



**AMERICAN COMMUNITY SURVEY DECISION MEMORANDUM SERIES**

No. ACS-D-201403

Date: August 18, 2014

**MEMORANDUM FOR:** The Record

From: James B. Treat  
Chief, American Community Survey Office

Subject: Decisions Supporting the American Community Survey (ACS)  
Content Review Analytics

**Purpose**

The purpose of this memorandum is to document design decisions for Phase I of the ACS Content Review. Because of its complexity, the review will be conducted in phases. Phase I provides the results based solely on the decision criteria and weighting approach discussed below.

**Overview**

In August 2012, the Interagency Council of Statistical Policy (ICSP) Sub-Committee for the American Community Survey (ACS) was formed to advise the Office of Management and Budget's Chief Statistician and the Census Bureau's Director from the perspective of the federal statistical system. The Sub-Committee is co-chaired by the Director of the Census Bureau and by the Chief Statistician of OMB. The Sub-Committee's mission includes overseeing and managing the content and respondent burden of the ACS. The Sub-Committee charter (*Charter of the Interagency Council on Statistical Policy: Subcommittee on the American Community Survey*, August 10, 2012) states "Each year there will be an annual review of questions to consider any deletion or addition of questions." In 2013, the Census Bureau initiated the first comprehensive examination of every question on the ACS form, resulting in the initial ACS Content Review project.

**Issue**

The ACS Content Review is a complex initiative to determine the cost and benefits of the questions contained in the ACS questionnaire. The review includes 19 decision criteria (13

benefits and 6 costs), which require 9 streams of information (See below and see the *ACS Content Review Decision Criteria*, August 2013). The information streams do not consistently capture the required information at the same level. For example, federal agencies use cost of utilities at the topic level while the interviewer survey captures information only at the specific sub-question level. Additionally, some questions are too new to have 5- year estimates; other questions' responses must be considered in tandem with additional questions. Industry coding, for example, requires combining several question to determine the appropriate code. Additionally, federal agency data use responses may not be complete and/or we may not receive validation from the Office of General Counsel in time for the analysis phase. Within these potentially incomplete or non-validated responses and inconsistent information among data streams, there is necessarily missing and/or non-validated information. To ensure transparency and objectivity, a priori analytic decisions must be made before data are run through the final decision-making algorithm. These decisions ensure that we meet project deadlines, but may introduce uncertainty into the data.

### **Analytic Component Decisions**

We are conducting the ACS Content Review in phases. Phase I is primarily data driven. It is based on the nine data streams that directly feed into the Decision Criteria Algorithm. The output will be the questions that are potential candidates for removal from the ACS questionnaire, along with those that will stay and ones needing further research. Phase II will consider supplemental information about all the questions, such as whether there are administrative records or other data sources that could be used in lieu of asking respondents. Phase III will include any research results or other changes that affect the content.

The following table includes the decisions needed to ensure a transparent, statistically defensible, and comprehensive assessment of ACS content from a benefits and cost perspective. The purpose of this memorandum is to provide a single document cataloging analytic decisions needed prior to testing, preparing, and running the data through the Phase I ACS Content Review Decision Criteria Algorithm. The Director of the Census Bureau confirmed that we should use Option 2 as the weighting algorithm.<sup>1</sup>

Note: To review the specific questions referred to in the table by their number (e.g., P30), see the ACS 2014 questionnaire.

| <b>Criteria</b>       | <b>Decision</b>                                                                                                                                                                                                                                                                                                                                                     |
|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>General Design</b> |                                                                                                                                                                                                                                                                                                                                                                     |
| Level of Analysis     | Not all data streams collect information at the sub-question level so the detail is uneven. We will report findings at the lowest level (i.e. question subpart). If a tool/analysis can only provide information at the higher level (i.e. question), we will assume that the information applies equally to the subparts, and distribute to each question subpart. |

