

**American Community Survey
Fiscal Year 2014 Content Review
Median County-Level
Allocation Rates of ACS Data**

FINAL REPORT

Sarah Heibel, Decennial Statistical Studies Division

(This page intentionally blank)

Table of Contents

1	Introduction.....	1
2	Methodology.....	2
2.1	Data Collection Mechanism/Tool	2
2.2	Universe	2
2.3	Research Questions of Interest.....	3
2.4	Analysis Design.....	3
3	Assumptions and Limitations	4
3.1	Assumptions.....	4
3.2	Limitations	4
4	Results.....	6
5	Summary.....	8

1 Introduction

The American Community Survey Office (ACSO) is conducting a review of the content on the American Community Survey (ACS). The ACS Content Review examines multiple facets of the ACS, reviewing key survey measures and conducting meetings with the survey's stakeholders and the general public.

This component of the ACS Content Review examines how often data had to be allocated, an important measure of data quality related to the frequency with which respondents leave a question blank while completing the survey. Allocation rates assess how reliable an estimate is; it is one of the Census Bureau's standard measures of measuring statistical quality.¹

The Census Bureau uses imputation methods to correct for item nonresponse or inconsistent/contrary responses. The methods either use rules to determine acceptable answers or use answers from similar housing units or people who provided the item information. "Assignment" involves logical imputation where a response to one question implies the value for a missing response to another question. For example, first name can often be used to assign a value to sex, if sex was left blank. Assignment also edits answers that are inconsistent, such as a six-year-old being in 10th grade. Allocation, on the other hand, uses statistical procedures, such as within-household or nearest neighbor matrices populated by donors, to impute for missing values.

The allocation rate per ACS question was calculated for each of the 3,143 counties (or county equivalents) in the United States (excluding Puerto Rico). The median county-level rate is reported in this paper.

The median allocation rate per ACS question across all counties is one of nineteen metrics being used in the ACS Content Review. These results will be combined with the eighteen other metrics to inform a decision about the future of questions on the ACS. The nineteen metrics were designated as either being a 'cost' (i.e., the question has a high number of complaints from respondents to Census, or the question takes a relatively long time for the respondent to answer) or a 'benefit' (i.e., the question is used to provide estimates that are statutorily required or is used for establishing the frame of another federal survey). The allocation rates were considered in the 'cost' category of the ACS Content Review; estimates with high county-level median allocation rates are more costly in terms of data quality.

¹ More information on allocation rates can be found at http://www.census.gov/acs/www/methodology/item_allocation_rates_definitions/index.php.

This report directly supports the findings described in the American Community Survey (ACS) Fiscal Year 2014 Content Review Results Final Report. For more information on the ACS Content Review, see the Content Review web page:

http://www.census.gov/acs/www/about_the_survey/acs_content_review/

2 Methodology

2.1 Data Collection Mechanism/Tool

The ACS is an annual survey that provides current demographic, social, economic, and housing information about America's communities. There were 125 distinct questions identified on the ACS questionnaire for inclusion in the Content Review. These included person-level questions such as age, relationship status, level of educational attainment, place of birth, and marital status. Housing-level questions included the type of building where the respondent resides, how many rooms exist in the home, and what type of fuel is used for heating. Subsequent discussion of the questions identifies them by number (as found on the paper questionnaire), with either a P or H prefix to designate whether a person or housing question.

For the ACS Content Review, the county-level allocation rate for each question was determined. From the 3,143 counties, the median value was identified for each question, resulting in 125 allocation rates that were the median value of each question on the ACS. These median rates were grouped into one of three clusters: low, medium or high allocation rates. The ACS questions with the highest value medians were the most costly, while the lowest medians were the least costly. SAS® was used to perform the clustering and found natural breaks in the data, specifically using the PROC FASTCLUS procedure. These clusters are shown in the final results of this paper.

2.2 Universe

ACS questionnaires from the 2008-2012 five-year data collection were included in this research. Specifically, eligible returns were those received from January 1, 2008 until December 31, 2012. A small number of allocation rates had to be determined using alternate years of collection, which will be discussed under Limitations in Section 3.2.

