THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

Federal Statistical Products Based on the Census Bureau's American Community Survey: A Reference Guide

July 17, 2015

[Draft for Comment]

Andrew Reamer with
Amy Hsieh and Eric Stokan

Contents

Introduction: Leveraging Mr. Madison

The Census Bureau's American Community Survey (ACS) is the current iteration of a federal government effort that began on the floor of the House of the Representatives in January 1790. Representative James Madison, a lead author of the Constitution and future President, asked the House to approve his suggestion that the first decennial census be augmented to collect information on the nation's households so that Congress could shape legislation that addresses their interests:

> Mr. Madison Observed that they had now an opportunity of obtaining the most useful information for those who should hereafter be called upon to legislate for their country if this bill was extended so as to embrace some other objects besides the bare enumeration of the inhabitants; ***it would enable them to adapt the public measures to the particular circumstances of the community.*** (italics added) In order to know the various interests of the United States, it was necessary that the description of the several classes into which the community was divided, should be accurately known; on this knowledge the legislature might proceed to make a proper provision for the agricultural, commercial and manufacturing interests, but without it they could never make their provisions in due proportion.
>
> This kind of information, he observed, all legislatures had wished for; but this kind of information had never been obtained in any country. He wished, therefore, to avail himself of the present opportunity of accomplishing so valuable a purpose. If the plan was pursued in taking every future census, it would give them an opportunity of marking the progress of the society, and distinguishing the growth of every interest.[1]

Mr. Madison's amendment was approved by the House and became part of the Census Act of 1790 signed by President Washington. The type of information Mr. Madison sought was collected in each decennial census from 1790 through 2000 and, since 2005, has been gathered through the continuously fielded ACS.

Through his action, Mr. Madison served as the catalyst for America's having, throughout the last two centuries, the world's most advanced system of national government statistics. Mr. Madison set in motion four types of innovations that bring the nation's statistical efforts to where they are today.

First is the innovation he explicitly sought—the very idea of gathering information on "the particular circumstances of the community" to guide policy, rather than rulers' whims, was largely unknown to human experience. The second innovation is institutional—the creation of a series of government statistical agencies, such as the Census Bureau, that successfully fulfill congressional mandates to gather that information fully, accurately, on time, and on budget. The third innovation is technical—the development of evermore sophisticated techniques for data collection and analysis that result in accurate portrayals of the American community's

---

[1] *Gales and Seaton's History of Debates in Congress*, January 25, 1790, p. 1115.

circumstances. Machine and computer calculations, statistical sampling, statistical estimation, and continuous measurement are all achievements of the Census Bureau that were on the cutting edge of the field.

The fourth innovation is represented in the contents of this document—the utilization of one statistical data series, the ACS, to serve as the foundation for multiple other federal statistical efforts. The ACS, like the long form of the decennial census before it, is unique in the size of the sample and the breadth of the questions, courtesy of Mr. Madison. As a result, the ACS becomes a springboard, a form of leverage, which enables numerous other federal efforts to gather information on the community's circumstances to guide public policy and, in a way that Mr. Madison would appreciate but did not envision, improve the workings of private markets.

The purpose of this document, then, is to identify and describe the federal statistical products—we found 85—that rely on the ACS for the following types of uses:

- Estimation
  - Input to estimation formula (arithmetic)
  - Input to modeled estimates (regression)
  - Weights or controls
  - Adjustment factors
  - Validation
- Sampling
  - Sampling frame
  - Sampling stratification
  - Sampling weights
- Special tabulations of ACS data
- Adoption of ACS questions
- Development of occupational profiles
- Contextual variables

For a federal data product to be included in this study, it should be a regularly updated data resource available to public users. Thus, reports that simply analyze publicly available ACS data are excluded, as are non-unique ACS-based measures created to guide the allocation of federal grant funds (such as the local unemployment rate). The types of products included here are data series (e.g., employment), data repositories and packages for analysis and regulation (e.g., in the fields of transportation, aging, banking, and education), indices (e.g., costs of living), indicators (e.g., program performance), and classification systems (e.g., occupations, metropolitan areas).

This document is intended to be a reference in two ways. The first is to help the many stakeholders in the ACS more fully comprehend its role as an input, platform, or resource for other federal statistical efforts. The audience includes Census Bureau and Commerce Department leadership, the Office of Management and Budget, the Interagency Council on Statistical Policy (ICSP), members of and staff to Congress and its committees, and the

nonfederal users of ACS statistics. In particular, it is hoped that the document will be useful in the processes of ACS content review as mandated by congressional directive and the ICSP.[2]

The document's second purpose is to assist staff in the Census Bureau's ACS Office (ACSO) in better understanding the needs of one important group of ACS customers, federal agencies.

Each statistical product overview is in a standard format—nature and purpose, responsible unit, authorization, frequency of data publication, timing of data release (i.e., how soon after collection), overall methodology, use of the ACS in the methodology, ACS questions utilized, and major uses of the data product.

To a very large extent, text are provided verbatim from footnoted sources. Periodically, that text has been shortened for clarity and conciseness.

The tables in the Appendix provide an overview of the statistical products in the document:

- Table 1 indicates the ACS questions relied on by each data product, as interpreted by the authors on the basis of the methodology documentation.
- Table 2 shows the nature of each federal statistical product's use of the ACS.
- Table 3 notes the subject of each federal statistical product reliant on the ACS.
- For each federal statistical product reliant on the ACS, Table 4 identifies product type, product access, intended primary purposes, and intended primary users.

To be continued . . .

---

[2] "Charter of the Interagency Council on Statistical Policy, Subcommittee on the American Community Survey," August 10, 2012.

I.   Executive Office of the President

A. Office of Management and Budget

1. *Core Based Statistical Areas*[3]

Purpose of the Measure: Metropolitan and micropolitan statistical areas, or "Core Based Statistical Areas" (CBSA), are geographic entities delineated by the Office of Management and Budget. The purpose of this classification is to provide a nationally consistent set of delineations for collecting, tabulating, and publishing Federal statistics for geographic areas.

Responsible Unit: Office of Information and Regulatory Affairs, Office of Management and Budget

Authorization: Under 44 U.S.C. 3504(e)(3) and 31 U.S.C. 1104(d) and Executive Order No. 10253 (June 11, 1951), OMB delineates Metropolitan Statistical Areas, Metropolitan Divisions, Micropolitan Statistical Areas, Combined Statistical Areas, and New England City and Town Areas for use in Federal statistical activities.

Frequency of Data Publication: OMB had been updating the delineations every year until 2008. It has shifted to a schedule of updating once every five years, the most recent being in 2013. The next scheduled update is 2018.

The standards for delineating the areas are reviewed and revised once every ten years, prior to each decennial census. Standards were last revised in 2010.

Timing of Data Release: Not applicable.

Modes to Access Data: Current Lists of Metropolitan and Micropolitan Statistical Areas and Delineations, Census Bureau website (xls and pdf)

Methodology: OMB delineates CBSAs and their components on the basis of the following definitions:

- Core Based Statistical Area (CBSA)—A statistical geographic entity consisting of the county or counties associated with at least one core (urbanized area or urban cluster) of at least 10,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties with the counties containing the core. Metropolitan and Micropolitan Statistical Areas are the two categories of Core Based Statistical Areas.
- Central county—The county or counties of a Core Based Statistical Area containing a substantial portion of an urbanized area or urban cluster or both, and to and from which commuting is measured to determine qualification of outlying counties.
- Employment interchange measure—A measure of ties between two adjacent entities. The employment interchange measure is the sum of the percentage of

---

[3] Information also is available at "Metropolitan and Micropolitan Statistical Areas Main," Census Bureau website.

workers living in the smaller entity who work in the larger entity and the percentage of employment in the smaller entity that is accounted for by workers who reside in the larger entity.

- Main county—A county that acts as an employment center within a Core Based Statistical Area that has a core with a population of at least 2.5 million. A main county serves as the basis for delineating a Metropolitan Division.
- Metropolitan Statistical Area—A Core Based Statistical Area associated with at least one urbanized area that has a population of at least 50,000. The Metropolitan Statistical Area comprises the central county or counties containing the core, plus adjacent outlying counties having a high degree of social and economic integration with the central county or counties as measured through commuting.
- Micropolitan Statistical Area—A Core Based Statistical Area associated with at least one urban cluster that has a population of at least 10,000, but less than 50,000. The Micropolitan Statistical Area comprises the central county or counties containing the core, plus adjacent outlying counties having a high degree of social and economic integration with the central county or counties as measured through commuting.
- Outlying county—A county that qualifies for inclusion in a Core Based Statistical Area on the basis of commuting ties with the Core Based Statistical Area's central county or counties.
- Secondary county—A county that acts as an employment center in combination with a main county or another secondary county within a Core Based Statistical Area that has a core with a population of at least 2.5 million. A secondary county may serve as the basis for delineating a Metropolitan Division, but only when combined with a main county or another secondary county.

A county qualifies as an outlying county of a CBSA if it meets the following commuting requirements:

- At least 25 percent of the workers living in the county work in the central county or counties of the CBSA; or
- At least 25 percent of the employment in the county is accounted for by workers who reside in the central county or counties of the CBSA.[4]

Application of ACS Data in Methodology: That ACS provides employment and commuting estimates that OMB uses to delineate CBSAs. In 2013, OMB delineated CBSAs based on 2010 Census data and 2006– 2010 American Community Survey 5-year estimates. It will issue the next round of delineations in 2018 using Census Bureau Population Estimates Program total population estimates as well as the 2011–2015 American Community Survey 5-year commuting and employment estimates.[5]

ACS Questions Utilized (as suggested by the methodology): P29-30

---

[4] OMB, "2010 Standards for Delineating Metropolitan and Micropolitan Statistical Areas; Notice," *Federal Register*, Vol. 75, No. 123, June 28, 2010, pp. 37246-37252.

[5] *Ibid.*

Major Uses: The CBSA classification provides a nationally consistent set of delineations for collecting, tabulating, and publishing Federal statistics for geographic areas. These delineations are used throughout the economy by state and local governments, businesses, and researchers. Delineations of CBSAs are an essential aid to understanding the dynamics of American society and economy.

II.   U.S. Department of Agriculture

A. Economic Research Service

2. *Commuting Zones*

Nature and Purpose of the Measure: A local economy and its labor market are bounded not by the nearest county line, but by interrelationships between buyers and sellers of labor. If we are to understand the diversity of nonmetro America, we need a geographic standard capturing variations in local economic and labor force activities. The objective of Commuting Zones (CZ's) is to develop a geographic unit that better captures the economic and social diversity of nonmetro areas.[6]

Responsible Unit: Economic Research Service (ERS), U.S. Department of Agriculture (USDA)

Authorization: 7 CFR 3700.3

Frequency of Data Publication: Once a decade or longer. ERS prepared CZ's on the basis of the 1980, 1990, and 2000 decennial censuses. It is contracting with Pennsylvania State University to prepare a new set of CZ's based on the American Community Survey.

Timing of Data Release: Not known

Modes of Data Access: "Commuting Zones and Labor Market Areas," ERS website (xls)

Methodology: The local labor market is viewed as a set of relationships between employers and workers. These relationships exist in a space bounded by places of work and residence. As such, this spatial conception of labor markets dictates our methods, data sources, and procedures.

Hierarchical cluster analysis is used along with the Census Bureau's journey to work data to group counties into CZ's. CZ's are designated without regard to a minimum population threshold and are intended to be a spatial measure of the local labor market. In 2000, there were 709 commuting zones delineated for the U.S.

In 1980, existing geographic schema used to represent labor market areas were deemed unsatisfactory for rural labor market analysis. A plan was developed for a new geographic specification that included all U.S. counties and county equivalents, used uniform criteria for designating labor market areas, employed the most recent journey-to-work data, did not require each area to have an urban center, and could meet prevailing U.S. Census confidentiality standards.

Data for New England MCDs were aggregated to the county level. Similarly, Virginia independent city data were combined with counties in which the cities are located. Commuter flows were then organized in frequency and proportional flow matrices. Frequency matrices are flows of commuters within and between the sample counties. Because there are wide variations in county populations, we converted absolute commuting flows in the frequency

---

[6] Text source: "Commuting Zones and Labor Market Areas—Documentation," ERS website.

matrices to proportional measures. For counties i and j, the proportional flow measure is defined as the sum of shared commuters divided by the smaller of the two resident labor forces:

> ((commuters from county i) + (commuters from county j))/(resident labor force of smaller county)

We employed a hierarchical cluster analytic technique which indicates the strength of association among combinations of units beginning with the strongest pair and ending with one large cluster of all units. [7]

<u>Application of ACS in the Methodology</u>: For the new CZ's, the matrices will be based on journey-to-work data from the ACS.

<u>ACS Questions Utilized (as suggested by the methodology)</u>: P30

<u>Major Uses</u>: According to ERS, the procedures utilized in the original delineation and the resulting geography greatly assisted research on rural and urban employment issues. The detailed commuting zones and broader labor market areas were principally intended to be statistical units for analysis of nonmetropolitan labor market performance and employment problems. The geography and procedures have been adapted to a variety of other issues, however, including research on small business development, health service areas, detailed journey-to-work patterns, and basic demographic processes. The development of the new geography permits longitudinal analyses of stability and change in local labor markets across the nation.[8]

---

[7] Text source: Charles M. Tolbert and Molly Sizer, "U.S. Commuting Zones and Labor Market Areas: A 1990 Update," Rural Economy Division, Economic Research Service, Staff Paper No. AGES-9614, September 1996.
[8] Text source: *Ibid.*

3. *County Economic and Policy Typology Codes*

Nature and Purpose of the Measure: An area's economic and social characteristics have significant effects on its development and need for various types of public programs. To provide policy-relevant information about diverse county conditions to policymakers, public officials, and researchers, the Economic Research Service (ERS) has developed a set of county-level typology codes that captures differences in economic and social characteristics.

ERS County Typology Codes classify all U.S. counties according to six non-overlapping categories of economic dependence and eight overlapping categories of policy-relevant themes:

- Economic types – farming, mining, manufacturing, services, Federal/State government, and unspecialized counties.
- Policy types – housing stress, low education, low employment, persistent poverty, population loss, nonmetro recreation, retirement destination, and persistent child poverty.[9]

Responsible Unit: Economic Research Service (ERS), U.S. Department of Agriculture (USDA)

Authorization: 7 CFR 3700.3

Frequency of Data Publication: County Typology Codes were first created in 1979 and were updated in 1986, 1989, and 2004. They are in the midst of being updated in 2014-2015. The persistent poverty county classification has been completed.

Timing of Data Release: Not applicable.

Modes of Data Access: "County Typology Codes," ERS website (xls)

Methodology: Although ERS coded all U.S. counties in 2004, the thresholds for the economic and policy types were set using nonmetro counties only. Most thresholds were roughly set at the nonmetro mean plus one standard deviation. ERS used counties that met the 2003 definition of nonmetro (micropolitan and noncore combined) in analyzing the means.

Labor and proprietors' earnings by place of work are the basis for the economic dependence categories. Each industry's earnings were calculated as a percent of total labor and proprietors' earnings in the county. County-level estimates of earnings by place of work used to measure economic dependence came from the Bureau of Economic Analysis' (BEA) Regional Economic Information System (REIS).

Selection of the industries ERS classified was guided by regional economics theory. Farming, mining, manufacturing, and Federal/State government industries produce goods or services for export outside the local economy. Exporting industries are termed 'basic' in regional economics and are often shown to be sources of larger growth in local economies (or declines during economic downturns) than industries that produce for the local market. Service industries may

---

[9] Text source: "County Typology Codes—Overview," ERS website.

either produce for the local or export economies. ERS set a high service earnings threshold to help assure that the counties we classified as services-dependent do have service industries that serve more than the local population. These economic dependence categories are mutually exclusive.

In 2004, the housing stress, low-education, low-employment, persistent poverty, population loss, and retirement destination classifications were based on Census 2000 long-form data. The recreation classification was based on data from BEA, County Business Patterns, the Economic Census, and Census 2000.[10]

Application of ACS in the Methodology: In the 2014-2015 revision of the policy typologies, the American Community Survey replaces the decennial long form as the primary data source.

For the 2014 classification of persistent poverty counties, ERS defined counties as being persistently poor if 20 percent or more of their populations were living in poverty over the last 30 years (measured by the 1980, 1990 and 2000 decennial censuses and 2007-11 American Community Survey 5-year estimates).[11]

In 2004, the following measures were computed on the basis of the decennial long form:

- Housing stress counties—30 percent or more of households had one or more of these housing conditions in 2000: lacked complete plumbing, lacked complete kitchen, paid 30 percent or more of income for owner costs or rent, or had more than 1 person per room.
- Low-education counties —25 percent or more of residents 25-64 years old had neither a high school diploma nor GED in 2000.
- Low-employment counties —less than 65 percent of residents 21-64 years old were employed in 2000.
- Persistent child poverty counties—20 percent or more of related children under 18 years old were poor as measured by each of the last 4 censuses, 1970, 1980, 1990, and 2000.
- Population loss counties—the number of residents declined both between the 1980 and 1990 censuses and between the 1990 and 2000 censuses.
- Retirement destination counties--number of residents 60 and older grew by 15 percent or more between 1990 and 2000 due to inmigration.

For each of these measures, the forthcoming updates will rely on the ACS.

ACS Questions Utilized (as suggested by the methodology): number in household, P4, P11, P29, P35, P47-48, H7-8, H14, H18, H20-24

---

[10] Text source: "County Typology Codes—Documentation," ERS website.
[11] Text source: "Persistence of Poverty," ERS website.

: The County Typology Codes enable analysts and researchers to provide policy-relevant information about diverse county conditions to policymakers, public officials, and the public.

4. *Current Population Survey – Food Security Supplement*

5. *Food Access Research Atlas*

<u>Nature and Purpose of the Dataset</u>: The Food Access Research Atlas:

- Presents a spatial overview of food access indicators for low-income and other census tracts using different measures of supermarket accessibility;
- Provides food access data for populations within census tracts; and
- Offers census-tract-level data on food access that can be downloaded for community planning or research purposes. [12]

Indicators can meaningfully be grouped by:

- General tract characteristics,
- Low-access and distance measures,
- Vehicle availability,
- Group quarters,
- Low-income and low-access measures, and
- Low access by population subgroups. [13]

<u>Responsible Unit</u>: Economic Research Service (ERS), U.S. Department of Agriculture (USDA)

<u>Authorization</u>: 7 CFR 3700.3

<u>Frequency of Data Publication</u>: ERS released the Food Desert Locator in May 2011. The updated and revised Atlas replaced it in March 2013.

<u>Timing of Data Release</u>: The 2011 Food Desert Locator was based on Census 2000 data. The 2013 Food Access Research Atlas was based on the 2006-2010 American Community Survey.

<u>Modes of Data Access</u>: "Food Access Research Atlas," ERS website (maps, xls)

<u>Methodology/Application of ACS in the Methodology</u>: A full discussion of the definitions of terms and indicators used in the Food Access Research Atlas is available online.[14]

Population data, including age and residence in group quarters, are from the 2010 Census of the Population and downloaded at the census-block level before being allocated to ½-kilometer-square grid cells. Urban or rural designation was also provided by the 2010 Census at the block level. Data on income and vehicle availability are from the 2006-10 American Community Survey and were downloaded at the block-group level for assignment to ½-kilometer-square grid cells.

Two 2010 lists of supermarkets, supercenters, and large grocery stores (food stores selling all major categories of food and having annual sales of at least $2 million) were combined to produce a comprehensive list of stores that represent sources of affordable and nutritious food.

---

[12] Text source: "Food Access Research Atlas—Overview," ERS website.
[13] "Food Access Research Atlas—Documentation," ERS website.
[14] *Ibid.*

One store list contains stores authorized to receive Supplemental Nutrition Assistance Program (SNAP) benefits. The second list is from Trade Dimensions TDLinx (a Nielsen company), a proprietary source of individual supermarket store listings.[15]

ERS's Food Access Research Atlas was built using Environmental Systems Research Inc. (ESRI) ArcGIS Server technology. The background topographic and satellite maps, as well as the address locator service, were also provided by ESRI.

The original version of the Food Desert Locator implemented a single measure of food deserts—low-income areas where a significant number or share of residents is far from a supermarket, where "far" is more than 1 mile in urban areas and more than 10 miles in rural areas.

In the new Food Access Research Atlas, food access indicators for census tracts using ½-mile and 1-mile demarcations to the nearest supermarket for urban areas, 10-mile and 20-mile demarcations to the nearest supermarket for rural areas, and vehicle availability for all tracts are estimated and mapped. Users of the Atlas can view census tracts by food access indicators using these different measures, including the original food desert measure, to see how the map changes as the distance demarcation or inclusion of vehicle access changes.[16]

ACS Questions Utilized (as suggested by the methodology): P47-48, H12

Major Uses: The Food Access Research Atlas provides researchers, city planners, non-profit organizations, and policymakers at all levels a more nuanced view of the factors that impact whether people across the country can attain nutritious and varied diets.[17]

A number of indicators from the Food Access Research Atlas are included in the ERS Food Environment Atlas, the objectives of which are to assemble statistics on food environment indicators to stimulate research on the determinants of food choices and diet quality, and to provide a spatial overview of a community's ability to access healthy food and its success in doing so.

---

[15] *Ibid.*
[16] Text source: "Food Access Research Atlas—About the Atlas," ERS website.
[17] Text source: Paula Dutko, "Updated Web Tool Maps U.S. Food Access in Greater Detail," USDA blog, March 11, 2013.

6. *Rural-Urban Commuting Area Codes*

Nature and Purpose of the Measure: Rural-Urban Commuting Area (RUCA) Codes are a classification of census tracts using measures of population density, urbanization, and daily commuting. The classification contains two levels. Whole numbers (1-10) delineate metropolitan, micropolitan, small town, and rural commuting areas based on the size and direction of the primary (largest) commuting flows. These 10 codes are further subdivided based on secondary commuting flows, providing flexibility in combining levels to meet varying definitional needs and preferences.

Responsible Unit: Economic Research Service (ERS), U.S. Department of Agriculture (USDA)

Authorization: 7 CFR 3700.3

Frequency of Data Publication: RUCA Codes were created in 1990 and updated in 2000 and 2010.

Timing of Data Release: The 2010 RUCA Codes were released at the end of 2013.

Modes of Data Access:

- For RUCA Codes by census tract, "Rural-Urban Commuting Area Codes—Overview," ERS website (xls)
- For RUCA Codes by ZIP Code, "Temporary Zip RUCA 3.10 File Access Page," University of North Dakota (csv)

Methodology/Application of ACS in the Methodology: The rural-urban commuting area (RUCA) codes, a detailed and flexible scheme for delineating sub-county components of rural and urban areas, have been updated using data from the 2010 decennial census and the 2006-10 American Community Survey (ACS). RUCA codes are based on the same theoretical concepts used by the Office of Management and Budget (OMB) to define county-level metropolitan and micropolitan areas. We applied similar criteria to measures of population density, urbanization, and daily commuting to identify urban cores and adjacent territory that is economically integrated with those cores. We adopted OMB's metropolitan and micropolitan terminology to highlight the underlying connectedness between the two classification systems. However, the use of census tracts instead of counties as building blocks for RUCA codes provides a different and more detailed geographic pattern of urban and rural areas.

Census tracts are used because they are the smallest geographic building block for which commuting flow estimates are available from the U.S. Census. Tract-to-tract commuting flow files were constructed from ACS data as part of a special tabulation for the Department of Transportation—the Census Transportation Planning Package. To derive estimates for small geographic units such as census tracts, information collected annually from over 3.5 million housing units was combined across 5 years (2006-10).

The classification contains 10 primary and 21 secondary codes. Few, if any, research or policy applications need the full set of codes. Rather, the system allows for the selective combination of codes to meet varying definitional needs.

- The primary codes designate the single largest, commuting share.
- Metropolitan cores (code 1) are defined as census tract equivalents of urbanized areas. Micropolitan and small town cores (codes 4 and 7, respectively) are tract equivalents of urban clusters. Tracts are included in urban cores if more than 30 percent of their population is in the urbanized area or urban cluster.
- High commuting (codes 2, 5, and 8) means that the largest commuting share was at least 30 percent to a metropolitan, micropolitan, or small town core. Many micropolitan and small town cores themselves (and even a few metropolitan cores) have high enough out-commuting to other cores to be coded 2, 5, or 8; typically these areas are not job centers themselves but serve as bedroom communities for a nearby, larger city.
- Low commuting (codes 3, 6, and 9) refers to cases where the single largest flow is to a core, but is less than 30 percent. These codes identify "influence areas" of metro, micropolitan, and small town cores, respectively, and are similar in concept to the "nonmetropolitan adjacent" codes found in other ERS classification schemes (Rural-Urban Continuum Codes, Urban Influence Codes).
- The last of the general classification codes (10) identifies rural tracts where the primary flow is local or to another rural tract.
- Secondary flows may indicate other connections among rural and urban places. Thus, the primary RUCA codes are further subdivided to identify areas where classifications overlap, based on the size and direction of the secondary, or second largest, commuting flow. For example, 1.1 and 2.1 codes identify areas where the primary flow is within or to a metropolitan core, but another 30 percent or more commute to a larger metropolitan core. Similarly, 10.1, 10.2, and 10.3 identify rural tracts for which the primary commuting share is local, but more than 30 percent also commute to a nearby metropolitan, micropolitan, or small town core, respectively.

A link to the ZIP code approximation of the 2010 RUCA codes is available on the Center for Rural Health website. It is based on a crosswalk between census tracts and ZIP code areas, not on a separate analysis of population and commuting data unique to the ZIP code geographic unit.

ACS Questions Utilized (as suggested by the methodology): P30

Major Uses: RUCA codes are many, but permit stricter or looser delimitation of metropolitan, micropolitan, and small town commuting areas. This classification scheme provides an alternative to county-based systems for situations where more detailed geographic analysis is feasible. It identifies areas of emerging urban influence and areas where urban-rural classifications overlap, thus providing an exhaustive system of statistical areas for the country.

7. *Rural-Urban County Classification Codes*

Nature and Purpose of the Measure: Economic Research Service (ERS) researchers and others who analyze conditions in "rural" America most often study conditions in nonmetropolitan (nonmetro) areas, defined on the basis of counties. Counties are the standard building block for collecting economic data and for conducting research to track and explain regional population and economic trends.

In addition to conducting research that uses the basic metro-nonmetro dichotomy, ERS has developed multi-level county classifications to measure rurality in more detail and to assess the economic and social diversity of nonmetro America. [18] They include:

- Rural-Urban Continuum Codes – a classification scheme that distinguishes metropolitan counties by the population size of their metro area, and nonmetropolitan counties by degree of urbanization and adjacency to a metro area. Each county in the U.S. is assigned one of the 9 codes.
- Urban-Influence Codes – a classification scheme that distinguishes metropolitan counties by population size of their metro area, and nonmetropolitan counties by size of the largest city or town and proximity to metro and micropolitan areas. There are two metro and 10 nonmetro categories, resulting in a 12-part county classification.

Responsible Unit: Economic Research Service (ERS), U.S. Department of Agriculture (USDA)

Authorization: 7 CFR 3700.3

Frequency of Data Publication: For the most part, the rural classification codes are updated after the completion of a decennial census.

- Rural-Urban Continuum Codes were originally developed in 1974 and updated in 1983, 1993, 2003, and 2013.
- Urban Influence Codes were developed in 1993 and updated in 2003 and 2013.

Timing of Data Release: The rural classification codes are updated between three and five years after the decennial year.

Modes of Data Access: "Rural Classifications," ERS website (xls)

Methodology: To create the 2013 Rural-Urban Continuum Codes:

- All U.S. counties and county equivalents were first grouped according to their official metro-nonmetro status, defined by the Office of Management and Budget (OMB) as of February, 2013. OMB determined current metropolitan status by applying population and worker commuting criteria to the results of the 2010 Census and the 2006-2010 American Community Survey (ACS).

---

[18] Text source: "Rural Classifications," ERS website.

- For the Continuum Codes, metro counties are divided into three categories according to the total population size of the metro area of which they are part: 1 million people or more, 250,000 to 1 million people, and below 250,000. Nonmetro counties are classified along two dimensions.
  - First, they are divided into three urban-size categories (an urban population of 19,999 or more, 2,500 to 20,000, and less than 2,500) based on the total urban population in the county.
  - Second, nonmetro counties in the three urban-size categories are sub-divided by whether or not the county is adjacent to one or more metro areas. A nonmetro county is defined as adjacent if it physically adjoins one or more metro areas, and has at least 2% of its employed labor force commuting to central metro counties. Nonmetro counties that do not meet these criteria are classed as nonadjacent.[19]

The 2013 Urban Influence Codes divide the 3,143 counties, county equivalents, and independent cities in the United States into 12 groups:

- Metro counties are divided into two groups according to the population size of the metro area--those in "large" areas have at least 1 million residents and those in "small" areas have fewer than 1 million residents.
- Nonmetro counties include all counties outside metro areas and are delineated as micropolitan or noncore using OMB's classification.
  - Nonmetro micropolitan counties are divided into three groups distinguished by metro size and adjacency: adjacent to a large metro area, adjacent to a small metro area, and not adjacent to a metro area.
  - Nonmetro noncore counties are divided into seven groups distinguished by their adjacency to metro or micro areas and whether or not they contain a town of at least 2,500 residents.
- The 2013 Urban Influence Codes are based on the OMB metropolitan classification announced in February 2013, which in turn are based on population data from the 2010 Census of Population and commuting data from the 2006-2010 American Community Survey (ACS).
- Nonmetro counties are defined as adjacent if they abut a metro area (or if nonmetro noncore counties abut a micro area) and have at least 2% of employed persons commuting to work in the core of the metro area (or in the micro area).[20]

Application of ACS in the Methodology: For the 2013 rural-urban county classification codes, 2006-2010 ACS journey-to-work data were used to determine nonmetro county adjacency to a metro area.

ACS Questions Utilized (as suggested by the methodology): P30

---

[19] Text source: "Rural-Urban Continuum Codes—Documentation," ERS website.
[20] Text source: "Urban Influence Codes—Documentation," ERS website.

Major Uses: The ERS rural-urban county classification schemes are used for research and policy analysis and to determine eligibility for Federal programs that assist rural areas.

B. Food and Nutrition Service

8. *Estimates of WIC Eligibles and Program Reach*

Nature and Purpose of the Measure: The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) provides supplemental nutritious foods, nutrition education (including breastfeeding promotion and support), and referrals to health care and other social services at no charge. WIC serves low-income pregnant, postpartum, and breastfeeding women, infants, and children up to age 5 who are at nutritional risk. Because WIC is a Federal grant program for which Congress authorizes a specific amount of funds each year, the Food and Nutrition Service (FNS) requires estimates of the total number of individuals eligible for the program to anticipate funding needs.

This report provides estimates of the population that met WIC eligibility criteria in 2011. National eligibility is shown for each categorical subpopulation: infants, children age 1 through 4 by single year of age, pregnant women, postpartum women who are breastfeeding, and postpartum women who are not breastfeeding.

The ratio of program participants to eligibles, defined as the coverage rate, is provided for all WIC participants and for these subpopulations. The report also shows trends in the estimated growth in WIC eligible and coverage rates from 2000 through 2011. Estimates of WIC eligibility are provided for the seven FNS regions and each State for 2011.[21]

Responsible Unit: Office of Research and Analysis (ORA), FNS, U.S. Department of Agriculture (USDA)

Authorization: Section 17 of the Child Nutrition Act of 1966 (42 USC 1786)

Frequency of Data Publication: Annual

Timing of Data Release: Approximately 24 months after the end of the reference calendar year.

Modes of Data Access: WIC Studies, ORA, FNS (pdf)

Methodology: The estimation procedures used in this report build on the methodology recommended by the Committee on National Statistics of the National Research Council (CNSTAT) in 2003. National eligibility estimation requires nationally representative data and numerous assumptions that take into account program certification periods, individuals' enrollment in other programs, and mothers' breastfeeding choices. The 2011 national estimates use the 2012 Current Population Survey, Annual Social and Economic Supplement (the CPS-ASEC, formerly referred to as the March supplement), as originally recommended by CNSTAT. The State estimates use the 2011 American Community Survey (ACS) and are converted to shares of the national estimates to produce State-specific eligibility estimates

---

[21] Text source: U.S. Department of Agriculture, Food and Nutrition Service, Office of Research and Analysis. "National and State-Level Estimates of Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) Eligibles and Program Reach, 2011," by Paul Johnson, Linda Giannarelli, Erika Huber, and David Betson. Project Officer: Grant Lovellette. Alexandria, VA: March 2014, p. i.

consistent with national totals. The number eligible in the territories is based on data from the ACS for Puerto Rico (the 2011 PRCS) and estimates of the population in other territories. Standard error of the estimates are calculated for national, regional, State, and Puerto Rico estimates.[22]

Application of ACS in the Methodology: The State estimates begin with the ACS data collected during 2011. The 2011 ACS captures a combination of 2010 and 2011 income. This is not ideal for estimation of 2011 WIC eligibility; but the ACS is nevertheless the best data source for determining State shares of WIC eligibility due to its very large sample sizes in all States.

The CNSTAT Panel recommended that all members of a household related by blood, marriage, or adoption be considered as one family unit for the purposes of determining WIC eligibility. However, the only information the ACS provides on family relationships is each individual's relationship to the reference person (householder); for people not related to the householder, their relationships to each other are unknown. Since the Minnesota Population Center's IPUMS-USA provides researchers with educated conjectures about the relationships between persons not related to the householder, we use the ACS with these imputations. For each State, the numbers of infants and children who are income-eligible or adjunctively-eligible for WIC (enrolled in SNAP, TANF, or public health insurance) are estimated using the ACS data.

The process for estimating State-level eligibility from the ACS data involves the use of adjustment factors. State-specific data were used in two of the adjustments—for population and for breastfeeding. Using Census Bureau State population estimates by age, the 2011 ACS person weights for infants and children are proportionally adjusted so that the sums of the persons by age are equal to the Census population estimates for each State.

For other adjustment factors used in estimating WIC eligibility for infants and young children—addressing annual-to-monthly income differences and nutritional risk—the national factors were used in each State. Estimates for pregnant and postpartum women are derived from the infant estimates as with the national estimates, with the exception that the breastfeeding adjustments incorporate State variation in breastfeeding rates.

These procedures produce ACS-based WIC eligibility estimates for each State and the District of Columbia; however, the sum of these estimates is not the same as the national estimate produced from the CPS-ASEC data. The CPS-ASEC has been judged as the better source for a national WIC eligibility estimate, due to the fact that the CPS-ASEC has more complete income and program participation data. Also, the CPS asks respondents for their income during the calendar year, while the ACS surveys households throughout the year and asks about income in the twelve months prior to the interview. Consequently, the ACS is less likely to detect increases in eligibility as the economy falters or decreases in eligibility when the economy improves.

To create a consistent set of national and State WIC eligibility estimates, a top down approach is used. Specifically, we compute each State's share of the total ACS-based eligibility estimate,

---

[22] Text source: *Ibid*, p. i.

and then allocate the national estimates computed from the CPS-ASEC according to those State shares. With this approach, the accepted methodology for producing national estimates and a consistent time series of estimates can be maintained.

We calculate State shares for each subgroup, which are applied to the CPS-ASEC national estimates for each subgroup. This produces estimates by subgroup at the FNS regional level and total WIC eligibility estimates at the State level. While estimates for subpopulations help to build total WIC eligibility variation across the States, they are not sufficiently reliable to publish individually, as eligibility subgroups are relatively small in many States.

The standard errors for the State-level estimates were derived using a generalized variance model described in the Census Bureau's annual ACS report based on one year accuracy of the data.

The territorial estimates use the Puerto Rico Community Survey (PRCS) to directly estimate the number of eligibles in Puerto Rico and the Census Bureau International Data Base to estimate WIC eligibility in other island territories.[23]

ACS Questions Utilized (as suggested by the methodology): P2, P4, P16, P24, P47-48, H15

Major Uses: Each year's estimates of WIC eligibles and program reach are used to assess changes over time in these variables by seven participant groups (pregnant women, breast-feeding women, non-breast-feeding women, and children ages one, two, three, and four). The analyses are used to determine the appropriate level of budget requests and appropriations for the WIC program as well as to guide program outreach efforts.

---

[23] Text source: *Ibid.*, pp. 12-14, 47.

9. *SNAP Program Access Index*

Nature and Purpose of the Measure: The Program Access Index (PAI) is designed to indicate, by state, the degree to which low-income people have access to benefits provided through the administration of the Supplemental Nutrition Assistance Program (SNAP). The Food and Nutrition Service (FNS), U.S. Department of Agriculture, uses the PAI to reward States for high performance in the administration of SNAP.[24]

Responsible Unit: Office of Policy Support, Food and Nutrition Service, USDA

Authorization: Section 4120, Farm Security and Rural Investment Act of 2002 and 7 CFR 275.24 (High Performance Bonuses)

Frequency of Data Publication: Annual

Timing of Data Release: By law, the SNAP performance bonuses are to be distributed by the end of the fiscal year that starts in the PAI reference calendar year. So, for instance, the PAI for calendar year 2012 was released by September 30, 2013.

Modes of Data Access: Annual "Step by Step Guide" to calculating the PAI (pdf)

Methodology: In general, the PAI is a simple index of the average monthly number of SNAP participants over the course of a calendar year to the number of people with income below 125 percent of the official poverty line in each State. FNS computes average monthly participation over a calendar year – rather than the Federal fiscal year – to better align the participation count with the annual poverty measure.

FNS makes an adjustment to the counts of participants, the numerator of this index, to better reflect State performance in the administration of SNAP. FNS subtracts out estimates of the number of people who received SNAP disaster assistance.

FNS makes two adjustments to the poverty counts in the denominator of the index. Participants in the Food Distribution Program on Indian Reservations (FDPIR) are subtracted from the number of people with income below 125 percent of poverty in each State. In addition, poor Supplemental Security Income (SSI) recipients in California are subtracted from the number of people with income below 125 percent of poverty. By law, SSI recipients in California are ineligible for SNAP because they receive a State-funded cash supplement for food.[25]

Application of ACS in the Methodology: The PAI uses the American Community Survey (ACS) for the initial state poverty count estimates (index denominator). FNS did a comparative analysis of the merits of the American Community (ACS) and Current Population (CPS) surveys and

---

[24] Text source: Food and Nutrition Service, Office of Policy Support, "Calculating the Supplemental Nutrition Assistance Program (Snap) Program Access Index: A Step-By-Step Guide for 2012," February 2014, p. 1.
[25] Text source: *Ibid*, pp. 1-2.

determined that the ACS provided a better basis for the calculation of the Program Access Index beginning with the 2005 high performance bonuses.[26]

<u>ACS Questions Utilized (as suggested by the methodology)</u>: P2, P4, P47-48

<u>Major Uses</u>: FNS awards performance bonuses to the four States with the highest PAI and to the four States with the most improved PAI, provided that a State is eligible to receive such a bonus.

Independent of its use in allocating bonuses, the PAI is used by program managers, state legislatures, and researchers as an indicator of state SNAP program performance.

---

[26] Text source: *Ibid,* p. 1.

III.   U.S. Department of Commerce

A. Bureau of Economic Analysis

10. *Local Employment*

Purpose of the Data Series:

Responsible Unit: Bureau of Economic Analysis, Department of Commerce

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data: Interactive Data Application, Regional Data, GDP and Personal Income, Bureau of Economic Analysis (xls, csv, pdf)

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

11. *Local Personal Income*

Purpose of the Data Series: The annual statistics of personal income for local areas measure the income received by or on behalf of the residents of the area. BEA prepares statistics for counties, metropolitan areas, micropolitan areas, metropolitan divisions, combined statistical areas, and BEA economic areas. These statistics provide a consistent framework for analyzing and comparing individual local area economies.

BEA also prepares annual statistics on per capita personal income for local areas, which is an indicator of the economic well-being of the residents of an area.[27]

Responsible Unit: Bureau of Economic Analysis, Department of Commerce

Authorization: Commerce Department Departmental Organization Order 35-1A (Dec 13, 1982)

Frequency of Data Publication: Annual

Timing of Data Release: In November, the estimates of local area personal income for the previous year are prepared, and the estimates for the two years before that are revised. The local area personal income estimates are controlled to the annual state personal income estimates released in September each year.[28]

Modes to Access Data: Interactive Data Application, Regional Data, GDP and Personal Income, Bureau of Economic Analysis (xls, csv, pdf)

Methodology: Personal income consists of the income of individuals, nonprofit institutions serving individuals, private noninsured welfare funds, and private trust funds. Personal income also includes transfer receipts, several types of imputed incomes, employer contributions to health and pension plans, and all of the interest received by individuals. Personal income excludes personal contributions for government social insurance, realized capital gains and losses, and pension and annuity benefits from private and government employee pension plans.

