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: September 3, 2009

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau

Table of Contents

List of Tables

Executive Summary

As part of its efforts to improve the usability of American Community Survey (ACS) data products, the Census Bureau examined two related issues associated with the presentation of information on sampling error for ACS estimates. The first issue concerned the determination of data user preference for the measure of sampling error, margin of error or confidence interval, to be displayed with estimates in ACS data products. To explore this issue, we distributed a Web survey requesting information about each user's background in statistics, their use of ACS data products, and their preference for displaying sampling error. The sampling frame was the *ACS Alert* email list and subjects were recruited to take the survey via an email. The eligible response rate for the full list was 9.7 percent (324 out of 3,331 possible respondents). The results of the survey show that of those respondents who expressed a preference and do not work for the Census Bureau, 61.2 percent favored the confidence interval and 33.6 favored the margin of error. The second issue was the need for revised definitions of the terms "standard error", "confidence interval", and "margin of error" for use in American Factfinder (AFF). For this purpose, we examined underlying concepts for these terms as they appear in textbooks and throughout statistical literature in an effort to develop alternative definitions that were clearer and more succinct. It is anticipated that improved definitions will contribute to a better understanding of ACS data products in AFF.

1.0 Introduction to Survey on Display of Sampling Error for ACS Data Products

Although this research originally had a broader scope in addressing questions about the display and calculation of measures of sampling error for ACS data products, after several months we decided to focus our resources on a specific issue, "Which sampling error measures are preferred by ACS data users?" Options included standard error, margin of error, coefficient of variation, and confidence bounds. For a more complete description of the questions not covered by the present research, please see the original project proposal (Weidman, 2007).
We planned to elicit user feedback through questionnaires and mock-ups of display options. Auxiliary information obtained would include a user's statistical sophistication, type of job, and purposes for using ACS estimates.

The first task was to design a survey questionnaire. It included only the two choices of displaying sampling error, confidence interval bounds or margin of error, since these are the two types of sampling error that have been presented in past ACS data products. Confidence bounds were presented in the ACS data products for the 2004 ACS and previous years, and margins of error have been presented in the 2005 ACS and subsequent year data products. In addition, the standard error and coefficient of variation can be easily calculated from the margin of error and we did not want to confuse less statistically-sophisticated users by asking them to choose among too many closely-related measures.

1.1 Determining the Type of Survey and Sampling Frame

While the questionnaire was being developed, we discussed what would be its appropriate target audience and sampling frame. Initially, we thought that we would concentrate on novice ACS

users or those unfamiliar with the ACS by bringing people into the Census Bureau for cognitive interviewing. Three sources of potential interview subjects were available from the Communication, Information, and Education Staff of the American Community Survey Office (ACSO): the *ACS Alert* email list (which is an internally-maintained list of self-selected ACS data users who signed up for email updates about ACS data products), Maryland librarians, and local educators. However, we would only be able to do a limited number of cognitive interviews and the people selected from these sources might have little interest in using sampling error. So, in order to contact more people with a potentially broader range of statistical sophistication, we decided to do an on-line survey.

We considered possible sources for sampling frames that would include people with different degrees of statistical sophistication and familiarity with ACS data products. This would allow us the flexibility to analyze display preferences separately for different levels of statistical sophistication. Sources we considered included the Association of Public Data Users (APDU), the Population Association of America, (PAA), Council of Professional Associations on Federal Statistics (COPAFS), and the *ACS Alert* email list. Factors considered in deciding whether to use the *ACS Alert* list or one of the external sources included (a) the *ACS Alert* list was the most accessible of these options, (b) the other organizations might not allow us to use their membership lists, and (c) many of the people on those lists were likely to be unfamiliar with ACS data products or their use. We decided to use the *ACS Alert* email list as our source due to its availability and because most of the people on the list should presumably be familiar with the ACS.

1.2 Survey Instrument and Email Messages

The survey instrument posed questions about the statistical sophistication of respondents, the ACS data products they use, how often they use these products, what type of organization they work for, what measures of sampling error they use, their preferred method of sampling error display, and the reason for this preference. The instrument went through several rounds of revision to more accurately ask these questions. Drafts of the survey were distributed by the Statistical Research Division (SRD) usability staff to individuals with varying statistical backgrounds for informal feedback during the survey's development. The final survey questions are provided in Appendix A. We developed messages for the initial emails that notified sample members of their selection for the survey and how to respond to it, and reminder emails to be sent to the full sample after a given number of weeks. See Appendix C for these messages. We obtained clearance from Office of Management and Budget (OMB) for all of these materials.

1.3 Sample Design and Response Results

The *ACS Alert* email list, which we used as our sampling frame, contained 5911 addresses. The target number of completed interviews was 100, a manageable number of responses to transcribe and analyze with available resources. (It turned out that Customer Liaison and Marketing Services Office (CLMSO) provided an Excel file of the individual responses, so that this concern about handling the number of responses was unnecessary.) Based on an expected response rate of 1/3, a sample of size 299 was obtained by using a sampling rate of 1 in 20 addresses within each stratum defined by email domains (Table 1). (Type of address is the first 2 or 3 characters

following the last period of the email address.) No information was collected about the respondents' Internet protocol (IP) addresses used when responding to the survey, so we were not able to process the responses for possible duplication. We also did not have information available about the number of potential respondents who accessed the survey but did not attempt or complete it.