<sup>1</sup> ACS Memorandum for the Record Selection of ACS Content Review Weighting Algorithm, August 18, 2014.

| <b>Criteria</b>                                                       | <b>Decision</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-----------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Related Questions                                                     | There are questions that are used in conjunction with each other to create a single topic or concept. For example, Industry is determined through the use of three separate questions (P42-P44). We will identify related questions (flag/link in analysis), but evaluate them separately at the question level. Once we evaluate each of the related questions, we will examine them as a whole to see if any issues emerge. Potentially affected topics include: P29 a & b; P35,P36, & P37; P39a & b; P42, P43, & P44.                                                                                                                                   |
| Related Questions with only one Allocation Rate, CV, and IQR Estimate | When there are multiple elements/items collected for a question, we identified a single element/item for analysis. For example, based on input from subject matter experts, all address components for the person's residence one year ago (P15b) are collected (such as city, county, and state), but only one allocation rate, CV, and IQR estimate could be used for this question in the ACS Content Review project. <i>State</i> of last residence was chosen as the unique estimate to represent this question.                                                                                                                                      |
| Results Schedule                                                      | We will run the Phase I ACS Content Review Decision Criteria Algorithm on August 19 and present results between August 22 and mid-September 2014. Formal documentation will be completed by October 31, 2014.                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Results Approach                                                      | The 9 input data sets will be normalized into a common format. They will be run through the Decision Criteria Algorithm using Option 2 weights. The results will be displayed in a scatterplot among four quadrants. The quadrants are: (1) High Benefit/Low Cost; (2) High Benefit/High Cost; (3) Low Benefit/Low Cost; and (4) Low Benefit/High Cost. We will use a scale from 1 to 100 with the quadrants evenly divided by the 50 X axis and 50 Y axis. We decided to use this scale because we will not have all the validated uses when we run the Phase I algorithm, so we do not have the information needed to better divide the quadrant values. |
| Candidate Question Results                                            | Upon running the algorithm, we will consider the Low Benefit /Low Cost and Low Benefit/High Cost quadrant questions as potential candidates for Phase I removal.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Federal Register Notice                                               | The October 2014 Federal Register Notice permits comments on all questions; however, the Census Bureau is particularly interested in the questions that have been identified as potential candidates for removal based on the Phase I Content Review analysis. Prior to this formal publication, we will alert federal agencies to the notice as well as provide a general notice on our external ACS web page.                                                                                                                                                                                                                                            |
| Short Form Questions                                                  | We will analyze the short form questions (age, sex, relationship, race, Hispanic origin, tenure), but they will not be considered for removal.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Metrics for P16h (Other Health Insurance)                             | We dropped P16h "Any Other Type of Health Insurance or Health Coverage plan" from the Content Review evaluation. Once received it is promptly recoded to one of the other sources listed in 16a through 16g. There is no estimate and no allocation rate since it disappears in the edits.                                                                                                                                                                                                                                                                                                                                                                 |
| Use of Averaged Weights for Design Criteria                           | Each member of the ICSP Sub-Committee provided a recommendation for how to best allocate the 200 total points to the 19 criteria. The Sub-Committee provided each member's recommendation for the 200 total points assigned to each design criteria. We used the average of the 5 members' points as the criteria weight and divided by half to be consistent with our quadrant analysis approach.                                                                                                                                                                                                                                                         |
| <b>Estimates</b>                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Questions Lacking 5-year Estimates                                    | Internet usage, Internet Subscription, Computer Usage, and Field of Degree are too new to have 5-year estimates. For Internet usage, we will use the 1-year estimate based on the likely stability of the estimates and standard errors. For Field of Degree,                                                                                                                                                                                                                                                                                                                                                                                              |

