



**UNITED STATES DEPARTMENT OF COMMERCE**  
**Economics and Statistics Administration**  
**U.S. Census Bureau**  
Washington, DC 20233-0001

**ACS INFORMATION SERIES MEMORANDUM No.: 2015-06**

Date: July 24, 2015

To: The Record

From: Deborah Stempowski (**signed 07/24/2015**)  
Chief, American Community Survey Office

Subject: Release of the Plan, "Evaluating the Availability and Suitability of External Data Sources For Use in the American Community Survey"

This memo announces the release of the plan, "Evaluating the Availability and Suitability of External Data Sources For Use in the American Community Survey." This plan provides information as to how the Census Bureau and the American Community Survey program is conducting research into the use of administrative records to minimize burden for ACS respondents while still allowing the survey to be responsive to emergent issues, keeping content current, and maintaining high quality data.

This report is public-facing, available on the new 2015-16 Survey Enhancements webpage at: <http://www.census.gov/programs-surveys/acs/operations-and-administration/2015-16-survey-enhancements.html>.

Please refer any questions concerning the plan to David Raglin, ACSO, at 301-763-4226 or [david.a.raglin@census.gov](mailto:david.a.raglin@census.gov).

# **EVALUATING THE AVAILABILITY AND SUITABILITY OF EXTERNAL DATA SOURCES FOR USE IN THE AMERICAN COMMUNITY SURVEY**



## Table of Contents

Background .....	1
First Step: Identifying and Evaluating Administrative Records Data for Use in the ACS .....	2
Second Step: Using Administrative Records and Third-Party Data in the ACS .....	4
Project Schedule .....	5
Conclusion.....	5

## Background

The Census Bureau is working to minimize burden for survey respondents while enabling the survey to be responsive to emergent issues, keeping content current, and maintaining the high quality of data that our country demands and deserves. Our ability to make the American Community Survey (ACS) agile in the face of constantly changing times ensures that we are continuing to deliver deep public value to our nation’s communities. Research on the use of administrative records, data collected by other Federal, state and local agencies as well as third party sources, is a key component of that commitment.

A National Academies of Sciences panel commissioned by the Census Bureau recently issued a report entitled “Realizing the Potential of the American Community Survey: Challenges, Tradeoffs, and Opportunities,”<sup>1</sup> which recommended the Census Bureau “should coordinate efforts across units on research related to the potential use of administrative records”. To support this effort early on, the Census Bureau explored how administrative records could replace or improve ACS content<sup>2</sup>. Building on this initial effort, this document describes the previous research already completed, research already underway, and the tasks necessary to further assess the utility of administrative records for specific questions on the ACS.

This plan describes the need to identify and acquire external data sources, match them to ACS data, and evaluate the coverage and quality of each data source and the resulting match. Using an agile approach and coordinating adaptation across program areas, including the 2020 Census Program, the Census Bureau will work to develop strategies for obtaining the desired records, and resolving policy and legal issues associated with their use. The research will identify matching issues, and the challenges associated with securing external data for all cases. It will integrate external data into ACS, and compare distributions between ACS data and external data sources for each topic. Researchers will also

---

<sup>1</sup> <http://www.nap.edu/catalog/21653/realizing-the-potential-of-the-american-community-survey-challenges-tradeoffs>

<sup>2</sup> [http://www.census.gov/content/dam/Census/library/working-papers/2015/acs/2015\\_Ruggles\\_01.pdf](http://www.census.gov/content/dam/Census/library/working-papers/2015/acs/2015_Ruggles_01.pdf)

document measurement issues, such as definitional differences and reference period alignment, as well as the expected stability of data elements (i.e., whether we should expect changes to external data sources over time), and the ability to get quality administrative data throughout the country.

## First Step: Identifying and Evaluating Administrative Records Data for Use in the ACS

The first step in this plan is to look at the various topics to document the coverage, quality, and availability of external data sources for potential ACS integration. The Census Bureau has a division, the Center for Administrative Records Research and Applications (CARRA), that has extensive experience collecting and researching administrative records . They have done work matching ACS records to other sources. For example, the American Community Survey Match Study,<sup>3</sup> conducted using 2010 ACS data, provided the critical research foundation needed to document findings and develop the next steps in necessary research. CARRA has also worked with outside researchers who have matched ACS records to other sources.

Based on this work, we developed a prioritized list of ACS question topics to study, based on the availability of administrative record data sources and likelihood of successful matching. This research will enable us to evaluate the potential of the data sources, identify challenges, and provide direction for further research. The table below contains the Priority 1 and 2 topics that will be the focus for the research. . Priority 1 topics are those we believe we have the highest likelihood of finding a suitable replacement using another records source, while Priority 2 topics are the next most likely group of topics. Administrative data for these topics include data sources already present at the Census Bureau and several that CARRA is still pursuing. Along with many existing files, the quality and coverage of these new acquisitions will need to be evaluated prior to operational use. Federal files and programs are prioritized in this initial research period.

The table shows the question number on the ACS paper questionnaire and two items from the 2014 ACS Content Review research, the estimated median seconds to complete the question in the ACS electronic instruments<sup>4</sup>, and whether the question was identified as sensitive or cognitively difficult in a survey of ACS telephone and personal visit interviewers<sup>5</sup>.