Per capita personal income is calculated as the personal income of the residents of a given area divided by the resident population of that area. In computing per capita personal income for states and counties, BEA uses the Census Bureau's annual midyear (July 1) population estimates.

The state and county personal income estimates are based primarily on administrative records data. In addition some survey and census data are used. The administrative records data are a byproduct of the administration of various federal and state government social insurance programs and tax codes. They may originate either from the recipients of the income or from the payer of the income. Some of the more important of these programs and taxes (and the agencies compiling the data) are:

---

[27] Text source: *Measuring the Nation's Economy: A Guide to the Bureau of Economic Analysis*, p. 18.
[28] Text source: BEA, "Local Area Personal Income Methodology," November 2014, p. I-12.

- State unemployment insurance programs (Bureau of Labor Statistics, U.S. Department of Labor)
- State Medicaid programs and the federal Medicare program (Centers for Medicare and Medicaid Services, U.S. Department of Health and Human Services)
- Social Security (Social Security Administration)
- Federal veterans' programs (U.S. Department of Veterans Affairs)
- State and federal income tax codes (Internal Revenue Service, U.S. Department of the Treasury and Bureau of the Census, U.S. Department of Commerce)

The census data are mainly collected from the recipients of the income. The most important sources of census data for the personal income estimates are the Census of Population and Housing, conducted by the Bureau of the Census every ten years and the Census of Agriculture, conducted by the U.S. Department of Agriculture (USDA) every five years, and the Census of Governments.

The survey data are collected from both the payers and the recipients of the income. The more important surveys include the Annual Survey of Public Pensions and the American Community Survey conducted by the Census Bureau and the monthly Current Employment Statistics survey conducted by the Bureau of Labor Statistics.

The estimates of farm proprietors' income rely principally on the USDA's estimates of the income of all farms. The USDA uses sample surveys along with census data and administrative-records data to develop its estimates. The estimates of military compensation and employment rely principally on tabulations of data provided by the U.S. Department of Defense.

Using administrative records data and census data to measure local area personal income has both advantages and disadvantages. By using these data, BEA can prepare detailed annual and quarterly estimates at a relatively low cost and without increasing the reporting burden on businesses and households. However, because the source data often do not precisely match the concept being estimated, they must be adjusted to compensate for differences in definitions, coverage, timing, and geographic detail.[29]

Application of ACS Data in Methodology: Journey-to-work data from the ACS are used to adjust state government wages and salaries in six states. The geographic coding of BLS Quarterly Census of Employment and Wages data for the non-education component of state government payrolls attributes too much of the payrolls to the counties of the state capitals in six states. Therefore, wage data from the decennial Census of Population and more recently, 5-year estimates from the annual American Community Survey (ACS), are used in the preparation of the county estimates of the wages and salaries for these employees.

The county estimates of the wages and salaries for the non-education component, for Illinois, Michigan, New Jersey, Rhode Island, Tennessee, and Wisconsin are based on wage data derived from unpublished decennial tabulations of journey-to-work (JTW) data from the Census for the

---

[29] Text source: *Ibid*, pp. I-9-I-11.

years before 2001. Starting with 2008, the ACS JTW employment data was used, with 2001-2007 interpolated between the two JTW distributions.[30]

State estimates of contributions for veterans' life insurance are allocated to counties using the 2006-2010 American Community Survey 5-year estimates of the veteran population, centered on 2008. This allocator is held constant for subsequent years. State estimates of contributions for veterans' life insurance for prior years are allocated to counties using the veteran population from the decennial Census of Population. Linear interpolation is used for other years.[31]

The ACS offers journey-to-work data that will enable BEA to convert components of local area personal income that are recorded on a place-of-work basis to a place-of-residence basis. Currently, BEA relies on dated information from the 2000 decennial census for making these adjustments.[32] BEA plans to incorporate the information from the 2005-2009 ACS (5-year file) in its local area personal income statistics for 2001-2014, scheduled to be released in September 2015. The ACS data will provide the most up to date information on interstate commuting available in the United States and thereby help improve the accuracy of the BEA's local area personal income statistics.[33]

ACS Questions Utilized (as suggested by the methodology): P26-28, P30, P41-43, P47

Major Uses: The local area statistics provide detailed information by type of income that is comparable across all local areas and with the state statistics. Statistics on compensation and on earnings by place of work indicate the economic activity and industrial structure of business and government within the area. Statistics on personal income by place of residence provide a measure of the fiscal capacity of the area.[34]

The local area estimates of personal income and its components, per capita personal income, employment, and GDP by metropolitan area are widely used by both the public and the private sectors to track economic well-being over time and make comparisons across counties and metropolitan areas in the level and composition of economic activity and the value added by an area's industries. These estimates provide a framework for the analysis of local area economies, and they serve as a basis for decision making.

Federal agencies use these estimates in econometric models, such as those used to project energy and water use. They also use the estimates as a basis for allocating funds and for determining matching grants. For example, in fiscal year 2013, $308.0 billion in federal funds were distributed on the basis of BEA's state and local personal income and GDP statistics.

---

[30] Text source: *Ibid*, pp. II-3.
[31] Text source: *Ibid*, pp. VI-2.
[32] *Ibid*, Chapter VIII: Residence Adjustment.
[33] Text source: Personal communication with Joel Platt, Associate Director for Regional Economics, Bureau of Economic Analysis, February 27, 2015.
[34] Text source: *Measuring the Nation's Economy: A Guide to the Bureau of Economic Analysis*, p. 19.

In addition, as part of its program for small area income and poverty estimation (SAIPE), the Census Bureau uses the estimates of county per capita personal income as a predictor variable in the preparation of its county estimates of median household income.

The SAIPE program provides updated estimates of income and poverty statistics for the administration of federal programs and the allocation of federal funds to local jurisdictions.

State governments use the estimates of personal income and GDP to measure the economic base of planning areas. They also use the estimates in econometric models for various planning purposes and to project tax revenue and the need for public utilities and services.

University schools of business and departments of economics use the estimates for theoretical and applied economic research. Some of these schools distribute the estimates in abstracts or similar reports to various state and local government agencies, regional councils of governments, private research groups, businesses, and libraries.

Businesses use the estimates for planning activities, such as evaluating markets for new or established products and determining areas for the location, expansion, and contraction of their activities.[35]

---

[35] Text source: BEA, "Local Area Personal Income Methodology," November 2014, p. I-3.

12. *Regional Price Parities*

Purpose of the Data Series: Regional price parities (RPPs) are regional price levels expressed as a percentage of the overall national price level for a given year. The price level is determined by the average prices paid by consumers for the mix of goods and services consumed in each region. RPPs allow for comparisons of the purchasing power of personal income across regions.[36]

Responsible Unit: Bureau of Economic Analysis.

Authorization: Commerce Department Departmental Organization Order 35-1A (Dec 13, 1982)

Frequency of Data Publication: First release was April 2014, and the second release is set for April 2015. This measure is likely to be released annually in April.

Timing of Data Release: Sixteen months after the end of the reference calendar year.

Modes to Access Data: Interactive Data Application, Regional Data, GDP and Personal Income, Bureau of Economic Analysis (xls, csv, pdf)

Methodology/Application of ACS Data in Methodology: The RPPs are calculated using price quotes for a wide array of items from the Consumer Price Index (CPI), which are aggregated into broader expenditure categories (such as food, transportation, or education). Data on rents are obtained separately from American Community Survey. The expenditure weights for each category are constructed using the BLS' Consumer Expenditure Survey and BEA's Personal Consumption Expenditures. The broader categories and the data on rents are combined with the expenditure weights using a multilateral aggregation method that expresses a region's price level relative to the US.

Detailed CPI price data are adjusted to obtain average price levels for BLS-defined areas. These are allocated to counties in combination with direct price and expenditure data on rents from the ACS. County data are then aggregated to states and metropolitan areas. To estimate RPPs, CPI price quotes are quality adjusted and pooled over 5 years. The ACS rents are also quality adjusted, and in the case of the metropolitan areas, pooled over 3 years. The expenditure weights are specific to each year.[37]

---

[36] Text source: "Real personal income and Regional Price Parities for States and Metropolitan Areas, 2007-2011," Survey of Current Business (August 2013).

[37] Text source: *Ibid*.

ACS Questions Utilized (as suggested by the methodology): H1, H7, H17-18

Major Uses: RPPs enable individual Americans to compare inflation-adjusted incomes across states and metropolitan areas to better understand how their personal income may be affected by a job change or move. Businesses considering relocating or establishing new plants can use the RPP to obtain a comprehensive and consistent measure of differences in the cost of living and the purchasing power of consumers nationwide.[38]

---

[38] Text source: *Ibid*.

13. State Employment

Purpose of the Data Series: State Employment estimates measure the number of jobs in a state. Generally, employment can be measured either as a count of workers or as a count of jobs.

In the former case, an employed worker is counted only once; in the latter case, all jobs held by the worker are counted. The State Employment estimates are a count of the number of jobs, so that, as with the earnings estimates, a worker's activity in each industry and location of employment is reflected in the measure. These estimates consist of the number of wage and salary jobs, sole proprietorships, and general partners.[39]

Responsible Unit: Bureau of Economic Analysis

Authorization: Commerce Department Departmental Organization Order 35-1A (Dec 13, 1982)

Frequency of Data Publication: Annual

Timing of Data Release: Nine months after the reference year

Modes to Access Data: Interactive Data Application, Regional Data, GDP and Personal Income, Bureau of Economic Analysis (xls, csv, pdf)

Methodology: BEA's estimates of state employment consist of the number of wage and salary jobs, sole proprietorships, and general partners. The estimates are available annually beginning with 1969.

The state employment estimates are a complement to the place-of-work earnings estimates. Earnings are estimated on both a place-of-work basis by industry, and on a place-of-residence basis for the sum of all industries. The employment estimates are designed to conform conceptually and statistically with the place-of-work earnings estimates; the same source data—generally from administrative records—are used for both the earnings and employment estimates whenever possible. The earnings estimates reflect the scale and industrial structure of a state's economy rather than the income of the state's residents. Therefore, the employment estimates measure the number of jobs in a state, instead of the number of workers who perform the jobs. The characteristics of the state employment estimates follow from this concept and from the characteristics and limitations of the available source data.

The estimates of wage and salary employment are annual averages of twelve monthly observations for the year. This gives a job which lasts only part of the year a lesser weight that a year-round job. In contrast, the estimates of nonfarm proprietors' employment are counts of the number of proprietors active during any portion of the year. This is because the available source data do not indicate the portion of the year that the businesses are in operation.[40]

---

[39] Text source: BEA, "State Personal Income and Employment: Concepts, Data Sources, and Statistical Methods," Chapter XI: Employment, September 2014.

[40] Text source: *Ibid*.

Application of ACS Data in Methodology: ACS data is used to estimate employment in industries not covered by unemployment insurance programs, i.e. private households. "For this largely non-covered industry—mainly domestic servants—the national employment estimates are allocated to states in proportion to the "3-year estimates" of place-of-work private household employment from the American Community Survey. Estimates for 2000 and earlier years are based on the Census of Population journey-to-work data."[41]

ACS Questions Utilized (as suggested by the methodology): P29-30, P41-46

Major Uses: The state estimates of personal income and its components, per capita personal income, disposable personal income, employment, and GDP by state are widely used by both the public and the private sectors to measure and track economic well-being over time and to make comparisons across states in the level and composition of economic activity and the value added by a state's industries. These estimates provide a framework for the analysis of state economies, and they serve as a basis for decision making. Federal agencies use these estimates in econometric models, such as those used to project energy and water use. They also use the estimates as a basis for allocating funds and for determining matching grants.

University schools of business and economics use the estimates for theoretical and applied economic research. Some of these schools distribute the estimates in abstracts or similar reports to various state and local government agencies, regional councils of governments, private research groups, businesses, and libraries. Businesses use the estimates for planning activities, such as evaluating markets for new or established products and determining areas for the location, expansion, and contraction of their activities.[42]

---

[41] Text source: *Ibid*.
[42] Text source: *Ibid*.

14. *State Personal Income*

Purpose of the Data Series: Statistics on State Personal Income measure the income received by or on behalf of the residents of the state. State disposable personal income provides a measure of the income available for spending and saving, and state per capita personal income is an indicator of the economic well-being of the residents of a state.

These measures are the counterpart of the national personal income statistics, and they provide a consistent framework for analyzing and comparing individual state economies. The statistics provide detailed information by type of income (such as wages and salaries, dividend income, and social security benefits) that is comparable across all states and with the nation as a whole. Statistics of compensation and of earnings by place of work indicate the economic activity and industrial structure of business and government within the state, while statistics of personal income by place of residence provide a measure of the fiscal capacity of the state.[43]

Responsible Unit: Bureau of Economic Analysis, Department of Commerce

Authorization: Commerce Department Departmental Organization Order 35-1A (Dec 13, 1982)

Frequency of Data Publication: Quarterly and annual

Timing of Data Release: Quarterly state personal income statistics are released three months after the end of the quarter. Preliminary annual state personal income statistics are released three months after the end of the reference year. Revised and more detailed annual statistics are released six months later.

Modes to Access Data: Interactive Data Application, Regional Data, GDP and Personal Income, Bureau of Economic Analysis (xls, csv, pdf)

Methodology: Personal income is defined as the income received by, or on behalf of, all the residents of an area (nation, state, or county) from all sources. It consists of the income received by persons from participation in production, from government and business in the form of transfers, and from government in the form of interest (which is treated like a transfer receipt). It does not include realized or unrealized capital gains or losses.

Alternatively, personal income can be defined as the sum of wages and salaries, supplements to wages and salaries, proprietors' income, dividends, interest, and rent, and personal current transfer receipts, less contributions for government social insurance. Because the personal income of an area represents the income that is received by, or on behalf of, all the persons who live in that area, and because the estimates of the earnings component of personal income is made on a place-of-work basis, state personal income includes an adjustment for residence. The residence adjustment represents the net flow of compensation (less contributions for government social insurance) of interstate commuters.

---

[43] Text source: *Measuring the Nation's Economy: A Guide to the Bureau of Economic Analysis*, pp. 17-19.

Estimates of state and county personal income should reflect the residence of the income recipients. However, some of the data that are used to estimate some components of personal income are reported by the recipient's place of work rather than by his place of residence. Therefore, these components are estimated on a place-of-work basis, the amounts aggregated, and the aggregate (called the income subject to adjustment) adjusted to a place-of-residence basis. Thus the combination of the components of personal income plus the residence adjustment yields personal income on a place-of-residence basis.

The estimates of wages and salaries, employer contributions for employee pension and insurance funds, and contributions for government social insurance (by employers and employees) are mainly derived from data that are reported by place of work. These data are reported by industry in the state and county in which the employing establishment is located.

The estimates of rental income of persons, personal dividend income, personal interest income, personal current transfer receipts, and contributions for supplementary medical insurance and for veterans' life insurance are derived from data that are reported by the place of residence of the income recipient.

The state personal income and employment estimates are based primarily on administrative-records data. In addition, some survey and census data are used. The administrative records data are a byproduct of the administration of various federal and state government social insurance programs and tax codes. They may originate either from the recipients of the income or from the payer of the income. Some of the more important of these programs and taxes (and the agencies compiling the data) are:

- State unemployment insurance programs (Bureau of Labor Statistics, U.S. Department of Labor)
- State Medicaid programs and the federal Medicare program (Centers for Medicare and Medicaid Services, U.S. Department of Health and Human Services)
- Social Security (Social Security Administration)
- Federal veterans' programs (U.S. Department of Veterans Affairs)
- State and federal income tax codes (Internal Revenue Service, U.S. Department of the Treasury and Bureau of the Census, U.S. Department of Commerce)

The census data are mainly collected from the recipients of the income. The most important sources of census data for the personal income and employment estimates are the Census of Population and Housing, conducted by the Bureau of the Census every ten years and the Census of Agriculture, conducted by the U.S. Department of Agriculture (USDA) every five years.

The survey data are collected from both the payers and the recipients of the income. The more important surveys include the Annual Survey of Public Pensions and the American Community Survey conducted by the Census Bureau and the monthly Current Employment Statistics survey conducted by the Bureau of Labor Statistics. The estimates of farm proprietors' income rely principally on the USDA's estimates of the income of all farms. The USDA uses sample surveys

along with census data and administrative-records data to develop its estimates. The estimates of military compensation and employment rely principally on tabulations of data provided by the U.S. Department of Defense.[44]

Application of ACS Data in Methodology: The national estimates of the cash wages paid to the employees of private households are based on data from the Current Population Survey. The state estimates of cash wages are based on ACS 3-year estimates of private household wages. The ACS wage data are extrapolated to the present by the annual change in household population. The extrapolated series for each year is adjusted proportionately to sum to the national estimate of cash wages.

The national estimates of pay-in-kind received by private household employees are distributed to the states using ACS private household employment data. The employment data are extrapolated to the present by the annual change in household population.

National estimates of imputed net rent from mobile homes for census years through 1990 were allocated to the states in proportion to the number of mobile homes from the Census of Housing. In 2000 it was allocated in proportion to the value of mobile homes from the Census of Housing. Intercensal estimates of imputed net rent were straight line interpolations of the census benchmarks. The 2000 estimates are extrapolated forward using the aggregate value of mobile homes from the American Community Survey.[45]

The ACS offers journey-to-work data that will enable BEA to convert components of state personal income that are recorded on a place of work basis to a place of residence basis. Currently, BEA relies on dated information from the 2000 decennial census for making these adjustments.[46] BEA plans to incorporate the information from the 2005-2009 ACS (5-year file) in its state personal income statistics for 2001-2014, scheduled to be released in September 2015. The ACS data will provide the most up to date information on interstate commuting available in the United States and thereby help improve the accuracy of the BEA's state personal income statistics.[47]

ACS Questions Utilized (as suggested by the methodology): P29-30, P41-47

Major Uses: The state estimates of personal income and its components, per capita personal income, disposable personal income, employment, and GDP by state are widely used by both the public and the private sectors to measure and track economic well-being over time and to make comparisons across states in the level and composition of economic activity and the value added by a state's industries. These estimates provide a framework for the analysis of state economies, and they serve as a basis for decision making.

---

[44] Text source: BEA, "State Personal Income and Employment: Concepts, Data Sources, and Statistical Methods," September 2014.

[45] Text source: Ibid.

[46] Ibid, Chapter VIII: Residence Adjustment.

[47] Text source: Personal communication with Joel Platt, Associate Director for Regional Economics, Bureau of Economic Analysis, February 27, 2015.

Federal agencies use these estimates in econometric models, such as those used to project energy and water use. They also use the estimates as a basis for allocating funds and for determining matching grants. For example, in fiscal year 2012, $320.8 billion in federal funds were distributed on the basis of BEA's state and local personal income and GDP statistics.

In addition, the Census Bureau uses the estimates of state per capita personal income as the key predictor variable in the preparation of state estimates of the mean annual income of four-person families.

State governments use the state estimates of personal income and GDP to measure the economic base of planning areas. They also use the estimates in econometric models for various planning purposes and to project tax revenue and the need for public utilities and services. Currently, 18 states have set constitutional or statutory limits on state government revenue and spending that are tied to state personal income or to one of its components. These states account for more than one-half of the population of the United States. A majority of the states use the quarterly estimates of state personal income to project tax collections.

University schools of business and economics use the estimates for theoretical and applied economic research. Some of these schools distribute the estimates in abstracts or similar reports to various state and local government agencies, regional councils of governments, private research groups, businesses, and libraries. Businesses use the estimates for planning activities, such as evaluating markets for new or established products and determining areas for the location, expansion, and contraction of their activities.[48]

---

[48] Text source: BEA, "State Personal Income and Employment: Concepts, Data Sources, and Statistical Methods," September 2014.

B. Census Bureau

15. *Citizen Voting Age Population*

Purpose of the Measure: The Citizen Voting Age Population (CVAP) is the number of persons in a geographic unit who are citizens and at least 18 years of age. The CVAP Special Tabulation by race/origin for small geographic areas is produced to examine whether voting districts comply with the Voting Rights Act, which prohibits development of voting districts that discriminate against potential voters on the basis of race and/or ethnicity. Data are published for the nation and by state, county, place, minor civil division, tract, and block group.

Responsible Unit: Redistricting Data Office (RDO), Census Bureau for the Civil Rights Division, Department of Justice

Authorization: Section 5 of the Voting Rights Act of 1965. Numerous court cases has indicated the need for CVAP data to determine compliance with Section 5.[49]

Frequency of Data Publication: CVAP was produced on the basis of Census 2000 long form data. With the advent of five-year ACS data in 2010, the RDO publishes an annual update of the CVAP based on the latest five-year ACS data, one each year between 2011 (2005-2009 ACS) and 2014 (2008-2012 ACS).

Timing of Data Release: The CVAP Special Tabulation of the 2008-2012 ACS was released in January, 2014.

Modes to Access Data: Voting Age Population by Citizenship and Race, RDO, Census Bureau (SAS and csv).

Methodology/Application of ACS Data in Methodology: The CVAP is a special tabulation of five-year ACS data. The CVAP uses racial categories different from those published by the ACS, see table on next page.[50]

ACS Questions Utilized (as suggested by the methodology): P4-6, P8

Major Uses: This data series enables the Civil Rights Division of the Justice Department, the courts, and other interested parties to assess of compliance of legislative district design with the Voting Rights Act.

---

[49] Jorge Chapa, Ana Henderson, Aggie Jooyoon Noah, Werner Schinkiv, and Robert Kengle, "Redistricting: Estimating Citizen Voting Age Population," The Chief Justice Earl Warren Institute on Law and Social Policy, Berkeley Law School, University of California, Research Brief, September 2011.
[50] *Ibid.*

TABLE 1	Racial Categories Reported in ACS and DOJ Tabulations	
DOJ Racial Categories	**ACS Racial Categories**
Total	Total
Not Hispanic or Latino	Not Hispanic or Latino
American Indian or Alaska Native Alone (non-Latino)	American Indian or Alaska Native Alone (including Latinos)
Asian Alone (non-Latino)	Asian Alone (including Latinos)
Black or African American Alone (non-Latino)	Black or African American Alone (including Latinos)
Native Hawaiian or Other Pacific Islander Alone (non-Latino)	Native Hawaiian or Other Pacific Islander Alone (including Latinos)
White Alone (non-Latino)	White Alone (including Latinos)
	White Alone (non-Latino)
Not Applicable	Some Other Race (including Latinos)
American Indian or Alaska Native and White (non-Latino)	
Asian and White (non-Latino)	
Black or African American and White (non-Latino)	Two or More Races (including Latinos)
American Indian or Alaska Native and Black or African Am (non-Latino)	
Reminder of Two or More Race Responses (non-Latino)	
Hispanic or Latino	Hispanic or Latino

16. *Current Population Survey – Annual Social and Economic Supplement*

Nature and Purpose of the Measure: The monthly Current Population Survey (CPS) is the primary source of employment and unemployment statistics for the population of the United States. An important component of the CPS program is the Annual Social and Economic Supplement (ASEC), conducted in conjunction with the February, March, and April CPS. Additional questions asked through the ASEC include:

- Household and family characteristics
- Marital status
- Geographic mobility
- Foreign-born population
- Income from the previous calendar year
- Poverty
- Work status/occupation
- Health insurance coverage
- Program participation
- Educational attainment

The ASEC provides supplemental data on work experience, income, noncash benefits, and migration. Comprehensive work experience information is given on the employment status, occupation, and industry of persons 15 years old and over. Additional data for persons 15 years old and older are available concerning weeks worked and hours per week worked, reason not working full time, total income and income components. Data on employment and income refer to the preceding year, although demographic data refer to the time of the survey.

This file also contains data covering nine noncash income sources: food stamps, school lunch program, employer-provided group health insurance plan, employer-provided pension plan, personal health insurance, Medicaid, Medicare, CHAMPUS or military health care, and energy assistance. Characteristics such as age, sex, race, household relationship, and Hispanic origin are shown for each person in the household enumerated.[51]

Responsible Unit: Social, Economic and Housing Statistics Division, Demographic Programs, Census Bureau, with the Bureau of Labor Statistics (BLS)

Authorization: Title 13, United States Code, Section 182, and Title 29, United States Code, Sections 1-9.

Frequency of Data Publication:

The following Census Bureau data products based on the ASEC are produced annually:

- Person Income

---

[51] Text source: Census Bureau, "Current Population Survey, 2014 Annual Social and Economic (ASEC) Supplement," September 2014.

- Poverty
- Family Income
- Household Income
- Health Insurance Coverage

In addition, BLS annually publishes "A Profile of the Working Poor" based on the ASEC.

Timing of Data Release: The Census Bureau releases reports based on the ASEC in September of the calendar year in which the data were collected. The income, poverty, and health insurance data reported are for the prior calendar year.

Modes to Access Data:

- CPS ASEC Reports, Census Bureau website
- CPS Table Creator (customized web tables)
- BLS, "A Profile of the Working Poor" (pdf)

Methodology: The Basic CPS universe is 114 million households. From the universe of 114 million, we select a sample of approximately 60,500 households. With the SCHIP general expansion, this increased to 72,000 households each month. Of these, approximately 60,000 households will be eligible for interview; and we will actually interview approximately 55,000 households.

The ASEC sample expands upon the Basic CPS sample by adding select households. Approximately 6,000 Hispanic households interviewed in the previous November CPS will be added. We expect to interview approximately 4,500 of these households. Additionally, we will add approximately 8,800 minority and White (with children) households that were interviewed in November. We expect to interview about 6,500 of these households. Finally, we will conduct the ASEC to selected minority and White (with children) households during February and April. These households will be "borrowed" from the February outgoing and the April incoming rotation groups. We expect to select approximately 12,800 such households, with about 12,000 actually being interviewed. This brings us to a total of approximately 78,000 households planned for interview in fiscal year 2014.

The CPS sample is a probability sample based on a stratified two-stage sampling scheme: selection of sample primary sampling units (PSUs) and selection of sample housing units within those PSUs. In general, the CPS sample is selected from lists of addresses obtained from the Master Address File (MAF) with updates from the United States Postal Service (USPS) twice a year.

In the first stage of sampling, PSUs are selected. These PSUs consist of counties or groups of contiguous counties in the United States, and are grouped into strata. The CPS is a state-based design. Therefore, all PSUs and strata are defined within state boundaries and the sample is allocated among the states to produce state and national estimates with the required reliability, while keeping total sample size to a minimum.

The variables chosen for grouping PSUs in each state into strata reflect the primary interest of the CPS in maximizing the reliability of estimates of labor force characteristics. Basically, the same set of stratification variables, from the 2010 Decennial Census and the American Community Survey (ACS), are used for each state: unemployment statistics by gender; number of families maintained by a woman; and the proportion of occupied housing units with three or more people. In addition, the number of persons employed in selected industries and the average monthly wage for selected industries are used as stratification variables in some states. The industry-specific data are averages over the period 2000 through 2008 and are obtained from the Quarterly Census of Employment and Wages program of the BLS.

The 2010 sample design comprises three frames: unit, coverage improvement (CI) and group quarters (GQ). Within these sampling frames, housing units are sorted based on characteristics of the ACS and geography. Then, from each frame, a systematic sample of addresses within the sample PSUs is obtained.

Each sample is divided into eight approximately equal panels, called rotation groups. A rotation group is interviewed for four consecutive months, temporarily leaves the sample for eight months, and then returns for four more consecutive months before retiring permanently from the CPS (after a total of eight interviews).[52]

This survey's estimation procedure adjusts weighted sample results to agree with independently derived population estimates of the civilian noninstitutionalized population of the United States and each state (including the District of Columbia). These population estimates, used as controls for the CPS, are prepared monthly to agree with the most current set of population estimates that are released as part of the Census Bureau's population estimates and projections program.

The population controls for the nation are distributed by demographic characteristics in two ways:

- Age, sex, and race (White alone, Black alone, and all other groups combined).
- Age, sex, and Hispanic origin.

The population controls for the states are distributed by race (Black alone and all other race groups combined), age (0-15, 16-44, and 45 and over), and sex. The independent estimates by age, sex, race, and Hispanic origin, and for states by selected age groups and broad race categories, are developed using the basic demographic accounting formula whereby the population from the 2010 Decennial Census data is updated using data on the components of population change (births, deaths, and net international migration) with net internal migration as an additional component in the state population estimates.[53]

---

[52] Text source: Census Bureau, "Supporting Statement, Annual Social and Economic Supplement to the Current Population Survey (OMB Control Number 0607-0354)," Part B, January 31, 2014.

[53] Text source: Census Bureau, "Current Population Survey, 2014 Annual Social and Economic (ASEC) Supplement," September 2014.

Application of ACS Data in Methodology: As indicated above, the ACS is used in three ways to stratify the survey sample:

- Primary sampling units (PSUs) are designed in part on the basis of select ACS geographic characteristics—unemployment statistics by gender, female householder, and the number of people in a housing unit.
- The ACS provides the housing unit sampling frame within those PSUs.
- Sample housing units are selected from that frame on the basis of specific ACS characteristic microdata.[54]

In addition, ACS data on international migration are used to develop the population estimates that serve as the population controls for the ASEC estimation procedure. (See the chapter of this report on Population Estimates for further information on methodology.)

ACS Questions Utilized (as suggested by the methodology): number of people, P3-6, P15, P29, P35-39

Major Uses: The analysis of the ASEC greatly informs policy, business, and research:

- The work experience items provide a unique measure of the dynamic nature of the labor force as viewed over a 1-year period. These items produce statistics that show movements in and out of the labor force by measuring the number of periods of unemployment experienced by people, the number of different employers worked for during the year, the principal reasons for unemployment, and part-/full-time attachment to the labor force. We can make indirect measurements of discouraged workers and others with a casual attachment to the labor market.

- The income data from the ASEC are used by social planners, economists, government officials, and market researchers to gauge the economic well-being of the country as a whole and of selected population groups of interest.
  - Government planners and researchers use these data to monitor and evaluate the effectiveness of various assistance programs.
  - Market researchers use these data to identify and isolate potential customers.
  - Social planners use these data to forecast economic conditions and to identify special groups that seem to be especially sensitive to economic fluctuations.
  - Economists use ASEC data to determine the effects of various economic forces, such as inflation, recession, recovery, and so on, and their differential effects on various population groups.
  - A prime statistic of interest is the classification of people in poverty and how this measurement has changed over time for various groups. Researchers evaluate ASEC income data not only to determine poverty levels but also to determine whether government programs are reaching eligible households.

---

[54] Census Bureau, "Supporting Statement, Annual Social and Economic Supplement to the Current Population Survey (OMB Control Number 0607-0354)," Part B, January 31, 2014.

- The following publications are released annually based on the ASEC data:

  o Income, Poverty, and Health Insurance Coverage in the United States
  o Alternative Income Estimates in the United States
  o America's Families and Living Arrangements
  o Geographic Mobility
  o Educational Attainment (since 1947).
  o In addition to the Census Bureau publications listed above, BLS has published yearly results from the ASEC supplement (educational attainment, marital history, and work experience) in news releases, articles in the Monthly Labor Review, and special bulletins.

- The questions (Q80 and Q83) concerning school lunches identify households with children who participated in the Federal School Lunch Program during the previous year. These data allow for more effective analyses of the program and, along with the food stamp data, show where the client groups of the two food nutrition programs overlap.

- The questions (Q85 and Q86) on public housing identify households that are owned by a housing authority or other public agency and households where a government agency is paying for part of the rent. These data allow for more effective analysis of public housing programs and measure the extent of participation among eligible families. The question (SPHS8) determines the specific source of housing assistance.

- The questions (Q87 through Q94) on food stamps identify which household members received food stamps during the previous year. These data, along with the detailed supplemental income data, allow for a more comprehensive evaluation of the food stamp program and measure the extent of participation among eligible families.

- Items Q93-Q94 provide a measurement of the number of households that have received government heating or cooling assistance in the previous calendar year, and the amount of such assistance. State agencies provide the only other available data relevant to the energy assistance program in the form of summary counts of households receiving this type of assistance, cross-classified by their annual income levels. The ASEC is the only current data set capable of cross-tabulating fuel assistance recipients/nonrecipients with their socioeconomic characteristics. This allows analysts to determine whether funds are reaching eligible households.[55]

---

[55] *Ibid.*

17. *Current Population Survey – Basic Demographics and Supplements*

Nature and Purpose of the Dataset: The Current Population Survey (CPS) is a monthly survey designed primarily to produce national and state estimates of labor force characteristics of the civilian noninstitutional population (CNP) 16 years of age and older. Conducted by the Bureau of Census for the Bureau of Labor Statistics (BLS), the CPS provides a comprehensive body of data on the labor force, employment, unemployment, persons not in the labor force, hours of work, earnings, and other demographic and labor force characteristics.[56]

The demographic information collected in the CPS provides a unique set of data on selected characteristics for the civilian noninstitutional population. Demographic information collected include age, marital status, gender, Armed Forces status, education, race, origin, and family income.[57]

The CPS serves as the foundation for a series of supplemental data collections on the U.S. population and specific population subsets. Supplements carried out for the Census Bureau's own purposes include:

- Housing vacancy
- Fertility (with BLS)
- School enrollment (with BLS and the National Center for Education Statistics)
- Voting and registration
- International migration

Other federal agencies sponsor CPS supplements on the following topics:

- Child support (Office of Child Support Enforcement)
- Unbanked and underbanked (Federal Deposit Insurance Corporation)
- Public participation in the arts (National Endowment for the Arts)
- Tobacco use (National Cancer Institute and Food and Drug Administration)
- Computer and internet use (National Telecommunications and Information Administration)
- Food security (Food and Nutrition Service)[58]

BLS also funds a number of CPS supplements, which are described elsewhere in this report.

Responsible Unit: Social, Economic and Housing Statistics Division, Demographic Programs, Census Bureau, with the Bureau of Labor Statistics (BLS)

Authorization: Title 13, United States Code, Section 182, and Title 29, United States Code, Sections 1-9

---

[56] Text source: "Labor Force Statistics from the Current Population Survey," BLS website.

[57] Text source: Census Bureau, "Supporting Statement, Current Population Survey (CPS) Basic Demographics (OMB Control Number 0607-0049)," Part A, April 2, 2014.

[58] "Current Population Survey: Supplemental Surveys," Census Bureau website.

- Monthly – housing vacancy
- Annual – supplements on fertility (June) educational enrollment (October), food security (December), computer and internet use, and participation in the arts
- Biennial – supplements on child support, voting and registration, the unbanked and underbanked,
- Every three-four years – supplement on tobacco use
- One time (so far) – supplement on international migration (2008)

Timing of Data Release: Varies by study

Modes to Access Data: The Census Bureau publishes CPS-based studies as part of its series of *Current Population Reports.*

- [Population Characteristics (P20)](#) – subjects include America's Families and Living Arrangements, Asian and Pacific Islander Population, Black Population, Educational Attainment, Fertility of American Women Foreign-Born Population, Geographical Mobility, Hispanic Population, Household and Family Characteristics Marital Status and Living Arrangements, Older Population, School Enrollment, and Voting and Registration.
- [Special Studies (P23)](#)
- [Consumer Income Reports (P60)](#)

[Basic monthly CPS data](#) are available for download (ftp).

Methodology: The universe for the CPS is 114 million households. From this universe, the Census Bureau selects a sample of approximately 72,000 households each month. Of these, approximately 59,000 households will be eligible for interview.

The CPS sample is a probability sample based on a stratified two-stage sampling scheme: selection of sample primary sampling units (PSUs) and selection of sample housing units within those PSUs. In general, the CPS sample is selected from lists of addresses obtained from the Master Address File (MAF) with updates from the United States Postal Service (USPS) twice a year.

In the first stage of sampling, PSUs are selected. These PSUs consist of counties or groups of contiguous counties in the United States, and are grouped into strata. The CPS is a state-based design. Therefore, all PSUs and strata are defined within state boundaries and the sample is allocated among the states to produce state and national estimates with the required reliability, while keeping total sample size to a minimum.

The variables chosen for grouping PSUs in each state into strata reflect the primary interest of the CPS in maximizing the reliability of estimates of labor force characteristics. Basically, the same set of stratification variables, from the 2010 Decennial Census and the American Community Survey (ACS), are used for each state: unemployment statistics by gender; number

of families maintained by a woman; and the proportion of occupied housing units with three or more people. In addition, the number of persons employed in selected industries and the average monthly wage for selected industries are used as stratification variables in some states. The industry-specific data are averages over the period 2000 through 2008 and are obtained from the Quarterly Census of Employment and Wages program of the BLS.

The 2010 sample design comprises three frames: unit, coverage improvement (CI) and group quarters (GQ). Within these sampling frames, housing units are sorted based on characteristics of the ACS and geography. Then, from each frame, a systematic sample of addresses within the sample PSUs is obtained.

Each sample is divided into eight approximately equal panels, called rotation groups. A rotation group is interviewed for four consecutive months, temporarily leaves the sample for eight months, and then returns for four more consecutive months before retiring permanently from the CPS (after a total of eight interviews).[59]

This survey's estimation procedure adjusts weighted sample results to agree with independently derived population estimates of the civilian noninstitutionalized population of the United States and each state (including the District of Columbia). These population estimates, used as controls for the CPS, are prepared monthly to agree with the most current set of population estimates that are released as part of the Census Bureau's population estimates and projections program.

The population controls for the nation are distributed by demographic characteristics in two ways:

- Age, sex, and race (White alone, Black alone, and all other groups combined)
- Age, sex, and Hispanic origin

The population controls for the states are distributed by race (Black alone and all other race groups combined), age (0-15, 16-44, and 45 and over), and sex. The independent estimates by age, sex, race, and Hispanic origin, and for states by selected age groups and broad race categories, are developed using the basic demographic accounting formula whereby the population from the 2010 Decennial Census data is updated using data on the components of population change (births, deaths, and net international migration) with net internal migration as an additional component in the state population estimates.[60]

Application of ACS Data in Methodology: As indicated above, the ACS is used in three ways to stratify the survey sample:

---

[59] Text source: Census Bureau, "Supporting Statement, Current Population Survey Basic Demographics (OMB Control Number 0607-0049)," Part B, April 2, 2014.

[60] Text source: Census Bureau, "Current Population Survey, December 2013, Food Security File," Technical Documentation, CPS-13, December 2013.

- Primary sampling units (PSUs) are designed in part on the basis of select ACS geographic characteristics—unemployment statistics by gender, female householder, and the number of people in a housing unit.
- The ACS provides the housing unit sampling frame within those PSUs.
- Sample housing units are selected from that frame on the basis of specific ACS characteristic microdata.[61]

ACS Questions Utilized (as suggested by the methodology): number of people, P3-6, P15, P29, P35-39

While it is possible that additional ACS questions could be used in sample stratification and estimation procedures for a particular topical supplement, the report authors have not found any examples so far.

Major Uses of the Dataset: We use the CPS data on household size and composition, age, education, ethnicity, and marital status to compile monthly averages or other aggregates for national and sub-national estimates. We use these data in four principal ways: in association with other data, such as monthly labor force or periodic supplement publications; for internal analytic research; for evaluation of other surveys and survey results; and as a general purpose sample and survey.

The demographic data are central to the publication of all labor force data in the BLS' monthly report *Employment and Earnings*. The data set that results from combining the monthly labor force data with the demographic data provides analysts with the ability to understand labor force patterns of many subpopulation groups. This is particularly important since the federal government often directs initiatives at special groups that historically have not conformed to general labor force participation patterns.

Analysts also use the demographic data in association with all supplement publications. For example, publications that use these data are *Fertility of American Women, School Enrollment-- Social and Economic Characteristics of Students and Educational Attainment in the United States* (Series P-20). Comparably, researchers are able to characterize the population within the subject area of the many supplements conducted in conjunction with the CPS. For instance, the Annual Social and Economic Supplement identifies which subpopulation groups, as established by the demographic variables, experience the highest incidence of poverty. While we collect and support independently the demographic variables, the labor force data, and the supplement inquiries, their use as a combined data set enhances the utility of each.