The notification emails were sent on June 24, 2008 and the reminder emails on July 17, 2008. Table 2 shows that we received only 23 completed questionnaires. A major reason for this is that the *ACS Alert* email list is out-of-date, with 51.5 percent of these sample addresses being undeliverable. In order to obtain a sufficient number of responses to analyze, we decided to include the rest of the *ACS Alert* email list in the sample. Emails were sent to the remaining addresses on this list on August 28, 2008. One of the recipients forwarded her email to at least one listserv so that people not in our sample could take the survey. We immediately closed down the survey so that no one could respond. We then prepared modified notification emails that included "IMPORTANT: Please do not forward the survey link in this email message to others." in bold at the top of the message. These revised letters were sent to the remaining list on September 8, reminder emails were sent on September 23, and the last day to respond was September 30, 2008. Column 3 of Table 2 shows the response results, which are similar to those of the initial sample. There were 324 total responses from both rounds of sampling available for analysis.

Table 1. Stratification of domain names

Type of Address	Domain	Number
gov + mil	Government + Military	1259
edu	Educational Institutions	1101
com	Commercial	1025
org	Organizations	990
us	United States	914
Other country*	Other Country	91
net + others	Internet + Remaining Types	531
TOTAL		5911

* Other country = al, as, au, az, ca, ch, de, dk, ee, es, fi, gu, il, jp, kr, lt, nc, ne, nl, nz, pl, pr, sa, se, si, th, tw, ua, uk, vi

Table 2. Outcomes for Email Addresses

Outcomes	Original Sample	Remainder	All Addresses
Emails Sent	299	5612	5911
Delivery Failures	154	2426	2580
Delivery Failure Rate	51.5%	43.2%	43.6%
Number of Possible Respondents	145	3186	3331
Number of Respondents	23	301	324
Eligible Response Rate	15.9%	9.4%	9.7%
Overall Response Rate	7.7%	5.4%	5.5%

For the most part we received responses to all questions for the completed interviews. If a question was left blank, the respondent was excluded from the analysis of that question.

2.0 Analysis of Survey Responses

Complete tables of results are provided in Appendix B. Only Tables 3 and 11 include the 31 respondents who said in question 1 that they worked for the Census Bureau. Those respondents are excluded from the rest of the tables since the purpose of the survey was to find out about how data users outside the Census Bureau use ACS tables and measures of sampling error. Percentages in the tables are calculated based on the number of respondents who answered each question. For those questions that asked for multiple responses, the percentages sum to greater than 100 percent. This is an analysis of the responses received only. Due to the large number of delivery failures and nonrespondents, we make no attempt to perform statistical comparisons or infer the results to the full population of ACS data users. All results are based on self-reported answers from respondents and have not been otherwise verified.

2.1 Characteristics of the Study Universe

Question 1 asked, "In what capacity do you use ACS data?" Respondents could choose more than one category, therefore the results sum to more than 100 percent. Survey respondents included data users from a broad set of professions. As shown in Table 3, the largest fractions of respondents included state or local government employees (34.0%), academic researchers (22.8%), other federal employees (12.0%) and Census Bureau employees (9.6%). Respondents also included individuals who are contractors, students, businesspersons, journalists, and other professions such as researchers for non-profits.

Most respondents have taken some statistics classes, with 68.5 percent having taken advanced undergraduate or graduate courses. The majority (57.2%) rate themselves as having intermediate statistical expertise and only 14.0 percent rate themselves as novices. Tables 6 and 7 summarize these results.

2.2 Uses of ACS Data Products

Table 4 shows that 35.4 percent of the non-Census Bureau respondents use ACS data daily or weekly and 36.4 percent reported using ACS data less than once a month. This result suggests that about two-thirds of the survey respondents are very familiar with at least one ACS data product. Question 3, "Which ACS data products do you use?" allowed more than one response, with the 293 non-Census Bureau respondents using an average of about 3 of data products. Table 5 summarizes these results with respondents most likely to use detailed tables (73.7%), data profiles (61.8%), and selected population profiles (52.6%). They were least likely to use other products (10.9%) and thematic maps (18.1%). Table 11 includes responses to question 3 for both Census Bureau and non-Census Bureau respondents, with the most notable difference being that non-Census Bureau respondents reported the use of detailed tables 1.76 times as frequently as did Census Bureau employees.

2.3 Use of Estimates of Sampling Error

Question 6 asked, "For what purposes do you use estimates of sampling error?" Multiple responses were allowed with a total of 555 responses provided. (Two respondents gave 'Don't use' as a response in addition to other responses, so are not included in any of the question 6 tabulations.) Overall results are presented in Table 8. Respondents most often use the sampling error measures for getting a general sense of the data (68.0%), to look for statistical significance of differences (50.5%), and to make comparisons (38.8%). Also, 16.2 percent of respondents publish estimates of sampling error. Although this is not the most frequently reported purpose of sampling error use, correct and accurate publication of estimates and their associated measures of error is of primary interest to the Census Bureau (Ashenfelter, Beck, and Murphy, 2008; Ashenfelter, Beck, & Murphy, 2009; Whitford & Weinberg, 2008).

When these purposes were analyzed by capacity of use (Table 12), several important differences were found. Other federal employees and journalists were more likely than all other types of data users to report using sampling errors to look for statistical differences. A smaller percentage of academic researchers report using sampling errors for making comparisons than do other federal employees, but a larger percentage report using them than members of the remaining groups do.

Tables 14, 16, and 18 summarize the purposes of use by data product, level of statistical training, and level of expertise with statistics. Data users reporting that they use "Other" data products were more likely than users of all other data products to look for statistically significant differences (67.6%) and publish sampling error (35.3%). The percent of users not using sampling error decreases for respondents with more than just an introductory statistics course, whereas the percent making comparisons and looking for statistically significant differences increases. Respondents classifying themselves as "experts" in statistics were more likely to acknowledge using sampling errors to make comparisons and look for statistically significant differences, and they were most likely of all the respondents to publish sampling errors. These two sets of results are not surprising. Table 20 shows that there are no major differences in use of data by type of sampling error used, except for "Other."