| <b>Criteria</b>                                                        | <b>Decision</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                        | we will use the 3-year estimate, again, due to the stability of the estimates and standard errors. For the estimates based on 1-year and 3-year data, the standard error has been adjusted (either by dividing by $\sqrt{5}$ or $\sqrt{5/3}$ , respectively).                                                                                                                                                                                                                                                                                    |
| <b>Federal Agency Data Uses (#1-9 Benefits and Data Stream #1)</b>     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| DOC OGC Completed Validation Definition                                | The Department of Commerce (DOC) Office of General Counsel (OGC) use validation will be considered complete when the OGC has reviewed and adjudicated all agency uses with minimal information as seen below.                                                                                                                                                                                                                                                                                                                                    |
| Minimal Federal Agency Use Data required for Legal Analysis            | The minimum information that must be populated in each federal agency data usage collection tool to complete a legal analysis includes: (1) Central Uses (Column B); (2) Complete statutory/Legal Citation (Column C); and, (3) Identification of questions connected to the use and citation (Columns T to end). (See Federal Agency Data Collection Tool.)                                                                                                                                                                                     |
| DOC OGC Legal Review of non-validated agency OGC federal use responses | If we have the data use information but not the affirmation of validation from the agency OGC, the DOC OGC may conduct a validation of Mandatory, i.e., required by statute for ACS or the decennial census to collect the data.                                                                                                                                                                                                                                                                                                                 |
| Designations of Mandatory, Required, or Programmatic Uses              | Until validated by the DOC OGC, all federal data uses will be considered Programmatic. If a statute does not state explicitly to use the ACS or the decennial census, then the use is no higher than Required. The use would never be Mandatory.                                                                                                                                                                                                                                                                                                 |
| Unusual but Valid Census Geography                                     | In some cases, federal agencies included geographies not in our data collection form, e.g., Congressional Districts. We have input from a MAF expert and will code as follows: <ul style="list-style-type: none"> <li>• School district: (decision - Tract)</li> <li>• PUMA: (decision - MSA)</li> <li>• Congressional District: (decision - MSA)</li> <li>• American Indian Area/Alaska Native Area/Hawaiian Home Land: (decision – Place)</li> <li>• County remainder: (decision - County)</li> <li>• Town/City: (decision – Place)</li> </ul> |
| Federal Data Collection Completion Definition                          | For purposes of completing the Phase I ACS Content Review, data collection is considered complete when we have received the relevant agencies' OGC letter and have validated at least one Mandatory or Required use for each question.                                                                                                                                                                                                                                                                                                           |
| OGC Validation Approach                                                | To prioritize validation, we conducted a test run and analyzed the results. Based on the results, we identified the questions, the uses, and the statutes for immediate validation that will support question retention. All uses will be validated by the OGC after the priorities needed for our results reporting.                                                                                                                                                                                                                            |
| Criteria for Question Removal                                          | Questions will be candidates for removal should they fall into the Low Benefit/Low Cost or Low Benefit/High Cost quadrants and: (1) have no Mandatory Uses and (2) have no Required Uses at the Block Group/Tract level and (3) have no Required Uses at the Place/County/MSA level.                                                                                                                                                                                                                                                             |
| <b>Other Data Uses (#10 Benefit and Data Stream #2)</b>                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Existence of Other Data Sources                                        | When blank, we are assuming that there are no other data sources.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <b>CVs (1 Benefit (#11 Benefit and Data Stream #3)</b>                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Identifying the Estimates and Selecting the CV                         | Subject matter analysts in POP and SEHSD agreed upon a set of estimates that represents each question even if such an estimate is not currently available on AFF,                                                                                                                                                                                                                                                                                                                                                                                |