The Census Bureau also uses the demographic data extensively for internal analytic work. For example, we use these data to develop estimates of family and household types and metropolitan and nonmetropolitan populations. We use these estimates to identify population trends between decennial censuses and to analyze the growth and distribution of various racial and ethnic groups. We may then use this information in preparing reports on these subjects or in determining the accuracy of population controls used throughout the Census Bureau. As is

---

[61] Census Bureau, "Overview of CPS Sample Design and Methodology," April 2, 2014.

noted below, we use the demographic data to improve our postcensal population estimates (that is, the components of emigration and undocumented immigration).

Also, we use the CPS as a source for other survey samples. A household remains in the CPS sample for 16 months. Other surveys conducted by the Census Bureau may use a CPS sample when it is no longer part of the CPS. In 2006, the National Survey of Fishing, Hunting, and Wildlife-Associated Recreation, sponsored by the Department of the Interior, used retired cases from the CPS sample. The ongoing American Time Use Survey, sponsored by the BLS uses expired CPS sample. By using the CPS demographics to select their samples, other surveys have been able to avoid screening samples and to obtain accurate estimates by demographics.

Another use of the demographic data is in evaluating other survey results. For example, analysts control the results of the National American Housing Survey to the CPS monthly averages of households. Similarly, in order to determine the plausibility of the results of the Survey of Income and Program Participation (SIPP), analysts continuously compare the data on household and family composition from the SIPP to the CPS monthly household and family composition data.

The Census Bureau often uses the CPS as a model and resource for improving the efficiency and quality of other surveys. For example, the Census Bureau designed some items for the SIPP from the CPS. Academicians and researchers have historically used the CPS to better understand the many complexities associated with sample surveys and household interviews in general.

In addition to the collection of demographic and labor force data, the CPS is also a major vehicle for the collection of supplemental questions on various socio-economic topics. In most months of the year we ask supplemental questions after asking the basic labor force questions of all eligible people in a household, thereby maximizing the utility of the CPS sample. We also collect annual data on work experience, income, migration (Annual Social and Economic Supplement), and school enrollment of the population (October supplement). In addition we collect biennial, but separately funded, data on fertility and birth expectations of women of child-bearing age (June), voting and registration (November) and child support and alimony. The BLS, the Census Bureau, other government agencies, and private groups sponsor the supplements.[62]

---

[62] Text source: Census Bureau, "Supporting Statement, Current Population Survey (CPS) Basic Demographics (OMB Control Number 0607-0049)," Part A, April 2, 2014.

18. *Investing in Manufacturing Communities Data Tool*

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

19. Health Insurance Coverage in the United States

Nature and Purpose of the Report: This report presents statistics on health insurance coverage in the United States.

Responsible Unit: Social, Economic, and Housing Statistics Division, Census Bureau

Authorization: Title 13, U.S. Code, Section 182, and Title 29, U.S. Code, Sections 1-9. Title 42, U.S. Code, Section 1397ii requires the ACS to collect data on health insurance coverage for children.

Frequency of Data Publication: Annual. Up through coverage year 2003, the annual health insurance coverage report was issued separately from the annual income and poverty report. For 2004 to 2012, the two reports were combined into one. For 2013, and presumably going forward, the health insurance coverage report is stand-alone.

Timing of Data Release: Simultaneous with the release of the ACS one-year data. Data for calendar year 2013 were released in September 2014.

Modes of Data Access: Health Insurance Main, Census Bureau (pdf, xls)

Methodology (for the 2013 report):

- The majority of the data in this report are from the 2014 Current Population Survey Annual Social and Economic Supplement (CPS ASEC) and were collected in the 50 states and the District of Columbia.
- The estimates in this report are controlled to independent national population estimates by age, sex, race, and Hispanic origin for March 2014; these population controls are based on the 2010 Census.
- Additional data in this report are from the American Community Survey (ACS) 1-year data, 2008 through 2013.
- Researchers have questioned the validity of the health insurance estimates in the previous version of the CPS ASEC. In particular, the estimate of the uninsured in the previous calendar year was consistently higher than in other federal surveys, indicating that the CPS ASEC did not capture as much health insurance coverage in comparison. Additionally, these concerns extended to the Medicaid undercount and general misreporting of the source and timing of health insurance coverage. To address these concerns, the Census Bureau implemented redesigned health insurance questions in the CPS ASEC.
- Due to the redesign of the health insurance section of the CPS ASEC, its estimates of health insurance coverage are not directly comparable to estimates from prior years of the survey. Consequently, this report presents an overall estimate of the year-to-

year change in health insurance coverage rates based on the ACS, which has collected data on health insurance since 2008.[63]

Application of ACS in the Methodology: For the 2013 report, the following data were based on ACS data:

- Health Insurance Coverage Status and Type of Coverage by State and Age for All People: 2013
- Health Insurance Coverage Status by State for All People: 2013
- Number and Percent of Children Under 19 Below 200% of Poverty by Health Insurance Coverage and State: 2013
- People Without Health Insurance Coverage by Race and Hispanic Origin: 2012-2013
- Number and Percentage of People Without Health Insurance Coverage by State: 2012-2013
- Comparison of Uninsured Rates Between States: 2013
- Low Income Uninsured Children by State: 2013
- Low Income Children by Type of Health Insurance by State: 2013

ACS Questions Utilized (as suggested by the methodology): P3-6, P16, P47-48,

Major Uses: The report guides health insurance policy research and analysis.

---

[63] Text source: Jessica C. Smith and Carla Medalia, "Health Insurance Coverage in the United States: 2013," U.S. Census Bureau, Current Population Reports, P60-250, September 2014.

20. Job-to-Job Flows

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

21. *On the Map*

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

22. On the Map for Emergency Management

Purpose of the Data Series: OnTheMap (OTM) for Emergency Management is a public data tool that provides an intuitive web-based interface for accessing U.S. population and workforce statistics, in real time, for areas being affected by natural disasters. Users can easily retrieve reports containing detailed workforce, population, and housing characteristics for hurricanes, floods, wildfires, winter storms, and federal disaster declaration areas.

To provide users this information for rapidly changing hazard event areas, OTM for Emergency Management automatically incorporates real time data updates from the National Weather Service's (NWS) National Hurricane Center, Department of Interior (DOI), Department of Agriculture (DOA), and the Federal Emergency Management Agency (FEMA)[64]

Responsible Unit: Center for Economic Studies, Census Bureau, Department of Commerce

Authorization: Title 13, Section 4 and Title 15, Section 1516

Frequency of Data Publication: A new version of OTM for Emergency Management is released each June. Emergency management event data are updated every six hours.

Timing of Data Release: OTM for Emergency Management 4.0, released in June 2014, includes the 2008-2012 American Community Survey estimates.

Modes to Access Data: OnTheMap for Emergency Management (mapping tool)

Methodology/ Application of ACS Data in Methodology: OTM for Emergency Management automatically incorporates real time data updates every six hours from the National Weather Service's (NWS) National Hurricane Center, Department of Interior (DOI), Department of Agriculture (DOA), and the Federal Emergency Management Agency (FEMA).

- The hurricane data are supplied by feeds from the National Hurricane Center (NHC) and include several detailed GIS formatted representations of the storm area including: the current wind radii as well as the historical wind area (cumulative wind swath) and forecast area (uncertainty cone).
  - Current Wind Radii - Represents the forecasters' best estimates of the current location, intensity, and size of a tropical cyclone while the cyclone is ongoing.
  - Forecast Area (Cone of Uncertainty) - Represents the probable track of the center of a tropical cyclone.
  - Wind History (Cumulative Wind Swath) - Area representing how the size of the storm has changed and the areas potentially affected so far.
- The flood data include Significant River Flood Outlook areas that are supplied daily by the Hydrometeorological Prediction Center (HPC) at the National Weather Service. The data, provided in GIS format, include representations for areas of active (Occurring or Imminent) or likely flooding in the forecast period.

---

[64] Text source: "OnTheMap for Emergency Management: Overview," LEHD webpage.

- The Snow and Freezing Rain probability forecast data are supplied in GIS format daily by the Hydrometeorological Prediction Center (HPC) at the National Weather Service. The snowfall data displayed in OnTheMap for Emergency Management represent areas with a moderate probability of receiving more than 12 inches of snow or a high probability of receiving more than 4 inches within the 24-hour forecast period. The freezing rain data displayed represent areas with a moderate or high probability of receiving more than .25 inches of freezing rain within the 24-hour forecast period.
- The wildfire data include active wild fire perimeters provided by the Geospatial Multi-Agency Coordination Group (GeoMAC), a group composed of technical and subject matter experts from the Department of Interior and Department of Agriculture. The data are updated regularly and provided in GIS format for public use.
- The disaster declarations data are defined and provided by the Federal Emergency Management Agency (FEMA). The declaration areas displayed include areas eligible for Public Assistance and Individual Assistance. The data are updated as new declarations are made and provided in GIS format by means of RSS feeds.

Recent improvements have been made that advance the utility of the tool and its data offerings for users including newly added social, economic, and housing data from the American Community Survey (ACS), greater reporting flexibility to better analyze communities affected by disaster events, and a variety of user interface enhancements. [65]

OTM for Emergency Management pre-tabulates data on population and jobs for disaster and emergency event areas based on boundaries provided by other Federal agencies. These boundaries typically do not conform to standard Census Bureau publication geographies. In these cases, the application must construct a geographic approximation of the event boundaries that can be used to tabulate statistical data from one of several data sources.

OTM for Emergency Management provides tabulations of the following data sources: LEHD Origin-Destination Employment Statistics (LODES), American Community Survey (ACS) 5-year estimates, and the 2010 Decennial census. Data from LODES and the Decennial census are published at the census block level. Data from ACS are published at a variety of levels (depending upon the specific table and variable interactions), the smallest of which is the blockgroup.[66]

- The base maps displayed in OTM for Emergency Management are provided by Google through the Google Maps Application Programming Interface (API).
- The workforce related data made available in OTM for Emergency Management are provided by the LEHD Origin-Destination Employment Statistics (LODES) database. These data comprise a variety of statistics describing the U.S. workforce along with information on residential and workplace locations. These data may be useful for

---

[65] Text source: "OnTheMap for Emergency Management: Technical Overview," LEHD webpage.
[66] Text source: "OnTheMap for Emergency Management: Geographic Selection Methodology," LEHD webpage.

determining the number and location of workers living and working within affected areas or for identifying potential impact to various demographic groups and sectors of the economy.

- o Data include: Workers by Age, Workers by Educational Attainment, Workers by Earnings, Workers by NAICS Industry Sector, Workers by Race, Workers by Ethnicity, Workers by Sex, Workers by location.
- o The LODES data are a product of the Local Employment Dynamics (LED) partnership, a partnership between States and the Census Bureau that integrates existing data on workers and employers from administrative records with demographic, business, and survey data.
- The 2010 Census data are based on the Decennial Summary File 1 (SF 1) which contains data compiled from 2010 Census questions asked of all people and about every housing unit. The data provide a basic set of standard population and housing statistics that may be useful for determining the number and location of people living within affected areas or for identifying potential impact to various demographic groups and sectors of the population.
  - o Available variables include Population by Age, Population by Race, Population by Ethnicity, Population by Sex, Population by Housing, Housing Units.
  - o The data items utilized in OTM for Emergency Management are provided nationally at the U.S. Census block level and have been integrated directly into the application.
- The ACS data are based on the 2008-2012 5-year estimates, and include a selected set of demographic, social, economic, and housing measures that may be useful for determining the number, location, and potential impact on various segments of the population living within affected areas. These data provide a more recent and expanded set of characteristics than those of the 2010 Census. Data include:
  - o Household Type
  - o Race
  - o Hispanic or Latino Origin
  - o Ability to Speak English by Age
  - o Ability to Speak English by Language Spoken at Home
  - o Disability Status by Poverty Status
  - o Disability Status by Age
  - o Ratio of Income to Poverty Level in the Past 12 Months
  - o Poverty Status in the Past 12 Months
  - o Population 65 Years and Over Living Alone
  - o Households with One or More People 65 Years and Over
  - o Earnings in the Past 12 Months for Households
  - o Social Security Income in the Past 12 Months for Households
  - o Supplemental Security Income in the Past 12 Months for Households
  - o Supplemental Security Income (SSI) in the Past 1 Months for Households
  - o Public Assistance Income in the Past 12 Months for Households
  - o Retirement Income in the Past 12 Months for Households
  - o Households Receiving Food Stamps/SNAP in the Past 12 Months

- o Year Structure Built
- o Aggregate Value of Owner Occupied Housing
- o House Heating Fuel [67]

The diagram below provides a technical overview of the OTM for Emergency Management methodology.[68]
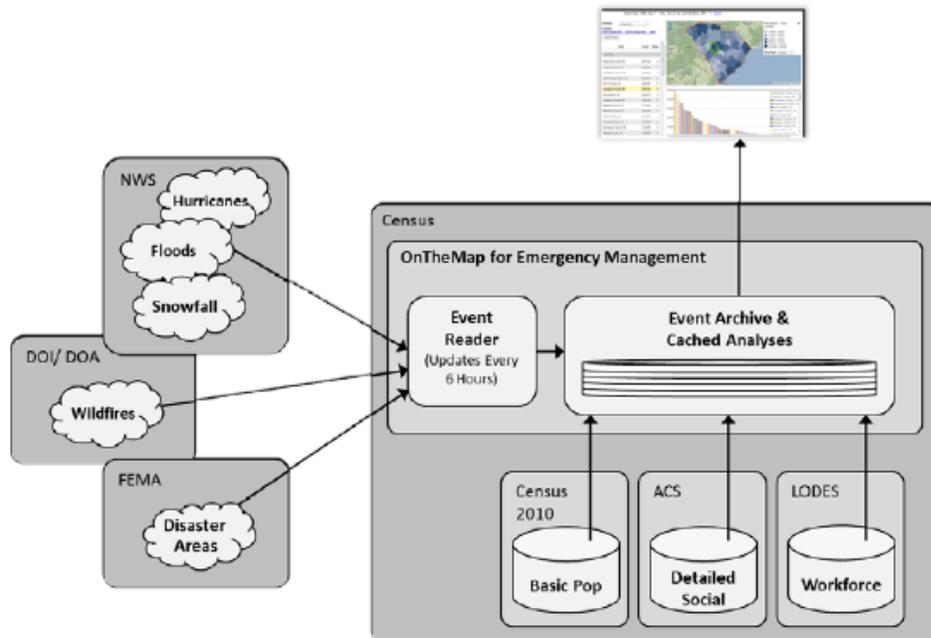
<u>ACS Questions Utilized (as suggested by the methodology)</u>: P2, P4-6, P14, P17-19, P47-48, H2, H13, H15, H19

<u>Major Uses</u>: OTM for Emergency Management is used throughout the country by federal, state, and local officials to more effectively respond to emergencies. FEMA encourages use of OTM for Emergency Management in its support, guidance, and training materials, such as through the Emergency Management Institute. Other national organizations that encourage the use of OTM for Emergency Services include the Red Cross, National Association of Counties, the National Association for Public Safety GIS Foundation, and NOAA's Office of Coastal Management.

---

[67] Text source: "OnTheMap for Emergency Management: Data Sources," LEHD webpage.
[68] Text source: "OnTheMap for Emergency Management: Technical Overview," LEHD webpage.

23. Population Estimates

Purpose of the Data Series: Population estimates are estimates of the resident population of the United States, its states, counties, cities, and towns, as well as for the Commonwealth of Puerto Rico and its municipios. The resident population includes all people currently residing in the United States.

Demographic components of population change (births, deaths, and migration) are produced at the national, state, and county levels of geography. Demographic characteristics of the resident population are estimated by single year of age, sex, Hispanic origin, and race. These estimates are also categorized by household, civilian, civilian non-institutional, and resident plus armed forces overseas.

Responsible Unit: Population Division, U.S. Census Bureau

Authorization: Title 13, Section 181(a)[69]

Frequency of Data Publication: Annual

Timing of Data Release: Population totals and components of change at a July 1 reference point are released for the nation in December of that year, for states in January of the next year, and for counties and metropolitan areas in March of the next year. Population totals for cities and towns are released in May of the next year. Data by age, sex, race, and Hispanic origin for all geographies are released in June of the next year.[70]

Modes to Access Data:

- American FactFinder (xls, csv)
- Population Estimates, Census Bureau website (xls, csv)

Methodology: The Census Bureau estimates the resident population for each year since the most recent decennial census by using measures of population change. The population estimate at any given time point starts with a population base (the last decennial census or the previous point in the time series), adds births, subtracts deaths, and adds net domestic migration and net international migration.[71] From this demographic balancing equation, a cohort component method is derived to produce estimates.

---

[69] "During the intervals between each census of population required under section 141 of this title, the Secretary, to the extent feasible, shall annually produce and publish for each State, county, and local unit of general purpose government which has a population of fifty thousand or more, current data on total population and population characteristics and, to the extent feasible, shall biennially produce and publish for other local units of general purpose government current data on total population. Such data shall be produced and published for each State, county, and other local unit of general purpose government for which data is compiled in the most recent census of population taken under section 141 of this title. Such data may be produced by means of sampling or other methods, which the Secretary determines will produce current, comprehensive, and reliable data."

[70] "Population Estimates," Census Bureau website.

[71] Text source: "Methodology for the United States Population Estimates: Vintage 2013," Census Bureau website.

These estimates are computed using data from a variety of federal and state-level sources. Data on births and deaths, or vital statistics, are from the National Center for Health Statistics and the Federal-State Cooperative for Population Estimates. Calculations of net domestic migration are based on data from the Internal Revenue Service tax exemptions, change in Medicare enrollment, and change in the group quarters population.[72] Net international migration is estimated in several parts: immigration of the foreign born, emigration of the foreign born, net migration between the United States and Puerto Rico, net migration of natives to and from the United States, and net movement of the Armed Forces population to and from the United States.[73] Estimates on net international migration rely on data from the American Community Survey, while estimates of net overseas movement of the Armed Forces population rely on data collected by Defense Manpower Data Center.

A key principle of the estimates is that all of the estimates must be consistent across geography and demographic characteristics, including age, sex, race and Hispanic origin. For example, the sum of the county total populations must equal the total national population, and the sum of a particular race group within a state's counties must equal the total of that particular race group in the state.[74]

Application of ACS Data in Methodology: ACS data are used to estimate net international migration. Immigration of the foreign-born population for Mexico and "All other countries" at the national level relies on the ACS question on residence one year ago. Information from the year of entry question from the ACS is used to create proxy universes to distribute the national-level total foreign-born immigration estimate by geographic and demographic detail.[75]

ACS Questions Utilized (as suggested by the methodology): P3-10, P13, P15, P24

Major Uses: The population estimates are used for federal funding allocations, as controls for major surveys including the Current Population Survey, for community development, to aid business, as indicators of recent demographic changes, and as denominators for statistical rates.

Among the Federal programs that use these intercensal population estimates to allocate funds are the Department of Health and Human Services' Medical Assistance Program (Medicaid), Social Service Block Grants/Entitlement Grants Program, and State Children's Health Insurance Program; the Department of Housing and Urban Development's Community Development Block Grant Program; and the Department of Labor's Employment and Training Administration – Workforce Investment Act Dislocated Workers Program. Hundreds of billions of dollars in

---

[72] The group quarters (GQ) population is also estimated every year, beginning with an estimates base derived from the previous decennial census, then applying updated data on GQ population change. The updated data comes from branches of the military, the Department of Veterans Affairs, and state agencies in the Federal-State Cooperative for Population Estimates.

[73] Text source: "International Migration Methodology," Census Bureau website.

[74] Text source: For more information, see the methodology section of the Population Estimates program on the Census Bureau's website.

[75] Ibid.

Federal funds are distributed to states and other areas based in some part on these estimates. The Congressional Budget Office uses the estimates throughout their reports and analysis.

These estimates of the geographic distribution of the population are also used for making decisions about state and local government services, planning utility services, redefining metropolitan areas, and locating retail outlets and manufacturing establishments. Federal time-series that are produced on a per capita basis, such as per capita income, births per capita, and cancer incidence rates per capita, rely on these estimates for their denominators. The intercensal population estimates are used as the controls for all household surveys conducted by the Census Bureau. These surveys have a major impact on the accuracy of the country's key measures such as unemployment, inflation, income, poverty, and health insurance coverage. Intercensal estimates are essential to controlling variability in the surveys that provide these measures.[76]

---

[76] Text source: Census Bureau, "US Census Bureau's Budget Estimates As Presented to Congress: Fiscal Year 2015," March 2014.

24. *Population Projections*

Purpose of the Data Series: Population projections are estimates of the national population for future dates. Whereas population estimates are for the past and present, population projections are based on assumptions about future demographic trends. Estimates generally use existing data collected from various sources, while projections make assumptions about what demographic trends will be in the future.

Responsible Unit: Population Division, U.S. Census Bureau, Department of Commerce

Authorization: Title 13, U.S. Code

Frequency of Data Publication: Every four years, the Census Bureau releases four sets of population projections, based on four differing assumptions regarding net international migration (low, middle, high, and constant). The most recent projections cover July 1, 2014 to July 1, 2060

Timing of Data Release: The 2014-2016 projections were released in December 2014.

Modes to Access Data: Population Projections, U.S. Census Bureau website (xls, csv)

Methodology: The population projections are produced using a cohort-component method beginning with an estimated base population. In this method, the components of population change (fertility, mortality, and net migration) are projected separately for each birth cohort (persons born in a given year). The base population is advanced each year by using projected survival rates and net international migration. Each year, a new birth cohort is added to the population by applying the projected fertility rates to the female population. The assumptions for the components of change were based on time series analysis of historical trends.

Age-specific fertility rates are calculated from birth registration data compiled by the National Center for Health Statistics (NCHS). Mortality rates are calculated from NCHS-compiled death registration data. Projections of net international migration consist of three components: foreign-born immigration; foreign-born emigration; and net native-born migration.[77]

Application of ACS Data in Methodology: Estimates of *foreign-born immigration* were developed using data from the 1990 and 2000 Censuses and the 2000-2010 single-year American Community Survey (ACS) data files. Using single-year ACS data, foreign-born immigration was measured as the foreign-born population who reported their *year of entry* to the United States as one year prior to the survey year. For example, if a foreign-born respondent in the 2009 ACS reported their year of entry as 2008, then that person would be counted in the 2008 estimate of foreign-born immigration.

The foreign-born immigration estimates and sending country population estimates and projections were categorized into four *country-of-birth* groupings:

---

[77] See the methodology section of the Population Estimates program of the Census Bureau website.

1. Europe, Central Asia, and the Middle East,
2. Asia and Pacific Islands,
3. Non-Spanish Caribbean and Sub-Saharan Africa, and
4. Spanish Caribbean and Latin America.

These groupings were devised to place migrants into categories that correspond to the *race* and *Hispanic origin* groups for which the population projections are produced. The race and Hispanic origin distributions that were used to determine the race and Hispanic origin categorization of the foreign-born immigrants in each of the four country-of-birth groupings were derived from 2006-2010 ACS 5-year estimates.

The foreign-born immigration projections were distributed by age, sex, race, and Hispanic origin detail based on the distributions of characteristics of immigrants within each of the four country of birth groupings from the 2006-2010 ACS. These distributions were held constant in all years of the projections.

Emigration of the foreign-born population from the United States was projected by first estimating a set of emigration rates and then applying those rates to the foreign-born population.

Foreign-born emigration rates were estimated using a residual method. These rates were held constant for all projected years. The rates were produced and applied by arrival cohort, age, sex, and Hispanic origin. Rates were produced for *three arrival cohorts*: (1) immigrants who arrived in the past 0-9 years, (2) immigrants who arrived in the past 10-19 years, and (3) immigrants who arrived 20 or more years ago.

The residual rates were estimated using Census 2000 as the base population and the 2010 ACS as the target population. A residual estimate was calculated by adding half of the annual immigrants to the initial population, surviving that population forward to the next year, and then adding the other half of the immigrants and half of the immigrants for the next period. This process was repeated until the target population of July 1, 2010 was reached. The result was the expected population, from which the target population provided by the 2010 ACS is subtracted to provide a residual estimate of emigration. This estimate of foreign-born emigration was converted into a rate by dividing the annual estimate by the number of person years lived during the period. The rates were smoothed using penalized least squares. For the ages where the rates become negative, they are modeled using mathematical curves between the non-negative points.

A preliminary level of foreign-born emigration was projected by applying the emigration rates to the foreign-born population. The same set of rates, by arrival cohort, age, sex, and Hispanic origin, were used for all projected years. For example, to estimate foreign-born emigration between 2010 and 2011, the emigration rates were applied to the foreign-born population from the 2010 ACS.[78]

---

[78] Text source: *Ibid.*

ACS Questions Utilized (as suggested by the methodology): P3-9, P15

Major Uses: Projections illustrate possible courses of population change and thus aid policy planning and development.

25. *Quarterly Workforce Indicators*

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

26. *Small Area Estimates of Disability*

Purpose of the Data Series: Provide reliable estimates of disability prevalence by state. (experimental measure)

Responsible Unit: U.S. Census Bureau

Authorization: Title 13, United States Code, Section 182

Frequency of Data Publication: Experimental

Timing of Data Release: NA

Modes to Access Data: NA

Methodology/Application of ACS Data in Methodology: Combine data from the six ACS disability questions (hearing, vision, cognitive, ambulatory, self-care and independent living difficulties) with data from the far more detailed questions of the wave 6 disability module of the 2008 SIPP to produce state estimates of disability consistent with the SIPP questions.[79]

ACS Questions Used (as suggested by the methodology): P3-6, P8, P11, P17-19, P26-29, P35-38

Major Uses: The measure is not available at present. If available, it would be useful in guiding the provision of state and local social services.

---

[79] Jerry J. Maples and Matthew Brault, "Improving Small Area Estimates of Disability: Combining the American Community Survey with the Survey of Income and Program Participation," Census Bureau, 2013.

27. *Small Area Health Insurance Estimates*

Purpose of the Data Series: Provide estimates of the numbers and proportions of those with and without health insurance coverage for demographic groups at the state and county levels. The demographic groups are defined by age, sex, and income. Race and ethnicity groups are available at the state-level. Income groupings are based on the income-to-poverty ratio (IPR).

Responsible Unit: Small Area Estimates Branch, Social, Economic and Housing Statistics Division, U.S. Census Bureau

Authorization: SAHIE is a reimbursable program funded by the Department of Health and Human Services.

Frequency of Data Publication: Annual

Timing of Data Release: Estimates for 2012 were issued in March 2014.

Modes to Access Data: Small Area Health Insurance Estimates (SAHIE), Census Bureau (txt, xls, csv)

Methodology: For estimation, SAHIE uses statistical models that combine survey data from the American Community Survey (ACS) with administrative records data and Census 2010 data. The models are "area-level" models because we use survey estimates and administrative data at certain levels of aggregation, rather than individual survey and administrative records. Our modeling approach is similar to that of common models developed for small area estimation, but with some additional complexities.

The published estimates are based on aggregates of modeled demographic groups. For states, we model at a base level defined by the full cross-classification of: four age groups, four race/ethnicity groups, both sexes, and five income groups. For counties, we model at a base level defined by the same age, sex, and income groups.

We use estimates from the Census Bureau's Population Estimates Program for the population in groups defined for state by age by race/ethnicity by sex, and for county by age by sex. We treat these populations as known. Within each of these groups, the number with health insurance coverage in any of the income categories is given by that population multiplied by two unknown proportions to be estimated: the proportion in the income category and the proportion insured within that income category. The models have two largely distinct parts - an "income part" and an "insurance part" - that correspond to these proportions. We use survey estimates of the number in the income groups and of the proportion insured within those groups. We assume these survey estimates are unbiased and normally distributed. We also assume functional forms for the variances of the survey estimates that involve parameters that are estimated. We treat supplemental variables that predict one or both of the unknown income and insurance proportions in one of two ways:

Some of these variables are used as fixed predictors in a regression model. There is a regression component in both the income and insurance parts of the model. In each case, a

transformation of the proportion is predicted by a linear combination of fixed predictors. Some of these predictors are categorical variables that define the demographic groups we model. Others are continuous. The continuous fixed predictors include variables regarding employment from the County Business Patterns data file, educational employment, and demographic population.

We also utilize random continuous predictors, which include data from 5-year ACS, Internal Revenue Service, Supplemental Nutrition Assistance Program, and Medicaid/Children's Health Insurance Program. These are not fixed predictors in the model. Instead, we treat them as random, in a way similar to survey estimates, but not as unbiased estimators of the numbers. Instead, we assume that their expectations are linear functions of the number in an income group or the number insured within an income group. We typically assume they are normally distributed with variances that depend on unknown parameters.

We formulate the model in a Bayesian framework and report the posterior means as the point estimates. We use the posterior means and variances together with a normal approximation to calculate symmetric 90-percent confidence intervals, and report their half-widths as the margins of error.[80]

Application of ACS Data in Methodology: Starting with 2008 SAHIE, we model single-year health insurance coverage as measured by the American Community Survey (ACS). . . . For 2007 SAHIE and prior, the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) was used. While the CPS ASEC is the official source for national estimates of health insurance coverage, the ACS has more precise state-level and sub-state estimates. Starting with 2008 ACS, the health insurance question was added to the ACS to enable the U.S. Department of Health and Human Services and other federal agencies to more accurately distribute resources and better understand state and local health insurance needs. This health insurance question had been part of the ACS Content Test administered in 2006.

The SAHIE modeling uses single-year ACS direct survey estimates from all counties and states regardless of population size. Single-year ACS estimates are published for counties and other places with population size 65,000 or larger, and three-year estimates are published for counties and other places with population size 20,000 or larger. Five-year ACS estimates are available for all counties, school districts, and other small geographic areas (e.g., census tracts or block groups). Since collection for the health insurance question started with 2008 ACS, the first five-year ACS health insurance estimates were released in December 2013.

The ACS health insurance question asks respondents whether they are currently covered at the time of interview, based on interviews conducted throughout the year. Respondents are considered insured if they are covered by any type of health insurance coverage, and they are considered uninsured if they are not covered by any type of health insurance. People with no coverage other than access to Indian Health Service are also considered uninsured. For more

---

[80] Text source: "SAHIE 2008 - 2012 Demographic and Income Model Methodology: Summary for Counties and for States," Census Bureau website.

information about the concept measured in ACS compared to that measured in CPS ASEC or Survey of Income and Program Participation, see the About Health Insurance page.

The standard ACS health-insurance data are for the civilian noninstitutionalized population, which excludes active-duty military and persons in prisons and nursing homes. However, the SAHIE data include only those for whom income is reported in the ACS (i.e., the "poverty universe"), which is a slightly more restrictive subset.

SAHIE also uses income data from the ACS in order to estimate the numbers of people in specific income-to-poverty ratio (IPR) categories. IPR is defined as total income of the family divided by the poverty threshold for that family size. For example, the SAHIE 200% IPR data provide estimates of the insured and uninsured who have income at or below 200% of the poverty threshold.

Starting with 2010 SAHIE, the following three SAHIE predictors are now created from five-year ACS: educational-attainment estimates, citizenship-status estimates, income-to-poverty ratio (IPR) estimates. For 2009 SAHIE and prior, these three predictors were created from the Census 2000 long-form sample. The SAHIE modeling by education, citizenship and IPR is described in the methodology documents available at Methodology. Importantly, we lag the five-year ACS predictors back by one year in order to avoid any overlap and/or correlation between the estimation years and the predictor years. For instance, for 2011 SAHIE, we model 2011 single-year ACS in conjunction with 2006-2010 five-year ACS predictors.[81]

ACS Questions Utilized (as suggested by the methodology): P3-6, P8, P11, P16, P47-48

Major Uses: SAHIE is only source of single-year health insurance coverage estimates for all U.S. counties.

This program is partially funded by the Centers for Disease Control and Prevention's (CDC), National Breast and Cervical Cancer Early Detection Program (NBCCEDP). The CDC uses SAHIE in fulfillment of its congressional mandate to provide screening services for breast and cervical cancer to low-income, uninsured, and underserved women through the NBCCEDP.[82]

More generally, SAHIE data are used to assess and address issues of health insurance coverage in counties with low populations. SAHIE provides single-year estimates for such entities, while comparable ACS data are five-year averages.[83]

---

[81] Text source: "SAHIE Data Inputs: American Community Survey," Census Bureau website.
[82] Text source: "About SAHIE," Census Bureau website.
[83] See "SAHIE Data Release Highlights" and "SAHIE Publications," Census Bureau websites.

28. *Small Area Income and Poverty Estimates*

Purpose of the Data Series: The U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program provides annual estimates of income and poverty statistics for all school districts, counties, and states. The main objective of this program is to provide estimates of income and poverty for the administration of federal programs and the allocation of federal funds to local jurisdictions. SAIPE estimates are specifically designed for use in annual Title I allocations to school districts.[84]

Responsible Unit: Small Area Estimates Branch, Social, Economic and Housing Statistics Division, U.S. Census Bureau

Authorization: Title I of the Elementary and Secondary Education Act of 1965 as amended (No Child Left Behind Act of 2001). Part 5 of Section 1001 of the Act notes that the purpose of the Act can be accomplished by "distributing and targeting resources sufficiently to make a difference to local educational agencies and schools where needs are greatest."

Frequency of Data Publication: Annual

Timing of Data Release: Data for 2012 were released in December 2013.

Modes to Access Data: Small Area Income and Poverty Estimates, Census Bureau (txt, xls, csv)

Methodology/Application of ACS Data in Methodology: SAIPE revises and improves methodology as time and resources allow. The details of the methodology differ slightly from year to year.[85]

Starting in 2004-2005, SAIPE began using data from the American Community Survey, rather than from the Annual Social and Economic Supplement to the Current Population Survey. The methodology combines the 1-year ACS estimates with other data sources to provide more timely, precise, and stable estimates than the 5-year ACS estimates alone.[86]

The SAIPE model has two distinct components: state and county estimates; and school district estimates. School district estimates sum to county estimates, county estimates sum to state estimates, and state estimates sum to the national estimate.[87]

State estimate techniques:

- Bayesian estimation techniques are applied to the Small Area Income and Poverty Estimates (SAIPE) program's models to combine regression predictions with direct estimates from the ACS in a way that varies the importance given to the direct ACS estimates from state to state depending upon their reliability.

---

[84] Text source: "About SAIPE," Census Bureau website.
[85] Text source: "SAIPE Methodology," Census Bureau website.
[86] Text source: "SAIPE FAQ," Census Bureau website.
[87] Text source: "SAIPE Methodology," Census Bureau website.

- The SAIPE program multiplies model-based estimates of state poverty ratios by demographic estimates of state populations to provide estimates of the numbers of people in poverty in each state for various age groups.
- The SAIPE program controls these state estimates of the number of people in poverty so that their total across states agrees with the direct ACS national estimate.
- The SAIPE program state models use data from the prior census (2000) to form a predictor variable in the regression models in one of two ways. The models either use the Census 2000 estimates directly, or use residuals from auxiliary cross-sectional regressions done with the Census 2000 estimates.
- Because the Department of Education requires estimates of the number of "related children age 5 to 17 in families in poverty", and not all children 5 to 17 are "related children", there are two sets of equations for children ages 5 to 17.
- The SAIPE program estimates the total number of people in poverty as the sum of estimates derived from a set of four age-specific equations.[88]

County estimation techniques:

- We estimate a regression model that predicts the number of people in poverty using single-year county-level observations from the American Community Survey as the dependent variable, and administrative records and census data as the predictors. A single year ACS sample estimate is used for every county, even those below 65,000 population for which the ACS data are unpublished. Although we use only the counties with nonzero reported poverty in the ACS to estimate the equation, we make regression "predictions" for all 3,142 county-level entities in the SAIPE universe (which excludes Kalawao, HI).
- The official, published direct ACS county estimates are single-year estimates only for sufficiently large counties (greater than 65,000 people); three-year or five-year accumulations of ACS data will be used in constructing estimates for smaller counties. Since modeling produces estimates with reduced sampling error, we feel we can use single-year ACS estimates for all counties in our models. We also feel it is important to do since primary uses of the SAIPE estimates (e.g., their use in allocation of federal funds) effectively involve comparing poverty estimates across places. For such uses, having all the estimates on a common basis is important, so that if we wanted to use multi-year ACS estimates for small counties, we should probably also use them for the large counties.
- The model is multiplicative; that is, we model the number of people in poverty as the product of a series of predictors which are numbers (not rates) and have unknown errors. When estimating the coefficients in the model, we take logarithms of the dependent and all independent variables. While we may omit reference to logs in the description, all variables in the county regression models for numbers of people in poverty are logarithmic.

---

[88] Text source: "2010-2012 State-Level Estimation Details," Census Bureau website.

- The ACS estimates for different counties are of different reliability because of the size of the sample in the counties. Our estimates take this factor into account.
- To use the information contained in the direct survey estimates for the counties in the ACS with nonzero reported poverty, we combine the regression predictions with these direct estimates using Empirical Bayes (or "shrinkage") techniques. The Empirical Bayes techniques weight the contribution of the two components (regression and direct estimates) based on their relative precision.
- We control the estimates for the counties of a given state to sum to the independently derived state estimate (which in turn has been controlled to sum to the ACS national estimate).
- We provide a confidence interval, which represents uncertainty from both sampling and from modeling, for each estimate.[89]

School district estimation techniques:

- The SAIPE program procedure for estimating poverty among relevant children ages 5-17 in families works with geographical units we call school district-county-pieces. These pieces are defined as the intersections of school districts and counties (i.e., all of a district if it does not cross county boundaries and each county part separately for districts that do). If a school district has territory in two counties, for example, we make estimates for the two parts separately and then combine them.
- The first part in making school district poverty estimates is to compute the school district piece tax-based child poverty rate, using federal tax information obtained from the IRS. The tax-based poverty rate for a school district piece is estimated by the product of the county related children ages 5-17 poverty rate and the ratio of school district piece to county share of "child tax-poor exemptions" over the share of total "child tax exemptions". For the 2010 school district estimates, the number of child tax exemptions and their corresponding poverty status are taken from tax year 2009 IRS income tax data. "Poor Child Exemptions" are defined as the number of child tax exemptions on returns whose adjusted gross income falls below the official poverty threshold for a family of the size implied by the number of exemptions on the form tabulated for each school district piece.
- Not all tax returns can be geocoded down to a specific school district piece. However, the total number of exemptions in a county is known. These exemptions will be called the non-geocoded exemptions. The tabulated child tax exemptions and child tax-poor exemption counts are adjusted to reflect the appropriate grade range of the school district piece because age of child is not included on the income tax form. The next step in calculating the tax-based shares is to estimate the school district piece to county share of relevant children age 5-17 and relevant children age 5-17 in poverty from ACS five-year sample, 2006-2010. The non-geocoded exemptions are also adjusted to reflect the target 5-17 year old population, then allocated to the school district pieces to minimize the difference between the tax-

---

[89] Text source: "2010 - 2012 County-Level Estimation Details," Census Bureau website.

based shares and the corresponding census-based shares using the Minimum Change algorithm. After allocating the non-geocoded exemptions, the tax-based poverty rate for a school district piece can be computed.

- The second part in creating the school district poverty estimates is to multiply the school district piece poverty rate to the official estimate of relevant child population for the school district piece. These estimates are then raked (ratio adjusted) to agree with the county estimates for number of children age 5-17 in poverty. Finally, the raked school district piece estimates are adjusted using "controlled rounding" to get results with the following properties.

  o The number of children in poverty in the pieces in a county adds up to the SAIPE program's model-based estimate of the number of children in poverty in their counties for the relevant income year.
  o The number of children in poverty in the school district pieces are integers.

- The final step is to reassemble the school district pieces into the school districts, simply by adding their controlled-rounded numbers of children in poverty together. [90]

ACS Questions Utilized (as suggested by the methodology): P4, P47-48

Major Uses: The main objective of this program is to provide estimates of income and poverty for the administration of federal programs, including Title I, and the allocation of federal funds to local jurisdictions. In addition to these federal programs, state and local programs use the income and poverty estimates for distributing funds and managing programs.

SAIPE estimates also can be used to assess and address poverty in counties with small populations. The SAIPE program annually produces single-year model-based estimates for all school districts, counties, and states, while one-year ACS estimates are annually available only for counties and school districts with populations 65,000 or greater.[91]

---

[90] Text source: "2010-2012 Overview of School District Estimates," Census Bureau website.
[91] "Why are SAIPE program poverty estimates recommended for use in Title I allocations?," Census Bureau website.

29. *Supplemental Poverty Measure*

Nature and Purpose of Measure: The Supplemental Poverty Measure (SPM) released by the U.S. Census Bureau, with support from the Bureau of Labor Statistics (BLS), extends the official poverty measure by taking account of many of the government programs designed to assist low-income families and individuals that are not included in the current official poverty measure. The development of the SPM is funded as part of the Current Population Survey program.