Table 10 summarizes responses to question 10 on the measures of sampling error that are used for individual's jobs. Of the different sampling error measures given as response options (confidence limits, margins of error, standard errors, and coefficients of variation), the coefficient of variation is used by the smallest percentage (16.7%) of respondents and the other three are each used by at least 47 percent of respondents.

2.4 Sampling Error Measure Preference

The overall result of interest is whether the preferred form of display for sampling error, as demonstrated in question 7, is the margin of error or the confidence interval. Table 9 shows that of the non-Census Bureau respondents who expressed a preference, 61.2 percent prefer the confidence interval and 33.6 percent the margin of error. (Although the survey instrument did not allow respondents to choose "prefer equally," some respondents wrote in for question 8 that they actually did not prefer one over the other. They were assigned the response "No

Preference" for question 7 in this analysis.) The written responses to question 8 were coded into eight categories as follows.

1. Easier to use, prefer, can more easily compare with other estimates
2. Easier to disseminate data table
3. Easier to read tables
4. Use in my job
5. More familiar with it
6. No need to calculate the margin of error or confidence interval bounds with this measure given
7. Easier to explain to others
8. Other

Table 23 shows the percentage of respondents who gave each of these as reasons based on the preference they identified in question 7. Of the 97 respondents who indicated a preference for margin of error, 86 provided a response to question 8 with a reason. The two reasons given most frequently were "easier to use, prefer, can more easily compare with other estimates" and "easier to read tables." Of the 177 respondents who indicated that they preferred confidence intervals, 158 provided a reason for this preference, with the most frequently given reasons being "no need to calculate…" and "easier to use, prefer, can more easily compare with other estimates." The largest reported difference between the preference groups was for the reason "no need to calculate…", where most of the responses were similar to "Because it saves me the work of calculating the bounds of the intervals." "Easier to explain to others" is the only other reason with a larger reported percentage difference, which was given by more respondents with a preference for the confidence interval. The reasons with larger percentage differences given by more respondents with a preference for the margin of error are "Easier to use" and "easier to read tables." "Easier to read tables" is reported only for those with a preference for the margin of error because they believed a single number representing sampling error was easier to read than two numbers were.

The two-way tables for question 7 on display preference indicate some relationships with respondent characteristics. When we look at the affiliation of the data user (Table 13), all of the user groups show a preference for confidence intervals. The greatest preference was seen for other federal employees and state/local government employees. Similarly, Table 15 analyzes user preference based on type of data product used. A preference for confidence intervals is seen regardless of which data products are used, with the exception that users of "Other data products" do not appear to have a preference. Tables 17 and 19 assess preferences by statistics classes taken and statistical expertise. Again, a preference for confidence intervals is seen across all groups with a slightly greater preference for confidence intervals found in individuals with the greatest expertise in statistics. As would be expected, the largest percentage of respondents who prefer the margin of error comes from those who use it in their jobs (question 10). Table 21 shows that regardless of the sampling error measure used as part of their jobs, respondents prefer the presentation of a confidence interval.

When looking at the responses for questions 6 and 7 cross-classified (Table 22), all purposes of use of sampling error show a preference for the confidence interval. So this analysis of the two-way tables shows that over all sets of characteristics, with perhaps the exception of those who use "Other" data products, the confidence interval is favored over the margin of error.

3.0 Development of Simplified Definitions of Standard Error and Related Measures

In addition to determining user preferences in sampling error display in ACS data products, we also wanted to develop more accurate and usable definitions of standard error, confidence interval, and margin of error by exploring concepts underlying these definitions in statistical literature.

3.1 Current Definitions

Following are the definitions as they currently appear on American Factfinder.

- **Standard error (American Community Survey)**
  The standard error is a measure of the deviation of a sample estimate from the average of all possible samples.

- **Confidence interval (ACS)**
  The sample estimate and its standard error permit the construction of a confidence interval that represents the degree of uncertainly about the estimate. Each American Community Survey and Census 2000 Supplementary Survey estimate is accompanied by the **upper and lower bounds** of the 90 percent confidence interval. A 90 percent confidence interval can be interpreted roughly as providing 90 percent certainty that the true number falls between the upper and lower bounds.

- **Margin of error**
  A margin of error is the difference between an estimate and its upper or lower confidence bounds. Confidence bounds can be created by adding the margin of error to the estimate (for an **upper bound**) and subtracting the margin of error from the estimate (for a **lower bound**). All published margins of error for the American Community Survey are based on a 90 percent confidence level.

When looking at these definitions, thoughts concerning their inadequacy were that
(1) standard error is the centerpiece that is used to determine the other two measures and should have a more complete definition,
(2) standard error is not used in defining either of the other measures, and
(3) formulas should be included that show the relationship among these measures.

Therefore, we concentrated on developing a clear and concise definition of standard error that would have the flavor of statistical exactness without relying on confusing statistical jargon, and clearly describing how the other measures are related to it. Also, the original definitions were developed for presentation in AFF as footnotes, so our revised definitions will face similar space

constraints. Novice users without prior knowledge about basic statistical concepts may find a more complete discussion of these concepts in the ACS Accuracy of the Data documentation, e.g. U.S. Census Bureau (2009).

3.2 Alternative Definitions

Our starting point for developing simplified definitions was the following set of definitions developed by SRD, using ideas from Moore & McCabe (2004).