| <b>Criteria</b>                                                        | <b>Decision</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clusters                                                               | such as percent of households without toilets. The estimates for each ACS question are to be used for the median county-level coefficient of variation (CV). We will determine how to best cluster the CVs into three clusters by using SAS PROC FASTCLUS                                                                                                                                                                                                                                                                                                                                                                                             |
| <b>IQRs (#12 Benefit and Data Stream #4)</b>                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Calculating Interquartile Ranges (IQRs)                                | The estimates for each ACS question (which were used in the CV analysis) are also to be used for the county-level interquartile range (IQR) decision criteria for the content review. The estimates are based on different measures—percentages, ratios, and medians. Therefore, it is not possible to get a comparable median IQR measure for every estimate. About 90 percent of the estimates are percentages. We will compute the IQRs for the estimates that are percentages and cluster those results. Because this benefit metric has such a small design effect, we will assign the remaining estimates a neutral score of 3 for this metric. |
| Calculating IQR Clusters                                               | To produce the IQR clusters for the percentages (see above), we used SAS PROC FASTCLUS. However, because there was an extreme outlier in the ranges, we removed it before running the SAS PROC FASTCLUS procedure and then hardcoded it into the results.                                                                                                                                                                                                                                                                                                                                                                                             |
| <b>Frame for Federal Survey (#13 Benefit and Data Stream #5)</b>       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Identifying Cluster Approach                                           | Each ACS question was designated as: <ul style="list-style-type: none"> <li>• Primary (P) = 5</li> <li>• Secondary (S) = 3</li> <li>• Not Required (N) = 0</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Applying Primary or Secondary to Sub-Parts                             | If a sub-part on its own was deemed not required, but it was a primary or secondary source when paired with other estimates, then that trumped the individual status. For instance, the occupation recode (which combines responses from P45 and P46) is a primary source for the frame for another federal survey, but P45 or P46 on its own is not a source. As a result, P45 and P46 were considered primary sources and given the maximum benefit score.                                                                                                                                                                                          |
| <b>Interviewer Survey (# 1-3 Costs and Data Stream #6)</b>             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| ACS Content Review Interviewer Survey Cluster Analysis Criteria        | The Interviewer Survey addressed three separate costs: Cognitive Burden, Survey Sensitivity, and Overall Difficulty. To group the results of the Interviewer Survey into High/5, Medium/3, and Low/1, we will use more than 60 percent, 20-60%, and less than 20% respectively. We identified these criteria in August 2013 and will use these to cluster and then assess the interviewer responses.                                                                                                                                                                                                                                                  |
| <b>Adjusted Median Seconds to Answer (#4 Costs and Data Stream #7)</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Paper v Automated Instrument Data Collection Levels                    | Some questions on the automated questionnaire are at a lower level than the same questions on the paper questionnaire. To normalize, we will roll-up the questions on the automated questionnaire to the lowest level on paper questionnaire.                                                                                                                                                                                                                                                                                                                                                                                                         |
| Calculating Adjusted Median Seconds to Answer Clusters                 | To produce the Median Seconds to Answer clusters we used SAS PROC FASTCLUS.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Limitations on Returns Included in Sample                              | The sample of returns used in this research is not a random sample. Returns received by mail, Telephone Questionnaire Assistance (TQA), Failed Edit Follow-up (FEFU), and Group Quarter interviewing were not in scope for this project. This research provides estimates of response time based on three months of survey responses from Internet, CATI, and CAPI. Thus, results may not reflect the experience of the entire                                                                                                                                                                                                                        |