Responsible Unit: Census Bureau

Authorization: 13 USC 182 "The Secretary may make surveys deemed necessary to furnish annual and other interim current data on the subjects covered by the censuses provided for in this title."

Frequency of Data Publication: Annual

Timing of Data Release: The 2013 SPM estimates were released in October 2014, one month after the release of the official poverty estimates.

Modes to Access Data: Supplemental Poverty Measure, Census Bureau (pdf, SAS, STATA, PUMS)

Methodology: The measures presented in this study use the 2014 Current Population Survey Annual Social and Economic Supplement (CPS ASEC) income information that refers to calendar year 2013 to estimate SPM resources. These are the same data used for the preparation of official poverty statistics. The SPM thresholds for 2013 are based on out-of-pocket spending on basic needs (FCSU). Thresholds use 5 years of quarterly data from the Consumer Expenditure Survey (CE); the thresholds are produced at the BLS.

The key concepts for the SPM, in comparison to the official poverty measure, are as follows:[92]

---

[92] Kathleen Short, "The Supplemental Poverty Measure: 2013," *Current Population Reports*, P60-251, October 2014. A detailed description of the methodology is on pp. 18-22.

Poverty Measure Concepts: Official and Supplemental		
	Official Poverty Measure	Supplemental Poverty Measure
Measurement Units	Families and unrelated individuals	All related individuals who live at the same address, and any coresident unrelated children who are cared for by the family (such as foster children) and any cohabiters and their relatives
Poverty Threshold	Three times the cost of a minimum food diet in 1963	The mean of the 30th to 36th percentile of expenditures on food, clothing, shelter, and utilities (FCSU) of consumer units with exactly two children multiplied by 1.2
Threshold Adjustments	Vary by family size, composition, and age of householder	Geographic adjustments for differences in housing costs by tenure and a three-parameter equivalence scale for family size and composition
Updating Thresholds	Consumer Price Index: all items	Five-year moving average of expenditures on FCSU
Resource Measure	Gross before-tax cash income	Sum of cash income, plus noncash benefits that families can use to meet their FCSU needs, minus taxes (or plus tax credits), minus work expenses, minus out-of-pocket medical expenses and child support paid to another household

Application of ACS Data in Methodology: The American Community Survey (ACS) is used to adjust the food, clothing, shelter, and utilities (FCSU) thresholds for differences in prices across geographic areas. The geographic adjustments are based on 5-year ACS estimates of median gross rents for two-bedroom apartments with complete kitchen and plumbing facilities. Separate medians were estimated for each of the 264 metropolitan statistical areas (MSAs) large enough to be identified on the public-use version of the CPS ASEC file. This results in 358 adjustment factors. For each state, a median is estimated for all nonmetro areas (48), for each MSA with a population above the CPS ASEC limit (264), and for a combination of all other metro areas within a state (46).[93]

In 2014, the Census Bureau explored the possibility of using the Regional Price Parities (RPPs) from the Bureau of Economic Analysis for the purposes of geographic cost-of-living adjustments. RPPs also use ACS data to estimate the relative cost of housing.[94]

ACS Questions Utilized (as suggested by the methodology): H1, H7-8, H18

Major Uses: The aim of the SPM is to allow researchers and policymakers to better understand the actual financial and experiential circumstances of U.S. households.[95]

---

[93] *Ibid*. Also, see Renwick, Trudi. "Geographic Adjustments of Supplemental Poverty Measure Thresholds: Using the American Community Survey Five-Year Data on Housing Costs," SEHSD Working Paper Number 2011-21, U.S. Census Bureau, July 2011.

[94] Trudi Renwick, Bettina Aten, Eric Figueroa and Troy Martin, "Supplemental Poverty Measure: A Comparison of Geographic Adjustments with Regional Price Parities vs. Median Rents from the American Community Survey," SEHSD Working Paper No. 2014-22, March 2014.

[95] See "Supplemental Poverty Measure Latest Research," Census Bureau website.

30. *Voting Rights Determination File*

Purpose of the File: The Voting Rights Determination File identifies the political subdivisions legally obligated to provide minority language assistance during elections for groups that are unable to speak or understand English adequately enough to participate in elections. The Director of the United States Census Bureau is obligated to make this determination.

Responsible Unit: Census Redistricting Data Office, Census Bureau

Authorization: According to Section 203 of the Voting Rights Act, the Director of the Census is required to determine where "more than five per centum of the citizens of voting age…" residing in a State or political subdivision "are members of a single language minority." (Title 42, U.S. Code, Section 1973aa–1a)

Frequency of Data Publication: Decennial

Timing of Data Release: In the year after the decennial census (to enable states to redistrict prior to the next congressional elections)

Modes to Access Data: 2011 Section 203 determination files and tables (zip, xls)

Methodology/Application of ACS Data in Methodology: The Census Bureau produces model-based estimates in order to make Section 203 determinations for each relevant language minority group within each political jurisdiction. Specifically, data inputs include citizenship status, limited-English proficiency status, and literacy status from five-year ACS estimates and voting age population from the 2010 Census.[96]

ACS Questions Utilized (as suggested by the methodology): P4-6, P8, P11, P13-14

Major Uses: The major use of this data is to ensure that each political subdivision is in compliance with Section 203 of the Voting Rights Act.[97]

---

[96] Text source: Joyce, P., Malec, D., Little, R., & A. Gilary (2012). "Statistical Modeling Methodology for the Voting Rights Act Section 203 Language Assistance Determinations." Research Report Series (Statistics #2012-02). Center for Statistical Research & Methodology, Research and Methodology Director, US Census Bureau. Washington, DC. In particular, see "Table 5: Definitions of ACS Variables of Interest," p. 24.

[97] 28 CFR 55, Implementation of the Provisions of the Voting Rights Act Regarding Language Minority Groups

C. Economic Development Administration

31. *StatsAmerica*

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

D. National Oceanographic and Atmospheric Administation

32. *Coastal Community Social Vulnerability Indicators*

Nature and Purpose of the Measure: In assessing impacts of climate change to oceans and coasts, one key component is impacts to coastal communities, especially communities that depend specifically on the ocean for meeting economic, social and cultural needs. Fishing (commercial, recreational, and subsistence), coastal tourism and recreation, and spiritual or cultural practices centered on marine locations or species are three examples.

Existing levels of social vulnerability affect the level of impact that a community experiences from stressors, including climate change. Factors affecting vulnerability include levels of access to resources and power (political, cultural, economic and social) and of susceptibility to harm or loss. Today these communities are experiencing impacts of multiple stressors: economic, social and ecological. Therefore, identification and monitoring of socially vulnerable populations in the coastal zone is a critical aspect of understanding the impacts of climate change and other stressors. Similarly, levels of dependence on and use of ocean-related resources and conditions create greater or lesser likelihood of specific kinds of impacts. Further, coastal gentrification may be an indication of community vulnerability to development that may transform coastlines thus increasing their vulnerability to the impacts of extreme weather conditions that can result from climate change.

The National Marine Fisheries Service (NOAA Fisheries) Community Social Vulnerability Indicators (CSVIs) are statistical measures of the vulnerability of communities to climate change and other events such natural hazards, disasters, and regulatory changes to fisheries. The CSVIs currently serve as indicators of social vulnerability, gentrification pressure vulnerability, and commercial and recreational fishing dependence.[98]

Responsible Unit: Office of Science and Technology, National Marine Fisheries Service, National Atmospheric and Oceanographic Administration (NOAA), Department of Commerce

Authorization: Magnuson–Stevens Fishery Conservation and Management Act, Section 404 (16 U.S.C. 1881c)

Frequency of Data Publication: The initial set of CSVIs was made available in 2013. NMFS indicates it will update these on an annual basis.

Timing of Data Release: Indices based on the 2005-2009 ACS were released in 2013.

Modes of Data Access:

- [Community Social Indicators](), NOAA Fisheries (map)
- Jepson, Michael and Lisa L. Colburn, ["Development of Social Indicators of Fishing Community Vulnerability and Resilience in the U.S. Southeast and Northeast]()

---

[98] Text source: ["Implications of Climate Change on the Fishing Communities of the U.S. NES LME,"]() Northeast Fisheries Science Service website.

Regions," U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-F/SPO-129, April 2013 (pdf)

Methodology: The CSVIs are constructed annually, using demographic data from the U.S. Census' American Community Survey (ACS) five-year rolling estimates, NOAA Fisheries annual commercial fisheries and Marine Recreational Information Program (MRIP) data, as well as a small number of publically available but non-government online databases. Ongoing data collection will allow the CSVIs to be continually updated to show long term trends.[99]

The initial set of CSVIs were created as follows:

We developed a suite of social indicators for use in fisheries social impact assessment. Data from more than 2,900 coastal communities in 19 states from Maine to Texas were used to create 14 social vulnerability and fishing dependence indices.

This project was developed in two phases. Prior to undertaking our research, a group of regional fisheries experts and social scientists familiar with social indicators and fishing communities were convened in 2010 to assist in the development of an approach to create measures of fishing community well-being. A second workshop held in 2011, shortly after the initial development of indicators, was to review and suggest revisions to the indicators.

The approach selected for constructing our indices of community vulnerability closely follows that developed by Jacob et al. Building on work by Cutter et al. and Jepson and Jacob, this approach utilizes several indices to examine different aspects of social vulnerability and over time resilience for each location, concentrating on those relevant to the coastal economy and fishing communities.

We collected data from both public and private sources for over 2,900 communities in coastal counties in 19 states from Maine to Texas. Estimate data at the place level had recently become available through the Census Bureau's American Community Survey (ACS 2006-2009) [sic]. Census data collected on line through the American Factfinder constituted the bulk of the community demographic component. Community data for crime and hazardous weather variables collected from nongovernmental websites augmented our demographic profile. We collected fisheries data, such as number of permits and volume and value of landings, from both the Northeast and Southeast Science Centers. One hundred and twenty variables were selected for the analysis.

We assembled the data for every Census Designated Place (CDP) from a set of predetermined coastal counties along the Eastern and Gulf coasts. The criterion for coastal county designation was that it has some connection with the ocean, through a coastline, river, bay or estuary. This criterion was chosen as we envisioned communities within these counties to have comparable economies and experiencing similar vulnerabilities that come from having ocean front property or beaches, inlets and bays with access to the ocean and many of the amenities that make coastal living such a desirable destination to so many. Communities were not chosen based

---

[99] *Ibid*.

upon their fishing activity, but once selected fishing activity was placed within the context of that community.

Factor analysis is a data reduction technique that allows for the construction of indices that represent the latent structure of a conceptual variable. Latent concepts are unobservable, but variables taken together can represent the concept, much as the concept of inflation is measured by the consumer price index. Here, our constructs relate to several concepts of well-being and fishing activity and correspond to vulnerability to social change at the community level.

We initiated our index construction using the variables originally chosen by Jacob et al. The factor analysis process consisted of a principal component analysis with a varimax rotation. Using a varimax rotation allows one to determine which variables are loading the highest onto the factor and would more likely result in a one-factor solution if included when a single factor is not achieved. When we were unable to achieve a one-factor solution with a particular set of variables for an index, we substituted comparable variables that had high factor loadings within the overall principal component analysis until we found a satisfactory one-factor solution. We also used substitution of the mean for missing data to ensure each community would receive a scale score and also because we had relatively few missing data.

We placed the indices under one of three categories within the larger context of vulnerability. The three main categories are social, gentrification, and fishing dependence vulnerability. While gentrification is both social and economic in its effect, the process itself has special significance as it creates different vulnerabilities and is thought to deserve consideration as a discrete process from other social and economic vulnerabilities.[100] Indices by category include:

- Social Vulnerability Indices
  - Personal Disruption Index represents factors that disrupt a community member's ability to respond to change because of personal circumstances affecting family life or educational levels or propensity to be affected by crime or poverty. A high rank indicates more personal disruption and a more vulnerable population.
  - Population Composition Index shows the presence of populations who are traditionally considered more vulnerable due to circumstances often associated with low incomes and fewer resources. A high rank indicates a more vulnerable population.
  - Poverty Index is a commonly used indicator of vulnerable populations. A high rank indicates a high rate of poverty and a more vulnerable population.
  - Labor Force Index characterizes the strength and stability of the labor force and employment opportunities that may exist. A high rank means likely fewer employment opportunities and a more vulnerable population.

---

[100] Text source: Jepson, Michael and Lisa L. Colburn, "Development of Social Indicators of Fishing Community Vulnerability and Resilience in the U.S. Southeast and Northeast Regions," U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-F/SPO-129, April 2013.

- o Housing Characteristics Index is a measure of infrastructure vulnerability and includes factors that indicate housing that may be vulnerable to coastal hazards. A high rank means a more vulnerable infrastructure and a more vulnerable population. On the other hand, the opposite interpretation might be that more affordable housing could be less vulnerability for some populations.
  - o Housing Disruptions Index
- Gentrification Vulnerability Indices
  - o Housing Disruption Index represents factors that indicate a fluctuating housing market where some displacement may occur due to rising home values and rents. A high rank means more vulnerability for those in need of affordable housing and a population more vulnerable to gentrification.
  - o Retiree Migration Index characterizes areas with a higher concentration of retirees and elderly people in the population. A high rank indicates a population more vulnerable to gentrification as retirees seek out the amenities of coastal living.
  - o Urban Sprawl Index describes areas experiencing gentrification through increasing population and higher costs of living. A high rank indicates a population more vulnerable to gentrification.
  - o Natural Amenities Index represents factors common to coastal areas experiencing gentrification through the presence of more water coverage, seasonal housing, and water related amenities. A high rank indicates a population more vulnerable to gentrification.
- Fishing Engagement and Reliance Indices – Recreational Fishing Reliance Index, Recreational Fishing Engagement Index, Commercial Fishing Reliance Index, Commercial Fishing Engagement Index[101]

We selected communities for this discussion using a methodology for classifying coastal communities for sampling purposes (Smith et al.). All communities showed some level of involvement in either commercial or recreational fishing. Using a community's factor scores for each discrete index, a K-means cluster analysis was used to create a taxonomy of fishing communities (Pollnac). This resulted in a typology of 35 communities [sic] clusters. Each cluster of communities has a unique set of shared characteristics for each index included in the analysis. For example, scores for one cluster may indicate a high degree of involvement in commercial fishing and be characterized by a strong labor force and low risk of gentrification while another cluster may show the opposite. This method made it possible to select communities with a wide range of involvement in commercial and/or recreational fishing and exhibiting varying degrees of social vulnerability.

As NMFS (NOAA Fisheries) began profiling fishing communities it became obvious that decisions had to be made regarding the geographical boundaries for data collection purposes. The decision to use a particular geographic boundary was a regional decision based upon the characteristics surrounding a particular community and its participation within its region's

---

[101] Text source: "Social Indicator Definitions," NOAA Fisheries webpage.

fisheries. In most cases, CDP was chosen to represent community boundaries. Although CDP does not always correspond to an incorporated place, it is used by the Census Bureau for the purposes of data collection and does include incorporated cities, towns and villages.

The advantage of this approach is the ability to use secondary data rather than rely on primary data collection. It provides a significant timesaving compared with conducting fieldwork or implementing a survey, an important consideration given the sometimes-short timeframe in which social impact assessments are conducted. These types of measures can be created for communities and other geographic scales and offer the ability to compare across regions. The analysis is comparatively straightforward and can be modified to adapt to changing circumstances and differential regional availability of specific variables within topical areas without compromising the results. Because they are based on existing standard time series data, these measures can also be developed retroactively to allow valuable analyses of change within a community over time as proxies of resilience.

While these indices are sufficiently robust for the assessment of fishing community vulnerability, additional indices and data on social capital, critical for understanding social networks and social cohesion, would strengthen the analysis. Data on fishing community infrastructure would improve our assessment of community dependence on fishing, both recreational and commercial. Additional variables and indices will also be needed to assess well-being more broadly. Currently, databases for most of these variables are inadequate at the community level. With regard to sea level rise, community level measures of coastal hazards are an area to be explored. However, as the level of analysis progresses from community, to county, and to state there is increasingly more secondary data; measures of health and well-being, for instance, are more attainable at county and state levels than at the community level and should be researched.. This may allow addition of new variables and indices at different geographic levels that could complement and contextualize the community indicators. [102]

Application of ACS in the Methodology: The baseline ACS data covers the years 2005 to 2009 and will be compared to 2010 to 2014 estimates once they are available. The 2006 to 2010 ACS data were used to construct the social and gentrification pressure vulnerability indicators.[103]

ACS Questions Utilized (as suggested by the methodology): P2-3, P6, P11, P14, P20, P29, P35-37, P41, P43-44, P47-48, vacancy rate, H1, H7, H18-19, H22-23[104]

Major Uses: We have developed a set of social indicators using secondary data that can improve the analytical rigor of fisheries social impact assessment. The majority of these data are readily accessible and can be compiled quickly to create measures of social vulnerability and to update community profiles. Because we know these communities exist within a larger coastal economy, the ability to profile the context of vulnerability to social factors outside of fishing is critical to understanding how regulatory change will be absorbed into these

---

[102] Text source: Jepson and Colburn, *op.cit.*

[103] Text source: "Implications of Climate Change on the Fishing Communities of the U.S. NES LME," Northeast Fisheries Science Service website.

[104] Based on list at http://nefsc.noaa.gov/ecosys/climate_change/figures/thumb-tab2.jpg.

multifaceted places. Creating social indicators of vulnerability for fishing communities provides a pragmatic approach toward standardization of data and analysis for assessment of some of the long term effects of management actions.

This research forms the initial step in developing more empirical measures to enhance NMFS' ability to understand the dynamics of fishing communities and their ability to recover from disruptive events, whether they are man-made, such as regulatory change, or natural. Lessons learned from this assessment can guide further research and assist in the development of social impact assessment best practices for management actions.

33. *U.S. County Social Vulnerability Index*

Nature and Purpose of the Measure: The Social Vulnerability Index (SoVI) measures the social vulnerability of U.S. counties to environmental hazards. The index is a comparative metric that helps users examine differences in social vulnerability among counties. SoVI graphically illustrates the geographic variation in social vulnerability. It shows differences in capacity for preparedness and response, as well as areas where resources might be used most effectively to reduce vulnerability. SoVI is also useful as an indicator in determining the level of recovery from disasters.

The index synthesizes 29 socioeconomic variables, which the research literature suggests contribute to reduction in a community's ability to prepare for, respond to, and recover from hazards.[105]

Generally defined, vulnerability is the potential for loss of life or property due to hazards. The hazards-of-place model combines the biophysical vulnerability (physical characteristics of hazards and environment) and social vulnerability to determine an overall place vulnerability. Social vulnerability is represented as the social, economic, demographic, and housing characteristics that influence a community's ability to respond to, cope with, recover from, and adapt to environmental hazards.[106]

Responsible Unit: Office for Coastal Management, National Oceanographic and Atmospheric Administration (NOAA), with NOAA's Sea Grant Program, funded the University of South Carolina Hazards & Vulnerability Research Institute to prepare SOVI.

Authorization: Coastal Zone Management Act

Frequency of Data Publication: SoVI has been published in three editions—2000, 2005-2009, and 2006-2010.

Timing of Data Release: The 2006-2010 SoVI was released in May 2012.

Modes of Data Access:

- SoVI, Digital Coast, Office of Coast Management (shp)
- Social Vulnerability Index for the United States, Hazards and Vulnerability Research Institute, University of South Carolina (html, xls)

Methodology: SoVI 2006-10 marks a change in the formulation of the SoVI metric from earlier versions. New directions in the theory and practice of vulnerability science emphasize the constraints of family structure, language barriers, vehicle availability, medical disabilities, and healthcare access in the preparation for and response to disasters, thus necessitating the inclusion of such factors in SoVI. Extensive testing of earlier conceptualizations of SoVI, in addition to the introduction of the U.S. Census Bureau's five-year American Community Survey

---

[105] Text source: "Social Vulnerability Index for the United States - 2006-10," HVRI webpage.
[106] Text source: "Social Vulnerability Index I Frequently Asked Questions," HVRI webpage.

(ACS) estimates, warrants changes to the SoVI recipe, resulting in a more robust metric. These changes, pioneered with the ACS-based SoVI 2005-09 carry over to SoVI 2006-10, which combines the best data available from both the 2010 U.S. Decennial Census and five-year estimates from the 2006-2010 ACS.[107]

Table 3: List of twenty-nine SoVI® 2006-10 variables with descriptions. Daggers notate new variables in SoVI® 2006-10.

VARIABLE	DESCRIPTION
QASIAN	Percent Asian
QBLACK	Percent Black
QHISP	Percent Hispanic
QNATAM	Percent Native American
QAGEDEP†	Percent of Population Under 5 Years or 65 and Over
QFAM†	Percent of Children Living in Married Couple Families
MEDAGE	Median Age
QSSBEN	Percent of Households Receiving Social Security
QPOVTY	Percent Poverty
QRICH200K	Percent of Households Earning Greater Than $200,000 Annually
PERCAP	Per Capita Income
QESL†	Percent Speaking English as a Second Language with Limited English Proficiency
QFEMALE	Percent Female
QFHH	Percent Female Headed Households
QNRRES	Percent of Population Living in Nursing and Skilled-Nursing Facilities
HOSPTPC	Hospitals Per Capita (County Level ONLY)
QNOHLTH†	Percent of Population Without Health Insurance (County Level ONLY)
QED12LES	Percent with Less Than 12th Grade Education
QCVLUN	Percent Civilian Unemployment
PPUNIT	People Per Unit
QRENTER	Percent Renters
MDHSEVAL†	Median House Value
MDGRENT†	Median Gross Rent
QMOHO	Percent Mobile Homes
QEXTRCT	Percent Employment in Extractive Industries
QSERV	Percent Employment in Service Industry
QFEMLBR	Percent Female Participation in Labor Force
QNOAUTO†	Percent of Housing Units with No Car
QUNOCCHU	Percent Unoccupied Housing Units

*Note: QSPNEEDS (Percent of Population with a Disability) was included in SoVI® 2005-09 but excluded from SoVI® 2006-10 because estimates were not available for all counties.

Source: "Changes and Improvements in the SoVI® Formulation for 2006-10," HVRI webpage.

---

[107] Text source: "Social Vulnerability Index for the United States - 2006-10," HVRI webpage.

After obtaining the relevant data, a principle components analysis is used to reduce the data into set of components. Slight adjustments are made to the components to ensure that the sign of the component loadings coincide with each individual population characteristic's influence on vulnerability. Components include race and poverty, wealth, age, ethnicity (Hispanic), ethnicity (Native American), nursing home residents, and employment in service industries. Analysis is done at the county and tract levels.

All components are added together to determine a numerical value that represents the social vulnerability for each county.

The SoVI was created as a comparative index at a county-level for the entire United States. Therefore, the SoVI scores need to be displayed in relation to each other. Generally, the SoVI is classified using standard deviations. Social vulnerability scores that are greater than 2 standard deviations above the mean are considered the most socially vulnerable, and scores below 2 standard deviations less than the mean are the least vulnerable.[108]

Application of ACS in the Methodology: SoVI is compiled from the five-year ACS, except in instances when 2010 Census data are available.

ACS Questions Utilized (as suggested by the methodology): number in household, P2-6, P11, P13-14, P16, P20, P35-37, P43-44, P47-48, H1, H12, H17-19, Group Quarters type

Major Uses: SoVI is used as a guide to the development of state and local hazard mitigation plans. It also can be used to track recovery after disasters.

---

[108] Text source: "Text source: "Social Vulnerability Index I Frequently Asked Questions," HVRI webpage.

E. National Telecommunications and Information Administration

34. Digital Nation

Nature and Purpose of the Data Reports: The National Telecommunications and Information Administration (NTIA) analyzes broadband Internet usage in the United States based on data gained through an annual nationwide survey commissioned from the U.S. Census Bureau. In a series of "Digital Nation" reports, NTIA publishes findings on broadband adoption in America, including why people do not subscribe and which populations are lagging in usage.[109]

Responsible Unit: Office of Policy Analysis and Development, NTIA

Authorization: 47 U.S. Code § 1303(d) - Improving Census data on broadband

> The Secretary of Commerce, in consultation with the Federal Communications Commission, shall expand the American Community Survey conducted by the Bureau of the Census to elicit information for residential households, including those located on native lands, to determine whether persons at such households own or use a computer at that address, whether persons at that address subscribe to Internet service and, if so, whether such persons subscribe to dial-up or broadband Internet service at that address.

Frequency of Data Publication: Annual, beginning in 2010.

Timing of Data Release: "Digital Nation" is published about one year after the collection of the data.

Modes of Data Access: Internet and Computer Use Studies and Data Files, NTIA (pdf, zip)

Methodology: "Digital Nation" uses data from the Current Population Survey ("CPS"), a monthly survey of a representative sample of the U.S. noninstitutional population that provides data on labor force participation, income, and demographic characteristics of households. It includes data from the October CPS Computer and Internet Use Supplement, a special supplement to the CPS periodically commissioned by NTIA.

The October CPS asked each surveyed household whether someone in that household used or owned a computer, as well as who in the household used the Internet, and the devices and locations from which they did so. In addition, the survey asked the household which of the following technologies members utilized to connect to the Internet from home: dial-up service, DSL, cable modem, fiber optics, satellite, mobile broadband, or some other Internet connection technology. The October CPS Supplement included a group of questions primarily concerning Internet and mobile phone usage habits.

About 53,600 household records comprise the sample, representing 122 million American households. NTIA analyzes computer and Internet use at the household and person levels and

---

[109] Text source: "Digital Nation Reports," NTIA webpage.

their association with characteristics such as age, family income, household size and composition, and geographic location.[110]

<u>Application of ACS in the Methodology</u>: As mandated by Congress, the Census Bureau initiated ACS collection of the data on computer and internet use in 2013 (questions H9-11), with the first data release in September 2014. NTIA is assessing its options for using the new ACS data to augment the CPS-based analysis in "Digital Nation" and to create new data series.

<u>ACS Questions Utilized (as suggested by the methodology)</u>: number in household, P2-6, P11, P17-19, P29, P35-37, P47-48, H9-11

<u>Major Uses</u>: "Digital Nation" informs efforts to close the digital divide, helping more Americans compete in the 21st century economy and improving overall quality of life.

---

[110] Text source: NTIA, "Exploring the Digital Nation: Embracing the Mobile Internet," Appendix A: Data and Methodology, October 2014, p. 43.

35. *National Broadband Map*

Nature and Purpose of the Measure: The National Broadband Map (NBM) is a searchable and interactive website that allows users to view broadband availability across every neighborhood in the United States. The NBM was created by the National Telecommunications and Information Administration (NTIA), in collaboration with the Federal Communications Commission (FCC), and in partnership with 50 states, five territories and the District of Columbia. The NBM is part of NTIA's State Broadband Initiative.

Responsible Unit: NTIA, Department of Commerce

Authorization: The primary goal of the Broadband Mapping Program, as mandated by section 6001(l) of the Recovery Act, is to develop and maintain a comprehensive, interactive, and searchable nationwide inventory map of existing broadband service capability and availability in the United States that depicts the geographic extent to which broadband service capability is deployed and available from a commercial or public provider throughout each state. Furthermore, Division A of the Recovery Act authorizes NTIA to expend up to $350 million pursuant to the Broadband Data Improvement Act (BDIA) and for the development and maintenance of the national broadband map.[111]

Frequency of Data Publication: The NBM will be updated approximately every six months and was first published on February 17, 2011.

Timing of Data Release: NA

Modes of Data Access: National Broadband Map, NTIA (maps, xml, csv)

Methodology: Each State Broadband Initiative grantee's data gathering process began by contacting the potential broadband providers within a state. Although participation is voluntary, most providers chose to support this effort. The success of this program rests, in part, on that support, and we appreciate their efforts to participate in this program. In sum, grantees identified and contacted more than 4,600 potential broadband providers nationwide and collected data for approximately 3,400 actual providers operating across the country. When rolled up into parent companies for the national dataset, this work resulted in data for more than 1,650 broadband providers.

Broadband providers submitted data in a variety of formats, and in a number of cases, the grantees also conducted technical assistance to support the efforts of smaller providers to participate in this effort. Grantees performed many different types of analysis and verification methods, from drive testing wireless broadband service across their highways to meeting with community leaders to receive feedback. Many met with broadband providers, large and small, to confirm data, or suggest changes to service areas. In areas where they needed more

---

[111] Text source: "State Broadband Data and Development Program (Broadband Mapping Program): Frequently Asked Questions," NTIA, August 2009.

information, a number of grantees also went into the field looking for infrastructure to validate service offerings.

Before submitting data to NTIA, grantees integrate the data from each provider into a single dataset using a Data Model. NTIA and the FCC then integrate each of these datasets into the single National Broadband Map dataset.

After integration we review each grantee's process and also review each NBM record by comparing it to other government and third-party datasets. Comparisons to other existing datasets help to identify the extent to which the data collected under this effort matches availability and speed information that have been collected elsewhere. Multiple matches can help solidify confidence in a given result, but because data this granular has never been collected before, non-matches do not indicate that the information is inaccurate.

Using the data above, we designed, developed and implemented a website for the NBM that is searchable and interactive. It displays broadband data at multiple levels of geography and in multiple formats. All of the data on the NBM is also made available for download.[112]

Application of ACS in the Methodology: Demographic data from the ACS provide community context for place-specific broadband statistics. Topics include age, race, ethnicity, education, citizenship, and income. Data, ACS and otherwise, are available by state, county, metro area, congressional district, census places, and state legislative districts.

ACS Questions Utilized (as suggested by the methodology): P4-8, P11, and P47-48

Major Uses: The NBM was created to encourage economic growth by facilitating the integration of broadband and information technology into state and local economies.

---

[112] Text source: "National Broadband Map—About: Technical Overview," NTIA webpage.

IV.   U.S. Department of Defense

A.   Office of the Under Secretary of Defense for Personnel and Readiness

36. Basic Allowance for Housing for the Uniformed Services

Nature and Purpose of the Measure: Basic Allowance for Housing (BAH) is a U.S.-based allowance that provides uniformed Service members equitable housing compensation based on housing costs in local housing markets when government quarters are not provided.[113]

Responsible Unit: Strategic Planning and Policy Division, Defense Travel Management Office (DTMO), Office of Military Personnel Policy, Office of Assistant Secretary of Defense for Readiness and Force Management, Office of the Under Secretary of Defense for Personnel and Readiness

Authorization: 37 U.S. Code § 403(b) - Basic allowance for housing inside the U.S.

Frequency of Data Publication: Annual

Timing of Data Release: DTMO issues BAH rates in the December immediately before the reference calendar year.

Modes to Access Data:

- BAH Calculator, DTMO (html)
- BAH Component Breakdown for All Locations, DTMO (pdf)
- BAH Rates for All Locations, DTMO (pdf, ascii)

Methodology/Application of ACS Data in Methodology: The purpose of the Basic Allowance for Housing (BAH) program is to provide fair housing allowances to service members. The goal of the program is to help members cover the costs of housing in the private sector; therefore, rental-housing costs in the private sector are the basis for the allowance. Members receive a housing allowance when government quarters are not available. DoD determines an equitable housing allowance to enable members to afford suitable rental housing near their duty location. The allowance is set based on geographic duty location, pay grade, and dependent status.

The Department of Defense and the Services designed the Basic Allowance for Housing program to provide accurate housing allowances based on the market price of rental housing rather than member-reported rents. The BAH program measures rental-housing costs in the civilian market rather than measuring how much members spend on housing. This method ensures a more accurate correlation between allowance payments and rental prices.

The six standard housing profiles used as anchor points for BAH are: 1 Bedroom Apartment, 2 Bedroom Apartment, 2 Bedroom Townhouse/Duplex, 3 Bedroom Townhouse/Duplex, 3 Bedroom Single Family Detached House, and 4 Bedroom Single Family Detached House. BAH

---

[113] Text source: DTMO, "Basic Allowance for Housing" fact sheet.

distinguishes between with-dependents and without-dependents, not the number of dependents.

In computing BAH, we include local price data for: rental housing; utilities (including electricity, gas, oil, water, & sewer); and renter's insurance.

We employ a contractor to collect the data annually for approximately 300 Military Housing Areas (MHAs) in the United States, including Alaska and Hawaii. DoD and the Services define these MHAs by sets of ZIP Codes. Nation-wide data collection occurs in the spring and summer when housing markets are most active. Rental costs are collected for apartments, townhouses/duplexes, and single-family rental units of varying bedroom sizes. The different types of units are referred to as "profiles" or "anchor points." DoD uses housing profiles to link rental costs with particular paygrades.

Current, valid rental costs are crucial to accurate BAH rates. We use data from multiple sources to provide a "checks and balances" approach. This ensures reliability and accuracy. We obtain current residential vacancies from local newspapers and real estate rental listings. We also contact apartment and real estate management companies to identify units for rental pricing. We consult with real estate professionals in each MHA to confirm market rental prices and obtain additional data. Where available, we also contact fort/post/base housing referral offices and installation leadership. We tap the local housing office knowledge and gain insights into the concerns of our members.

We gather enough data to attain a statistical confidence level of 95% to ensure the estimated median rent is within 5% of the true median rent.

We use data from the Census Bureau's annual American Community Survey (ACS) to determine average expenditures for utilities specific to each dwelling type in each MHA. All data is sensitive to local housing conditions, geography, and climate.

The renter's insurance portion of BAH covers the value of household contents. We use data from major insurance companies in each MHA to determine average renter's insurance. These values are correlated with selected incomes and dwelling types.[114]

ACS Questions Utilized (as suggested by the methodology): H1, H7, H14

Major Uses: BAH enables Service members to live off-base comparably to their civilian counterparts.[115]

---

[114] Text source: DTMO, "A Primer on the Basic Allowance for Housing (BAH) For the Uniformed Services," October 2013.
[115] Text source: DTMO, "Basic Allowance for Housing" fact sheet.

V.   U.S. Department of Education

B.   National Center for Education Statistics

37. [School District Demographics System](#)

Nature and Purpose of the Dataset: The School District Demographics System (SDDS) is a Web-based resource operated by the National Center for Education Statistics (NCES) of the U.S. Department of Education. SDDS enables access to school district demographic and related geographic data that is not available in any other form. SDDS was initially developed to provide access to the 2000 School District Special Tabulation (STP2) and corresponding 1990 Census SDST. SDDS has evolved as more recent data have become available through the American Community Survey School District Special Tabulations. SDDS includes only summary data tabulated for geographic areas from public use data files. SDDS includes no data for/about individuals.

The purpose of SDDS is to help community leadership, educators, researchers and analysts, libraries and information providers, students, and the public to access and use demographic data to help them better understand current demographic characteristics, patterns and change taking place, as well as plan for improved educational programs and opportunities.

SDDS data include:

- Characteristics of children pre-primary, primary and secondary school age by:
  - All children under 18 years of age and not high school graduates (ages 18-19)
    - for those enrolled
    - for those enrolled in public school
    - for those enrolled in private school
    - for those not enrolled
- Education provider resource requirements (how many classrooms where, etc.) and facilitate development of planning
- Characteristics of children enrolled in public versus private school
- Characteristics of school age children not enrolled in school[116]

The components of SDDS include:

- Data Tables interactively access school district demographics from the 1990 Census, Census 2000, and more recent American Community Surveys (five-year).   The focus of this section is to provide data for a school district, county, American Indian Area, state or the U.S. Selection of the tabulation universe refers to total population universe or one of the School District Special Tabulation universes.
- Snapshot Reports: Most of the SDDS features described above are focused on accessing demographic and related geographic data. There are many other related data that are collected by the Department of Education by school district and can

---

[116] Text source: "[What is SDDS?,](#)" NCES website.

sometimes be useful in presenting a more holistic view of the school district demographic-economic characteristics.

- The Map Viewer provides access to interactive mapping. Use this feature to view school district boundaries in the context of a region, other school districts and other types of geography. The Map Viewer can also display demographic data, available elsewhere in SDDS in tabular format, as thematic patterns. For example, view median household income by school district for all school districts in a region. The thematic pattern shows levels of median household income (or any subject item selected) as different colors, hatch patterns, or gradations of a selected color. Use the Map Viewer to visually examine the set of school districts in a state/region for which the American Community Survey (ACS) data are available.
- The Download Data feature is similar in scope and operation to the Data Tables operation. The difference is that Download Data enables the selection of up to 20 tables and the option to download/save these data on your computer for further local processing.
- The Profiles feature displays a "comparative analysis" view of two selected school districts for any one of the statistical datasets supported. Unlike the Data Tables, the Profile data display is comprised of a set of pre-selected subject matter items from the corresponding statistical program. Use the Profiles feature when seeking a school district demographic overview for one district or to compare these summary items for two school districts. The Profiles feature also enables viewing a multi-year profile for a selected school district. The multi-year profile shows a table with data for the school district for each year for which the data are available for the SDDS demographic datasets.
- Download Maps. Use the Download Maps feature to download map files of school district boundaries in shapefile format.[117]

Responsible Unit: National Center for Education Statistics (NCES), Department of Education

Authorization: National Education Statistics Act of 1994, 20 USC 9003 (a) (4).This section of the US Code describe the duties of the Center as including a provision for the "collecting, analyzing, cross-tabulating, and reporting, to the extent feasible, so as to provide information by gender, race, socioeconomic status, limited-English proficiency, and other population characteristics when such disaggregated information would facilitate educational and policy decisionmaking."

Frequency of Data Publication: Annual. Currently, four sets of five-year ACS data are available (2005-2009 through 2008-2012).

Timing of Data Release: Within a year of the Census Bureau's release of the five-year ACS data.

Modes to Access Data: School District Demographics System, NCES (html, xls)

---

[117] *Ibid.*

Methodology/Application of ACS Data in Methodology: Tabulations are prepared from ACS public use files that correspond to school district boundaries.

ACS Questions Utilized (as suggested by the methodology): all questions[118]

Major Uses: The purpose of SDDS is to help community leadership, educators, researchers and analysts, libraries and information providers, students, and the public to access and use demographic data to help them better understand current demographic characteristics, patterns and change taking place, as well as plan for improved educational programs and opportunities.[119]

---

[118] See "List of Tables, Reference, American Community Survey 2006-2010," NCES website.
[119] Text source: "What is SDDS?," NCES website.

VI.   U.S. Department of Energy

A.   Energy Information Administration

38. *[Residential Energy Consumption Survey](#)*

Purpose of the Data Series: Residential Energy Consumption Survey (RECS) gathers information on the energy consumption characteristics and patterns of residential housing units nationwide to a nationally representative sample of housing units. The information is combined with data from energy suppliers to these homes to estimate energy costs and usage for heating, cooling, appliances and other end uses. Estimates are for the nation and each state.

Responsible Unit: Energy Information Administration, Department of Energy

Authorization: The RECS is authorized under the Federal Energy Administration Act of 1974 (Public Law 93-275) as amended, and the Energy Policy Act of 1992. P.L. 93-275 says that the Secretary of Energy shall "collect, evaluate, and assemble, and analyze energy information on reserves, production, demand, and related economic data."

Frequency of Data Publication: Every four years, most recently in 2009. It does not appear a 2013 survey was conducted.

Timing of Data Release: Data from the 2009 RECS was released in 2013.

Modes to Access Data: [2009 RECS Survey Data](#), Energy Information Administration (xls)

Methodology: RECS is a multi-year effort consisting of a Household Survey phase, data collection from household energy suppliers, and detailed consumption and expenditures estimation.

- The Household Survey collects data on energy-related characteristics and usage patterns of a nationally representative sample of housing units. For renters that do not directly pay for their energy usage, a supplementary Rental Agent Survey is conducted.
- The Energy Supplier Surveys collect data on how much electricity, natural gas, propane/LPG, fuel oil, and kerosene were consumed in the sampled housing unit during the reference year. It also collects data on actual dollar amounts spent on these energy sources.
- EIA uses a non-linear statistical model to produce consumption and expenditures estimates for heating, cooling, refrigeration and other end uses in all housing units occupied as a primary residence in the United States.

RECS uses a combination of Computer-Assisted Personal Interview, internet, and mail to collect data for the Household and Energy Supplier Surveys. Data collection for the 2009 RECS took place in 2010.

The scope and purpose of RECS differ slightly from similar EIA products that report "residential" energy data. For one, RECS samples homes occupied as a primary residence, which excludes secondary homes, vacant units, military barracks, and common area in apartment buildings. RECS estimates, therefore, do not represent sector level estimates, but are best suited for those wishing to compare across different characteristics of homes within the residential sector. RECS "totals" also exclude consumption and expenditures estimates for which data are difficult to obtain, such as estimates for biomass (wood) or solar consumption.