> Three frequently used (but related) measures of uncertainty about results from sampling are (1) **standard error of the estimator**, (2) **margin of error of the estimator**, and (3) **confidence interval for the parameter**.
>
> In general, there are many possible samples that could be observed from a given population. In practice, we observe only one of these samples, and that sample leads to one estimate. Another sample would give another estimate, etc. Consider the average of all possible estimates. We hope that this average of all possible estimates is close to the unknown value of the parameter. (It is very rare that we would ever know with certainty either this average or the value of the parameter!) Consider a measure of the distance that one estimate for one sample is from the average of all possible estimates, a measure of the distance that a second estimate for a second sample is from the average of all possible estimates, etc.
>
> The concept of the **standard error of the estimator** is based roughly on an average of all of these distances. Small values of the standard error of the estimator suggest that the different estimates are each close to the average of all possible estimates and hence (we hope) that many of the different estimates are each close to the value of the parameter, especially the specific estimate for the one sample we actually observe. In practice, we actually use the data from the sample to estimate the standard error of the estimator.
>
> A **90% confidence interval for a parameter** is an interval computed from sample data by a method that has probability 90% of producing an interval containing the true value of the parameter. For a sample with a large number of units, a 90% confidence interval for the parameter is approximated by
>
> $$\text{estimate} \pm 1.65 \times (\text{standard error of the estimator}).$$
>
> The quantity $1.65 \times$ (standard error of the estimator) is called the **margin of error of the estimator at a 90% confidence level**.

The definition of standard error is complete but seems somewhat lengthy and possibly confusing to those with little statistical sophistication. Our goal was to shorten it without losing any important points. We presented a shortened version to two statisticians in SRD to review and modified it further based on their comments. The resultant definition of standard error and slightly modified definitions of a 90% confidence interval and margin of error were sent to the Decennial Statistical Studies Division (DSSD) for additional comments. Again, some modifications were made based on those comments, and this new version was given to the same SRD statisticians as before, who had no additional comments. The final modified definitions sent to ACSO and DSSD follow.

3.3 Final Modified Definitions

- **Standard Error**
  There are many possible samples that could be observed from a given population. In practice, we observe only one of these samples and one estimate of a population *parameter* is produced from it. Consider the average of these estimates from all possible samples. The **standard error** is an estimate of the *variability or how much these possible sample estimates differ from* their average. Smaller standard errors suggest that the estimates, including the one for the observed sample, tend to be closer to the population parameter.

- **Confidence Interval**
  A **90% confidence interval for a parameter** in the ACS is
  estimate $\pm$ 1.645 x (standard error).
  Over all possible samples, 90% of the intervals produced by this method contain the population value of the parameter.

- **Margin of Error**
  The quantity 1.645 x (standard error) is called the **margin of error of the estimator at a 90% confidence level**. It is one-half the width of a 90% confidence interval for a parameter.

4.0 Summary

Most of the respondents to the Web survey were state or local government employees, academic researchers, and other federal employees, and most of the respondents had taken some statistics classes. The majority of respondents also reported an intermediate level of expertise with statistics. The results of the survey on the preferred method of displaying sampling error in ACS data products show that, overall, respondents prefer confidence intervals over the margin of error. Future research on how sampling error is displayed in ACS data tables may build upon the results of this survey.

The revised AFF definitions of standard error, confidence interval, and margin of error sent to DSSD and ACSO will correct some existing inaccuracies and should more clearly and succinctly explain these statistical concepts to data users.

In general, the results of this work should lead to improving data users' experience with ACS data and with AFF statistical definitions. Enhancing data user understanding of sampling variance and its relevance to the data tables will also help to ensure that the estimates from these tables will be used more appropriately.

5.0 References

Ashenfelter, K., Beck, J., & Murphy, E. (December 2008). "Final Report for First-Round Usability Testing of Data-Reliability Indicator Prototypes", *HCI Memorandum 131*.

Ashenfelter, K., Beck, J., and Murphy, E. (2009). "Final Report for First-Round Usability Testing of Data-Reliability Indicator Prototypes", *Statistical Research Division Report (SSM2009-01)*.

Moore, D.S. and McCabe, G.P. (2004), *Introduction to the Practice of Statistics* (5th Edition), New York: W.H. Freeman and Company.

Weidman, L. (2007). "What Measures of Variability Should be Displayed on ACS Data Products?", A*CS Research and Evaluation Project Proposal D00170_2008011*, http://cww.acs.census.gov/Meeting%20Notes%20Presentations/R_E/Project_Documents/2008/ACSRE_proposal_D00170_20080111.doc

U.S. Census Bureau (2009). "Accuracy of the 2005-2007 3-Year ACS Data", http://www.census.gov/acs/www/Downloads/ACS/Accuracy2005-2007ACS3-Year.pdf

Whitford, D. and Weinberg, D. (2008). "Proposal to Highlight American Community Survey Data with a Data Confidence Indicator", 05/01/08.

Appendix A: Survey in Live Form Online, April 2009

Survey on the Display of Sampling Error in American Community Survey (ACS) Data Products March 2008

The U.S. Census Bureau is conducting a study for two purposes:

- To obtain information about the use of the measure of sampling error we present with American Community Survey (ACS) estimates
- To gain insight into ACS data users' preference, if any, for a particular measure of sampling error

The results of this survey will aid the Census Bureau in determining the measure(s) of sampling error to provide with the ACS data tables.

We estimate that the survey should require 10 minutes of your time. Your answers will only be used to help us make informed decisions about possible revisions to the way we display sampling error. Your responses are voluntary, and we ensure your confidentiality under the provision of Title 13 USC Section 9.

We will report findings in summary form only. Findings will be posted when available on the U.S. Census Bureau's Web site at: http://www.census.gov/mso/www/npr/mktbrfs.html.

The Office of Management and Budget approval number for this survey is 0607-0760. Without this number we could not conduct this survey or ask for your participation. Approval expires November 30, 2010.