---

<sup>3</sup> [https://www.census.gov/srd/carra/2010\\_American\\_Community\\_Survey\\_Match\\_Study.pdf](https://www.census.gov/srd/carra/2010_American_Community_Survey_Match_Study.pdf)

<sup>4</sup> [http://www2.census.gov/programs-surveys/acs/operations\\_admin/2014\\_content\\_review/methods\\_results\\_report/Response\\_Time\\_per\\_Item\\_Report.pdf](http://www2.census.gov/programs-surveys/acs/operations_admin/2014_content_review/methods_results_report/Response_Time_per_Item_Report.pdf)

<sup>5</sup> [http://www2.census.gov/programs-surveys/acs/operations\\_admin/2014\\_content\\_review/methods\\_results\\_report/Interviewer\\_Survey\\_Results\\_Report.pdf](http://www2.census.gov/programs-surveys/acs/operations_admin/2014_content_review/methods_results_report/Interviewer_Survey_Results_Report.pdf)

Table 1. Priority 1 and 2 Topics to be Studied for Replacement by Data Sources

Topic	Question Number	Estimated Seconds to Complete	Sensitive or cognitively difficult?
<i>Priority 1:</i>			
Phone Service	H8g	1	
Year Built	H2	11	Difficult
Part of Condominium	H16	4	
Tenure	H17	11	
Property Value	H19	11	Difficult
Real Estate Taxes	H20	9	Difficult
Have mortgaged/mortgage amount	H22a and H22b	11	
Second mortgage/HELOC and payment	H23a and H23b	5	
Sale of Agricultural Products	H5	1	
Social Security	P47d	10	Sensitive
Supplemental Security Income (SSI)	P47e	8	Sensitive
Wages	P47a	41	Sensitive
Self Employment Income	P47b	8	
Interest/dividends	P47c	20	Sensitive
Pensions	P47g	8	Sensitive
Residence 1 year ago and Address	P15a	18	
<i>Priority 2:</i>			
Number of Rooms and Bedrooms	H7a and H7b	13	Difficult
Facilities	H8a, H8b, H8c, H8d, H8e, H8f	6	Sensitive
Fuel type	H13	14	
Acreage	H4	5	
TOTAL		214 Seconds (~3.5 mins)	

Analyses on the housing items use third-party data that the Census Bureau has acquired or is in the process of acquiring. For phone service, we will use the Contact Frame that is constructed from vendor files containing addresses, landline and cellphone data. We will explore items relating to the structure - year built, real estate tax, property value, and acreage - using public property tax records compiled by a

data aggregator. We will compare the property owner to the ACS reference person or spouse to evaluate housing tenure and assess information from public deeds records relative to ACS questions on mortgages and home equity lines of credit. The Census Bureau will conduct research to determine whether they can acquire information about buildings that register as condominiums.

Analyses on the income items use data from the Social Security Administration (SSA) and Internal Revenue Service (IRS). Census is acquiring SSA data on Old Age, Survivors, and Disability Insurance beneficiaries and Supplemental Security Income recipients. The extent of IRS data varies by income type. We have information returns (e.g., 1099-INT) which reflect whether a person received the income type, but we only acquire income amounts for two types (wages and pensions). We are pursuing additional information on these IRS income types.

Census will also assess data from the Department of Agriculture for the sale of agricultural products question and data from the IRS to explore the residence one year ago question.

## Second Step: Using Administrative Records and Third-Party Data in the ACS

Next, we will create teams for each ACS topic identified as a potential candidate for records usage based on the results from the first phase of research. Each team will include statistical researchers, subject matter experts and data processors who together can identify and research issues related to records usage.

The teams will make recommendations on whether each question is a good candidate for removal with the use of external data sources in its place. This recommendation will be based on an assessment of the implications of implementing such a change, considering data quality, reliability, alignment of reference periods, break in series, and the limitations of the data source affecting the suitability for use. The team will document and evaluate various options for integrating the records. For instance, for some topics, records may be better suited in assisting with imputation whereas for other topics the records may be used for direct substitution of a survey question (for all or a subset of the ACS respondent pool).

The teams will consider the following types of issues and potential research questions:

- Quality of the Source – In some cases, the quality of the records may be more accurate than the respondent's recollection (e.g., W2 information for wages). In some cases, we may not be able to decipher whether data from records are superior or inferior to response data. What research is needed to study quality in those cases? Does the source capture all aspects of a survey question? Since the ACS publishes data for geographic areas throughout the country, we need to know the consistency of quality of the administrative data. Implications to estimates – Will the move to records cause a break in series? Do the reference periods between the survey item and records align?
- Timing – Is the data available for the specific data collection period? How quickly can we access data for the survey period?

- Data products – what impact does the use of external sources have on the data products, particularly Public Use Microdata Sample (PUMS) files?

Each team will need to perform an assiduous review for each topic of the comparability between what is reported on the data source and what is collected in the ACS. The team will document the alternatives and issues, and run simulations to study the impact on estimates and standard errors using external data in different capacities for each topic. The teams will also document a general timeline for implementing the various options.

From Fall 2015 through Summer 2016, Census Bureau staff will present at public conferences and meetings (such as the Association of Public Data Users, Population Association of America, American Association of Public Opinion Research, etc.) to seek input on the potential uses of administrative records and third party data that have been identified, and to share the research results to date. If staff identifies strong external data sources to replace or supplement ACS questions, a Federal Register Notice would solicit public input on the proposed uses of administrative records and third party data in the ACS, and the potential removal of associated survey questions from the questionnaire.

## Project Schedule

Initial research on the ACS concepts identified for FY2015-2016 is underway. The milestones documented below have been set to manage this work.

Milestone	Date
Census Bureau delivers this action plan outlining the next steps in pursuing administrative and commercial data sources	July 2015
First Step: Preliminary findings on availability, coverage and quality (flow by topic)	September 2015 – October 2016
Launch Topic Teams	July 2015
Second Step: Recommendations from topic-specific research (flow by topic)	March 2016 to March 2017

## Conclusion

The Census Bureau is committed to fully understanding and applying alternate data sources that help reduce respond burden while preserving data quality. We feel this research will help us understand reach that goal. Research findings and future plans will be implemented and revised to ensure the goals of the ACS are met in a timely and cost effective manner.