All returns from Puerto Rico were excluded. Questionnaires from Group Quarters were included in the person-level analysis as applicable; for instance, no question on relationship status is asked of persons in Group Quarters.

Completed returns are grouped by their county or county equivalent. For instance, parishes in Louisiana are treated as counties. All references to 'counties' in this documentation also includes these county equivalents.

2.3 Research Questions of Interest

The research questions of interest are:

- 1) What was the median county-level allocation rate for each question on the ACS questionnaire?
- 2) How do the median allocation rates for each question compare to each other?

Allocation rate is used by the Census Bureau as the key measure of item nonresponse error. A value is allocated for a question if it is left blank by the respondent and cannot be assigned using information provided in an answer to another question (such as assigning a child's race based on the parent's race).

In the ACS Content Review analysis, it is considered a 'cost' to have a high allocation rate. Final estimates can be adversely impacted when item nonresponse and allocation rates are high; bias can be introduced if the characteristics of the nonrespondents differ from those reported by respondents.

2.4 Analysis Design

The allocation rate is defined as:

number of persons (or cases) that needed a value allocated for a given question
number of persons (or cases) that were expected to answer the question

Allocation rates are calculated using final ACS weights.

Allocation rates are only relevant for those persons or cases that were to have answered the question, according to the questionnaire skip patterns. For instance, only persons at least fifteen years of age were asked their marital status. Children under fifteen did not need to answer the question on marital status and so are not considered part of the denominator. Universes for each question can be found online at

http://www.census.gov/acs/www/methodology/item_allocation_rates_data/.

There are some situations where the denominator of the allocation rate for a county is zero; that is, no one sampled for ACS in the county fits the criteria to have been asked a given question on the survey. For instance, if there were no persons who were foreign-born in the ACS sample for a county, then no responses would have answered the question of when the person entered the United States. The allocation rate for that county would be irrelevant, and mathematically undefined. There would then be one less county allocation rate to ultimately use in deterring the median allocation rate for that question. Thus, for some ACS questions, there were fewer than 3,143 counties contributing an allocation rate to a given question. The table of results in Section 4.1 contains a column with the number of counties contributing to the median for a given ACS question.

3 Assumptions and Limitations

All assumptions or limitations that apply to the ACS production estimates also apply to this research, as mentioned in the documentation found at

http://www.census.gov/acs/www/data_documentation/documentation_main/

In addition, the following assumptions and limitations that are unique to this research are given below.

3.1 Assumptions

The definition of what constituted a question was determined in the larger ACS Content Review process.

3.2 Limitations

For some ACS questions, it was not possible to follow the prescribed methodology without some modification. The following modifications were applied:

- Question H3: “When did Person 1 move into this house?” – collects both month and year
 - There is one allocation rate available for the month of move in, and a separate allocation rate available for the year of move in. In order to understand how often both had to be allocated, additional programming would have been required and seemed unnecessary since the month is not as essential information as the year. The allocation rate for year alone was also larger than the allocation rate for month alone, so the allocation rate for year is used to represent H3.
- H9, H10, and H11: Computer and internet access
 - These questions were added to the ACS questionnaire in 2013, so only have one year of data collected. Allocation rates were calculated using the 1-year data instead of the 5-year data.
- H23a: “Do you have a second mortgage or home equity loan on this property?”
 - Typically, two allocation rates are computed from this one question; an allocation rate for second mortgage and an allocation rate for home equity loan. For the purposes of the ACS Content Review however, there could only be one allocation rate associated with the question. Additional programming was performed to obtain one comprehensive allocation rate.
- P4: Age and Date of Birth
 - An allocated answer to question P4 is necessary when a respondent does not answer either the age or the date of birth question. Date of birth is not allocated however; so the allocation rate for age was used without regard to responses for date of birth.
- P12: Field of Degree