Finish Later | Next >>

USCENSUSBUREAU
Helping You Make Informed Decisions

QuestionWeb.com Disclaimer

Powered By Apian QuestionWeb

Survey on the Display of Sampling Error in American Community Survey (ACS) Data Products March 2008

Completed

1. In what capacity do you use American Community Survey (ACS) data? (Check all that apply).

- ☐ State/local government employee
- ☐ Census Bureau employee
- ☐ Other federal employee
- ☐ Contractor
- ☐ Student
- ☐ Journalist
- ☐ Businessperson
- ☐ Academic researcher
- ☐ Other, please specify: _____

2. How often do you use ACS data products? (Select only one).

- ○ Daily
- ○ Weekly
- ○ Monthly
- ○ Less than once a month

3. Which ACS data products do you use? (Check all that apply).

- ☐ Data Profiles
- ☐ Selected Population Profiles
- ☐ Ranking Tables
- ☐ Subject Tables
- ☐ Detailed Tables
- ☐ Geographic Comparison Tables
- ☐ Thematic Maps
- ☐ Other, please specify: _____

4. What statistics classes have you completed?

- ○ No statistics course(s) completed
- ○ Introductory statistics course(s) only
- ○ Advanced undergraduate/beginning graduate level statistics courses only
- ○ Advanced graduate level statistics courses

5. Please rate your level of expertise with statistics.

- ○ Novice (just beginning to use statistics or rarely use them)
- ○ Intermediate (moderate experience with statistics)
- ○ Expert (a great deal of experience with and/or frequent use of statistics)

<< Back | Finish Later | Next >>

USCENSUSBUREAU
Helping You Make Informed Decisions

QuestionWeb.com Disclaimer

Powered By Apian QuestionWeb

Survey on the Display of Sampling Error in American Community Survey (ACS) Data Products March 2008

Completed

Estimates of sampling error, such as the margin of error and confidence intervals, are provided with all published ACS data.

6. For what purposes do you use estimates of sampling error? (Check all that apply).

- ○ Do not use sampling error
- ○ To get a general sense of the data
- ○ To make comparisons
- ○ To look for statistically significant differences
- ○ To publish sampling error
- ○ Other, please specify: _____

<< Back | Finish Later | Next >>

USCENSUSBUREAU
Helping You Make Informed Decisions

QuestionWeb.com Disclaimer

Powered By Apian QuestionWeb

Survey on the Display of Sampling Error in American Community Survey (ACS) Data Products March 2008

Completed

Imagine that for <u>your job</u> you are writing a report about young people and find two relevant ACS Data Tables.

Table 1:		
Prince George's County, Maryland		
Age Group	Estimated Total	90 % Margin of Error
Under 5	66,183	+/- 422
5-9	57,917	+/- 3,624
10-14	65,199	+/- 3,624
15-19	57,851	+/- 946

Table 2:			
State of Wyoming			
		90% Confidence Interval	
Age Group	Estimated Total	Lower Limit	Upper Limit
Under 5	31,029	30,125	31,933
5-9	28,392	26,293	30,491
10-14	32,677	30,668	34,686
15-19	33,123	31,876	34,370

7. Which of these tables would you find easier to use for your job?

- ○ Table 1 (with margin of error displayed)
- ○ Table 2 (with confidence intervals displayed)

8. Why do you find this table easier to use?

9. Is there any other information that would be useful to you?

- ○ Yes
- ○ No

If yes, please elaborate.

<< Back | Finish Later | Next >>

USCENSUSBUREAU
Helping You Make Informed Decisions

QuestionWeb.com Disclaimer

Powered By Apian QuestionWeb

Survey on the Display of Sampling Error in American Community Survey (ACS) Data Products March 2008

Completed

10. For your job, which measures of sampling error are useful to you when working with ACS data? (Check all that apply).

- ☐ Confidence limits
- ☐ Margin of error
- ☐ Standard error
- ☐ Coefficient of variation
- ☐ Other, please specify:

11. Based on your experience using ACS products, please give any other comments you have about how sampling error is presented?

<< Back | Finish Later | Next >>

USCENSUSBUREAU
Helping You Make Informed Decisions

QuestionWeb.com Disclaimer

Powered By Apian QuestionWeb

Survey on the Display of Sampling Error in American Community Survey (ACS) Data Products March 2008

Thank you for taking the time to help us in completing this survey!

<< Back | Finish Later | Send Answers

USCENSUSBUREAU
Helping You Make Informed Decisions

QuestionWeb.com Disclaimer

Powered By Apian QuestionWeb

Appendix B: Tables of Responses

Only Tables 3 and 11 include respondents who said in question 1 that they worked for the Census Bureau. They are excluded from the rest of the tables since the purpose of the survey was to find out about the use of ACS tables and measures of sampling error by non Census Bureau data users. The number of respondents included in each table is shown in parentheses following the table title. Percentages in Tables 2 through 10 are based on the table totals; percentages in Tables 11 and 23 are based on row totals; and percentages in Tables 12 through 22 are based on column totals. (Two respondents gave 'Don't use' as a response to question 6 in addition to other responses, so are not included in any of the question 6 tabulations.)