As in previous survey years, the 2009 RECS used a multistage area probability sample design, where the universe was broken up into successively smaller, statistically selected areas starting from counties and ending with individual housing units. The universe for the 2009 RECS sample design included all housing units that are occupied as primary residences in the 50 States and the District of Columbia. Because EIA benchmarks to occupied housing totals from the U.S. Census Bureau's American Community Survey, the RECS uses the Census Bureau definition of a housing unit. It includes single-family homes, units in multi-family buildings, and mobiles. The RECS excludes vacant, seasonal or vacation homes, and group quarters such as prisons, military barracks, dormitories, and nursing homes.

The sample design for the 2009 RECS expanded on the design used for the previous round. It retained the structural definitions in the 2005 RECS design but substantially added to the 2005 selections to achieve an almost triple sample size; 12,083 complete interviews compared to 4,382 in 2005.

All housing units in the 50 States and the District of Columbia that are occupied as primary residences are eligible to be included in the RECS sample. Sample selection begins by randomly choosing counties. The selected counties are then sub-divided into groups of Census blocks called segments and a sample of segments is randomly drawn from the selected counties.

Within each selected segment, a list of housing units (sample frame) was created from a combination of the Delivery Sequence File (DSF) and field listing. Addresses in approximately 88% of selected segments were obtained from the DSF. Field listing was used in the remaining segments, which were mostly rural areas where the DSF contained many post office boxes and rural route boxes that not suitable for an in-person survey sample frame.

The final sample of housing units is randomly selected from the housing unit frame constructed from the selected area segments. This type of sampling is called a multi-stage area probability design. Its proper application ensures that the selected sample represents the entire population of occupied housing units in the United States.

The number of counties, segments, and housing units to be selected are carefully controlled so that RECS produces estimates of average energy consumption at specified levels of precision within the following geographic levels, called domains: National, Census Region, Census Division, and individual states or group of states within Census Divisions

The RECS Household Survey was conducted on a voluntary basis with respondents. The 2009 RECS again used trained professional interviewers to administer a standardized survey

instrument programmed in Blaise software on laptops. The interview was conducted in person, allowing for interviewers to also measure the dimensions of the housing unit, data which is used to produce estimates of total, heated and cooled square footage.

For renters that did not pay some or all of their energy bills directly, follow-up surveys were attempted with a rental agent or landlord to get more information about the housing unit. Respondents in the Rental Agent Survey were asked about the equipment and fuels used by tenant households for space heating, water heating, air conditioning, and cooking, as well as the method of bill payment. The interviews with rental agents were conducted concurrently with the household survey. The Household Survey interviews spawned 829 Rental Agent Surveys cases, of which 583 (70%) were successfully completed.

During the 2009 RECS household interviews, respondents provided names and contact information for their energy suppliers. The list of suppliers accumulated during the household survey served as the frame for each of the four Energy Supplier Surveys (ESS); Electricity (Form EIA-457E), Natural Gas (EIA-457F, LPG/Propane (EIA-457D), and Fuel Oil/Kerosene (EIA-457G). The frame for the 2009 RECS ESS included over 21,000 unique requests for household billing data from more than 1,500 energy suppliers. Fuel accounts for which household respondents did not pay energy bills directly (for example, a renter who pays electricity as part of monthly rent) were excluded from the frame.

After energy supplier data were collected and edited, a process called annualization needed to occur. This statistical process used the billing data to estimate a housing unit's consumption and expenditures between January 1, 2009 and December 31, 2009. An energy end-use model was used to break down the total, annualized consumption and expenditures for each sampled case into portions used for space heating, air-conditioning, water heating, refrigerators, appliances, and other uses. There are separate models for electricity, natural gas, fuel oil, LPG/propane, and kerosene.[120]

Application of ACS Data in Methodology: EIA benchmarks to occupied housing totals from the American Community Survey, using the Census Bureau definition of a housing unit. It includes single-family homes, units in multi-family buildings, and mobiles. The RECS excludes vacant, seasonal or vacation homes, and group quarters such as prisons, military barracks, dormitories, and nursing homes.

ACS Questions Utilized (as suggested by the methodology): None. The RECS benchmarks to the occupied housing estimates provided by the ACS.

Major Uses: RECS data users rely on quality household consumption data from EIA to determine appliance standards, impacts of new building technologies, and effects of energy policies. RECS stakeholders include:

---

[120] Text sources: "Residential Energy Consumption Survey (RECS) 2009 Technical Documentation-Summary," Energy Information Administration, January 2013, and "How does EIA estimate energy consumption and end uses in U.S. homes?," EIA website.

- National Energy Modeling System (NEMS)—Office of Integrated Analysis and Forecasting (OIAF), EIA
- Low Income Home Energy Assistance Program (LIHEAP)-U.S. Department of Health and Human Services, Administration for Children and Families (HHS/ACF): LIHEAP distributes almost two billion dollars per year to assist low-income households in meeting the costs of home heating and cooling. Specifically RECS data are used for the LIHEAP Home Energy Notebook and Annual Report to Congress. Since 1981, HHS/ACF has supported the RECS by funding a set of questions added to the household interview and, in some years, by funding additional sample cases in low-income areas. These questions allow HHS/ACF to identify LIHEAP households in the RECS data and track their use of energy and the burden of energy costs.
- Policy Formulation and Analysis—Department of Energy (DOE): The Office of Policy (PO) extensively uses the RECS data in a variety of analytical studies. These studies have used RECS data to arrive at national estimates of energy savings for various policy options within DOE; to assess the amount of energy used by heating and cooling equipment when setting efficiency standards; and to assess the potential for fuel switching and cogeneration.
- Office of Energy Efficiency and Renewable Energy (EERE) Programs, DOE:
  - EERE's Appliances and Commercial Equipment Standards Program develops test procedures and minimum efficiency standards for residential appliances and commercial equipment. As an example, the program uses the annual RECS end-use consumption estimates to determine whether efficiency improvements have an adequate payback time for consumers. The Energy Star and Building America are other EERE-sponsored programs.
  - The EERE Buildings Energy Data Book provides a consistent and readily accessible set of core residential building data that are used by a wide range of energy analysts in their analysis of energy use in residential sectors. RECS data are used as key input for the consumption, expenditure, and household characteristics tables.
  - The Weatherization and Intergovernmental Program uses RECS consumption data to track the efficiencies of newly-constructed housing units, and other weatherization data track longitudinal changes in conservation measures as well as the energy burden for low-income households.
  - National Renewable Energy Laboratory (NREL) and the Residential Buildings Program within EERE developed the Building America Research Benchmark in consultation with the Building America industry teams. The RECS data are used to analyze relationships between various household characteristics and energy consumption.
  - RECS data supports the development of Building Codes. DOE works with other government agencies, state and local jurisdictions, national code organizations, and industry to promote stronger building energy codes and help states adopt, implement, and enforce those codes.
- U.S. Census Bureau: EIA has provided extensive analysis of RECS data to the Census Bureau as part of an investigation into the use of consumption-based measures of

poverty using expenditures and other indicators of material well-being. The U.S. Census Bureau also uses the RECS data to adjust for the over reporting of electricity and gas costs by American Housing Survey respondents.

- Lawrence Berkeley National Laboratory (LBNL)-Appliance Standards: RECS data are used by LBNL for analyzing impacts from possible energy efficiency standards for common household appliances, such as refrigerators and dishwashers, and emerging efficiency technologies in home electronics. LBNL relies on RECS for information about the age, size and usage of appliances and electronics.
- National Renewable Energy Laboratory (NREL): NREL uses RECS data to understand the adoption rates and impact of the new technologies, building designs, and energy-efficient equipment they test and promote. The laboratory reports it needs much larger RECS sample sizes to perform necessary multivariate analyses.
- National Association of Home Builders (NAHB): NAHB indicates that a larger RECS sample will improve the accuracy of estimates showing much consumption is explained by householder behavior versus that which the builder can control.
- U.S. Bureau of Labor Statistics (BLS)-Consumer Price Index: The Bureau of Labor Statistics (BLS) uses RECS data in the preparation of the Consumer Price Index (CPI). BLS uses the RECS micro-data file to develop equations for imputing utility costs for renters whose utility costs are included in their rent. BLS has automated this imputation process for approximately one-fifth of the renter sample used for the CPI. As a result of this process, BLS does not have to field an additional survey to collect these important data.
- Environmental Protection Agency (EPA)-Energy Star: EPA uses RECS data to support their programs and identify new products that have the potential as an Energy Star products. As an example, EPA used RECS data for Energy Star market analysis on programmable thermostats.
- Public utilities, interest groups, trade associations, state and local governments, equipment manufacturers, media, and the general public are also major users of RECS data.[121]

---

[121] Energy Information Administration, "Supporting Statement for Forms EIA-457 A-G Residential Energy Consumption Survey," September 2009.

VII. U.S. Department of Health and Human Services

A. Administration for Children and Families

39. Low Income Home Energy Assistance Program – Special Tabulations

Purpose of the Dataset: The Office of Community Services (OCS) in the Administration for Children and Families (ACF), Department of Health and Human Services (HHS), administers the Low Income Home Energy Assistance Program (LIHEAP), which provides financial assistance to low-income families for their home energy bills, energy crises, and energy-related minor home repairs. The Census Bureau provides OCS with two custom tabulations of ACS data that allow it to equitably allocate LIHEAP funds to Native American tribal grantees.

Responsible Unit: OCS, ACF, HHS

Authorization: Section 2604(d) of the LIHEAP statute [42 U.S.C. § 8623(d)]

Frequency of Data Publication: The special ACS tabulations were prepared for the first time in 2014, supplanting measures drawn from the SF3 file of the 2000 Census.[122] Presumable, OCS will ask the Census Bureau to prepare a set of new tabulations each year, but this is not spelled out in the OCS materials currently online.

Timing of Data Release: The OCS memo describing the new data, based on special tabulations of 2007-2011 five-year ACS estimates, was published April 3, 2014.

Modes to Access Data: The special ACS tabulations do not appear on the Internet. However, their application is in the OCS memo indicating the allocation of LIHEAP funds to tribes and tribal organizations.[123]

Methodology/Application of ACS Data in Methodology: The U.S. Census Bureau used the American Community Survey to complete two custom tabulations for the Office of Community Services (OCS), using five-year estimates for 2007-2011:

- the number of LIHEAP income-eligible households living in each of the 50 States and the District of Columbia; and
- the number of Native American LIHEAP income-eligible households living on reservations and off-reservation trust lands.

The five-year series was chosen because it provides the lowest margin of error for counting small communities, such as populations living on reservations. Prior to 2014, the OCS relied on the 2000 Census SF3 file.

---

[122] Text source: OCS, "LIHEAP Information Memoranda 2014 American Community Survey Tabulation," April 3, 2014.

[123] OCS, "LIHEAP Initial CR Release of Block Grant Funds to Indian Tribes and Tribal Organizations 2015," October 10, 2014.

To determine the LIHEAP allocation for tribes or tribal organizations that wish to become directly-funded LIHEAP grantees, the OCS first bases their decision on agreement between the tribe and the state in which the tribe's reservation is located. The agreement represents a mutually-agreed upon portion of the state's gross LIHEAP allocation that would be granted by the U.S. Department of Health and Human Services directly to the tribe.

In the absence of such an agreement, OCS relies on ACS special tabulation data to determine the LIHEAP allocation to a tribe. Specifically, it divides the number of Native American LIHEAP income-eligible households living on the tribe's reservation by the total number of LIHEAP income-eligible households residing within the state in which the reservation is located. This calculation results in a percentage, which is applied to the state's gross LIHEAP allocation to determine the LIHEAP award for the tribe.

For the purposes of this calculation, LIHEAP income-eligibility is defined as the greater of 150% of the 2012 Federal Poverty Guidelines published by HHS or 60% of each State's median income, as published by OCS for optional use by LIHEAP grantees in FY 2013 and mandatory use in FY 2014.[124]

OCS draws its state median income series from three-year ACS data.[125]

ACS Questions Utilized (as suggested by the methodology): P47-48

Major Uses: The ACS special tabulations are used to determine LIHEAP funding allocations to Native American tribes and tribal organizations.

---

[124] Text source: OCS LIHEAP ACS memorandum, April 3, 2014.
[125] OCS, "State Median Income Estimates for Optional Use in Federal Fiscal Year 2014 LIHEAP Programs and Mandatory Use in Federal Fiscal Year 2015 LIHEAP Programs," July 23, 2014.

B. Administration on Aging

40. AGing Integrated Database (AGID) – Special ACS Tabulations

Purpose of the Database: The Administration on Aging (AoA) compiles statistics on the nation's older population, namely persons aged 65 years or older. These statistics are made available through the AGing Integrated Database (AGID) system, an on-line query system that provides access to AoA-related program performance results, surveys and other data files. The purpose of the system is to:

- Provide a single, user friendly source for a variety of information on AoA supportive services and comprehensive systems of care for older people and their caregivers
- Allow users to quickly produce tables, maps, and other summary information from AoA-related data files and surveys, supplemented by Census-based population and demographic characteristics
- Provide users full access to results from national surveys of recipients of Older Americans Act services and AoA special tabulations produced by the Census Bureau, including of the ACS[126]

Responsible Unit: Administration on Aging (AoA), Department of Health and Human Services

Authorization: Older Americans Act of 1965, as amended, Section 202. In Section 202(2), it is noted that "It shall be the duty and function of the Administration to collect and disseminate information related to problems of the aged and aging." Section 202(5) notes that "It shall be the duty and function of the Administration to develop plans, conduct and arrange for research in the field of aging, and assist in the establishment and implementation of programs designed to meet the needs of older individuals…" Finally, Section 202(16) notes that "It shall be the duty and function of the Administration to collect for each fiscal year, for fiscal years beginning after September 30, 1988, directly or by contract, statistical data regarding programs and activities carried out with funds provided under this Act…"

Frequency of Data Publication: As new data resources become available. ACS data updates are released annually.

Timing of Data Release: As announced in the Release Notes, AGing Integrated Database website. Each ACS special tabulation is produced one or two calendar years after the regular ACS release on which it is based. For instance, the 2008-2012 special ACS tabulation was released in November 2014.

Modes to Access Data: AGing Integrated Database, AoA (xls, xml, csv, ascii). Modes of access to ACS data specifically include:

- Data-at-a-Glance (ACS)
- State Profiles
- American Community Survey (ACS) Demographic Data Custom Tables

---

[126] Text source: "About AGID: Resources," AoA website.

- [Special Tabulation Files](#)

Methodology: The AGID has four options or paths that provide different levels of focus and aggregation of the data – from individual data elements within Data-at-a-Glance, to State-level summaries in State Profiles, to detailed, multi-year tables in Custom Tables, and finally, to full database access within Data Files.

- Data-at-a-Glance: Single data element at a time access to all of AGID's state-level databases. Excellent tool for data mining or exploration of the various databases and their content, along with producing quick tables and geographical representations of key data elements.
- State Profiles: Provides pre-populated tables of key data elements from OAA Programs for the selected state. The user also has the ability to make comparisons between one state and another state or one state and the total U.S. In addition, the location of the State Unit on Aging (SUA), Area Agencies on Aging (AAA), and Tribal Organizations (where applicable) are displayed in both map and tabular form.
- Custom Tables: Powerful tool for producing detailed, multi-year tables. Users have the ability to select only those data elements applicable to their needs, and to further refine their results based on demographic stratifiers or geographic locations that are meaningful to their application. In addition, multiple years of data can be selected to analyze trends across time, while simple sorting tools have been incorporated to rank individual data elements across both time and geography.
- Data Files: Provides access to the AoA survey databases and AoA Special Tabulations conducted by the Census Bureau. The individual survey data files are provided in CSV or SAS format and are supplemented by survey instruments, reporting requirements documents, codebooks with variable listings and frequency counts and percentages of all individual data elements, and SAS programming statements for loading and processing the data in SAS. AoA Special Tabulation results are provided in both Excel and XML format.

The databases that are currently available in the system are listed below. These files are updated on a periodic basis as new data become available.

- AOA-Related Files
  - State Program Reports (SPR)
  - National Ombudsman Reporting System (NORS)
  - Title VI Services by Tribal Organization
  - National Survey of Older Americans Act (OAA) Participants
  - National Survey of Area Agencies on Aging (AAA)
- Census Files
  - State Level Population Estimates Data
  - County and PSA Level Population Estimates Data
  - Decennial Census 2010 Summary File 1 (SF1) Data
  - American Community Survey (ACS) Public Use Microdata Sample (PUMS) 1-Year Files

- o AoA Special Tabulation: American Community Survey 2008-2012
- o AoA Special Tabulation: American Community Survey 2007-2011
- o AoA Special Tabulation: American Community Survey 2009-2011, Disability Tables
- o AoA Special Tabulation: Decennial Census 2010
- o AoA Special Tabulation: American Community Survey 2005-2009
- o AoA Special Tabulation: American Community Survey 2008-2010, Disability Tables[127]

Application of ACS Data in Methodology: Each year, the Census Bureau provides AGID with two special ACS tabulations.

- Five-year ACS Special Tabulation on Aging – Population Characteristics (latest on AGID is 2008-2012). The data are customized for geography, in that they include data by AoA's planning and service areas (PSAs). Content includes detailed person-level and family-level tables at some or all geography levels, including diverse topics such as age, sex, and race/ethnicity; living alone, marital status, householder relationship/family type; education, poverty, employment, income, and Social Security income; grandparents responsible for grandchildren; and citizenship, migration, and English language proficiency.
- Three-year ACS Special Tabulation on Disability and Health Insurance Coverage (latest is 2009-2011). Data are provided for civilian non-institutionalized population age 60 years and over with sensory, cognitive, and physical disabilities at the national, state, and county levels.

ACS Questions Utilized: P3-8, P11, P14-20, P25-26, P29, P35-37, P41-42, and P47-48.

Major Uses: ACS special tabulation data on aging are used to serve as a component in OAA allocation formulas, to plan programs and services for older adults, and to compare populations within and across state boundaries.

---

[127] Text source: "About AGID," AoA website.

C. Agency for Toxic Substances and Disease Registry

41. *Social Vulnerability Index*

Nature and Purpose of the Measure: The Social Vulnerability Index (SVI) is a tool to help emergency response planners and public health officials identify and map the communities that will most likely need support before, during, and after a hazardous event. It uses U.S. Census data to determine the social vulnerability of every Census tract. Census tracts are subdivisions of counties for which the Census collects statistical data. The SVI ranks each tract on 14 social factors, including poverty, lack of vehicle access, and crowded housing, and groups them into four related themes—socioeconomic status, household composition, race/ethnicity/language, and housing/transportation. Each tract receives a separate ranking for each of the four themes, as well as an overall ranking.[128]

The index helps state, local, and tribal disaster management officials identify the locations of their most vulnerable populations. It is particularly intended for use by Office of Public Health Preparedness and Response (OPHPR)-funded state partners in all phases of the disaster cycle.

Responsible Unit: Geospatial Research, Analysis & Services Program (GRASP), Division of Toxicology and Human Health Sciences (DTHHS), Agency for Toxic Substances and Disease Registry (ATSDR), Centers for Disease Control and Prevention (CDC), Department of Health and Human Services (HHS), in cooperation with CDC's National Center for Environmental Health and its Office of Public Health Preparedness and Response (OPHPR).

The responsibility of OPHPR is to prepare the nation for and respond to urgent threats to the public's health. OPHPR's mission is to safeguard health and save lives by providing a platform for public health preparedness and emergency response.

Authorization:

- ATSDR: Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA), Resource Conservation and Recovery Act of 1976 (RCRA), and Superfund Amendments and Reauthorization Act of 1986 (SARA)
- OPHPR: 42 USC § 247d-6 (Public Health Emergencies)

Frequency of Data Publication: The SVI has had two releases, for 2000 and 2010.

Timing of Data Release: The 2010 SVI was released in 2012.

Modes of Data Access:

- The SVI data are in Esri's geodatabase format (mdb). These files may be used in Microsoft Access; Quantum GIS (QGIS), a free, Open Source Geographic Information System (GIS); or in ArcGIS 10.0 or higher.
- SVI Mapping Dashboard, ATDSR website

---

[128] Text source: ATSDR, "The Social Vulnerability Index (SVI)."

Methodology/Application of ACS in the Methodology:

2010 tract level data—Census 2010 100% count data (SF1) for the following variables:

- Persons aged 65 and older
- Persons aged 17 and younger
- Single parent households with children under 18
- Minority status (i.e. Total population minus white, non-Hispanic population)
- Persons living in Group Quarters

American Community Survey (ACS), 2006-2010 (5-year) data for the following variables/estimates:

- Persons below the poverty level
- Civilian unemployed
- Per capita income
- No high school diploma for persons aged 25 and older
- Persons who speak English "less than well"
- Housing units with 10 or more units in the structure
- At the household level, more people than rooms
- Mobile homes
- No vehicle access

Raw data estimates for each variable, for each tract, are included in the database. In addition, the margins of error (MOEs) for each estimate are also included.

The US Census Bureau did not collect tract level disability data, included in SVI 2000, for either the 2010 Census or the 2006-2010 ACS. Therefore, a disability variable is not included in SVI 2010.

We processed 2010 Census SF1 data as follows.

- We calculated the proportion for each variable for each tract (e.g. proportion of persons aged 65 and older) and included these proportions in the database.
- We used appropriate SF1 variables as denominators (e.g. total population to calculate proportion of persons aged 17 and younger).

Because of estimate error, the ACS data include additional data fields.

- Margins of error (MOEs) are included for each estimate, including derived estimates. We calculated MOEs for derived estimates using Census specifications. The confidence level is at the Census standard of 90%.
- We used appropriate ACS estimates as denominators (e.g. total population estimate to calculate the proportion of persons who speak English "less than well").
- Note: Confidence intervals can be calculated by subtracting the MOE from the estimate (lower limit) and adding the MOE to the estimate (upper limit).

We ranked Census tracts within each state and the District of Columbia, to enable mapping and analysis of relative vulnerability in individual states. We also ranked tracts for the entire United States against one another, for mapping and analysis of relative vulnerability in multiple states, or across the U.S. as a whole. Tract rankings are based on percentiles, as for SVI 2000. Percentile ranking values range from 0 to 1, with higher values indicating greater vulnerability.

For each tract, we generated percentile rankings for 1) the fourteen individual variables, 2) the four themes, and 3) an overall ranking.

Theme rankings: For each of the four themes, we summed the percentiles for the variables comprising each theme. We ordered the summed percentiles for each theme to determine theme-specific percentile rankings.

Overall tract rankings: We summed the sums for each theme, ordered the tracts, and then calculated overall percentile rankings. Please note; taking the sum of the sums for each theme is the same as summing individual variable rankings.

Tracts in the top 10%, i.e. at the 90th percentile of values, are given a value of 1 to indicate high vulnerability. Tracts below the 90th percentile are given a value of 0.

For a theme, the flag value is the number of flags for variables comprising the theme. We calculated the overall flag value for each tract as the number of all variable flags. [129]

ACS Questions Utilized (as suggested by the methodology): number in household, P4, P11, P14, P29, P35-37, P47-48, H1, H7, H12

Major Uses: The SVI helps public health officials and local planners better prepare for and respond to emergency events like hurricanes, disease outbreaks, or exposure to dangerous chemicals. The SVI is used to:

- Estimate the amount of needed supplies like food, water, medicine, and bedding.
- Help decide how many emergency personnel are required to assist people.
- Identify areas in need of emergency shelters.
- Plan the best way to evacuate people, accounting for those who have special needs, such as people without vehicles, the elderly, or people who do not understand English well.
- Identify communities that will need continued support to recover following an emergency or natural disaster.[130]

State, local, and tribal agencies are most knowledgeable about the people in their communities. The social vulnerability index is designed to aid them in their efforts to ensure the safety and well-being of their residents. The components of the SVI can assist state and local personnel concerned with all phases of the disaster cycle. Knowing the location of

---

[129] Text source: ATSDR, "SVI 2010 Documentation," September 2014.

[130] Text source: ATSDR, "The Social Vulnerability Index (SVI)."

socially vulnerable communities, planners can more effectively target and support community-based efforts to mitigate and prepare for disaster events. Responders can plan more efficient evacuation of those people who might need transportation or special assistance, such as those without vehicles, the elderly, or residents who do not speak English well. Local governments can identify neighborhoods that may need additional human services support in the recovery phase or as a mitigating measure to prevent the need for the costs associated with post-response support. The Katrina case study illustrated how the SVI can be used as part of the risk equation in the response and recovery phases. The elderly were particularly vulnerable during this event. Moreover, areas that are slower to recover include those that were heavily flooded and those with socioeconomically vulnerable populations. Future case studies will explore how the SVI can be used as part of the equation in the preparedness and mitigation phases to aid in targeting disaster management interventions.[131]

---

[131] Text source: Flanagan, Barry E.; Gregory, Edward W.; Hallisey, Elaine J.; Heitgerd, Janet L.; and Lewis, Brian (2011) "A Social Vulnerability Index for Disaster Management," Journal of Homeland Security and Emergency Management: Vol. 8: Iss. 1, Article 3, p. 14.

D. Center for Medicare and Medicaid Services

42. *Culturally and Linguistically Appropriate Standards*

Nature and Purpose of the Measure: The Public Health Service Act requires that certain information on group health plans and health insurers be available in a "culturally and linguistically appropriate manner." The threshold percentage for the group and individual markets is that at least 10 percent of the population in the claimant's county speak a particular non-English language and speak English less than "very well" as determined by American Community Survey (ACS) data published by the United States Census Bureau.

Each year, the Center for Medicare and Medicaid Services (CMS) publishes a list of counties which meet or exceed the 10 percent threshold (rounded to the nearest percent).[132] Languages include Spanish, Chinese, Tagalog, and Navajo.

Responsible Unit: CMS, Department of Health and Human Services (HHS)

Authorization: Public Health Service Act, Sections 2715 and 2719

Frequency of Data Publication: Annual

Timing of Data Release: The Culturally and Linguistically Appropriate Standards (CLAS) County Data are published immediately after the annual release of the five-year ACS data (typically each December).

Modes of Data Access: County Data for Culturally and Linguistically Appropriate Services, CMS

Methodology/Application of ACS in the Methodology: CMS analyzes publicly available five-year ACS data by county to identify counties that meet or exceed the linguistic threshold.

ACS Questions Utilized (as suggested by the methodology): P14

Major Uses: The CLAS County Data allow health insurers to determine the counties in which they have to provide information in a particular non-English language.

---

[132] Text source: CMS, "Updated Instructions for Calculating County Level Estimates Pertaining to the Culturally and Linguistically Appropriate Standards Set Forth in the Internal Claims and Appeals and External Review Processes under the Affordable Care Act," July 24, 2013.

43. *Geographic Practice Cost Index—Medicare Physician Fee Schedule*

Nature and Purpose of the Measure: The Centers for Medicare and Medicaid Services (CMS) pays physicians for their services according to the Physician Fee Schedule (PFS), which specifies a set of allowable procedures and payments for each service. Each procedure is interpreted as being produced by a combination of three categories of inputs: physician work (PW), practice expense (PE), and malpractice insurance (MP). The particular blend of PW, PE, and MP inputs assessed to produce a service specifies its composition of relative value units (RVUs). A payment for a procedure depends on its assigned RVUs and the input prices assessed for each RVU component.

As mandated under, CMS must establish geographic indices as part of the Resource-Based Relative Value Scale (RBRVS) method for paying physicians. The geographic practice cost indices (GPCIs) account for geographic variation in the price of the PW, PE, and MP classes of inputs.

GPCIs measure geographic differences in input prices. Paralleling the RVU structure, GPCIs are split into three parts: PW, PE, and MP. Each of these three GPCIs adjusts its corresponding RVU component. In essence, GPCIs increase the price associated with an RVU in high cost regions and decrease the price associated with an RVU in low-cost regions. GPCIs are budget neutral and do not affect aggregate payment levels; rather, they reallocate payment rates by locality to reflect regional variation in relative input prices.

The three GPCIs are calculated for 89 localities. The localities are defined alternatively by state boundaries (e.g., Wisconsin), metropolitan statistical areas (MSAs) (e.g., Metropolitan St. Louis, MO), portions of an MSA (e.g., Manhattan), or rest-of-state area which exclude metropolitan areas (e.g., Rest of Missouri).[133]

Responsible Unit: CMS, Department of Health and Human Services (HHS)

Authorization: Section 1848(e) of the Social Security Act

Frequency of Data Publication: CMS first implemented the GPCIs as part of the Medicare PFS in 1992 and requires the GPCIs to be updated at least every three years.

Timing of Data Release: The Seventh Update of the GPCI took effect on January 1, 2015.[134]

Modes of Data Access: Physician Fee Schedule Federal Register Notices, CMS (xls, txt)

Methodology: The components and corresponding data sources of the three GPCIs are shown in the tables below.[135]

---

[133] Text source: MaCurdy et al., "Revised Final Report on the CY 2014 Update of the Geographic Practice Cost Index for the Medicare Physician Fee Schedule," Acumen, LLC, June 2014, pp. i-ii.

[134] CMS, "Medicare Program; Revisions to Payment Policies Under the Physician Fee Schedule, Clinical Laboratory Fee Schedule, Access to Identifiable Data for the Center for Medicare and Medicaid Innovation Models & Other Revisions to Part B for CY 2015," *Federal Register*, Vol. 79, No. 219, November 13, 2014.

[135] MaCurdy, *op.cit.*, pp. iii-iv.

Table 1: Breakdown of GPCIs into Current Component Indices

GPCI	Component Index	Measures Geographic Differences in:
Physician Work	Single Component	Physician wages
Practice Expense	*Employee Wage*	Wages of clinical and administrative office staff
	Purchased Services	Cost of contracted services (e.g., accounting, legal)
	Office Rent	Physician cost to rent office space
	Equipment, Supplies, and Other	Practice expenses for inputs such as chemicals and rubber, telephone use and postage
Malpractice	Single Component	Cost of professional liability insurance

Table 2: Overview of Updated Data Sources for the CY 2014 Update

GPCI Component Index	Current Regulation	CY 2014 Update
Physician Work GPCI	2006-2008 BLS Occupational Employment Statistics	2009-2011 BLS Occupational Employment Statistics
Practice Expense GPCI		
Employee Wage	2006-2008 BLS Occupational Employment Statistics	2009-2011 BLS Occupational Employment Statistics
Purchased Services	2006-2008 BLS Occupational Employment Statistics	2009-2011 BLS Occupational Employment Statistics
	CMS Labor-Related Classification	CMS Labor-Related Classification
Office Rent	2006-2008 American Community Survey	2008-2010 American Community Survey
Equipment, Supplies, and Other	1.00 for All Counties	1.00 for All Counties
Malpractice GPCI	2006-2007 Malpractice Premiums	2011-2012 Malpractice Premiums
Cost Share Weights	2006 MEI Weights	Reclassification of 2006 MEI Weights
County RVU Weights	2009 RVUs	2011 RVUs

Application of ACS in the Methodology: As is shown in Table 2, three-year average ACS data on two-bedroom apartment rent are used as a proxy measure for physician office rent, one component of the Practice Expense GPCI. The Census Bureau does a special tabulation for CMS to produce the needed numbers by the 89 localities.

The Sixth GPCI Update used 2006-2008 ACS data; the Seventh uses 2008-2010 ACS data.

ACS Questions Utilized (as suggested by the methodology): H7, H18

Major Uses: The GPCIs are used to adjust Medicare payments to physicians to reflect local costs relative to the nation.

E. Health Resources and Services Administration

44. *Area Health Resources Files*

Nature and Purpose of the Dataset: The Area Health Resources Files (AHRF) provide a comprehensive set of data offering a broad range of health resources and socioeconomic indicators which impact demand for health care. The AHRF provides current as well as historic data for more than 6,000 variables for each of the nation's counties and states. The AHRF contains information on health facilities, health professions, measures of resource scarcity, health status, economic activity, health training programs, and socioeconomic and environmental characteristics. In addition, the basic file contains geographic codes and descriptors which enable it to be linked to many other files and to aggregate counties into various geographic groupings.

The state and national AHRF include in-depth demographic, workforce, employment and training data for 50 health care professions. These data are not available for smaller geographic areas, but provide a broad perspective on the health resources available from a state and national perspective.

Key health and demographic data from the AHRF are available using several AHRF web tools:

- The Health Resources Comparison Tools (HRCT) provide users with an easy method to compare county and state measures of health resources and health care demand based on criteria and specific areas of interest to the user. Two different tools are available:
  - The county HRCT identifies similar counties for the user, and permits refinement of counties selected based on key demographic indicators.
  - The state HRCT provides comparison of selected state health resources and indicators of health care demand for up to five states selected by the user.
- The AHRF Map Tool displays select health resource data as well as demographic and environmental statistics impacting health status.[136]

Responsible Unit: National Center for Health Workforce Analysis (NCHWA), Bureau of Health Workforce, Health Resources and Services Administration, Department of Health and Human Services (HHS)

Authorization: 42 U.S. Code § 294n - Health professions workforce information and analysis. The NCHWA was created by Congress "to provide necessary information for decision-making regarding future directions in health professions and nursing programs in response to societal and professional needs."

Frequency of Data Publication: The AHRFs are released annually. The 2013-2014 AHRF was released in September 2014 and is the latest release. In addition to the county AHRF, annual release of the state and national AHRF began in September 2013.

---

[136] Text source: "AHRF Overview," NCHWA webpage.

Timing of Data Release: The difference between the time data are collected and the time they are available in the AHRFs varies by data source:

Modes of Data Access:

- [AHRF Download](#), NCHWA (ascii, accdb)
- [Health Resources Comparison Tool](#), NCHWA (html)
- [AHRF Mapping Tool](#), NCHWA

Methodology: The AHRF integrates data from numerous primary data sources including: the American Hospital Association, the American Medical Association, the American Dental Association, the American Osteopathic Association, the Bureau of the Census, the Centers for Medicare and Medicaid Services (formerly Health Care Financing Administration), the Bureau of Labor Statistics, the CDC/National Center for Health Statistics, the Environmental Protection Agency and the Veteran's Administration.[137]

Application of ACS in the Methodology: Data directly derived from the ACS include:

- Percent Females Divorced and Numbered Divorced Females
- Foreign Born Population and Percent Foreign Born
- Non-English Speaking Persons by age
- Employment and Labor Force (employment status, journey to work mode, veterans status, industry, work location, mean travel time)
- Persons and Families below Poverty Level
- Estimates of Persons in Poverty; Estimate of Persons Age 0-17 in Poverty; and Estimate of Children 5-17 in Families in Poverty
- Persons in Deep Poverty
- Ratio of Income to Poverty Level
- Median Family Income
- Household Income and Median Household Income
- Persons with and without Health Insurance, and Percent without Health Insurance by age, gender, and income-to-poverty ratio
- Disability by Age, Employment Status and Veteran Status
- Households with Social Security Income
- Households with Supplemental Security Income (SSI) and Households with Public Assistance Income
- Persons age 25 years or more and persons aged 25 years or more with less than a high school diploma, with high school diploma or more and with four years of college or more
- Housing Data (type of housing, median home value, gross rent)

---

[137] Detailed methodological documentation is available here: http://ahrf.hrsa.gov/download.htm.

The AHRFs also include a number of federal statistical products based in whole or part on the ACS, such as USDA Economic Research Service county typologies and the Department of Veterans Affairs' veterans population projections.

ACS Questions Utilized (as suggested by the methodology): P2-7, P11, P14-20, P26, P29-31, P34-37, P41-44, P47-48, H1, H8, H13, H18-19

Major Uses: In terms of policy analysis, many potential applications are available using the AHRF, including: determining the availability of primary care services in a county (not only practitioners but FQHCs and CHCs), considering medical school graduates in a state with shortages to anticipate the future availability of physicians, considering the availability of specialty care for aged populations (e.g., the number of rehabilitation facilities in and near a county of interest).

With regard to workforce research, many aspects of the workforce can be addressed; for example, the aging of the primary care physician workforce is a matter of concern, in this case, the AHRF contains data on the distribution of physicians and other health professionals by age in each county, by specialty.

In terms of preparing grant applications, fundamental to any such application is the establishment of a need and a baseline for evaluation of performance - the AHRF can be used to establish a need (e.g. the number of RNs in an area, the number of primary care practitioners, the number of hospital beds per population), identify the recent experience in the topic area identified, and propose the use of the annual AHRF to assess performance given the grant resources. Also, the extensive demographic information in the AHRF can be used to accurately determine the population of a catchment area across the years to coincide with other data available to users (e.g., if the topic is infectious diseases in children, household income may be needed as a proxy variable at a particular point in time for a certain type of analysis).[138]

---

[138] Text source: "AHRF FAQs," NCHWA webpage.

45. *Health Professional Shortage Areas*

Nature and Purpose of the Measure: An Health Professional Shortage Area (HPSA) is a geographic area, population group, or health care facility that has been designated by the Federal government as having a shortage of health professionals. There are three categories of HPSAs: primary care (shortage of primary care clinicians), dental (shortage of oral health professionals), and mental health (shortage of mental health professionals).

HPSAs are designated using several criteria, including population-to-clinician ratios. This ratio is usually 3,500 to 1 for primary care, 5,000 to 1 for dental health care, and 30,000 to 1 for mental health care.[139]

As of June 19, 2014, there were currently approximately 6,100 designated Primary Care HPSAs, 4,900 Dental HPSAs, and 4,000 Mental Health HPSAs.[140]

HPSA designation is used to determine eligibility to a variety of federal health care programs. Each year, the U.S. Department of Health and Human Services prepares listings of designated HPSAs and areas that need to be updated to maintain their designation. The listings are sent to each State Primary Care Office (PCO). Copies are also sent to the Primary Care Associations and other interested parties. The PCOs have a few months to submit designation updates for their States. After review and consideration of all comments, the Secretary designates Health Professional Shortage Areas (HPSAs) and withdraws the designations of areas determined to no longer meet the criteria for designation.[141]

Responsible Unit: Shortage Designation Branch, Division of Policy and Shortage Designation, Bureau of Health Workforce (BHW), Health Resources and Services Administration (HRSA), Department of Health and Human Services (HHS)

Authorization: 42 U.S. Code § 254e - Health professional shortage areas

Frequency of Data Publication: Annual. In 2014, the HPSA list was current as of May 23.

HPSAs are currently updated on an ongoing basis based on the identification of new areas, population groups, facilities, and sites that meet the eligibility criteria or that no longer meet eligibility criteria and/or are being replaced by another type of designation. As such, additional HPSAs may have been designated by letter since that date. The appropriate agencies and individuals have been or will be notified of these actions by letter. These newly designated HPSAs will be included in the next publication of the HPSA list and are currently included in the daily updates posted on the HRSA Web site at http://www.hrsa.gov/shortage/find.html.[142]

---

[139] Text source: "Frequently Asked Questions," BHW webpage.
[140] "Shortage Designation: Health Professional Shortage Areas & Medically Underserved Areas/Populations," BHW webpage.
[141] Text source: "Health Professional Shortage Areas (HPSAs)," BHW webpage.
[142] Text source: HRSA, "Lists of Designated Primary Medical Care, Mental Health, and Dental Health Professional Shortage Areas," *Federal Register*, Vol. 79, No. 122, pp. 36075-36076.

Timing of Data Release: HRSA publishes notice of the list of HPSAs each June in the *Federal Register*.

Modes of Data Access:

- [County and County Equivalent Listing](#) (pdf)
- [HPSA by State & County](#) (html)
- [HPSA by Address](#) (html)
- [HPSA Data Download](#) (zip)
- [HRSA Data Warehouse](#) (various tools)

Methodology: Section 332 of the Public Health Service Act provides that the Secretary of Health and Human Services shall designate HPSAs based on criteria established by regulation. The authority for designation of HPSAs is delegated to the Bureau of Primary Health Care's Office of Shortage Designation (OSD). Criteria and the process used for designation of HPSAs were developed in accordance with the requirements of Section 332. HPSA designation is a prerequisite for participation in a number of Federal programs, including National Health Service Corps approved sites.

The HPSA criteria require three basic determinations for a geographic area request:

1. the geographic area involved must be rational for the delivery of health services,
2. a specified population-to- practitioner ratio representing shortage must be exceeded within the area, and
3. resources in contiguous areas must be shown to be overutilized, excessively distant, or otherwise inaccessible.