- This question was added to the ACS questionnaire in 2009. Since the 2009-2013 ACS five-year data have not been released yet, an allocation rate was calculated from the most recent three-year ACS data (2010-2012).
- P13: Ancestry
 - Responses for ancestry are not allocated. An item nonresponse rate for ancestry was calculated instead of an allocation rate; this included uncodeable responses as well as missing responses.
- P15b: Migration
 - All address components for the person's residence a year ago are collected and allocated (such as city, county, and state), but only one allocation rate can be used for this question in the ACS Content Review project. The rate at which state of last residence had to be allocated was chosen by experts as the unique allocation rate to represent this question.
- P30: Place of work, county
 - All address components for the person's place of work are collected and allocated (such as city, county, and state), but only one allocation rate can be used for this question in the ACS Content Review project. The rate at which county of employment had to be allocated was chosen by experts in SEHSD as the unique allocation rate to represent this question.
- Employment status recode
 - A number of questions are used jointly by analysts to determine a person's employment status. As a result, the individual questions do not have their own allocation rates, but all share the allocation rate for the employment status recode. These questions are P29, 35, 36, 37, and P39.
- Industry recode
 - Three questions are used jointly by analysts to determine a person's industry of employment. As a result, the individual questions do not have their own allocation rates, but all share the allocation rate for the industry recode. These questions are P42, P43, and P44.
- Occupation recode
 - Two questions are used jointly by analysts to determine a person's occupation. As a result, the individual questions do not have their own allocation rates, but all share the allocation rate for the occupation recode. These questions are P45 and P46.

4 Results

In the following section, we answer the research questions that were posed in Section 2.3.

Table 1 shows the allocation rates per question, sorted with the highest allocation rate cluster on top.

Table 1. Median County-Level Allocation Rates per ACS Question, sorted by median value

ACS Question Number	Question Topic	Number of counties included	Median county rate	Cluster
H11e	Internet subscription: Broadband plan	3143	25.60	High
H21	Home insurance	3142	24.55	High
P48	Total person income	3143	21.31	High
H24	Mobile home costs	3133	19.51	High
H20	Real estate taxes	3142	16.45	Middle
H23B	Second mortgage amount	3104	15.81	Middle
H2	Year residence was built	3143	14.94	Middle
P13	Ancestry	3143	14.41	Middle
P47A	Amount of wages earned	3143	14.10	Middle
H19	Home value	3142	13.66	Middle
H14B	Cost of gas for home	3143	11.42	Middle
P25C	Length of time grandparent caring for child	3079	11.15	Middle
H14D	Cost of fuel for home	3143	10.88	Middle
H22D	Mortgage includes insurance	3141	10.70	Middle
P33	Time of departure for work	3143	10.54	Middle
H22C	Mortgage includes taxes	3141	9.90	Middle
P23	Year last married	3143	9.58	Middle
P14A	Foreign Language at home, yes/no	3138	9.46	Middle
P25B	Grandparent responsible for child	3109	9.26	Middle
H22B	Monthly mortgage payment	3141	9.23	Middle
H18A	Monthly rent	3143	8.79	Middle
P16G	Indian Health Service	3143	8.74	Middle
P12	Undergraduate Field of Degree	3142	8.52	Middle
P9	Year of entry to the U.S.	3138	8.42	Middle
P47D	Amount received from Social Security	3143	8.40	Middle
P16E	TRICARE or military insurance	3143	8.35	Middle
P16F	Health insurance through VA	3143	8.22	Middle
P47C	Income from Interest	3143	8.10	Middle
P16D	Medicaid	3143	8.02	Middle
P34	Minutes from work to home	3143	7.89	Middle
P15B	State of residence one year ago	3142	7.43	Middle
H14C	Cost of water for home	3143	7.12	Middle
P45&46	Occupation	3143	6.93	Low
P16B	Health insurance from insurance company	3143	6.72	Low
P40	Hours worked per week	3143	6.62	Low
P42-44	Industry of employment	3143	6.61	Low
P47G	Income from retirement	3143	6.59	Low