| Criteria                                              | Decision                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                       | ACS sample.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Computing Address Median Response                     | All parts of the address are necessary to compute location accurately. Therefore, all sub-parts (a-f) for P30 must be used to calculate the median response time. This is consistent with the calculation of median response time for P15b.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>Allocation Rates (#5 Costs and Data Stream #8)</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Allocation Rates – Special Cases                      | <p>There are three special cases we addressed with regard to allocation rates:</p> <ol style="list-style-type: none"> <li>1. Missing allocation rates – Questions P13 (Ancestry) and P16h (Other health insurance) are not allocated. An item nonresponse rate for ancestry was calculated using uncodeable and missing responses.</li> <li>2. Multiple allocation rates available for a single question – In certain cases, such as question P15b (Address 1 Year Ago), an allocation rate is available for each component of the address, which is considered a single question. In these instances, the subject matter experts provided guidance on which allocation rate to use to represent the question. In this case, the allocation rate for the ‘State’ field was selected.</li> <li>3. Single allocation rate that represents multiple questions – There are instances where one allocation rate represents multiple questions because the questions are not allocated individually and are used in conjunction with one another to define a concept or topic (e.g., questions P42, 43, and 44 which comprise Industry).</li> </ol> |
| <b>Complaints (#6 Costs and Data Stream #9)</b>       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Complaint Data Stream Timeframe and Scope             | <p>We will use January-May 2014 to time box complaint data collection to coincide with the implementation of performance measures associated with complaint information. This helps to ensure standardized responses. The scope includes:</p> <ul style="list-style-type: none"> <li>• Emails: Direct, GovDelivery, Non-Controlled</li> <li>• Letters: CQAS congressional, CQAS non-congressional</li> <li>• Phone Calls: Call Center, Direct, Other Sources</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Lack of Common Definition of “Complaint”              | <p>Because of high volume, several people analyzed the complaint data. We did not provide a clear operating definition of complaint so the results from the three sources may not always be consistent. We believe that the number of reported complaints may be too high; however, due to the high volume and short timeframe, we have decided to accept the results as is.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Computing the Complaint Clusters                      | <p>Complaint clusters were pre-specified in the design criteria as follows:</p> <ul style="list-style-type: none"> <li>• No complaints: scale = 0, points added to cost score = 0.</li> <li>• Some complaints: scale = 3, points added to cost score = 7.8</li> <li>• Top 3 number of complaints: scale = 5, points added to cost score = 13</li> </ul> <p>Note: We ended up with a tie for the Top 3 so we included four questions in this top tier.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

The major Phase I milestones included in the project are available for reference on the ACS Content Review internet site: [http://www.census.gov/acs/www/about\\_the\\_survey/acs\\_content\\_review/](http://www.census.gov/acs/www/about_the_survey/acs_content_review/).

cc: Frank Vitrano, AD2020  
Stephen Buckner, ADCOM  
Jeannie Shiffer, ADCOM  
Enrique Lamas, ADDP  
Eloise Parker, ADDP  
Burton Reist, 20RPO  
Patricia McGuire, 20RPO  
ACS Division Chiefs' Forum  
Barbara M LoPresti, TMO  
Brian Monaghan, FLD  
David Hackbarth, NPC  
Victoria Velkoff, SEHSD  
Fern Bradshaw, FLD  
Jeffrey D Sisson, DSCMO  
Marian E Brady, DSCMO  
Michael R Ratcliffe, GEO  
Patrick J Cantwell, DSSD  
Robert A Kominski, SEHSD  
Ruth Ann Killion, DSMD  
Steven Werner Tornell, TMO  
Timothy F Trainor, GEO  
Timothy P Olson, DIR  
Tommy Wright, CSR  
Victoria A Velkoff, Acting POP  
W Neil Tillman, PIO  
ACS Senior Staff  
Agnes S Kee, ACSO  
Anthony G Tersine Jr, DSSD  
Arumugam Sutha, ACSO  
Cheryl V Chambers, ACSO  
Dameka M Reese, ACSO  
David A Raglin, ACSO  
Deborah L Ambill, ACSO  
Donna M Daily, ACSO  
Gail M Denby, ACSO  
Gary B Chappell, ACSO  
Grace L Clemons, ACSO  
Jennifer Guarino Tancreto, DSSD  
Judy G Belton, ACSO  
Kai T Wu, ACSO  
Karen E King, DSSD  
Kenneth B Dawson, ACSO  
Mark E Asiala, DSSD  
Matthew A Zimolzak, ACSO  
Nicholas M Spanos, ACSO  
Steven P Hefter, DSSD  
Tasha R Boone, ACSO  
Todd R Hughes, ACSO