These criteria have been defined for shortage of primary medical care physicians, dentists, and mental health professionals. The particular level used to indicate primary medical care, dental, and mental health shortage is referenced in the Criteria for Designation of HPSAs, codified at [42 CFR Chapter 1, PART 5 - DESIGNATION OF HEALTH PROFESSIONAL(S) SHORTAGE AREAS, 10-1-93 edition.](#)

Where a geographic area does not meet the shortage criteria, but a population group within the area has access barriers, a population group designation may be possible. In such cases the population group and the access barriers must be defined/described, and the ratio of the number of persons in the population group to the number of practitioners serving it must be determined. These ratios are also referenced in the Criteria for Designation of HPSAs.

In some cases, facilities may be designated as HPSAs. This applies to correctional facilities and to State mental hospitals. In addition, public and non-profit private facilities located outside designated HPSAs may receive facility HPSA designation if they are shown to be accessible to and serving a designated geographic area or population group HPSA.

BHW manages the designation of HPSAs through the new Shortage Designation Management System (SDMS), which came into use in 2014.

Application of ACS in the Methodology: The American Community Survey (ACS) version 2012 is used in shortage designation analysis and scoring process. ACS data points are used by the Shortage Designation Management System (SDMS) to evaluate demographic and economic characteristics for proposed rational service areas (RSAs) and the disparity analysis for contiguous areas. ACS data points are also used to determine if a proposed RSA meets population eligibility requirements. Calculations are based on five-year ACS estimates. Data points utilized include:

- Total Resident Civilian Population
- % of Population that is:
    - aged 18 or under
    - over 65
    - female 15-44 years
    - at or below 100% of the Federal Poverty Level
    - at or below 200% of the Federal Poverty Level
    - Linguistic isolation
    - African American
    - American Indian/Eskimo/Alaska Native
    - Asian
    - White
    - Hispanic
    - Pacific Islander[143]

ACS Questions Utilized (as suggested by the methodology): P3-6, P14, P47-48

Major Uses: Federal Programs using HPSA Designations include:

- National Health Service Corps (Section 333 of the Public Health Service Act) - provides for assignment of federally-employed and/or service- obligated physicians, dentists, and other health professionals to designated HPSAs
- National Health Service Corps Scholarship Programs (Section 338A) - provides scholarships for training of health professionals who agree to serve in designated HPSAs through the NHSC or the private practice option
- National Health Service Corps Loan Repayment Program (Section 338B) - provides loan repayment to health professionals who agree to serve in the NHSC in HPSAs selected by the Secretary
- Rural Health Clinics Act (Public Law 95-210) - provides Medicare and Medicaid reimbursement for services provided by physician assistants and nurse-practitioners in clinics in rural HPSAs
- Medicare Incentive Payments for Physician's Services Furnished in HPSAs (Public Law 100-203, Section 4043, as amended) - CMS (formerly HCFA) gives 10 percent bonus payment for Medicare-reimbursable physician services provided within geographic HPSAs. This payment does not apply to population group HPSAs.

---

[143] Text source: Email to author from Kae Brickerd, Chief, Shortage Designation Branch, BHW, January 20, 2015.

- Higher "Customary Charges" for New Physicians in HPSAs (Public Law 100-203, Section 4047) - CMS (formerly HCFA) exempts new physicians opening practices in non-metropolitan geographic HPSAs from new Medicare limitations on "customary charges"
- Area Health Education Center Program (Section 781(a)(1)) - gives special consideration to centers that would serve HPSAs with higher percentages of underserved minorities; gives funding priority to centers providing substantial training experience in HPSAs
- Federal Employees Health Benefits Programs - provides reimbursement for non-physician services in States with high percentages of their population residing in HPSAs[144]
- Exchange Visitor Program for physicians with J-1 visas working in HPSAs.
- National Interest Waiver for Physicians (EB-2)

Numerous state health care programs use the HPSA designation as well.

---

[144] Text source: "Guidelines for Primary Medical Care/Dental HPSA Designation," BHW webpage.

46. *Health Workforce Projections*

Nature and Purpose of the Measure: Health workforce projections present the national supply of and demand for health professions by discipline using HRSA's Health Workforce Simulation Model (HWSM). The HWSM is an integrated microsimulation model that estimates the future demand for and supply of health care workers in multiple professions and care settings. It can provide state-level estimates and describe the effects of a policy option at any point in time within the projection period.

Responsible Unit: National Center for Health Workforce Analysis (NCHWA), Bureau of Health Workforce, Health Resources and Services Administration, Department of Health and Human Services (HHS).

Authorization: 42 U.S. Code § 294n - Health professions workforce information and analysis. The NCHWA was created by Congress "to provide necessary information for decision-making regarding future directions in health professions and nursing programs in response to societal and professional needs."

Frequency of Data Publication: The projections are a new activity. The first report, on primary care practitioners, was published in November 2013. Five reports were published in 2014—for nurses, pharmacists, therapy occupations, vision occupations, and non-primary care clinicians.

Timing of Data Release: NA

Modes of Data Access: U.S. Health Workforce Projections, NCHWA (pdf)

Methodology: While the nuances of modeling differ for individual health professions and medical specialties, the basic framework used within HWSM remains the same and consists of three components: 1) the model for supply of health professional; 2) the model for demand for health care services; and, 3) the staffing ratios that convert demand for services to demand for health care workers. Consistent with prevailing practice, the model assumes that supply equals demand in the base year.

To project the number and characteristics of future health care workers and service users, HWSM simulated individual-level data based on predicted probabilities estimated from the current or base year data. Depending on the predicted probabilities, individual records were simulated to age forward. The aged individual-level records were then aggregated to obtain the national or state- level projections. On the service use side, the current utilization rates by individual characteristics were applied to projected populations at the national and state levels

A number of elements contribute to the development of the model. To calculate supply, workforce decisions for future professionals are simulated based on provider characteristics (demographics), profession and specialty, and the features of the local or national economy (wages, unemployment rate). The major components of the supply model include:

1. A micro data file containing the characteristics of the current workforce in a given profession.

2. Estimates of the annual number and characteristics of newly trained workers entering a given profession.
3. Equations that describe workforce decisions, such as retirement and number of hours worked, based on current labor market factors. Predicted probabilities from these equations simulate labor supply decisions of future health care professionals.

HWSM simulates the demand for health care services based on individual characteristics of the U.S. population (demographics, socioeconomics, health behavior, and health status). Two major components of the demand model are:

1. A database that contains characteristics for each person in a representative sample of the current and projected population in each state through 2025.
2. Regression equations that relate health care use patterns by setting to a person's characteristics. Predicted probabilities from these equations are applied to simulate health care utilization of future populations.

The third component of HWSM is staffing ratios, which translate demand for health care services into projected demand for full time equivalent (FTE) providers in different care delivery settings. HWSM simulates demand for health care services in seven settings (emergency rooms, hospitals, provider offices, outpatient departments, home health, nursing homes, and residential facilities). Demands for specific services within a setting are combined with provider staffing ratios in that setting to estimate the demand for health care providers.

Consistent with recommended standards, HWSM consists of self-contained modules that describe different components of the health care system. HWSM runs using SAS (Statistical Application Software).

The model makes use of the most recent data available to date and can be updated with new data as they become available without changing the basic features of the model. [145]

Application of ACS in the Methodology: ACS data are used in multiple ways in the HWSM.

1) Modeling Supply of Health Professionals:

Workforce Participation: For individuals who were projected to still be alive, the probability that the person would be actively employed in the health occupation was estimated using ACS retirement rates. Since the ACS does not list the profession of individuals who have been retired for more than five years, profession-specific labor force participation rates were imputed for workers over age 50—many of whom may have retired more than five years ago. For these individuals who had been employed at some time during their adult life, activity rates were based on their level of education (less than baccalaureate, baccalaureate, or graduate degree). For professions where substantial portions fell into two or three of the education categories, weighted average of the participation rates were used. The probability that a person would be actively engaged in their profession was calculated as (1-Retirement Rate). People

---

[145] Text source: NCHWA, "Technical Documentation for Health Resources Service Administration's Health Workforce Simulation Model," 2014, pp. 1-3, 24.

sometimes change professions or further their education to enhance career opportunities. When this happens, HWSM treats these as exits from the original profession and entrants to the new profession. One limitation of the ACS is the inability to discern career changes. While this phenomenon is more common for professions with low barriers to entry and exit, data limitations did not allow this aspect to be built into HWSM. The only profession for which career progression is built into HWSM is the nursing profession.

Hours Worked and FTE Supply: Ordinary Least Squares regressions on 2006-2011 ACS data were used to derive the expected number of hours worked in a week by each individual active in the profession. Explanatory variables included age, sex, log of hourly earnings, the overall unemployment rate, and a year indicator.

2) Modeling Demand for Health Professionals:

The microsimulation approach—where demand for health care services is modeled for each individual—requires individual level (micro) data on the predictors of health care use for each person in a representative sample in a designated geographic area in HWSM, the state, or the nation. The core micro data file that forms HWSM's baseline population was the 2011 American Community Survey (ACS) combined with the Behavioral Risk Factor Surveillance System (BRFSS) and the National Nursing Home Survey (NNHS).

The HWSM population database used a statistical matching process that combined patient health information from the BRFSS and NNHS with the larger ACS file that had a representative population in each state (and for some sub-state levels). Using information on residence type, the ACS population was stratified into those residing in nursing facilities (matched to people in the NNHS), and those not residing in nursing facilities (matched to people in BRFSS).

For the non-institutionalized population, each individual in the ACS was matched with someone in the BRFSS from the same sex, age group, race, ethnicity, insured/uninsured status, household income level, and state of residence.

Developing demand forecasts for future years required the creation of micro data sets for future populations. This was done by assigning new sample weights to ACS respondents so that when these weights were used, the file produced population estimates that mirrored Census Bureau projections by demographic groups (age group, sex, race and ethnicity) at the national level and population projections estimated by state governments.

3) Application of HWSM to Project Supply and Demand for Specific Occupations

   a) Nurses

Estimates of the current supply of RNs and LPNs in each state came from the pooled 2006-2012 ACS files. Multiple years of data were combined to increase the sample size required to get stable estimate of the distribution of nurses by state, age, sex, and education level. The ACS sample weights for each nurse were recalibrated to sum to the national totals of RNs and LPNs in the 2012 ACS.

Labor force participation rates for nurses were calculated directly for individuals through age 50. Activity rates based on the highest educational attainment were used for nurses over age 50. ACS data were used to determine the highest level of education (less than baccalaureate degree, baccalaureate degree, and graduate degree) and to calculate labor force participation rates of each group of nurses. RNs over the age of 50 with baccalaureate education were assumed to have similar labor force participation rates as other women with a baccalaureate degree. For RNs educated at the associate level and for LPNs, the HWSM used activity rates of women educated at the associate level.

Forecasting equations related average hours worked to nurse age, sex, education level, state overall unemployment rate, and average wage in the profession. Data for all variables came from the ACS with the exception of average wage, which was obtained from the BLS.

b) Health Care Support and Technical Occupations' Model

The base year counts for the occupations in this section came from pooled 2006-2012 ACS. When small sample size in ACS resulted in unreliable estimates, information from BLS' Occupational Employment Statistics (OES) was used to calibrate ACS data. Data from multiple years of the ACS were pooled and calibrated to 2012 national estimates to provide more stable estimates of the age and sex distribution of workforce. Because these occupations were projected only at the national level, no other characteristics were attached to the ACS data file.

Using data from 2006-2012 ACS, the age-sex specific probability that individuals would remain active in their occupation was estimated from a logistic regression equation. In addition to age and sex, education attainment was used in the prediction equation for those over age 50. Since many of the health care support and technical occupations showed representation from multiple educational groups, weights were created in HWSM that blended the proportions of workers in each category to reflect the attrition rate for those over age 50. The predicted probabilities were applied to the starting year supply of professionals in those occupations to simulate individuals who were expected to leave the occupation over the year.

Data from 2006-2011 ACS were used to derive the number of hours each individual spent in professional activities. Explanatory variables included age, sex, unemployment rate, and expected hourly earnings. The BLS estimates of the average wage for each occupation and the overall unemployment rate in each year were incorporated in the model so that wages and unemployment rates varied by year.[146]

ACS Questions Utilized (as suggested by the methodology): P3-6, P11, P16, P29, P35-40, P43-48, GQ

Major Uses: The projections are used to ensure that the nation is producing a sufficient number of health care professionals to fully meet consumer demand.

---

[146] Text source: *Ibid.*

47. HRSA Data Warehouse

Nature and Purpose of the Measure: The HRSA Data Warehouse (HDW) serves as the enterprise repository for HRSA's data and makes that data available to the public. The data warehouse integrates this data with external sources, such as the U.S. Census Bureau, enabling users to gather relevant and meaningful information about health care programs and the populations they serve. The data are updated frequently. The data warehouse allows users to explore, interact with, and export data through charts, maps, reports, analyzer and locator tools, data downloads and data services, widgets, and the data portal. The HDW is used by the public, HRSA professionals, grantees, health care providers and other audiences interested in HRSA's public health services and information.[147]

HDW topics include:

- Active Grants
- Area Health Resources Files
- Health Care Service Delivery and Look-Alike Sites
- Health Professional Shortage Areas
- Health Professions
- Medically Underserved Area/Population
- National Health Service Corps
- National Sample Surveys of Nurses
- Primary Care Service Areas
- Rural Health Care

Responsible Unit: Division of Data and Information Services, Office of Information Technology, Health Resources and Services Administration (HRSA), Department of Health and Human Services (HHS)

Authorization: Section 301 of the Public Health Service Act.

Frequency of Data Publication/Timing of Data Release: The HDW contains data from multiple source systems within HRSA and from external sources. These data are refreshed as the source data are available. Information on data refresh dates, refresh cycle, and date of source data are available online.[148]

Modes of Data Access: The HDW offers data access through:

- Download (xls, csv)
- Map services – a web-based method for end users to access geographic (map) data

---

[147] Text source: "Welcome to the HRSA Data Warehouse" and "About the HRSA Data Warehouse," HRSA Data Warehouse webpages.
[148] See "Data Refresh Dates," HRSA Data Warehouse webpage.

- Web services – a standardized way of integrating web-based applications using Extensible Markup Language (XML), Simple Object Access Protocol (SOAP), and Web Services Description Language (WSDL)
- HRSA in Your State Data Explorer
- Data portal – enables users to select related datasets to create custom data views)
- Quick reports – preformatted reports that provide quick access to summary data.

Methodology: The HDW manages data from more than 30 HRSA and other federal government sources. Data are refreshed on a regular basis, with frequency varying by source.[149]

The Data Warehouse cleanses and standardizes data, applies established business rules to validate the data, and enriches and expands the data available from the sources.[150]

Application of ACS in the Methodology: The American Community Survey (ACS) is assembled and included in the HRSA Data Warehouse in order to provide context for HRSA's activities, and to provide input data for HRSA's analytical activities. The HDW makes use of five-year ACS estimates for the number of families in poverty, with income below 1.5 times the federal poverty level, and with income 2.0 times the federal poverty level. Poverty data are available by state and county.[151]

ACS Questions Utilized (as suggested by the methodology): P2, P47-48

Major Uses: The HDW is used by HRSA, grantees, care providers, the public, and other audiences interested in HRSA's public health services and information.

The HRSA Data Warehouse is the official repository for current HRSA data and promotes maximum operating efficiency through centralization, reconciliation, and standardization of data across HRSA's various transactional business systems. The Data Warehouse supports the Shared Services Strategy and is offered as a Shared Service to other HHS Operating Divisions.

The Data Warehouse promotes "Open Data" by providing HRSA and the general public with a single source of HRSA programmatic information, related health resources, demographic, and statistical data for analyzing and reporting on HRSA activities with easily accessible, readily available pre-designed tools, charts, maps, and reports.[152]

---

[149] See *Ibid*.
[150] HHS, "Fiscal Year 2015, Health Resources and Services Administration Justification of Estimates for Appropriations Committees," 2014, p. 399.
[151] HRSA, "Data Sources," 2014 plus review of HDW data tools.
[152] HHS, *op.cit.*

48. *Medically Underserved Areas/Populations*

Nature and Purpose of the Measure: The federal Medically Underserved Area (MUA) and Medically Underserved Population (MUP) designations identify areas and populations that have limited access to primary care services. MUAs include groups of census tracts that have a population-to-provider ratio indicating a shortage. Medically Underserved Populations (MUPs) may include groups of persons who face economic, cultural or linguistic barriers to health care and reside in a specific geographic area. MUA/P designations are used to qualify for state/local and federal programs aimed at increasing health services to underserved areas and populations.[153]

At latest count, there are 4195 MUA/P designations, of which 3580 are MUAs, 400 are MUPs, and 215 are Governor-designated.[154]

Responsible Unit: Shortage Designation Branch, Division of Policy and Shortage Designation, Bureau of Health Workforce (BHW), Health Resources and Services Administration (HRSA), Department of Health and Human Services (HHS)

Authorization: Section 330(b)(3) of the Public Health Service Act.

Frequency of Data Publication: The list of MUA/Ps is updated daily.[155]

Timing of Data Release: As updated.

Modes of Data Access:

- [Find Shortage Areas: MUA/P by State and County](#) (html)
- [Find Shortage Areas: HPSA & MUA/P by Address](#) (html)
- [Shortage Areas](#), HRSA Data Warehouse (html, pdf, csv)

Methodology: The Index of Medical Underservice (IMU) is applied to data on a service area to obtain a score for the area. The IMU scale is from 0 to 100, where 0 represents completely underserved and 100 represents best served or least underserved.

Under the established criteria, each service area or population group found to have an IMU of 62.0 or less qualifies for designation as a Medically Underserved Area or Medically Underserved Population.

---

[153] Text source: ["Medically Underserved Areas – Populations,"](#) Department of Health, Government of the District of Columbia webpage.

[154] Text source: ["Medically Underserved Areas/Populations (MUA/P): State Summary of Designated MUA/P,"](#) HRSA Data Warehouse, webpage. Regarding Governor-designated MUA/Ps,: "Under the provisions of Public law 99-280, enacted in 1986, a population group which does not meet the established criteria of an IMU less than 62.0 can nevertheless be considered for designation if 'unusual local conditions which are a barrier to access to or the availability of personal health services' exist and are documented, and if such a designation is recommended by the chief executive officer and local officials of the State where the requested population resides." Source: ["Medically Underserved Areas/Populations: Guidelines for MUA and MUP Designation,"](#) BHW webpage.

[155] ["Data Refresh Dates,"](#) HRSA Data Warehouse webpage.

The IMU involves four variables - ratio of primary medical care physicians per 1,000 population, infant mortality rate, percentage of the population with incomes below the poverty level, and percentage of the population age 65 or over. The value of each of these variables for the service area is converted to a weighted value, according to established criteria. The four values are summed to obtain the area's IMU score.[156]

For MUP designations, the IMU methodology is applied to data on an underserved population group within an area of residence to obtain a score for the population group. Population groups requested for MUP designation should be those with economic barriers (low-income or Medicaid-eligible populations), or cultural and/or linguistic access barriers to primary medical care services.

This MUP process involves assembling the same data elements and carrying out the same computational steps as stated for MUAs. The population is now the population of the requested group within the area rather than the total resident civilian population of the area.[157]

Application of ACS in the Methodology: Two of the standard IMU variables are drawn from the ACS – the percentage of the population below the poverty level and the percentage of the population age 65 years or older.

ACS data also are used in MUP designations for populations with barriers to care for economic reasons (e.g., low-income, uninsured, Medicaid-eligible) or due to cultural or linguistic isolation.

A review of the inventory of MUP designations (including Governor-designated MUPs) finds the following terms as defining population characteristics (in approximate order of frequency): poverty, low-income, Medicaid-eligible, medically indigent, migrant farm workers, Native American, Hispanic, homeless, homebound, and Asian-American.[158]

ACS Questions Utilized (as suggested by the methodology): P2, P4-8, P13-14, P16, P19, P41-48.

Major Uses: Major federal programs using MUA/MUP designation include:

- Consolidated Health Centers (Community Health Centers, Migrant Health Centers, Health Care for the Homeless, and Public Housing Primary Care), HRSA
- Systems of care which meet the definition of a community health center contained in Section 330 of the Public Health Service Act, but are not funded under that section, and are serving a designated MUA or MUP, are eligible for certification as a Federally Qualified Health Center (FQHC) and thus for cost-based reimbursement of services to Medicaid-eligibles.
- Health centers that are FQHC Look-Alike or have Health Care for the homeless status.

---

[156] Text source: "MUA/Ps: Index of Medical Underservice Data Tables," BHW webpage.
[157] Text source: "Medically Underserved Areas/Populations: Guidelines for MUA and MUP Designation," BHW webpage.
[158] "Preformatted Reports: Medically Underserved Areas/Populations (MUA/P)," BHW webpage.

- Clinics serving rural areas designated as MUAs are eligible for certification as Rural Health Clinics by the Centers for Medicare and Medicaid Services under the authority of the Rural Health Clinics Services Act (Public Law 95-210, as amended).
- PHS Grant Programs administered by BHW - gives funding preference to Title VII and VIII training programs in MUA/Ps.[159]
- Telehealth Network Grant Program (TNGP), HRSA
- Nurse Education, Practice Quality and Retention Grants, HRSA
- Nurse-Managed Health Clinics, HRSA
- Grants for Capital Development in Health Centers, HRSA
- Grants for New and Expanded Services under the Health Center Program, HRSA
- Primary Care Services Resource Coordination and Development Program, HRSA
- Special Programs for the Aging, Title III, Part D, Disease Prevention and Health Promotion Services, HRSA[160]
- J1 Visa Waivers
- National Interest Waiver for Physicians (EB-2)

---

[159] Text source: "Medically Underserved Areas/Populations: Guidelines for MUA and MUP Designation," BHW webpage.

[160] "Catalog of Federal Domestic Assistance," General Services Administration webtool.

49. *Primary Care Service Areas*

Nature and Purpose of the Measure: Primary Care Service Areas (PCSAs) are geographic areas of relatively self-contained markets for ambulatory primary care services that can be used to measure primary care for planning and evaluation. Primary care is the most localized medical service. PCSAs methods were designed to identify small areas that are relatively self-contained markets for primary care in which the residents are likely to seek care from within PCSA primary care providers.

Currently, there are 7,144 PCSAs, with mean population per PCSA of 43,217. [161]

Responsible Unit: National Center for Health Workforce Analysis (NCHWA), Bureau of Health Workforce (BHW), Health Resources and Services Administration (HRSA), Department of Health and Human Services (HHS). The NCHWA contracts with Dartmouth College to update and enhance the Primary Care Service Areas and to make the tool available to health professions training programs, HRSA-supported health centers, States and health workforce researchers.[162]

Authorization: Sections 761, 792, and 806(f) of the Public Health Service Act.

Frequency of Data Publication: PCSAs were first defined with Medicare primary care claims from 1996-97, and were redefined using data from 1999. PCSA Version 2.1 came out in 2003. PCSA Version 3.1, published in 2013, further updated the areas using 2009 claims.

The HRSA Data Warehouse (HDW) has available Primary Care Service Areas Download for 1999-2001, 2005-2006, 2007, 2009, and 2010. Data from 1999-2009 are on a ZIP Code basis; data for 2010 are on a census tract basis.

Timing of Data Release: PCSA Version 3.1 was released in 2013.

Modes of Data Access: Data Available for Download, HDW (dbf)

Methodology: Each version of PCSAs uses definition methods that relies on a set of small geographic areas linked to the residence location of Medicare beneficiaries. These geographic areas are assigned to primary care location based on patient travel for ambulatory non-consultative primary care services. The initial assignments are adjusted for contiguity, and then additional criteria are applied, depending on the PCSA version. A detailed description of the PCSA methodology is available online.[163]

Application of ACS in the Methodology: Data from the American Community Survey (ACS) used include:

- Travel time to work
- Means of travel to work
- Linguistic isolation

---

[161] Text source: David Goodman, "Primary Care Service Area Version 3.1 Methods," September 17, 2013.

[162] Text source: "Primary Care Service Areas," NCHWA webpage.

[163] Goodman, *op.cit.*

- Population in poverty
- Ratio of income to poverty level
- Families in poverty, by number of parents, gender of single parent, and ages of related children
- Civilian population by age and gender
- Veteran status
- Housing tenure
- Total occupied housing units
- Median household income
- Per capita income
- Median contract rent
- Median value of owner-occupied housing units
- Educational attainment by gender[164]

ACS Questions Utilized (as suggested by the methodology): number in household, P2-4, P11, P14, P20, P26, P31, P34, P47-48, H17-19

Major Uses: The following uses of PCSAs have been identified:

- Evaluating primary care physician resources for allocation, e.g. location of new or expanded rural health or community health centers
- Tracking changes in primary care health providers, facilities or utilization over time
- Determining the availability of primary care health resources for special populations
- Determining the effect of policy changes on population utilization and access to services
- Evaluating unmet primary care health care needs (e.g. 5 year requirement to develop a needs assessment for Maternal and Child Health Care Block Grants)
- Allocating grants and monies for underserved areas
- Responding to federal data requirements
- Developing bio-terrorism and Emergency Preparedness Plans
- Developing community health needs assessment and identifying priority projects
- National studies of primary care workforce availability
- Outcomes studies and comparisons of primary care practices or services[165]

---

[164] David Goodman et al., "Version 3.1 Data Dictionary," Center for Health Policy Research, The Dartmouth Institute for Health Policy and Clinical Practice, September 2013.

[165] Text source: Laura Danielson, "SAVI Primary Care Service Area (PCSA) Data User Guide," The Polis Center at IUPUI, June 2010.

50. *UDS Mapper*

Nature and Purpose of the Measure: It is essential that HRSA, BPHC, and health center organizations have access to tools and data that can assist in evaluating the geographic reach, penetration, and growth of the Section 330-funded Health Center Program (HCP) and its relationship to other federally-linked health resources. HCP funds are used for include Community Health Centers, Migrant Health Centers, Healthcare for the Homeless Programs, and Public Housing Primary Care Programs. HCP grantees provide comprehensive, culturally competent, quality primary health care services to medically underserved communities and vulnerable populations.[166]

The UDS Mapper is designed to help inform users about the current geographic extent of U.S. federal (Section 330) Health Center Program (HCP) grantees and look-alikes. The information available in the UDS Mapper includes estimates of the collective service area of these health centers by ZIP Code Tabulation Area (ZCTA), including the ratio of HCP grantee and look-alike patients reported in the Uniform Data System (UDS) to the target population, the change in the number of those reported patients over time, and an estimate of those in the target population that remain unserved by HCP grantees and look-alikes reporting data to the UDS (but may be served by other providers).[167]

Responsible Unit: Bureau of Primary Health Care (BPHC), Health Resources and Services Administration (HRSA), Department of Health and Human Services (HHS). The Robert Graham Center for Policy Studies in Family Medicine and Primary Care at the American Academy of Family Physicians manages the UDS Mapper under a grant from BPHC.

Authorization: Section 330 of the Public Health Act.

Frequency of Data Publication: The UDS Mapper is updated as new releases of its data sources become available.

Timing of Data Release: NA

Modes of Data Access: UDS Mapper (html, pdf, ftp)

Methodology: Patient data in the UDS Mapper are derived from the Uniform Data System (UDS) count of patients by ZIP Code for all health centers. UDS data are submitted to HRSA by Heath Center Program grantees and look-alikes (collectively referred to as "health centers") every calendar year. A variety of demographic data on the total population, including health, economic, and social factors, are also available in the UDS Mapper. The majority of data in the UDS Mapper are shown by ZIP Code Tabulation Area (ZCTA), using 2010 ZCTA boundaries. Health centers report patients by ZIP Code; however, ZIP Codes are simply driving routes, created by the US Postal Service. ZCTAs are an approximation of the area covered by ZIP Codes, created by the US Census Bureau. Data also are shown by Public Use Micro Area (PUMA).

---

[166] Information on federal health center programs is available at "What is a Health Center?," BPHC webpage.
[167] Text source: "About the UDS Mapper," UDS Mapper webpage.

Data are viewed and accessed through a variety of "Tools" and displays:

- "Explore Service Area" Tool
- "Main Maps" Tool
- "Population Indicators" Tool
- "Community Health View" Tool
- "Uninsurance Explorer" Tool
- ZCTA Rollovers
- Health Care Facility Rollovers[168]

Application of ACS in the Methodology: The UDS Mapper relies on the American Community Survey for data on the following variables:

- Total Population Count
- Age
- Poverty and Low Income
- Race and Ethnicity
- Uninsurance[169]
- Non-Employment
- Limited English Proficiency
- Less than High School Education

These ACS data are used in the Explore Service Area Tool (Rollovers and Data Table), Main Maps, Uninsurance Explorer (Rollovers and Data Table), Population Indicators, and View Analysis Results.[170]

UDS Mapper data by ZCTA relies primarily on five-year ACS estimates. Insurance coverage estimates are based on three-year ACS data. Data by PUMA is one-year ACS data.[171]

ACS Questions Utilized (as suggested by the methodology): P2-6, P11, P14, P16, P29, P35-37, P47-48

Major Uses: The UDS Mapper has become an essential tool for:

- Planning
- Grantwriting
- Service area definition and analyses
- Community health needs assessments

---

[168] Text source: "Data in the UDS Mapper," tutorial, UDS Mapper webpage, slides 2-3 and 5.
[169] See C. Fahey, "Methodology for ZCTA-Level Imputations of Uninsurance by Federal Poverty Level," September 13, 2013.
[170] Text source: *Ibid*, slide 7.
[171] A detailed description of the use of ACS data by indicator is available at "Data Sources & Definitions," UDS Mapper webpage.

Since launching in 2010, the UDS Mapper user base has grown to more than 10,000 registered users. Half of all users are from health centers, with the other half including users from Primary Care Organizations, Primary Care Offices, HRSA, other Federal government, State or local government, Rural Health Clinics, hospitals, and others.

- Project officers gain familiarity with health center portfolios, perform environmental scans, and review applications to examine potential service area overlap.
- Health centers learn whether there is potential unmet need in their service area, see where other safety net providers are in their area, and create service area maps.
- Primary Care Associations (PCAs) and Offices (PCOs) visualize areas of growth or loss of patients, identify communities with great unmet need, and pinpoint areas with little or no federal investment that may qualify for state investment.[172]

---

[172] Text source: Text source: "What is the UDS Mapper?," tutorial, UDS Mapper webpage, slides 5-6.

51. *U.S. Health Workforce Chartbook and State Profiles*

Nature and Purpose of the Measure: The U.S. Health Workforce Chartbook provides data on 35 health occupations, including a wide variety of health workers, from physicians, nurses, and dentists to counselors, physical therapists, laboratory technicians, nursing assistants, and others. The Chartbook includes detailed information on each occupation including estimated total numbers and detailed demographic information including sex, age, race and ethnicity, work setting and geographic distribution.

The U.S. Health Workforce: State Profiles 2014 are companion documents to the Chartbook. The State Profiles also provide data on 35 types of health workers.

The health occupations discussed in the Chartbook and State Profiles are classified based on the U.S. government's Standard Occupational Classification (SOC) system and included more than 14 million individuals in 2010, representing approximately 10 percent of the nation's workforce.

Responsible Unit: National Center for Health Workforce Analysis (NCHWA), Bureau of Health Workforce, Health Resources and Services Administration, Department of Health and Human Services (HHS).

Authorization: 42 U.S. Code § 294n - Health professions workforce information and analysis. The NCHWA was created by Congress "to provide necessary information for decision-making regarding future directions in health professions and nursing programs in response to societal and professional needs."

Frequency of Data Publication: It appears the Chartbook and State Profiles have been published once, the former in November 2013 and the latter in August 2014. Plans for updating are not provided on the web.

Timing of Data Release: NA

Modes of Data Access: The U.S. Health Workforce Chartbook, NCHWA (pdf)

Methodology/Application of ACS in the Methodology: Data sources for the Chartbook and State Profiles are as follows:

- The three-year 2008 to 2010 ACS PUMS file was analyzed to provide the estimated number of individuals within each selected health occupation, along with information pertaining to workforce settings (industry), and the demographic makeup of the occupation (i.e., sex, age, race, and ethnicity). The three-year ACS PUMS file was used, rather than the most recent single-year ACS PUMS file, to have sufficient sample sizes. While many of the health care occupations included in this report have substantial sample sizes, some occupations have relatively small sample sizes.
- The National Center for Education Statistics Integrated Postsecondary Education Data System (IPEDS) Completions Survey 2010 was used to report on the number of post-secondary degrees and certificates awarded.

- Data from IPEDS on individuals completing educational programs in the 2009 to 2010 academic year were used to estimate new graduates entering health occupations. The IPEDS data are used to estimate the number of students who, upon graduation, may be entering the occupation for the first time.
- The 2011 to 2012 Area Resource File (ARF) was used to report on three occupations for which ACS data were not available (i.e., nurse practitioners, nurse anesthetists, and nurse-midwives).

The occupations included in this report were selected according to the following criteria: (1) the occupation is among those with the largest number of jobs as identified by the SOC code, (2) the occupation is among the fastest-growing occupations as projected by the BLS, and/or (3) the occupation is among the top 35 occupations that have adequate data (i.e., sample sizes) available in the three-year 2008 to 2010 ACS PUMS file. These determinations were made based on data from the BLS Occupational Employment Statistics May 2010 National Employment and Wage Estimates.

The "workforce" is defined as individuals employed in the occupation and individuals whose last job was in the occupation and who are still seeking employment (e.g., individuals unemployed because they were laid off but are still seeking employment). The 2008 to 2010 ACS PUMS has a variable that distinguishes between non-working individuals in the workforce and those no longer in the workforce. Only working-age individuals in the workforce are included in the Chartbook. State workforce estimates are based on the reported location for place of work; when place of work is not reported, state of residence is used.

Demographic information presented in the Chartbook includes sex, race, ethnicity, and age. The race and ethnicity categories reported are White (non-Hispanic), Black/African American (non-Hispanic), Hispanic, Asian/Native Hawaiian/Pacific Islander (non-Hispanic), American Indian/Alaska Native (non-Hispanic), and Multiple/Other Race (non-Hispanic). Race and ethnicity data in the ACS are reported according to standards defined by the Office of Management and Budget. Some race and ethnicity categories were combined to allow sufficient data for reporting.

The U.S. population estimates of the age groups and the race and ethnicity and sex distributions were derived from the 2008 to 2010 ACS PUMS for the population 16 years of age or older to represent the working-age population. The population estimates from the 2008 to 2010 ACS PUMS represent the average annual population distributions for the three-year period, 2008 through 2010. This measure is consistent with the estimates for each occupation based on the 2008 to 2010 ACS.[173]

ACS Questions Utilized (as suggested by the methodology): P3-6, P29-30, P35-38, P41-46

---

[173] Text source: NCHWA, "The U.S. Health Workforce Chartbook: Technical Documentation," November 2013, pp. 1-4.

Major Uses: The Chartbook and State Profiles provide healthcare policymakers, analysts, researchers, educators, trainers, and provider organizations with detailed information on the size, nature, and dynamics of the U.S. health workforce, by profession.

F. National Center for Health Statistics

52. *Births in the United States*

Nature and Purpose of the Report: Each year, the National Center for Health Statistics (NCHS) publishes a statistical report on births in the United States according to a wide variety of characteristics. Data are presented for maternal age, live-birth order, race and Hispanic origin, marital status, attendant at birth, method of delivery, period of gestation, birthweight, and plurality. Birth and fertility rates are presented by age, live-birth order, race and Hispanic origin, and marital status. Selected data by mother's state of residence and birth rates by age and race of father also are shown. Trends in fertility patterns and maternal and infant characteristics are described and interpreted.[174]

Responsible Unit: Division of Vital Statistics, National Center for Health Statistics, Centers for Disease Control and Prevention, Department of Health and Human Services

Authorization:

Frequency of Data Publication: A preliminary report and a final report are produced annually.

Timing of Data Release: For each reference calendar year, the preliminary report is published in the middle and the final report is published at the end of the next calendar year.

Modes of Data Access:

- Data Briefs, NCHS website (pdf) – 2013 births final report
- National Vital Statistics Reports, NCHS website (pdf) – all previous reports

Methodology: Data shown in this report are based on 100% of the birth certificates registered in all states and DC. More than 99% of births occurring in this country are registered. Tables showing data by state also provide separate information for Puerto Rico, Virgin Islands, Guam, American Samoa, and Commonwealth of the Northern Marianas. These areas, however, are not included in totals for the United States.

Race and Hispanic origin are reported independently on the birth certificate. In tabulations of birth data by race and Hispanic origin, data for Hispanic persons are not further classified by race because the majority of women of Hispanic origin are reported as white. Most tables in this report show data for the categories of non-Hispanic white, non-Hispanic black, and Hispanic. Data for births are also presented in some tables for American Indian or Alaska Native (AIAN) and Asian or Pacific Islander (API); for specific Hispanic groups Mexican, Puerto Rican, Cuban, and Central and South American; and for other and unknown Hispanic. Data for AIAN and API births are not shown separately by Hispanic origin because the majority of these populations are non-Hispanic.

---

[174] Text source: Joyce A. Martin, Brady E. Hamilton, Michelle J.K. Osterman, Sally C. Curtin, and T.J. Mathews, "Births: Final Data for 2012," National Vital Statistics Reports, Vol. 62, No. 9, December 30, 2013, p. 1.

The 2003 revision of the U.S. Standard Certificate of Live Birth allows the reporting of more than one race (multiple races) for each parent in accordance with the revised standards issued by the Office of Management and Budget (OMB) in 1997.[175]

Application of ACS in the Methodology: Birth and fertility rates for specific Hispanic population groups (Mexican, Puerto Rican, Cuban, Central and South American, and other Hispanic populations) for 2010–latest year are based on population estimates derived from the American Community Survey (ACS), conducted by the U.S Census Bureau. (In other words, ACS-based population estimates serve as the denominator in calculating birth and fertility rates). Rates for specific Hispanic population groups prior to 2010 are based on population estimates derived from the Current Population Survey (CPS).[176]

ACS Questions Utilized (as suggested by the methodology): P4-6, P13

Major Uses: The NCHS birth reports are used by demographers and vital statistics researchers to ascertain demographic trends on which to base population projections and determine public health needs.

---

[175] *Ibid,* p. 3.
[176] *Ibid.*

53. *National Health Interview Survey – Family Disability Questions*

Nature and Purpose of the Measure: The National Health Interview Survey (NHIS) is the principal source of information on the health of the civilian noninstitutionalized population of the United States and is one of the major data collection programs of the National Center for Health Statistics (NCHS) which is part of the Centers for Disease Control and Prevention (CDC).

The main objective of the NHIS is to produce descriptive statistics about the health and health-related parameters of the civilian noninstitutionalized population of the United States, and to monitor change in these variables over time. The basic variables of interest are extent and nature of illness, both acute (e.g., disability days and work loss) and chronic (e.g., limitation of activity); extent and nature of disability (acute and chronic); incidence and prevalence of acute and chronic morbidity; utilization of health care services; health care expenditures; utilization of health facilities and health resources; and other health-related variables.[177]

The six questions in the 2013 Family Disability Questions section (FDB) of the NHIS asked about difficulties hearing, seeing, concentrating, walking, dressing/bathing, and doing errands.

Responsible Unit: National Center for Health Statistics (NCHS), Department of Health and Human Services (HHS)

Authorization: The National Health Survey Act of 1956 provided for a continuing survey and special studies to secure accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States and the services rendered for or because of such conditions.

Frequency of Data Publication: The survey referred to in the Act, now called the National Health Interview Survey, was initiated in July 1957.

The disability questions were initially asked from October 2008 through December 2009 on the Family Disability questionnaire at both the person level (FDB) and the family level (FDA), but the FDA was dropped in 2010. Disability questions were included in the 2011 and 2012 Sample Adult (ADB) and Sample Child (CDB) questionnaires, in addition to the Family Disability questionnaire (FDB), but the ADB and CDB were dropped in 2013.[178]

Timing of Data Release: The NHIS microdata files are made available in the first June after the reference calendar year. Preliminary data are available earlier.

---

[177] Text source: "Van L. Parsons, Chris Moriarity, Kimball Jonas, Thomas F. Moore, Karen E. Davis, and Linda Tompkins, "Design and Estimation for the National Health Interview Survey, 2006–2015," National Center for Health Statistics, Vital and Health Statistics reports, Series 2, Number 165, April 2014, p. 4.

[178] National Center for Health Statistics, "2013 National Health Interview Survey (NHIS) Public Use Data Release: Family Disability Questions (FDB) File Description," p. 1.