Table 3. Responses to Question 1: In what capacity do you use ACS data? (Check all that apply) (324)

Response	State/ Local Govt.	Census	Other Federal	Contractor	Student	Journalist	Business	Academic Researcher	Other
Count	110	31	39	12	6	5	17	74	77
Percentage	34.0	9.6	12.0	3.7	1.9	1.5	5.2	22.8	23.8

Table 4. Responses to Question 2: How often do you use ACS data products? (291)

Response	Daily	Weekly	Monthly	<1 time /month
Count	19	84	82	106
Percentage	6.5	28.9	28.2	36.4

Table 5. Responses to Question 3: Which ACS data products do you use? (Check all that apply) (293)

Response	Data Profiles	Selected Population Profiles	Ranking Tables	Subject Tables	Detailed Tables	Geographic Comparison Tables	Thematic Maps	Other
Count	181	154	79	111	216	99	53	32
Percentage	61.8	52.6	27.0	37.9	73.7	33.8	18.1	10.9

Table 6. Responses to Question 4: What statistics classes have you completed? (292)

Response	None	Intro stats	Advanced undergrad/ Beginning grad	Advanced grad
Count	29	63	85	115
Percentage	9.9	21.6	29.1	39.4

Table 7. Responses to Question 5: Please rate your level of expertise with statistics. (292)

Response	Novice	Intermediate	Expert
Count	41	167	84
Percentage	14.0	57.2	28.8

Table 8. Responses to Question 6: For what purposes do you use estimates of sampling error? (Check all that apply) (291)

Response	Don't use	General sense	Compare	Statistically significant difference	Publish SE	Other
Count	38	198	113	147	47	12
Percentage	13.1	68.0	38.8	50.5	16.2	4.1

Table 9. Responses to Question 7: Which of these tables would you find easier to use for your job? (289)

Response	Table 1: MOE	Table 2: CI	*No Preference
Count	97	177	15
Percentage	33.6	61.2	5.2

* The 'No Preference' category was determined from write-in entries to question 8.

Table 10. Responses to Question 10: For your job, which measures of sampling error are useful to you when working with ACS data? (Check all that apply) (293)

Response	Confidence Limits	MOE	SE	CV	Other
Count	191	177	139	49	15
Percentage	65.2	60.4	47.4	16.7	5.1

Table 11. Responses to Question 1 by Responses to Question 3 (Check all that apply for both) (324)

Q3 / Q1_CB*	Data Profiles	Selected Population Profiles	Ranking Tables	Subject Tables	Detailed Tables	Geographic Comparison Tables	Thematic Maps	Other	Total
Count									
No	181	154	79	111	216	99	53	32	293
Yes	21	18	13	13	13	10	9	6	31
Percentage									
No	61.8	52.6	27.0	37.9	73.7	33.8	18.1	10.9	
Yes	67.7	58.1	41.9	41.9	41.9	32.3	29.0	19.4	

*Q1_CB = Yes, if respondent is a Census Bureau employee on Question 1
= No, otherwise

Table 12. Responses to Question 1 by Responses to Question 6 (Check all that apply for both) (291)

Q1 / Q6	State/ Local Govt.	Other Federal	Contractor	Student	Journalist	Business	Academic Researcher	Other
Count								
Don't use	10	4	0	0	1	1	8	17
General sense	84	24	8	4	2	13	42	46
Compare	37	24	4	2	1	6	34	23
Stat Sig diff	53	27	7	3	4	7	37	31
Publish SE	19	7	3	0	0	5	15	8
Other	3	2	1	1	0	1	3	4
Percentage								
Don't use	9.2	10.3	0.0	0.0	20.0	5.9	11.1	23.3
General sense	77.1	61.5	66.7	80.0	40.0	76.5	61.1	63.0
Compare	33.9	61.5	33.3	40.0	20.0	35.3	47.2	31.5
Stat Sig diff	48.6	69.2	58.3	60.0	80.0	41.2	51.4	42.5
Publish SE	17.4	17.9	25.0	0.0	0.0	29.4	20.8	11.0
Other	2.8	5.1	8.3	20.0	0.0	5.9	4.2	5.5
Total	109	39	12	5	5	17	72	73

Table 13. Responses to Question 1 (Check all that apply) **by Responses to Question 7** (289)

Q1 / Q7	State/ Local Govt.	Other Federal	Contractor	Student	Journalist	Business	Academic Researcher	Other
Count								
Table 1: MOE	34	8	5	2	2	8	26	29
Table 2: CIs	69	26	7	3	3	9	42	40
***No Preference**	6	4	0	0	0	0	4	4
Percentage								
Table 1: MOE	31.2	21.1	41.7	40.0	40.0	47.1	36.1	39.7
Table 2: CIs	63.3	68.4	58.3	60.0	60.0	52.9	58.3	54.8
***No Preference**	5.5	10.5	0.0	0.0	0.0	0.0	5.6	5.5
Total	109	38	12	5	5	17	72	73

* The 'No Preference' category for question 7 was determined from write-in entries to question 8.

Table 14. Responses to Question 3 by Responses to Question 6 (Check all that apply for both) (291)

Q3 / Q6	Data Profiles	Selected Population Profiles	Ranking Tables	Subject Tables	Detailed Tables	Geographic Comparison Tables	Thematic Maps	Other
Count								
Don't use	16	20	10	8	18	11	9	3
General sense	137	112	59	82	157	70	32	22
Compare	74	64	35	54	90	46	21	15
Stat Sig diff	99	81	41	68	118	58	26	23
Publish SE	27	26	17	24	37	19	10	12
Other	5	6	3	5	10	2	5	3
Percentage								
Don't use	8.9	13.1	12.7	7.3	8.4	11.2	17.3	8.8
General sense	76.1	73.2	74.7	74.5	73.4	71.4	61.5	64.7
Compare	41.1	41.8	44.3	49.1	42.1	46.9	40.4	44.1
Stat Sig diff	55.0	52.9	51.9	61.8	55.1	59.2	50.0	67.6
Publish SE	15.0	17.0	21.5	21.8	17.3	19.4	19.2	35.3
Other	2.8	3.9	3.8	4.5	4.7	2.0	9.6	8.8
Total	180	153	79	110	214	98	52	34