ACS Question Number	Question Topic	Number of counties included	Median county rate	Cluster
P41	Class of worker	3143	6.56	Low
P30	County of place of work	3143	6.28	Low
P47H	Income from other sources	3143	6.23	Low
P16A	Health insurance through work	3143	6.10	Low
H14A	Cost of electricity	3143	5.97	Low
P32	How many people in carpool	3143	5.91	Low
P27	Period of military service	3143	5.83	Low
P47F	Income from public assistance	3143	5.72	Low
P39B	Weeks worked per year	3143	5.72	Low
P47E	Income from Supplemental Security Income	3143	5.68	Low
P7	Place of birth	3143	5.61	Low
P47B	Income from Self-employment	3143	5.51	Low
H11b	Internet subscription: DSL	3143	5.04	Low
H11c	Internet subscription: Modem	3143	5.04	Low
H11d	Internet subscription: Fiber-optic internet service	3143	5.04	Low
H11f	Internet subscription: Satellite	3143	5.04	Low
H11g	Internet subscription: Other internet service	3143	5.04	Low
H11a	Internet subscription: Dial-up	3143	5.02	Low
H7A	Number of rooms in house	3143	5.00	Low
P38	When last worked	3143	4.99	Low
P10B	Current grade in school	3142	4.82	Low
P31	Mode of transport to work	3143	4.81	Low
P16C	Medicare	3143	4.68	Low
H7B	Number of bedrooms in house	3143	4.54	Low
P29, 35-37, 39a	Employment status recode	3143	4.53	Low
P11	Educational attainment	3143	4.44	Low
H5	Sales from agriculture	3143	4.34	Low
P22	Number of times married	3143	4.32	Low
P14C	English fluency	3138	4.28	Low
P21B	Widowed in last year	3143	4.18	Low
P21C	Divorced in last year	3143	4.14	Low
H10	Internet access	3143	3.96	Low
P21A	Married in last year	3143	3.92	Low
H4	Acreage	3143	3.76	Low
P15A	Moved in last year	3143	3.50	Low
H9c	Own/use other type of computer	3143	3.43	Low
P28A	Presence of VA disability rating	3143	3.36	Low
H23a	Second mortgage	3141	3.19	Low
P26	Military service	3143	3.13	Low
P18B	Difficulty walking	3143	3.06	Low
P18C	Difficulty dressing	3143	3.04	Low
P19	Difficulty doing errands	3143	3.03	Low
P18A	Difficulty remembering	3143	3.03	Low
P17B	Difficulty seeing	3143	2.96	Low

ACS Question Number	Question Topic	Number of counties included	Median county rate	Cluster
H13	Type of heating fuel	3143	2.96	Low
P10A	Attended school recently	3143	2.94	Low
H9b	Own/use a handheld computer	3143	2.91	Low
P14B	Foreign language at home, which language	3143	2.86	Low
H8F	Have refrigerator	3143	2.81	Low
H9a	Own/use a laptop or desktop	3143	2.72	Low
P24	Given birth in last year	3143	2.68	Low
P17A	Difficulty hearing	3143	2.65	Low
H8E	Have stove	3143	2.65	Low
H3	Year moved in	3143	2.39	Low
H6	Business on property	3143	2.30	Low
P20	Marital status	3143	2.19	Low
P8	Citizen	3143	1.91	Low
H8C	Have bathtub	3143	1.87	Low
H8B	Have toilet	3143	1.81	Low
H8A	Have running water	3143	1.81	Low
H8D	Have sink	3143	1.79	Low
P5	Hispanic origin	3143	1.77	Low
H18B	Rent plus meals	3143	1.63	Low
H22A	Have mortgage	3142	1.52	Low
P6	Race	3143	1.02	Low
H12	Number of vehicles	3143	0.94	Low
H15	Food stamps	3143	0.90	Low
P2	Relationship	3143	0.88	Low
H17	Tenure	3143	0.87	Low
H8G	Have telephone	3143	0.86	Low
P4	Age	3143	0.86	Low
H1	Type of structure	3143	0.68	Low
P25A	Presence of grandchildren	3143	0.59	Low
H3	Month moved in	3143	0.51	Low
H16	Condo status and fee	3142	0.16	Low
P3	Sex	3143	0.10	Low
P28B	VA disability rating	3132	0.00	Low

5 Summary

Median county-level allocation rates ranged from 25.60% to 0.00%.

Four questions were determined to be in the highest allocation rate cluster; broadband internet subscription, home insurance, total person income, and mobile home costs. Another 28 questions were identified in the middle allocation rate cluster, ranging from 16.45% to 7.12% allocated. The remaining 93 questions were considered by the software to have low allocation rates.