Modes of Data Access:

- "NHIS Questionnaires, Datasets, and Related Documentation: 1997 to the Present," NCHS (SAS, SPSS, STATA, ascii)
- Vital and Health Statistics Series, Series 10: Data From the National Health Interview Survey, NCHS (pdf)

Methodology: The NHIS questionnaire has Core questions and Supplements. The Core questions remain largely unchanged from year to year and allow for trends analysis and for data from more than one year to be pooled to increase sample size for analytic purposes. The Core contains four major components: Household, Family, Sample Adult, and Sample Child.

The Supplements are used to respond to new public health data needs as they arise. The questionnaires are sometimes fielded only once or are repeated as needed. These questionnaires may be used to provide additional detail on a subject already covered in the Core or on a different topic not covered in other parts of the NHIS.

The 2013 NHIS consists of seven questionnaires - Household, Family, Family Disability Questions, Adult, Child, Cover, and Functioning & Disability.[179]

Application of ACS in the Methodology: The Family Disability Questions file contains a set of six core disability questions first developed and used on the American Community Survey (ACS) and subsequently adopted by several federal data collection systems.

The U. S. Census has a history of including questions about disability to satisfy a variety of stakeholder needs. The 2000 Census of Population and the 2000–2007 questionnaires of the Census Bureau's American Community Survey included six disability questions: a combined sensory (vision and hearing) question; separate mobility, self-care, and cognition questions; and two independent living questions (based on daily activities and working).

Before each decennial census, other agencies and researchers join with the Census Bureau to develop survey questions most appropriate for their needs and to determine if existing questions need modification. The user community had expressed some dissatisfaction with the disability measures on the 2000 Census and 2000–2007 ACS.

In response, the Census Bureau and other stakeholder agencies refined the disability questions to bring them into line with recent changes in the definition of disability and the conceptualization of the components of that definition. A work group was formed under the auspices of the Office of Management and Budget and led by the National Center for Health Statistics (NCHS). That work group researched the theoretical approach to the definitions based on the latest national and international ideas about disability, analyzed available data to test their conceptualizations, identified and examined agency mandates for collecting disability data, discussed question content and wording, and sponsored cognitive testing of a new question set. The new set included separate questions for vision and hearing, refined the

---

[179] Parsons et al., *op.cit.*, p. 3. The questions on the Functioning & Disability survey instrument are not attributed to the ACS.

mobility question, expanded the cognitive functioning question, continued the inclusion of a self-care question, and improved the question on independent living. The question about work, which had been included in earlier censuses, was not retained.

The 2006 ACS Content Test was a test of the modified questions, with the results of that testing used to inform the content for the 2008 ACS. The Content Test compared two sets of disability questions: the then current ACS set (called the "control questions"), and a set recommended by the work group (called the "test questions"). A formal evaluation of those two versions of the disability questions was completed and the results used to make the decision to adopt the modified questions for the 2008 ACS.[180]

ACS Questions Utilized (as suggested by the methodology): P17-19

Major Uses: The NHIS serves a critical role in providing information for monitoring and evaluating the performance of public health programs. The NHIS meets this need in several ways. First, the NHIS provides information on the overall health status of the U.S. population and its health needs, which serves as a background of "contextual" data against which program goals and performance measures are formulated and evaluated. Secondly, data from the NHIS serve as the national benchmark against which individual state monitoring efforts and other national surveys are compared.

The purposes of the NHIS are (1) to provide national data on an annual basis on the incidence of acute illness and accidental injuries, the prevalence of chronic conditions and impairments, the extent of disability, the utilization of health care services, and other health-related topics; (2) to provide more detailed information on selected topics periodically and on a one time basis; and (3) to provide a sampling frame for the Medical Expenditure Panel Survey and other followback surveys. It is also a main provider of data for the congressionally mandated Health U.S. report and provides the majority of indicators used in monitoring progress toward the Healthy People goals and for describing health disparities.

A major strength of the NHIS is its ability to display health characteristics by selected demographic and socio-economic characteristics of the U.S. civilian, noninstitutionalized population. The uses of NHIS data are generally in the areas of program planning and evaluation, public health education and health promotion and epidemiological research.[181]

In particular, NHIS data are used widely throughout the Department of Health and Human Services to monitor trends in illness and disability and to track progress toward achieving national health objectives. The data are also used by the public health research community for epidemiologic and policy analysis of such timely issues as characterizing those with various

---

[180] National Center for Health Statistics, "2013 National Health Interview Survey (NHIS) Public Use Data Release: Family Disability Questions (FDB) File Description," pp. 1-2.
[181] Text source: NCHS, "Supporting Statement A, Revision Request for Clearance, NATIONAL HEALTH INTERVIEW SURVEY (OMB No. 0920-0214)," December 4, 2014.

health problems, determining barriers to accessing and using appropriate health care, and evaluating Federal health programs.[182]

The NHIS disability questions are used to identify the subpopulation that is at a greater risk than the general population of experiencing restrictions in social participation, for example, restrictions in employment, education, or civic life. They are used to measure equalization of opportunities. Four basic domains of functioning (vision, hearing, mobility, and cognitive functioning) were identified that would define the largest component of the population of people with disabilities. These four domains are used individually or combined in order to assess equalization of opportunities for people with disabilities. Also, two more domains are identified that could be used for monitoring independent living and the need for services: the ability to take care of oneself (self-care, in particular, the ability to bathe and dress oneself), and the ability to move around the community (independent living, in particular, the ability to visit a doctor's office or go shopping).[183]

---

[182] Text source: "About the National Health Interview Survey," NCHS website.
[183] Text source: National Center for Health Statistics, "2013 National Health Interview Survey (NHIS) Public Use Data Release: Family Disability Questions (FDB) File Description," pp. 1-2.

54. [National Health Interview Survey – Native Hawaiian and Pacific Islander Survey](#)

Purpose of the Dataset: The National Health Interview Survey (NHIS) is the principal source of information on the health of the civilian non-institutionalized population of the United States. The main objective of the NHIS is to monitor the health of the United States population through the collection and analysis of data on a broad range of health topics. A major strength of this survey lies in the ability to display these health characteristics by many demographic and socioeconomic characteristics.

The Native Hawaiian and Pacific Islander National Health Interview Survey (NHPI NHIS) is a new component of the NHIS. The NHPI population is a distinct racial group that comprises 0.4% of the total U.S. population. This makes it a difficult group to include in sufficient numbers in most national surveys. To protect the privacy of NHPI survey participants, NHPI data has historically been combined with Asian data when statistics are shown by race. As a result, there is little nationally representative data on NHPI health.

To address this information gap, in 2014, about 4,000 additional households containing one or more NHPI residents were added to the NHIS so that indicators would be available on the health of Native Hawaiians and Pacific Islanders in all 50 states.

Responsible Unit: Survey Planning and Special Surveys Branch, Division of Health Interview Statistics, National Center for Health Statistics, Centers for Disease Control and Prevention, Department of Health and Human Services

Authorization: The National Health Survey Act of 1956 (42 USC 306) provides for a continuing survey and special studies to secure accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States and the services rendered for or because of such conditions.

Frequency of Data Publication: NHIS data are released annually.

Timing of Data Release: The full set of 2013 NHIS data were released in June 2014. Public data files from the NHPI NHIS will be available in fall 2015.

Modes to Access Data: [Questionnaires, Datasets, and Related Documentation: 1997 to the Present](#), NHIS, NCHS, CDC, HHS (txt).

Methodology: The NHIS is a cross-sectional household interview survey. Sampling and interviewing are continuous throughout each year. The sampling plan follows a multistage area probability design that permits the representative sampling of households and non-institutional group quarters. Approximately every ten years, the NHIS sampling plan is revised following the decennial census of the population. The current sampling plan started with the 2006 NHIS and is based on the 2000 decennial census. A new sampling plan is intended for 2016.

For the 2006 design, the basic NHIS sample contains 428 primary sampling units (PSUs), usually a county, a small group of counties, or a metropolitan statistical area, drawn from 1,838 PSUs that cover the 50 States and the District of Columbia. Within PSUs, second stage sampling units

called segments contain an expected 4, 8, 12 or 16 housing units. The sample assigned to each month is representative of the target population and the monthly samples are additive.

In a typical data collection year, if there are sufficient resources to fund the survey fully, the final NHIS sample will contain approximately 35,000 households and 87,500 persons. Data are collected through a personal household interview conducted by interviewers employed and trained by the U.S. Bureau of the Census according to procedures specified by the NCHS.

Application of ACS Data in Methodology: In 2014, NCHS introduced an NHPI nationally representative sample of 4,000 addresses identified from the 2012 ACS that will be administered the 2014 NHIS questionnaire. This represents a randomly selected half of the 2012 ACS NHPI households.

Results from the survey will be publically available for statistical purposes, after being thoroughly processed to ensure that the data records cannot be linked to individual persons or households.

ACS Questions Utilized (as suggested by the methodology): P6

Major Uses: NHIS data are used widely to monitor trends in illness and disability and to track progress toward achieving national health objectives. The data are also used by the public health research community for epidemiologic and policy analysis of such timely issues as characterizing those with various health problems, determining barriers to accessing and using appropriate health care, and evaluating Federal health programs.

The NHIS also has a central role in the ongoing integration of household surveys in DHHS. The designs of two major DHHS national household surveys have been or are linked to the NHIS. The National Survey of Family Growth used the NHIS sampling frame in its first five cycles and the Medical Expenditure Panel Survey currently uses half of the NHIS sampling frame. Other linkage includes linking NHIS data to death certificates in the National Death Index (NDI).

With the launch of the Native Hawaiian and Pacific Islander Survey of the NHIS, it will be possible for policymakers, community leaders, researchers and health professionals to understand NHPI health problems, strengths, and needs and to plan policies and programs to improve NHPI health and well-being.[184]

---

[184] References for this section include "Native Hawaiian and Pacific Islander (NHPI) National Health Interview Survey" and "About the National Health Interview Survey," NCHS website; NCHS, "Supporting Statements A and B: Revision Request for Clearance, National Health Interview Survey (OMB No. 0920-0214)," December 1, 2013.

55. *National Survey of Family Growth*

Purpose of the Dataset: The National Survey of Family Growth (NSFG) gathers information on family life, marriage and divorce, pregnancy, infertility, use of contraception, and men's and women's health.

Responsible Unit: The National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC), Department of Health and Human Services (HHS)

Authorization: NCHS is to collect and publish reliable national statistics on "family formation, growth, and dissolution," according to Sec. 306(b) of the Public Health Service Act.

Frequency of Data Publication: The NSFG was conducted as a periodic survey six times (1973, 1976, 1982, 1988, 1995, and 2002). The NSFG then shifted to continuous interviewing for the June 2006-June 2010 period. It is in the midst of a second four-year interview cycle (September 2011- September 2015).

Timing of Data Release: For the 2006-2010 cycle, the first public use data files were released in May 2010, followed by a second set of data files in October 2011. NCHS publications based on analysis of NSFG data are released on a flow basis.

Modes to Access Data:

- [2006-2010 NSFG: Public Use Data Files, Codebooks, and Documentation](#), National Survey of Family Growth, NCHS (SAS, SPSS, and STATA)
- Vital and Health Statistics Series, ["Series 23. Data From the National Survey of Family Growth,"](#) NCHS (pdf)

The NSFG data also includes contextual data files for the 1995, 2002, and 2006-2010 surveys, containing information on the context or community in which respondents live. These are available to the research community, and are accessible only through the NCHS Research Data Center due to the increased risk of deductive disclosure of respondents' identities when geographic variables are linked to survey data.

Methodology: The NSFG was done with periodic interviewing from 1973 to 2002. In the 1973 to 1995 surveys, the NSFG was based on national samples of women, and focused on factors affecting pregnancy and birth rates. In 2002, the NSFG began interviewing men as well as women, to obtain data on fatherhood involvement, behaviors related to HIV and other sexually transmitted diseases, and other closely related topics.

The interviews for any one survey were done in less than one year, but the surveys were 6-7 years apart. This interval between surveys was too long; data were needed more frequently. Continuous interviewing began in June of 2006 and stopped in June 2010. There was a cessation of interviewing while a new contract was awarded, and interviewing resumed in September 2011, after approval from OMB of a change request in July 2011. Under continuous interviewing, costs per case are lower, sample sizes are higher, and interviewing is more efficient (in terms of hours of interviewer labor per interview, and in terms of costs per

interview) than before. The design yielded about 5,600 interviews per year in 2006-2010, within budget--and with a 77% response rate.

In 4 years of data collection (September 2011-September 2015), about 20,000 interviews will be collected from a national sample in 117 Primary Sampling Units. With the NSFG's old periodic design, estimates were possible once every 7 years, but with the new design, estimates will be possible every 4 years—almost twice as often, with larger sample sizes per dataset, at approximately the same cost.

For the 2006-2010 NSFG, statistical design, interviewing, and data processing have been conducted by the University of Michigan's Institute for Social Research (ISR), under a contract with the National Center for Health Statistics (NCHS), in collaboration with the NCHS NSFG team led by William Mosher (Project Officer).

The NSFG interviewed a national sample of 22,682 men and women 15-44 years of age living in households in the United States. Interviews were done 48 weeks of every year for 4 years—from June, 2006 through June, 2010. The first public use data files were released in May 2010, and included 13,495 interviews. A second set of data files was released in October 2011, containing all 22,682 interviews—over 10,000 interviews with men and more than 12,000 interviews with women.

The contextual data files for the 2006-2010 NSFG, which include information on the context or community in which respondents live, are available to the research community. Contextual data files for the 1995, 2002, and 2006-2010 NSFGs are accessible only through the NCHS Research Data Center due to the increased risk of deductive disclosure of respondents' identities when geographic variables are linked to survey data.

There are two contextual data files for each respondent. These correspond to the respondent's address at two points in time: 1) at the date of interview and 2) on April 1, 2000 (the time of the 2000 U.S. Census). Geographic variables are provided at the state, county, tract, block group, and block level in these data files. Identifiers for each of these geographic units are also available that allow researchers to merge other, external data with the NSFG survey data.

The variables in the contextual data files are drawn from these sources:

- NSFG 2006-2010 sample data;
- Census 2000 Summary Files 1 and 3 (SF1 and SF3);
- American Community Survey 5-Year Estimates Summary File, 2005-2009 (ACS5);
- County Characteristics, 2000-2007 (ICPSR 20660);
- Centers for Disease Control and Prevention data on sexually transmitted diseases
- for the place-at-interview file, 2006 data (CDC06)
- for the file for residence on April 1, 2000, 2000 data (CDC00); and
- Guttmacher Institute data of abortion and family planning services
- for the place-at-interview file, 2005-2006 data (AGI05_06)
- for the file for residence on April 1, 2000, 2000-2001 data (AGI00_01)

Application of ACS Data in Methodology: The list and utilization of contextual variables from the American Community Survey 2005-2009 five-year estimates are in Tables 1.1 and 2.7 of the "Continuous National Survey of Family Growth, 2006-2010 Contextual Data File for Place of Interview Codebook." Levels of geography include county, tract, and block.

ACS Questions Utilized (as suggested by the methodology): P3-6, P8, P10-11, P20, P24, P29, P36, P47-48, H1, H7, H12, H17, H18

Major Uses: The National Survey of Family Growth responds to the congressional mandate for NCHS to collect and publish reliable national statistics on "family formation, growth, and dissolution" (Sec. 306(b), paragraph 1(H) of the Public Health Service Act) as well as vital statistics on births and deaths, and a number of aspects of health status and health care. The NSFG collects and publishes the most reliable, and in most cases the only, national data on such major topics as: adoption, unplanned births, contraceptive use and effectiveness, infertility and use of infertility services, pelvic infection and sexually transmitted disease, sterilization, expected future births, the sexually active population, and the use of and need for family planning services.

Other HHS agencies that fund and use the NSFG include:

- the Office of Family Planning, Office of Population Affairs (OPA), DHHS, under 42 U.S.C. 300a (Section 1009 of Title X of the Public Health Service Act);
- the Adolescent Family Life Program of the Office of Adolescent Pregnancy Programs, Office of Population Affairs, DHHS, under 42 U.S.C. 300z (Section 2001 of the Public Health Service Act);
- the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), of the National Institutes of Health (NIH), under 42 U.S.C. 241 (Section 301 of the Public Health Service Act);
- the Children's Bureau of the Administration for Children, Youth, and Families, Office of Human Development Services, under PL 96-272, the Adoption Assistance and Child Welfare Act of 1980 and other laws;
- the Office of the Assistant Secretary for Planning and Evaluation (OASPE), under Section 301 of the Public Health Service Act;
- the CDC's Division of HIV/AIDS Prevention (DHAP) of the National Center for HIV, Sexually Transmitted Disease, and Tuberculosis Prevention;
- the CDC's Division of Sexually Transmitted Disease Prevention, under Section 301 of the Public Health Service Act;
- the Administration for Children and Families' Office of Planning, Research, & Evaluation;
- the CDC's Division of Cancer Prevention and Control (DCPC), under the EARLY Act, a part of the Affordable Care Act; and
- the CDC's Division of Birth Defects and Developmental Disabilities (DBDDD), under Section 399 of the Public Health Service Act.

Examples of uses by HHS agencies include:

- The Office of Population Affairs uses NSFG data to estimate the characteristics of women who use Title X and other family planning services, as well as for research on factors affecting contraceptive use, unintended pregnancy, teenage sexual activity, and use of medical services for family planning and reproductive health.
- The Center for Population Research, NICHD, NIH, uses the data as a resource for research on marriage, cohabitation, fertility and infertility, contraceptive use and breast-feeding in the United States.
- The Children's Bureau, ACF, DHHS, has a special research interest in the data collected on children in foster care, especially as it relates to children leaving the foster care system through adoption.
- The Office of the Assistant Secretary for Planning and Evaluation, DHHS, makes use of NSFG data on father involvement with children, and studies marriage, divorce, and teenage sexual activity.
- The Administration for Children and Families, Office of Planning, Research, and Evaluation, DHHS, relies on NSFG data on fatherhood, marriage, and teen pregnancy prevention, for planning programs to improve the economic and social well-being of children and families.
- The Division of HIV/AIDS Prevention, CDC, undertakes research based on NSFG data on behaviors that affect the risk of transmission of HIV—including condom use, numbers of sexual partners, and others.
- The Division of Sexually Transmitted Disease Prevention, CDC, relies on data on sexual behavior and the uptake of new vaccines to research its programs.
- The Division of Cancer Prevention and Control, CDC, uses NSFG data on screening for cervical cancer, Human Papillomavirus, and breast cancer, which can be analyzed in relation to the NSFG's extensive data on pregnancy histories, sexual behavior, and reproductive health.
- The Division of Birth Defects and Developmental Disabilities, CDC, uses estimates of the number and characteristics of women at risk of an alcohol-exposed pregnancy that could lead to Fetal Alcohol Syndrome.

The NSFG is the only nationally representative household survey that is specifically focused on childbearing experience, family formation, sexual behavior, contraceptive use, and reproductive health of men and women in the entire childbearing age range (15-44 years of age). A few other surveys have obtained data related to topics covered in the NSFG, but most were more limited in the questions they ask, the population they represent, or both.[185]

In addition, the NSFG is used by scholars in the behavioral sciences (e.g., sociology, demography, and economics) to study marriage, divorce, fertility, and family life; by scholars in public health to study reproductive, maternal and infant health topics; by state and local governments to plan health and social service programs; by private-sector research

---

[185] NCHS, "Supporting Statement A for Request for Clearance: National Survey of Family Growth, 2012-2015 (OMB No. 0920-0314)," March 2012.

organizations which distribute the information to the public and to policy makers; and by the press, to prepare articles on a number of topics related to health and family life.

The impact of the NSFG goes well beyond the more than 600 journal articles, NCHS reports, and book chapters shown in our publication lists. The NSFG's impact includes behind-the-scenes policy discussions, briefings, and program planning at the federal, state, and local levels. The survey results are also used by people providing health and social services, through government agencies and in private groups.

The survey results supplements the information obtained on birth certificates collected through the National Vital Statistics System, and are used by the U.S. Department of Health and Human Services and other research and policy organizations to plan health services and health education programs, and to do statistical studies on the topics listed above.[186]

---

[186] Text source: "About the National Survey of Family Growth," NCHS website.

G. Office of the Assistant Secretary for Planning and Evaluation

56. Poverty Guidelines

H. Office on Women's Health

57. Quick Health Data Online

Nature and Purpose of the Measure: Quick Health Data Online is an interactive system that provides reliable and easily accessible health data to help assess needs, develop programs, and inform policies. The system is for anyone looking for US health data and is used by the public health community, policymakers, grant writers, researchers, and students.

The system provides state- and county-level data for all 50 states, the District of Columbia, and US territories and possessions. Data are available by gender, race and ethnicity and come from a variety of national and state sources. The system is organized into eleven main categories, including demographics, mortality, natality, reproductive health, violence, prevention, disease and mental health. Within each main category, there are numerous subcategories.[187]

Data are available in tables, charts, and maps.

Responsible Unit: Office of Women's Health, Office of the Assistant Secretary for Health, Department of Health and Human Services (HHS)

Authorization: Title III of the Public Health Act.

Frequency of Data Publication: The Office on Women's Health plans to provide quarterly updates.

Timing of Data Release: NA

Modes of Data Access:

- Quick Health Data Online, Office of Women's Health (html, csv, pdf)
- State fact sheets present a snapshot of demographic characteristics as well as a variety of health status indicators for females (pdf)

Methodology: Quick Health Data Online has organized a series of health indicators by state and county in the following categories:

- Access
- Demographics
- Disease
- Maternal Health
- Mental Health
- Mortality
- Prevention
- Reproductive Health
- Violence

---

[187] Text source: "Welcome to Quick Health Data Online," Office of Women's Health, HHS

- Healthy People 2020

Application of ACS in the Methodology: ACS-based data provided include:

- Divorced Females
- Education – educational attainment by gender
- Families/Households – number of families, married-couple families, female-head families, average family size, average household size
- Foreign Born/Linguistically Isolated – by race, ethnicity
- Income/Employment: SSI, public assistance, labor force status, in management, self-employed, median earnings, by gender
- Median Age – by race and gender
- Poverty – by gender
- Same Sex Partners
- Children by Household Type/Living Quarters – number of children in households and group quarters

ACS Questions Utilized (as suggested by the methodology): number in household, P2-7, P11, P14, P20, P29, P35-37, P41, P45-48, Group Quarters

Major Uses: Quick Health Data Online is used for geographic analysis of health indicators by county and state. The tool also is the basis for a number of online reports, including:

- The Women's Health and Mortality Chartbook is a statistical resource on women's health in each of the states, the District of Columbia, Guam, Puerto Rico and the US Virgin Islands. The chartbook was developed to provide readers with an easy-to-use collection of current jurisdiction data on critical issues of relevance to women. A total of 28 different health indicators are featured, which highlight some of the key issues related to women's health that are being measured regularly at the state level. (annual)
- Health Disparities Profiles examines key health indicators at the state level for different racial and ethnic populations in each of the 50 states, the District of Columbia, Guam, Puerto Rico and the US Virgin Islands. Twenty-two health indicators are presented, which highlight some of the key areas related to health disparities among different populations. It can be used as a reference for policymakers and program managers to identify areas where major health disparities exist in each state. (annual)
- The Women's Health Assessment Toolkit (WHAT) has been developed to foster the formation, growth and success of women's health networks throughout the region. The WHAT is organized in four main sections to provide help local community planners:
  o Identify existing assets and prioritize needs;
  o Learn about the women of their area;
  o Adopt measurable objectives;
  o Incorporate findings into a usable planning document.

VIII. U.S. Department of Homeland Security

A. Federal Emergency Management Agency

58. Community Preparedness and Participation Survey

Nature and Purpose of the Dataset: The Community Preparedness and Participation Survey measures the public's knowledge, attitudes, and behaviors relative to preparing for a wide range of hazards and disasters (flood, tornado, earthquake, hurricane, severe winter weather, chemical release, pandemics, and terrorism) and provides data to inform emergency preparedness at a national, community, and hazard-specific level.[188]

Responsible Unit: Individual and Community Preparedness Division, Federal Emergency Management Agency (FEMA), Department of Homeland Security

Authorization: Presidential Policy Directive-8 (PPD-8) directs the Secretary of Homeland Security to "coordinate a comprehensive campaign to build and sustain national preparedness, including public outreach and community-based and private sector programs to enhance national resilience, the provision of Federal financial assistance, preparedness efforts by the Federal Government, and national research and development efforts."[189]

Frequency of Data Publication: The Community Preparedness and Participation Survey was conducted in 2007, 2009, 2011, and 2012. The FEMA submission to OMB in September 2014 indicated the survey would be conducted annually over the following three years.

Timing of Data Release: Findings based on data collected in the reference year are released in the following calendar year.

Modes to Access Data: Citizen Preparedness Surveys, FEMA webpage (pdf)

Methodology: The potential respondent pool includes the entire civilian non-institutionalized U.S. adult population residing in telephone-equipped dwellings or owning a cell phone. This population does not include adults in penal, mental, or other institutions; adults living in dormitories, barracks, or boarding houses; adults living in a dwelling without a telephone; and/or adults who do not speak English or Spanish well enough to be interviewed.

The survey is conducted twice every year. The total number of respondents is 3,000 adults per administration, or 6,000 per year. The total number of respondents (3,000) for any particular administration will include a national level sample of about 1,000 respondents and four separate oversamples (of size 500 each) for four hazard specific areas. The hazard specific areas will be defined in terms of complete counties (or fipscodes). The selection of the hazard profiles to be surveyed in any specific administration will vary across different administrations of the survey. Hazards selected for the 2014 surveys are Flood, Hurricane, Tornado, Wildfire,

---

[188] FEMA, "Supporting Statement Part A, OMB Control Number: 1660- 0105, Community Preparedness and Participation Survey, FEMA Form 008-0-15," September 13, 2014.
[189] *Ibid.*

Earthquake and Winter Storm and Extreme Cold. In each administration, four hazard areas is surveyed while each of these hazards is covered in one or both administrations.

For the National sample, the target population is geographically stratified into four census regions (Northeast, Midwest, South, and West) and sampling is done independently within each stratum (region). Within each region, roughly 50 percent of the interviews is done from the cell phone sample while the rest (50%) is done from the landline sample.

The hazard area samples is selected independently following procedures similar to those used for the national sample described above. The target population for each hazard area survey will consist of groups of counties identified based on specific requirements for each hazard.

Hazard samples will also be stratified by census region and so each of the four strata for any hazard will consist of counties identified for that hazard in that particular census region. The sample allocation across strata is proportional to the size of the stratum derived as the estimated total adult population of the counties selected for that hazard in that particular stratum. As proposed for national sample, both landline and cell phone numbers is included in hazard area samples and the total number of completed surveys is roughly split equally between the landline and cell phone samples.

Each of the five samples (the national and the four hazard samples) is weighted independently. Once those weights are finalized, the sample data consisting of the national sample (1,000 completed surveys) and four hazard area level surveys (500 completes each) is combined (composite weighting) and then weighted (post-stratified) to generate estimates for unknown populations parameters at various levels (national, regional or for other subgroups of interest).[190]

Application of ACS Data in Methodology: Once the landline and cell samples are combined using the composite weight, a post-stratification weighting step will be carried out to simultaneously rake the combined sample to (i) known characteristics of the target population (adults 18 years of age or older) and (ii) an estimated parameter for relative telephone usage (landline-only, cell only, cell mostly, other dual users). The demographic variables to be used for weighting will include Age, gender, Race, Ethnicity (Hispanic/Non-Hispanic), and Education. The target numbers for post-stratification weighting will be obtained from [the American Community Survey].[191]

ACS Questions Utilized (as suggested by the methodology): P3-6 and P11

---

[190] FEMA, "Supporting Statement Part B, OMB Control Number: 1660- 0105, Community Preparedness and Participation Survey, FEMA Form 008-0-15," September 13, 2014.

[191] Ibid, p. 26. While the FEMA submission to OMB says "Current Population Survey" in this space, it is assumed FEMA meant to say the ACS, which is referenced as the source of weighting data in the annual preparedness reports. For example: "Weighting the data according to geography, age, gender, and race/ethnicity accounted for potential biases and adjusted the sample's demographic distributions to match the distribution in the American Community Survey (ACS)," as cited in FEMA, "Preparedness in America: Research Insights to Increase Individual, Organizational, and Community Action," updated August 2014, p. 59.

<u>Major Uses</u>: FEMA's Individual and Community Preparedness Division analyzes and uses data collected in the Community Preparedness and Participation Survey to identify progress and gaps in citizen and community preparedness and participation and to better understand the motivators and barriers to preparedness in general and about specific hazards (e.g., hurricanes, wildfires).

This information is used to tailor awareness and recruitment campaigns, specifically America's PrepareAthon!, messaging and public information efforts, community outreach and strategic planning initiatives to more effectively improve the state of citizen preparedness and participation across the country. This information is used to provide information to each of the FEMA regions and identified jurisdictions to develop targeted strategies for educating the public in their respective areas of responsibility. Findings are used to refine hazard-specific messaging (e.g., messaging specifically about hurricanes or wildfires).[192]

---

[192] FEMA, "Supporting Statement Part A, OMB Control Number: 1660- 0105, Community Preparedness and Participation Survey, FEMA Form 008-0-15," September 13, 2014.

B. Office of Immigration Statistics

59. *Estimates of Unauthorized Immigrants*

<u>Nature and Purpose of the Measure</u>: This report provides estimates of the size of the unauthorized immigrant population residing in the United States as of January of each year by period of entry, region and country of origin, state of residence, age, and sex.

An estimated 11.4 million unauthorized immigrants were living in the United States on January 1, 2012 compared to 11.5 million on January 1, 2011.

<u>Responsible Unit</u>: Office of Immigration Statistics (OIS), Policy Directorate, Department of Homeland Security

<u>Authorization</u>: Immigration and Nationality Act

<u>Frequency of Data Publication</u>: Annual (January 2005-January 2012). A report has not been published for January 2013.

<u>Timing of Data Release</u>: March of the calendar year after the reference year. The estimate for January 2012 was published in March 2013.

<u>Modes to Access Data</u>: Annual Population Estimates, OIS, DHS (pdf)

<u>Methodology</u>: Two populations are estimated in order to derive the unauthorized population estimates: 1) the total foreign-born population living in the United States on January 1, 2012 and 2) the legally resident population on the same date. The unauthorized population estimate is the residual when 2) is subtracted from 1). Foreign-born residents who entered the United States prior to 1980 were assumed to be legally resident since most were eligible for LPR status. Therefore, the starting point for the estimates was January 1, 1980.

- Data on the foreign-born population that entered during 1980–2011 by country of birth, state of residence, year of entry, age, and sex were obtained from the 2011 American Community Survey (ACS).
- Data on persons who obtained LPR status by country of birth, state of residence, age, sex, category of admission, and year of entry were obtained from DHS administrative records maintained in an application case tracking system of U.S. Citizenship and Immigration Services (USCIS).
- Data on refugees arriving in the United States by country of origin were obtained from the Department of State.
- Data on persons granted asylum by country of origin were obtained from USCIS for those granted asylum affirmatively and from the Executive Office for Immigration Review of the Department of Justice for those granted asylum defensively in removal proceedings.

- Data on nonimmigrant admissions by country of citizenship, state of residence, age, sex, and class of admission were obtained from I-94 arrival-departure records in the TECS system of the U.S. Customs and Border Protection.[193]

Application of ACS Data in Methodology: The foreign-born population is estimated as follows:

- Foreign-born population, entered 1980–2011: The estimated total foreign-born population that entered between 1980–2011 was obtained from the ACS's FactFinder. FactFinder is the Census-maintained online data portal for obtaining ACS estimates from the full sample for a particular year. Data on the distribution of the foreign born by country of origin, state of residence, year of entry, age, and sex were obtained from the 2011 Public Use Microdata Sample (PUMS). The overall FactFinder estimate for the total foreign-born population entering in the post-1979 period was reduced to remove PUMS estimates of the post-1979 Cuban-born population. Further, a three-year moving average was applied to PUMS data for year of entry to reduce heaping effects.
- Shift in reference date to January 1, 2012: The reference date for the 2011 ACS, the most recently available ACS data, was shifted from mid-year 2011 to January 1, 2012 by multiplying the population of 2011 entrants by 1.71, which is the average of three ratios: the ratio of the estimated population in the 2011 ACS that entered the United States during 2010 compared to the population in the 2010 ACS that entered in 2010 and the comparable ratios for the 2009 entrants in the 2009 and 2010 ACS surveys and the 2008 entrants in the 2008 and 2009 ACS surveys.
- Undercount of nonimmigrants in the ACS: Undercount refers to the number of persons who should have been counted in a survey or census, but were not. A rate of 10 percent was used to estimate the nonimmigrant undercount. This rate was used in DHS unauthorized population estimates for 2000 and 2005–2011.
- Undercount of LPRs, refugees, and asylees in the ACS: The undercount rate for LPRs, refugees, and asylees in the ACS was assumed to be 2.5 percent. This was the same rate used in DHS estimates for 2000 and 2005–2011.
- Undercount of unauthorized immigrants in the ACS: The undercount rate for unauthorized immigrants in the ACS was assumed to be 10 percent. This was the same rate used in previous DHS estimates for 2000 and 2005–2011.
- Estimated foreign-born population, January 1, 2012: The sum of the above is the estimated foreign-born population on January 1, 2012 that entered the United States during the 1980–2011 period.

ACS Questions Utilized (as suggested by the methodology): P3-4, P7-9

Major Uses: Annual estimates of unauthorized immigrants in the U.S. are used to inform DHS and homeland security enterprise strategic direction, integration, and decision-making.

---

[193] Bryan Baker and Nancy Rytina, "Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2012," *Population Estimates,* Department of Homeland Security, March 2013.

IX. U.S. Department of Housing and Urban Development

A. Office of Economic Resilience

60. *Location Affordability Index*

Nature and Purpose of the Measure: The Location Affordability Index (LAI) is intended to estimate combined housing and transportation cost burdens for any given household, analyze the relationships between the affordability landscape and other factors (e.g. transportation infrastructure, development pressure, etc.), and communicate this information to the public and stakeholders.

Because affordability is different for everyone, the Index is calculated for eight different household profiles, defined by different incomes, family sizes and number of commuters. Users can search for a location by address, intersection, city, county, state or zip code. In addition, users can view housing and transportation costs combined or separately, or view costs for renters and homeowners separately.[194]

Responsible Unit: The Partnership for Sustainable Communities, which includes the Office of Economic Resilience, Department of Housing and Urban Development; Office of Transportation Policy, the Department of Transportation; and Office of Sustainable Communities, Environmental Protection Agency.

Authorization: HUD, DOT and EPA Sustainable Communities Partnership Agreement, June 2009

Frequency of Data Publication: Annual, with each 5-year ACS release.[195]

Timing of Data Release: To date, two versions of the Location Affordability index have been released. Version 2, based on the 2008-2012 ACS, was released in July 2014.

Modes to Access Data: LAI data are made available through the Location Affordability Portal Version 2, which contain two tools: My Transportation Cost Calculator (html) for individuals and Location Affordability Index (html, csv, pdf) for professionals.

Methodology: The Index is expressed in terms of housing costs, transportation costs, and income. These three components are derived in three different ways. Housing costs are predicted using regression modeling and simultaneous equation modeling. Transportation costs are calculated by estimating how much transportation people in a given neighborhood use—i.e. how many cars they own, how much they drive, and how much they use public transit—and then multiplying each of these quantities by their unit cost (e.g. annual cost per car). To make the Index as concrete and useful as possible, housing and transportation costs are calculated for eight different household profiles characterized by income level, household size and number of commuters. These costs are then divided by the income for each profile to give a percentage of

---

[194] Text source: "Location Affordability Index," HUD website.
[195] See FAQs for the Location Affordability Portal v. 2.

a given family's income associated with a given location. The incomes used are based on the median income levels in each region covered by the Index.

The LAI draws from six different Federal data sources and Illinois state odometer readings:

- U.S. Census American Community Survey (ACS)
- U.S. Census TIGER/Line Files
- U.S. Census Longitudinal Employment-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES)
- National Transit Database, Department of Transportation
- Consumer Expenditure Survey, Bureau of Labor Statistics
- Illinois State odometer readings – as part of the smog check required for vehicles in the state's non-attainment areas (the Chicago and St. Louis metro areas), the Illinois Environmental Protection Agency records odometer readings. Odometer data for 2007 and 2009 were compared to determine how many miles had been driven by location. To validate the use of this data for entire country, it was compared to the 2009 National Household Travel Survey (maintained by the Federal Highway Administration).

Given currently available data, we are able to reliably estimate housing and transportation costs at the Census block-group level for the 94% of the U.S. population living in the 942 Metropolitan or Micropolitan Statistical Areas (i.e., Combined Base Statistical Areas or CBSAs). For non-metropolitan areas that account for the balance of the population living in the 50 states and the District of Columbia, the LAI is calculated at the county level.

To calculate the housing and transportation costs for a given location, we use data for demographics and features of the built environment that we know influence these costs: income, average household size, average commuters per household, population density, walkability, transit access, and employment access. Using these data and statistical regression – a widely used statistical technique that assesses the relationship between one or more input variables and an output variable – we generate mathematical models for the relationship between all of these data points and housing and transportation costs. By plugging data into these models, we can estimate components of housing and transportation costs - mostly at the Census block-group level - that can then be used to calculate the Index.

For the purposes of the LAI, we are interested in predicting the following outputs as they apply to households:

- Annual vehicle miles traveled
- Automobiles per household (renters and homeowners)
- Annual transit trips and transit expenditures (renters and homeowners)
- Gross Rent (includes rent and utilities)
- Selected Monthly Ownership Costs – SMOC (includes mortgage, utilities, and other costs related to homeownership like property taxes and condo fees)

These are predicted using a large number of input and output variables. Many advances in statistics have enabled the creation of more nuanced models for explaining complex phenomena like impact of demographics and the built environment on housing and transportation costs. One approach that has proved useful in urban planning studies is simultaneous (or structural) equation modeling (SEM). SEM allows set of regression models that are theoretically related to interact directly and influence each others' outputs.

Most of the input variables come from data that describe features of a neighborhood that are common to everyone who lives there: population density, walkability, transit access and quality, and employment access (these are all features of the built environment). For inputs that identify characteristics the residents themselves--household size, income, and number of commuters--using actual data for each block group wouldn't produce a very useful Index. Since people tend to live in places they can afford, using actual demographic data would produce a map where the majority of neighborhoods look more or less affordable. Instead, we have chosen eight household profiles—characterized by the number of family members, income, and number of commuters—that represent a wide range of American families, providing useful insight on affordability for a variety of different users, including consumers, planning agencies, real estate professionals, and housing counselors.[196]

Application of ACS Data in Methodology: ACS data is the basis for estimating Selected Monthly Ownership Costs and Gross Rent. Specifically, the following inputs are derived from ACS data:[197]

Gross Density	# of households (HH) / total acres
Fraction of Rental Units	Number of rental units as a percentage of total housing units
Fraction of Single Family Detached Housing Units	Number of single family detached housing units as a percentage of total housing units
Median Rooms/Owner HU	Median number of rooms in owner occupied housing units (HU)
Median Rooms/Renter HU	Median number of rooms in renter occupied housing units
Area Median Household Income	Determined using County median household income for rural areas or CBSA median household income for Metropolitan and Micropolitan Areas
Fraction of Median Income Owners	Median income for owners at the block group level as a percentage of either CBSA or County median income (County for rural areas / CBSA for Metropolitan and Micropolitan Areas)
Fraction of Area Median Income Renters	Median income for renters at the block group level as a percentage of either CBSA or County median

---

[196] Text source: "About the Portal: Data and Methodology," Location Affordability Portal, HUD website.
[197] Ibid.

	income (County for rural areas / CBSA for Metropolitan and Micropolitan Areas)
Average Household Size: Owners	Calculated from data on Tenure and Total Population in Occupied Housing Units by Tenure
Average Household Size: Renters	Calculated from data on Tenure and Total Population in Occupied Housing Units by Tenure
Average Commuters per Household Owners	Calculated using the total number of workers 16 years and over who do not work at home
Average Commuters per Household Renters	Calculated using the total number of workers 16 years and over who do not work at home
Median Selected Monthly Owner Costs	Includes mortgage payments, utilities, fuel, and condominium and mobile home fees where appropriate
Median Gross Rent	Includes contract rent as well as utilities and fuel if paid by the renter
Autos per Household Owners	Calculated from Aggregate Number of Vehicles Available by Tenure and Occupied Housing Units
Autos per Household Renters	Calculated from Aggregate Number of Vehicles Available by Tenure and Occupied Housing Units
Percent Transit Journey to Work Owners	Calculated from Means of Transportation to Work by Tenure
Percent Transit Journey to Work Renters	Calculated from Means of Transportation to Work by Tenure

ACS Questions Utilized (as suggested by the methodology): household size, P4, P29-31, P47-48, H1, H7, H12, H17-18, H22-23

Major Uses: The LAI is for direct use by researchers, developers, planners, and policymakers to help enhance their understanding of combined housing and transportation cost burdens, analyze the relationships between the affordability landscape and other factors (e.g. transportation infrastructure, development pressure, etc.), and communicate this information to the public and stakeholders. The LAI also helps homebuyers and renters estimate the cost of living in a particular city or neighborhood.[198]

---

[198] Text source: "About the Portal: Background," Location Affordability Portal, HUD website.