Table 15. Responses to Question 3 (Check all that apply) **by Responses to Question 7** (289)

Q3 / Q7	Data Profiles	Selected Population Profiles	Ranking Tables	Subject Tables	Detailed Tables	Geographic Comparison Tables	Thematic Maps	Other
Count								
Table 1: MOE	52	45	29	39	75	31	15	15
Table 2: CIs	117	99	45	69	127	61	36	14
***No Preference**	11	9	4	3	12	6	2	3
Percentage								
Table 1: MOE	28.9	29.4	37.2	35.1	35.0	31.6	28.3	46.9
Table 2: CIs	65.0	64.7	57.7	62.2	59.3	62.2	68.0	43.8
***No Preference**	6.1	5.9	5.1	2.7	5.6	6.1	3.8	9.4
Total	180	153	78	111	214	98	53	32

* The 'No Preference' category for question 7 was determined from write-in entries to question 8.

Table 16. Responses to Question 4 by Responses to Question 6(Check all that apply) (290)

Q4 / Q6	None	Intro stats	Advanced undergrad/ Beginning grad	Advanced grad
Count				
Don't use	10	12	7	9
General sense	13	44	64	77
Compare	5	18	37	53
Stat Sig diff	10	26	42	69
Publish SE	1	10	10	26
Other	1	2	3	6
Percentage				
Don't use	34.5	19.0	8.3	7.9
General sense	44.8	69.8	76.2	67.5
Compare	17.2	28.6	44.0	46.5
Stat Sig diff	34.5	41.3	50.0	60.5
Publish SE	3.4	15.9	11.9	22.8
Other	3.4	3.2	3.6	5.3
Total	29	63	84	114

Table 17. Responses to Question 4 by Responses to Question 7 (289)

Q4 / Q7	None	Intro stats	Advanced undergrad/ Beginning grad	Advanced grad
Count				
Table 1: MOE	11	21	31	34
Table 2: CIs	17	37	49	74
***No Preference**	1	5	4	5
Percentage				
Table 1: MOE	37.9	33.3	36.9	30.1
Table 2: CIs	58.6	58.7	58.3	65.5
***No Preference**	3.4	7.9	4.8	4.4
Total	29	63	84	113

* The 'No Preference' category for question 7 was determined from write-in entries to question 8.

Table 18: Responses to Question 5 by Responses to Question 6 (Check all that apply) (311)

Q6 \ Q5	Novice	Intermediate	Expert
Count			
Don't use	15	21	2
General sense	23	121	55
Compare	8	63	42
Stat Sig diff	11	79	57
Publish SE	6	22	19
Other	1	4	7
Percentage			
Don't use	36.6	12.7	2.4
General sense	56.1	72.9	66.3
Compare	19.5	38.0	50.6
Stat Sig diff	26.8	47.6	68.7
Publish SE	14.6	13.3	22.9
Other	2.4	2.4	8.4
Total	41	166	83

Table 19: Responses to Question 5 by Responses to Question 7 (289)

Q7 \ Q5	Novice	Intermediate	Expert
Count			
Table 1: MOE	15	60	22
Table 2: CIs	23	99	55
***No Preference**	3	7	5
Percentage			
Table 1: MOE	36.6	36.1	26.8
Table 2: CIs	56.1	59.6	67.1
***No Preference**	7.3	4.2	6.1
Total	41	166	82

* The 'No Preference' category for question 7 was determined from write-in entries to question 8.

Table 20: Responses to Question 10 by Responses to Question 6 (Check all that apply for both) (291)

Q10 / Q6	Confidence Limits	MOE	SE	CV	Other
Count					
Don't use	13	15	7	1	7
General sense	140	129	100	33	6
Compare	89	77	70	29	2
Stat Sig diff	111	94	86	33	5
Publish SE	33	29	31	11	4
Other	6	3	6	1	3
Percentage					
Don't use	6.8	8.6	5.0	2.0	46.7
General sense	73.7	73.7	71.9	67.3	40.0
Compare	46.8	44.0	50.4	59.2	13.3
Stat Sig diff	58.4	53.7	61.9	67.3	33.3
Publish SE	17.4	16.6	22.3	22.4	26.7
Other	3.2	1.7	4.3	2.0	20.0
Total	190	175	139	49	15

Table 21: Responses to Question 10 (Check all that apply)**by Responses to Question 7** (289)

Q10 / Q7	Confidence Limits	MOE	SE	CV	Other
Count					
Table 1: MOE	41	73	40	16	4
Table 2: CIs	136	90	87	24	10
***No Preference**	13	12	10	7	1
Percentage					
Table 1: MOE	21.6	41.7	29.2	34.0	26.7
Table 2: CIs	71.6	51.4	63.5	51.1	66.7
***No Preference**	6.8	6.9	7.3	14.9	6.7
Total	190	175	137	47	15

* The 'No Preference' category for question 7 was determined from write-in entries to question 8.

Table 22: Responses to Question 6 (Check all that apply) **by Responses to Question 7** (289)

Q6 / Q7	Don't use	General sense	Compare	Statistically significant difference	Publish SE	Other
Count						
Table 1: MOE	15	67	30	42	14	4
Table 2: CIs	21	117	72	93	31	8
***No Preference**	2	12	10	11	2	0
Percentage						
Table 1: MOE	39.5	34.2	26.8	28.8	29.8	33.3
Table 2: CIs	55.3	59.7	64.3	63.7	66.0	66.7
***No Preference**	5.3	6.1	8.9	7.5	4.3	0.0
Total	38	196	112	146	47	12

* The 'No Preference' category for question 7 was determined from write-in entries to question 8.

Table 23: Responses to Question 8 by Responses to Question 7 (263)

Q8 / Q7	Easier to use	Easier to disseminate	Easier to read tables	Use in my job	More familiar with	No need to calculate	Easier to explain	Other	Total
Count									
Table 1: MOE	37	2	24	10	4	1	1	7	86
Table 2: CI	43	10	0	8	4	66	18	9	158
Percentage									
Table 1: MOE	43.0	2.3	27.9	11.6	4.7	1.2	1.2	8.1	
Table 2: CI	27.2	6.3	0.0	5.1	2.5	41.8	11.4	5.7	

Appendix C: Email Message Content for Survey on the Display of Sampling Error in American Community Survey (ACS) Data Products

Initial Email, both Rounds of Sampling, 06/24/08 and 08/28/08:

Dear American Community Survey User:

You have been selected to participate in an important research study regarding the data tables in the American Community Survey (ACS). The U.S. Census Bureau is conducting this study for two purposes:

(a) To obtain information about the use of the measure of sampling error we present with ACS estimates
(b) To gain insight into ACS data users' preference, if any, for a particular measure of sampling error

The results of this survey will aid the Census Bureau in determining the measure(s) of sampling error to be displayed in future ACS data products.

We estimate that the survey should require 10-15 minutes of your time. Your answers will only be used to help us make informed decisions about possible revisions to the way we display sampling error. Your responses are voluntary, and we ensure your confidentiality under the provisions of Title 13 USC Section 9*.

You may access the survey with this link:

https://questionweb.com/39475

We thank you for your time and look forward to receiving your feedback.

Sincerely,

U.S. Census Bureau

*We will report findings in summary form only. Findings will be posted when available on the Census Bureau's Web site at: http://www.census.gov/mso/www/npr/mktbrfs.html. The collection of information involved in this survey has been reviewed and previously approved by the Office of Management and Budget (OMB) under OMB control number 0607-0760.

Reminder Email 07/17/08:

Dear American Community Survey User:

This is a reminder that you have been selected to participate in an important research study regarding the data tables in the American Community Survey (ACS). The U.S. Census Bureau is conducting this study for two purposes:

(a) To obtain information about the use of the measure of sampling error we present with ACS estimates
(b) To gain insight into ACS data users' preference, if any, for a particular measure of sampling error

The results of this survey will aid the Census Bureau in determining the measure(s) of sampling error to be displayed in future ACS data products.

We estimate that the survey should require 10-15 minutes of your time. Your answers will only be used to help us make informed decisions about possible revisions to the way we display sampling error. Your responses are voluntary, and we ensure your confidentiality under the provisions of Title 13 USC Section 9*.

You may access the survey with this link:

https://questionweb.com/39475

We thank you for your time and look forward to receiving your feedback.

Sincerely,

U.S. Census Bureau

*We will report findings in summary form only. Findings will be posted when available on the Census Bureau's Web site at: http://www.census.gov/mso/www/npr/mktbrfs.html. The collection of information involved in this survey has been reviewed and previously approved by the Office of Management and Budget (OMB) under OMB control number 0607-0760.

Second Round of Sampling Revised Email with Warning Message and New Link Added 09/08/08:

IMPORTANT: Please do not forward the survey link in this email message to others.

Dear American Community Survey User:

Recently, you received an email stating that you have been selected to participate in an important research study regarding the data tables in the American Community Survey (ACS).

When you respond please access the survey with the link

https://questionweb.com/98995

rather than the one specified in the previous email.

In order to maintain the integrity of the sample for this survey, please do not forward this link to others; it is intended for your use only.

We estimate that the survey should require 10-15 minutes of your time. Your answers will only be used to help us make informed decisions about possible revisions to the way we display sampling error. Your responses are voluntary, and we ensure your confidentiality under the provisions of Title 13 USC Section 9*.

We thank you for your time and look forward to receiving your feedback.

Sincerely,

U.S. Census Bureau

*We will report findings in summary form only. Findings will be posted when available on the Census Bureau's Web site at: http://www.census.gov/mso/www/npr/mktbrfs.html. The collection of information involved in this survey has been reviewed and previously approved by the Office of Management and Budget (OMB) under OMB control number 0607-0760.

Second Round of Sampling Reminder Email with Warning Message and New Link 09/23/08:

IMPORTANT: Please do not forward the survey link in this email message to others.

Dear American Community Survey User:

This is a reminder that you have been selected to participate in an important research study regarding the data tables in the American Community Survey (ACS). The U.S. Census Bureau is conducting this study for two purposes:

(a) To obtain information about the use of the measure of sampling error we present with ACS estimates
(b) To gain insight into ACS data users' preference, if any, for a particular measure of sampling error

The results of this survey will aid the Census Bureau in determining the measure(s) of sampling error to be displayed in future ACS data products.

In order to maintain the integrity of the sample for this survey, please do not forward this link to others; it is intended for your use only.

We estimate that the survey should require 10-15 minutes of your time. Your answers will only be used to help us make informed decisions about possible revisions to the way we display sampling error. Your responses are voluntary, and we ensure your confidentiality under the provisions of Title 13 USC Section 9*.

You may access the survey with this link:

https://questionweb.com/98995

The survey will be available until September 30, 2008.

We thank you for your time and look forward to receiving your feedback.

Sincerely,

U.S. Census Bureau

*We will report findings in summary form only. Findings will be posted when available on the Census Bureau's Web site at: http://www.census.gov/mso/www/npr/mktbrfs.html. The collection of information involved in this survey has been reviewed and previously approved by the Office of Management and Budget (OMB) under OMB control number 0607-0760.