B. Office of Policy Development and Research

61. *Assessment of Fair Housing Tool*

Nature and Purpose of the Datasets: The Department of Housing and Urban Development (HUD) is proposing a rule that would require local governments and States that receive Community Development Block Grants (CDBG), HOME Investment Partnerships (HOME), Emergency Solutions Grants (ESG), and Housing Opportunities for Persons With AIDS (HOPWA), as well as public housing agencies (PHAs) to affirmatively further the purposes of the Fair Housing Act.

HUD plans to create a web-based tool by which state and local governments that receive HUD grants would assess the availability of fair housing in their area. The Assessment of Fair Housing (AFH) focuses program participants' analysis on four primary goals:

- improving integrated living patterns and overcoming historic patterns of segregation;
- reducing racial and ethnic concentrations of poverty;
- reducing disparities by race, color, religion, sex, familial status, national origin, or disability in access to community assets such as education, transit access, and employment, as well as exposure to environmental health hazards and other stressors that harm a person's quality of life; and
- responding to disproportionate housing needs by protected class.

To facilitate this obligation, HUD would provide these program participants with guidance, data, and an assessment template from which they would complete an assessment of fair housing (the AFH). This assessment would then link to Consolidated Plans, PHA Plans, and Capital Fund Plans, meaningfully informing resulting investments and related policies to affirmatively further fair housing.

HUD would provide all program participants with nationally uniform data on these four areas of focus as well as outstanding discrimination findings. Data topics would include:

- patterns of integration and segregation;
- racially and ethnically concentrated areas of poverty (R/ECAPs);
- access to education, employment, low-poverty neighborhoods, transportation, and environmental health, among other critical community assets;
- disproportionate housing needs; and
- data on individuals with disabilities and families with children.

Once program participants have analyzed the HUD data, as well as local or regional information they choose to add, they would evaluate their present environment to assess fair housing

issues, identify significant determinants that influence or contribute to those issues, and set forth fair housing priorities and goals to address fair housing issues and determinants.[199]

Responsible Unit: Office of Policy Development & Research (PD&R), Department of Housing and Urban Development (HUD)

Authorization: Fair Housing Act (Title VIII of the Civil Rights Act of 1968)

Frequency of Data Publication: To be determined

Timing of Data Release: To be determined

Modes of Data Access: "[P]rogram participants will complete the assessment via a web-based system that will guide participants' through the data and required analysis."[200]

Methodology: HUD designed the Assessment Tool with three key objectives in mind. First, the Assessment Tool must ask questions that would be sufficient to enable program participants to perform a meaningful assessment of key fair housing issues and determinants and set meaningful fair housing goals and priorities. Second, the Assessment Tool must clearly convey the analysis of fair housing issues and determinants that program participants must undertake in order for an AFH to be considered acceptable to HUD. Third, the Assessment Tool must be designed so program participants would be able to use it to prepare an acceptable AFH without unnecessary burden.

Geographic Notes: Core data on race, ethnicity, and poverty is provided at two geographic levels: block-groups from the Census 2010, and census tracts from the American Community Survey 2006-2010 estimates. Where census 2010 data is available it is used in place of survey data to allay concerns about sampling error. Data that incorporates economic cross-tabulations is from the American Community Survey 2006-2010.

Racially/Ethnically-Concentrated Areas of Poverty: To assist communities in identifying racially/ethnically-concentrated areas of poverty (RCAPs/ECAPs), HUD PD&R has developed a census tract based definition for RCAP/ECAPs. The definition involves a racial/ethnic concentration threshold and a poverty test. The racial/ethnic concentration threshold is straightforward: RCAP/ECAPs must have a non-white population of 50 percent or more. Regarding the poverty threshold, Wilson (1980) defines neighborhoods of extreme povertyas census tracts with 40 percent or more of individuals living at or below the poverty line. Because overall poverty levels are much lower in many parts of the country, HUD supplements this with an alternate criterion. Thus, a neighborhood can be an RCAP/ECAP if it has a poverty rate that exceeds 40% or is three times the average tract poverty rate for the metro/micro area,

---

[199] Text sources: "HUD's Notice of Proposed Rulemaking on Affirmatively Furthering Fair Housing—Summary," HUD website and HUD, "Affirmatively Furthering Fair Housing Assessment Tool: Solicitation of Comment—60-Day Notice Under Paperwork Reduction Act of 1995," *Federal Register*, Vol. 79, No. 187, September 26, 2014, pp. 57949-57955.

[200] *Ibid*, p. 57951.

whichever threshold is lower. Census tracts with this extreme poverty that satisfy the racial/ethnic concentration threshold are deemed RCAPs/ECAPs.

Segregation: PD&R will construct several common social science indices that measure segregation. These metrics will allow program participants to identify whether their area features high, moderate, or low levels of segregation.

- A dissimilarity index represents a summary measure of the extent to which the distribution of any two groups (frequently racial or ethnic groups) differs across census tracts or block-groups. PD&R will provide dissimilarity indices at the jurisdiction-level for jurisdictions of similar size in the same census region, and for metropolitan/micropolitan level. At each level, PD&R calculates the index between the relevant racial/ethnic groups and the majority group, disabled populations by disability type relative to non-disabled persons, and the largest immigrant populations and the majority group in a jurisdiction. In addition, to help communities understand how their situation compares with others around the country, PD&R reports will provide a color-coding designation that will signify whether HUD interprets the value as high, moderate, or low.
- An isolation index compares a group's share of the overall population in a jurisdiction to the average neighborhood share for members of that group. The isolation index highly correlates with the dissimilarity index, and while conceptually very similar, it tends to provide a better characterization of residential segregation when minority populations are extremely small.
- For very small communities, there are generally too few census block-groups or minorities for statistical metrics such as a dissimilarity index or even the isolation index to be particularly informative. Instead, for these communities, PD&R calculates a predicted value for the racial/ethnic minority share for a jurisdiction and compares this to the actual composition. Predicted values are based on a metropolitan/micropolitan area's income distribution by race and ethnicity. For a jurisdiction, the metro-level racial share for each income category is multiplied by the number of households the jurisdiction has in that category. The totals are summed to determine the predicted number of minorities in a jurisdiction. This total is then compared with the actual number of minorities in a community by calculating a ratio of actual to predicted.

Community Assets: HUD has developed a two-stage process for analyzing disparities in access to community assets. The first stage involves quantifying the degree to which a neighborhood offers features commonly viewed as important community assets such as education, employment, and transportation, among others. This stage uses metrics that rank each neighborhood along a set of key dimensions. In the second stage, HUD combines these dimension rankings with data on where people in particular subgroups live to develop a measure of that group's general access or exposure to each asset dimension. These summary measures can then be compared across subgroups to characterize disparities in access to community assets. HUD considers community assets a multi-dimensional notion. To focus the analysis, HUD developed methods to quantify a select number of the important "stressors" and

"assets" in every neighborhood. These dimensions were selected because existing research suggests they have a bearing on a range of individual outcomes. In particular, HUD has selected six dimensions upon which to focus:

- The neighborhood school proficiency index uses school-level data on the performance of students on state exams to describe which neighborhoods have high-performing elementary schools and which have lower performing elementary schools.
- HUD created a simple poverty index to capture the depth and intensity of poverty in a given neighborhood. The index uses family poverty rate and public assistance receipt to operationalize both aspects. The index is a linear combination of two vectors: the family poverty rate and the percentage of households receiving public assistance.
- The job access index summarizes the accessibility of a given residential neighborhood as a function of its distance to all job locations, with distance to larger employment centers weighted more heavily. Specifically, a gravity model is used, where the accessibility (Ai) of a given residential block-group is a summary description of the distance to all job locations, with the distance from any single job location positively weighted by the size of employment (job opportunities) at that location and inversely weighted by the labor supply (competition) to that location.
- The labor market engagement index provides a summary description of the relative intensity of labor market engagement and human capital in a neighborhood. This is based upon the level of employment, labor force participation, and educational attainment in that neighborhood. Formally, the labor market engagement index is a linear combination of three standardized vectors: unemployment rate, labor-force participation rate, and percent with bachelor's or higher.
- HUD has constructed a health hazards exposure index to summarize potential exposure to harmful toxins at a neighborhood level. Potential health hazards exposure is a linear combination of standardized EPA estimates of air quality carcinogenic, respiratory, and neurological.
- HUD has constructed a transit access index where available data exists to support local analysis. HUD uses data on over 200 transit agencies to assess relative accessibility within metro areas (or balance of state).

Exposure to Community Assets: To identify disparities in access or exposure to community assets, HUD PD&R calculates exposure indices for each asset dimension across a range of subgroups, including protected classes as identified in the Fair Housing Act. The exposure index calculates a weighted average for a given characteristic. To account for differences in household income across groups, PD&R also provides these exposure indices across protected classes for persons in poverty. This assists jurisdictions in understanding whether there are differences in exposure to community assets across groups that cannot be explained by differences in income. PD&R provides these exposure calculations for each non-white group (overall and in poverty) and the disparity relative to the white population (overall and in poverty).

Disproportionate Housing Needs: To assist communities in describing disproportionate housing needs in their geography, HUD is providing data which identifies instances when the incidence of housing problems are measurably higher for members of racial or ethnic groups than for the population as a whole. These measures are the same as those required in the data analysis associated with the Consolidated Plan.

A disproportionate housing need is defined as a circumstance when the members of racial or ethnic group within an income level experience housing problems at least 10 percentage points more frequently than the entire population at that income level. Four housing problems are considered:

- Lacks complete kitchen facilities
- Lacks complete plumbing facilities
- More than one person per room
- Cost Burden - monthly housing costs (including utilities) exceed 30 percent of monthly income

This data will be provided for four income categories: 0-30 percent of area median family income (AMI), 30-50 percent of AMI, 50-80 percent of AMI, and 80-100 percent of AMI. Additionally, HUD will provide a subset of "severe" housing problems, defined as:

- Lacking complete kitchen facilities
- Lacking complete plumbing facilities
- More than 1.5 persons per room
- Severe Cost Burden - monthly housing costs (including utilities) exceed 50 percent of monthly income

This data will report on the percentage of households in each category that have housing problems by race. Those cases where the difference in percentage exceeds 10 percentage points are deemed to indicate disproportionate housing needs.[201]

Application of ACS in the Methodology: The five-year ACS will be the sole data source for the poverty index (family poverty rate, percent of households receiving public assistance) and the labor market engagement index (unemployment rate, labor force participation rate, and percent with a bachelor's degree or higher).

Data to identify disproportionate housing needs will be drawn from the Comprehensive Housing Affordability Strategy (CHAS) data, which demonstrate the extent of housing problems and housing needs, particularly for low-income households. The CHAS data are produced for HUD via custom tabulations conducted by the U.S. Census Bureau using Census products.[202]

ACS Questions Utilized (as suggested by the methodology): number in household, P11, P35-37, P47-48, H7-8, H14, H16-18, H20-24.

---

[201] Text source: PD&R, "AFF Data Documentation," draft, June 2013.
[202] *Ibid*, pp. 7 and 9.

<u>Major Uses</u>: Informed by lessons learned in localities across the country, HUD issues this proposed rule, which provides new tools now available to help guide communities in fulfilling the original promise of the Fair Housing Act. The proposed rule involves refining the fair housing elements of the existing planning process that states, local governments, insular areas, and public housing agencies (program participants) now undertake. The process proposed by this rule assists these program participants to assess fair housing determinants, prioritize fair housing issues for response, and take meaningful actions to affirmatively further fair housing.

As recognized by HUD staff, program participants, civil rights advocates, the GAO, and others, the fair housing elements of current housing and community development planning are not as effective as they could be, do not incorporate leading innovations in sound planning practice, and do not sufficiently promote the effective use of limited public resources to affirmatively further fair housing. The approach proposed by the rule addresses these issues and strengthens AFFH implementation. It does so by providing data to program participants related to fair housing planning, clarifying the goals of the AFFH process, and instituting a more effective mechanism for HUD's review and oversight of fair housing planning. The proposed rule does not mandate specific outcomes for the planning process. Instead, recognizing the importance of local decision-making, it establishes basic parameters and helps guide public sector housing and community development planning and investment decisions to fulfill their obligation to affirmatively further fair housing. In addition, it helps educate other public sector agencies in their planning and investment decisions, and provides relevant civil rights information to the community and other private and public sector stakeholders.[203]

---

[203] Text source: Office of the Secretary, "Affirmatively Furthering Fair Housing: Proposed Rule," *Federal Register*, Vol. 78, No. 139, July 19, 2013, pp. 43710-43743.

62. *Comprehensive Housing Affordability Strategy Data*

Purpose and Nature of the Database: The primary purpose of the Comprehensive Housing Affordability Strategy (CHAS) data is to demonstrate the number of households in need of housing assistance from the Department of Housing and Urban Development (HUD). This is estimated by the number of households that have certain housing problems and have income low enough to qualify for HUD's programs (primarily 30, 50, and 80 percent of median income). It also considers the prevalence of housing problems among different types of households, such as the elderly, disabled, minorities, and different household types. The CHAS data provide counts of the numbers of households that fit these HUD-specified characteristics in HUD-specified geographic areas.[204]

CHAS datasets are special tabulations of ACS data. They may be based on one-, three-, or five-year ACS estimates. The May 2014 release relied primarily on five-year ACS data and secondarily on three-year ACS data (to measure housing affordability of people with disabilities).

Responsible Unit: Office of Policy Development & Research (PD&R), HUD.

Authorization:

- Requirement for a local housing strategy: 42 USC Chapter 130, Section 12705, Special Consolidated Housing Affordability Strategy (CHAS).
- Use of ACS data: According to 24 CFR 91.205(a), "Housing data included in this portion of the plan shall be based on U.S. Census data, as provided by HUD, as updated by any properly conducted local study, or any other reliable source that the jurisdiction clearly identifies . . . ."

Frequency of Data Publication: The Census Bureau annually provides special CHAS tabulations based on ACS data.

Timing of Data Release: HUD releases the special tabulations about 24-29 months after the end of the tabulations' final calendar year. So, for instance, CHAS data based on the 2007-2011 and 2009-2011 ACS were released in May 2014.[205]

Modes to Access Data: CHAS Data Download Page (csv); CHAS Data Query Tool (html, xls, pdf).

Methodology/Application of ACS Data in Methodology: HUD first obtained the CHAS data after the 1990 Census, and again after the 2000 Census. Since 2005, the Census Bureau has administered the American Community Survey (ACS), which replaced the detailed survey portion of the decennial census. The ACS offers timely data for the period between censuses, allowing for a relatively current picture of local conditions. The transition from the Census long-

---

[204] Text source: "CHAS: Background," PD&R website.
[205] "CHAS Data Documentation," PD&R website.

form to the ACS forced HUD to revise the CHAS data, so the 1990 CHAS and 2000 CHAS differ in some important ways from the ACS-based CHAS data released in 2009 and later.

HUD has identified a large set of characteristics of interest to housing planners and policy-makers, and as a result the CHAS data can be quite unwieldy. To streamline the data and make it easier to use, HUD has created a series of "tables," which are grouped by theme.

Each of these tables contains certain "dimensions" (also referred to as variables). These dimensions can be combined in a number of ways, and the data files for each table present every possible combination of those dimensions. As an example, consider Table 11. Table 11 contains 3 dimensions: tenure, housing problems, and household income. Tenure has 2 options: owner-occupied or renter-occupied. Housing problems has 3 options: household has at least one housing problem, household has no housing problems, or household has no income (so cost burden could not be computed) but no other housing problems. Household income, in this particular table, has 13 options. Thus Table 11 essentially has 78 buckets (2*3*13=78), and every household belongs in one (and only one) of those buckets. In the CHAS data, we have counted the number of households in each of those buckets, for thousands of states, counties, cities, and neighborhoods.

CHAS relies on the following definitions, each of which is based on Census data or definitions:

- Household – We use the Census designation of households, which is all people living in a housing unit. Members of a household can be related (see family) or unrelated.
- Household Income – The CHAS tabulations use adjusted household income, which includes the income of all members of the household at the time of the survey.
- Family – We use the Census designation of family, which is related individuals living in the same household. The Census Bureau also tracks subfamilies.
- Housing Problems – There are four housing problems in the CHAS data: 1) housing unit lacks complete kitchen facilities; 2) housing unit lacks complete plumbing facilities; 3) household is overcrowded; and 4) household is cost burdened. A household is said to have a housing problem if they have any 1 or more of these 4 problems.
- Overcrowding – More than 1 person per room.
- Severe overcrowding – More than 1.5 persons per room.
- Cost burden – monthly housing costs (including utilities) exceed 30% of monthly income.
- Severe cost burden – monthly housing costs (including utilities) exceed 50% of monthly income.
- Elderly – HUD defines elderly as age 62 and up. Individuals age 75 and up are generally recognized as a population with different needs than those 62-74, so the CHAS data separates these groups. "Elderly" refers to individuals 62-74, while those 75 and up may be referred to as "extra elderly" or "frail elderly".
- Disabled – The Census asks a series of questions related to physical and mental handicaps. For the CHAS data, HUD defines disabled as having a "mobility or self-care limitation"—for example, being unable to run errands outside the house

without assistance. Disability questions on the ACS were modified between 2007 and 2008, so HUD is unable to provide tabulations of disability data spanning that break.[206]

Census prepares 24 tables of CHAS data. In the current version, 23 of these are based on five-year ACS data (2007-2011) and one (focused on disabilities) is based on three-year ACS data (2009-2011). HUD provides detailed spreadsheets of CHAS data dictionaries, organized by table.[207]

CHAS includes also HUD Area Median Family Income estimates, which are determined outside the CHAS process, also based on ACS data, and described in section 18 below.

ACS Questions Utilized (as suggested by the methodology): P4-6, P17-20, P47-48, H1-2, H7-8, H17-19, H22-23

Major Uses: The CHAS data are used by local governments to plan how to spend HUD funds, and may also be used by HUD to distribute grant funds.[208] In addition to estimating low-income housing needs, the CHAS data contribute to a more comprehensive market analysis by documenting issues like lead paint risks, "affordability mismatch," and the interaction of affordability with variables like age of homes, number of bedrooms, and type of building.[209]

---

[206] Text source: "CHAS Background," PD&R website.
[207] Text source: "CHAS Data Documentation," PD&R website.
[208] Text source: "Consolidated Planning/CHAS data," PD&R website.
[209] Text source: "CHAS Background," PD&R website.

63. *Fair Market Rents*

Nature and Purpose of the Measure: Fair Market Rents (FMRs) are gross rent estimates of specific geographic areas that determine the amount of housing assistance that lower-income families can receive to rent safe and decent housing. FMRs are used to determine payment amounts for federal housing assistance programs. The U.S. Department of Housing and Urban Development (HUD) annually estimates FMRs for 530 metropolitan areas and 2,045 nonmetropolitan county FMR areas.

HUD defines FMR areas as metropolitan areas and non-metropolitan counties. Most current Office of Management and Budget (OMB) definitions of metropolitan areas are used because of the generally close correspondence between them and housing market areas. FMRs are intended to be housing market-wide rent estimates that provide housing opportunities throughout the geographic area in which rental units are in direct competition.

As gross rent estimates, FMRs include the shelter rent plus the cost of all tenant-paid utilities, except telephones, cable or satellite television service, and internet service. HUD sets FMRs to assure that a sufficient supply of rental housing is available to program participants. To accomplish this objective, FMRs must be both high enough to permit a selection of units and neighborhoods and low enough to serve as many low-income families as possible. [210]

Responsible Unit: Office of Policy Development and Research (PD&R), Department of Housing and Urban Development

Authorization: Section 8(c)(1) of the United States Housing Act of 1937 and 24 CFR 888.113. According to 24 CFR 888.113 e(1), "HUD uses the most accurate and current data available to develop the FMR estimates and may add other data sources as they are discovered and determined to be statistically valid."

Frequency of Data Publication: Annual

Timing of Data Release: By law, the final FMRs for use in any fiscal year must be published and available for use at the start of that fiscal year. HUD's regulations at 24 CFR part 888 provide that HUD will develop proposed FMRs, publish them for public comment, provide a public comment period of at least 30 days, analyze the comments, and publish final FMRs. (See 24 CFR 888.115.) The notice of FY2015 Final FMRs was published on October 3, 2014.[211]

Modes to Access Data: Fair Market Rents Data Sets, HUD USER, PD&R (pdf, csv, web-based text)

Methodology: The level at which FMRs are set is expressed as a percentile point within the rent distribution of standard-quality rental housing units. The current definition used is the 40th percentile rent, the dollar amount below which 40 percent of the standard-quality rental

---

[210] Text source: PD&R, "Fair Market Rents for the Section 8 Housing Assistance Payments Program," July 2007.
[211] Text source: PD&R, "Final Fair Market Rents for the Housing Choice Voucher Program and Moderate Rehabilitation Single Room Occupancy Program Fiscal Year 2015," *Federal Register*, October 3, 2014.

housing units are rented. The 40th percentile rent is drawn from the distribution of rents of all units occupied by recent movers (renter households who moved to their present residence within the past 15 months). HUD is required to ensure that FMRs exclude non-market rental housing in their computation. Therefore, HUD excludes all units falling below a specified rent level determined from public housing rents in HUD's program databases as likely to be either assisted housing or otherwise at a below-market rent, and units less than two years old.

The vast majority of areas remain at the 40th percentile rent. However, certain areas are assigned the 50th percentile rent. Fiftieth percentile FMRs were established by a rule published on October 2, 2000, that also established the eligibility criteria used to select areas that would be assigned 50th rather than the normal 40th percentile FMRs.[212]

<u>Application of ACS Data in Methodology</u>: The methodology for estimating FMRs for metropolitan areas and non-metropolitan FMR areas relies heavily on ACS data. For Fiscal Year 2014, FMRs were developed as follows:

1. HUD used special tabulations of 5-year ACS data collected between 2008 through 2012. For FY 2015 FMRs, HUD updated the base rents set in FY 2014 using the 2007–2011 5-year data with the 2008–2012 5-year ACS data. For FY2014, 2007-2011 5-year American Community Survey (ACS) estimates of 2-bedroom adjusted standard quality gross rents were calculated for each FMR area. In areas where the 2007-2011 5-year ACS 2-bedroom adjusted standard quality gross rent estimate was less than its respective margin of error, the state non-metro estimate of 2-bedroom adjusted standard quality gross rent was used.

2. HUD historically based FMRs on gross rents for recent movers (those who have moved into their current residence in the last 24 months). However, due to the nature of the 5-year ACS data, HUD developed a new methodology for calculating recent-mover FMRs in FY 2012. As in FY 2012, HUD assigns all areas a base rent which is the estimated two-bedroom standard quality 5-year gross rent from the ACS. Because HUD's regulations mandate that FMRs represent recent mover gross rents, HUD continues to apply a recent mover factor to the standard quality base rents assigned from the 5-year ACS data. Calculation of the recent mover factor is described below.

3. Following the assignment of the standard quality two-bedroom rent described above, HUD applies a recent mover factor to these rents. The calculation of the recent mover factor for FY 2015 is similar to the methodology used in FY 2014, with the only difference being the use of updated ACS data. The following describes the process for determining the appropriate recent mover factor. In general, HUD uses the 1 year ACS-based two-bedroom recent mover gross rent estimate from the smallest geographic area encompassing the FMR area for which the estimate is statistically reliable to calculate the recent mover factor. HUD calculates some areas'

---

[212] Text source: *Ibid*.

recent mover factors using data collected just for the FMR area. However, HUD bases other areas' recent mover factor on larger geographic areas if this is necessary to obtain statistically reliable estimates. For metropolitan areas that are sub-areas of larger metropolitan areas, the order is FMR area, metropolitan area, aggregated metropolitan parts of the state, and state. Metropolitan areas that are not divided into subparts follow a similar path from FMR area, to aggregated metropolitan parts of the state, to state. In nonmetropolitan areas the recent mover factor is based on the FMR area, aggregated nonmetropolitan parts of the state, or if that is not available, on the basis of the whole state. HUD calculates the recent mover factor as the percentage change between the 5-year 2008–2012 standard quality two-bedroom gross rent and the 1-year 2012 recent mover two-bedroom gross rent for the recent mover factor area. HUD does not allow recent mover factors to lower the standard quality base rent; therefore, if the 5-year standard quality rent is larger than the comparable 1-year recent mover rent, the recent mover factor is set to 1.

4. HUD updates the ACS-based ''as of'' 2012 rent through the end of 2013 using the annual change in CPI from 2012 to 2013. As in previous years, HUD uses Local CPI data coupled with Consumer Expenditure Survey (CEX) data for FMR areas with at least 75 percent of their population within Class A metropolitan areas covered by local CPI data. HUD uses Census region CPI data for FMR areas in Class B and C size metropolitan areas and nonmetropolitan areas without local CPI update factors. Additionally, HUD is using CPI data collected locally in Puerto Rico as the basis for CPI adjustments from 2012 to 2013 for all Puerto Rico FMR areas. Following the application of the appropriate CPI update factor, HUD converts the ''as of'' 2013 CPI adjusted rents to ''as of'' December 2013 rents by multiplying each rent by the national December 2013 CPI divided by the national annual 2013 CPI value.

5. As in FY 2014, HUD continues to calculate the trend factor as the annualized change in median gross rents as measured across the most recent 5 years of available 1-year ACS data. The national median gross rent in 2007 was $789 and $884 in 2012. The overall change between 2007 and 2012 is 12.04 percent and the annualized change is 2.30 percent. Over a 15-month time period, the effective trend factor is 2.883 percent. HUD applies this trend factor to the ''as of'' December 2013 rents to produce FMRs that correspond to the middle of the 2015 fiscal year.

6. HUD calculates the primary FMR estimates for two-bedroom units. This is generally the most common sized rental unit and, therefore, the most reliable to survey and analyze. Formerly, after each decennial Census, HUD calculated rent relationships between two-bedroom units and other unit sizes and used them to set FMRs for other units. HUD did this because it is much easier to update two-bedroom estimates annually and to use pre-established cost relationships with other unit bedroom counts than it is to develop independent FMR estimates for each unit bedroom count. When calculating FY 2013 FMRs, HUD updated the bedroom ratio adjustment factors using 2006–2010 5-year ACS data using similar methodology to what was implemented when calculating bedroom ratios using 2000 Census data to

establish rent ratios. The bedroom ratios used in the calculation of FY 2015 FMRs remain the 2006–2010 based ratios applied to the two-bedroom FMR computed from the 2012 ACS data.

7. HUD established bedroom interval ranges based on an analysis of the range of such intervals for all areas with large enough samples to permit accurate bedroom ratio determinations. These ranges are: Efficiency (zero-bedroom) FMRs are constrained to fall between 0.59 and 0.81 of the two-bedroom FMR; one-bedroom FMRs must be between 0.74 and 0.84 of the two-bedroom FMR; three-bedroom FMRs must be between 1.15 and 1.36 of the two-bedroom FMR; and four-bedroom FMRs must be between 1.24 and 1.64 of the two-bedroom FMR. (The maximums for the three-bedroom and four-bedroom FMRs are irrespective of the adjustments discussed in the next paragraph.) HUD adjusts bedroom rents for a given FMR area if the differentials between unit bedroom-count FMRs were inconsistent with normally observed patterns (i.e., efficiency rents are not allowed to be higher than one-bedroom rents and four-bedroom rents are not allowed to be lower than three-bedroom rents). The bedroom ratios for Puerto Rico follow these constraints.

8. HUD further adjusts the rents for three-bedroom and larger units to reflect HUD's policy to set higher rents for these units than would result from using unadjusted market rents. This adjustment is intended to increase the likelihood that the largest families, who have the most difficulty in leasing units, will be successful in finding eligible program units. The adjustment adds 8.7 percent to the unadjusted three-bedroom FMR estimates and adds 7.7 percent to the unadjusted four-bedroom. FMR estimates. The FMRs for unit sizes larger than four bedrooms are calculated by adding 15 percent to the four-bedroom FMR for each extra bedroom. For example, the FMR for a five-bedroom unit is 1.15 times the four-bedroom FMR, and the FMR for a six-bedroom unit is 1.30 times the four-bedroom FMR. FMRs for single-room occupancy units are 0.75 times the efficiency FMR.

9. For low-population, nonmetropolitan counties with small or statistically insignificant 2006–2010 5-year ACS recent-mover rents, HUD uses state nonmetropolitan data to determine bedroom ratios for each unit bedroom count. HUD made this adjustment to protect against unrealistically high or low FMRs due to insufficient sample sizes.[213]

ACS Questions Utilized (as suggested by methodology): H1, H7, H18

Major Uses: Fair Market Rent estimates are used to set the payment standards of federal housing assistance programs.

---

[213] Text source: PD&R, "Final Fair Market Rents for the Housing Choice Voucher Program and Moderate Rehabilitation Single Room Occupancy Program Fiscal Year 2014," *Federal Register*, vol. 78 (no. 192), October 3, 2013.

- FMRs are used to determine payment standard amounts for the Housing Choice Voucher program, the Moderate Rehabilitation program, the project-based voucher program, and other housing assistance programs.
- Section 8 FMRs for the Voucher program (24 CFR 888 Subpart A) currently serve as the payment standard for approximately one million rental units. Under this program, HUD subsidizes the difference between the FMR payment standard and 30 percent of the incomes of participating households. Subsidy outlays associated with this program total over $5 billion annually, and outlays for any given unit are a function of the applicable FMR.[214]
- Based on Section 210 of Division L, Title II of the Consolidated Appropriations Act, 2014, Public Housing Authorities administering Public Housing must use these FMRs in calculating Flat Rents for public housing.

In addition, FMRs are contained in PD&R's Housing Affordability Data System (HADS), a set of housing-unit level datasets that measures the affordability of housing units and the housing cost burdens of households, relative to area median incomes, poverty level incomes, and Fair Market Rents. The purpose of these datasets is to provide housing analysts with consistent measures of affordability and burdens over a long period.[215]

---

[214] Text source: PD&R, "Supporting Statement for HUD Section 8 Fair Market Rent Random Digit Dialing Surveys," Submission to Office of Management and Budget Information Collection Review, November 10, 2011.
[215] David A. Vandenbroucke, "Housing Affordability Data System," PD&R, January 28, 2011.

64. *Median Family Income*

Purpose of the Measure: The Department of Housing and Urban Development (HUD) is required by law to set income limits that determine the eligibility of applicants for HUD's assisted housing programs. Median Family Income (MFI) estimates are the basis of calculating income limits for a variety of federal housing assistance programs, including the Public Housing program, the Section 8 Housing Choice Voucher program, Section 202 housing for the elderly program, Section 811 housing for persons with disabilities program, the Low Income Housing Tax Credit, and Multifamily Tax Subsidy Projects.[216]

Responsible Unit: Office of Policy Development & Research (PD&R), Department of Housing and Urban Development.

Authorization: The statutory basis for HUD's income limit policies is Section 3 of the U.S. Housing Act of 1937, as amended.

Frequency of Data Publication: Annual

Timing of Data Release: PD&R published the FY2014 MFI estimates in December 2013.

Modes to Access Data: Income Limits, HUD USER, PD&R (pdf).

Methodology/Application of ACS Data in Methodology: HUD updated the methodology to produce Median Family Income (MFI) estimates to take advantage of new data available from the Bureau of the Census' American Community Survey (ACS). In June, 2012, the Census Bureau released 5-year ACS and Puerto Rico Community Survey (PRCS) data aggregated from 2007 through 2011. The FY 2014 MFI estimates use the 2007 -2011 5-year ACS and PRCS data, augmented by the 2011 1-year ACS and PRCS information and updated with Consumer Price Index (CPI) data through the end of 2012. The factor used to trend the 2012 estimates to the midpoint of FY 2014 MFIs is 0.98 percent per year. This factor was calculated as the annualized change in national median family income as measured by the 2006 1-year ACS and the 2011 1-year ACS.

- The ACS, conducted annually, was designed to produce estimates similar to the longform sample survey previously conducted with the Decennial Census after 5 years of data became available to be aggregated together for a given area. Each year since full implementation of the survey in 2005, the Census Bureau collected an ACS sample sufficient to provide estimates of most survey items for areas with populations of 65,000 or more. After the 2007 ACS, the Census Bureau released data aggregated from the ACS samples collected over the three years, 2005, 2006, and 2007. This allowed the Census Bureau to release estimates for most items for areas with populations of 20,000 or more. FY 2010 MFIs reflected ACS survey data aggregated over 2006, 2007 and 2008. Since the 2009 ACS sample, the Census Bureau has sufficient data to release aggregated 5-year estimates. Five year

---

[216] Text source: PD&R, "FY 2014 HUD Income Limits Briefing Material," March 18, 2014.

estimates are designed to provide estimates for areas of all sizes relevant to MFI and income limit production.

- As mentioned above, the FY 2014 MFIs incorporated the 2007-2011 5-year ACS data into the calculation process. Specifically, for each metropolitan area, subarea of a metropolitan area, and non-metropolitan county, HUD used 5-year ACS data as the new basis for calculating MFI estimates. In areas with a valid 1-year ACS survey median family income result, HUD incorporated this data as well to take advantage of more recent survey information. By using both the 5-year data and the 1-year data, where available, HUD established a new basis for median family income estimates while also capturing the most recent information available.

- HUD begins by setting the base median family income equal to the 2007-2011 5-year ACS survey value. For areas with a valid 2011 1-year survey result, HUD uses the margin of error for the 1-year data to ensure that the 1-year MFI are statistically reliable. Where the 1-year data are statistically reliable (i.e., where the margin of error is lower than the estimate itself), HUD uses the 1-year survey result as the base value for median family income. In the few cases where the statistical confidence interval for the 5-year ACS estimate of median family income includes zero, HUD assigns the state nonmetropolitan median.

- For FY 2014, MFI estimates based on the 2011ACS data (the middle of 2011) were updated to the end of 2012 using CPI data. The national CPI-U is used in the CPI adjustment calculation. For Income Limits, FY 2014 40th percentile rents as calculated as part of the FY 2014 FMR calculation process are used to determine if a high housing cost adjustment is necessary and, if so, the level of that adjustment.

- MFI estimates are based on the most currently available data, but the delay in collecting and reporting the survey data mean that 2011 ACS income data is used for FY 2014 estimates that have an as-of date of April 1, 2014. A trend factor based on the most recent historic patterns of nominal income growth is used to inflate the estimate from the end of 2012 to April, 2014. As mentioned earlier, the trend factor for FY 2014 is 0.98 percent which is the annualized growth rate in national median family income as captured by the 1-year 2006 and 1-year 2011 ACS income data.

- Median family incomes start with the development of estimates of median family income for the metropolitan areas and non-metropolitan FMR/income limit areas (including U.S. territories). HUD uses 2007-2011 5-year ACS estimates of median family income calculated for the areas used for FMRs and income limits as the basis for FY 2014. In areas where there is also a valid 2011 1-year ACS estimate of median family income, HUD replaces the 5-year data with the 1-year data. A valid 1-year 2011 ACS estimate is one where the margin of error of the estimate is less than the estimate itself. Once the appropriate 2011 ACS data has been selected, the data are set as of December 2012 using the December 2012 national CPI value divided by the 2011 national annual average CPI value.

There are currently several legislated income limit standards (e.g., 30%, 50%, 60%, 65%, 80%, 95%, 100%, 115%, 125%) that were intended to have progressive relationships. To ensure that this occurs, the very low-income limits have been used as the basis for deriving other income

limits unless that relevant statutory language has no references or relationship to low and very low-income limits as defined by the U.S. Housing Act of 1937. [217]

ACS Questions Utilized (as suggested by the methodology): P47-48

Major Uses: ACS-based MFI estimates are the foundation for calculating income limits for a substantial number of HUD and other federal programs, including:

- HUD
  - Public Housing
  - All Section 8 Programs
  - Indian Housing (1996 Act)
  - Section 202 Elderly and Section 811 Handicapped programs
  - Section 235 (Homeownership program)
  - Section 236 (Rental program)
  - Section 221(d)(3) (BMIR)(Below Market Interest Rate) rental program
  - Community Planning and Development programs
  - HOME Investment Partnerships Act of 1990
  - National Homeownership Trust Act of 1990
  - Neighborhood Stabilization Program
  - Low-Income Housing Preservation and Resident Homeownership Act of 1990
- Rural Housing and Community Development Service – Rental and ownership assistance programs
- Treasury Programs
  - Multifamily Tax Subsidy Projects
  - Tax-exempt Mortgage Revenue Bonds for homeownership financing
  - "Difficult Development Area" Designation (Low-Income Housing Tax Credit)
  - "Qualified Census Tract" (Low-Income Housing Tax Credit Program Definition)
  - "Qualified Census Tract" (Mortgage Revenue Bond Program)
- Federal Deposit Insurance Corporation
  - Disposition of Multifamily Housing to Non-profit and Public Agencies
  - Disposition of Single Family Housing
- Federal Home Loan Banks
  - Rental program funding priorities
  - Homeownership funding priorities
- Federal Housing Finance Agency – Income-based Housing Goals of Freddie Mac and Fannie Mae
- Other Federal Banking Regulatory Provisions -- Targeting of loan funds to low-income households and areas
- Uniform Relocation Act – Reimbursement to households forced to relocate from their residence by Federal agency

---

[217] *Ibid*.

- Department of Veterans Affairs – Eligibility for disability income support payments to veterans[218]

---

[218] *Ibid*.

65. *Small Area Fair Market Rents*

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

66. *Special Tabulation of Households*

Nature and Purpose of Dataset: These special tabulations of households are statistical summaries of counts of households by tenure (owner/renter); by income intervals; by age of householder; by size of household; and by housing conditions. This system allows a user to extract data to conduct a longitudinal analysis of changes in a particular area. Beginning in 2010, the system uses data from the Census ACS 5-year survey.[219]

Responsible Unit: Economic and Market Analysis Division, Office of Policy Development and Research (PD&R), Department of Housing and Urban Development

Authorization: Title V of the 1970 Housing Act

Frequency of Data Publication: The data retrieval system is updated annually. Currently, five-year ACS data are available for the periods ending in 2010, 2011, and 2012.

Timing of Data Release: The database is updated within a year of Census publication of the five-year ACS estimates.

Modes to Access Data: Special Tabulations Data Retrieval System, PD&R (html)

Methodology/Application of ACS Data in Methodology: The Census Bureau produces a series of special tabulations according to PD&R specifications, with cross-tabulations by tenure, income, age of householder, and housing conditions for specific geographies (counties and county equivalents, places of with populations of 50,000, the Nation, each State, and metropolitan statistical areas).

ACS Questions Utilized (as suggested by the methodology): P4, P47-48, H7-8, H17-18, H22

Major Uses: These special tabulations data have been developed for the assistance and guidance of HUD in its operations. They are the most detailed available for a qualitative analysis of housing demand based on incomes and age of householder. The data are a key element in the allocation formulae for the Section 8 and the Section 202 rental assistance programs, as well as a key element in qualitative demand market analysis activities for review of program applications and multifamily mortgage insurance applications submitted to the Federal Housing Administration. They also are useful to planners, developers, appraisers, third party market analysts and mortgagees, and others concerned with local housing conditions.[220]

---

[219] Text source: "Special Tabulations of Households," PD&R website.
[220] *Ibid.*

X. U.S. Department of Justice

A. Bureau of Justice Statistics

67. *Crime Against Persons with Disabilities*

Nature and Purpose of the Measure: Presents estimates of nonfatal violent victimization (rape, sexual assault, robbery, aggravated and simple assault) against persons age 12 or older with disabilities. Findings are based on the National Crime Victimization Survey (NCVS). The report compares the victimization of persons with and without disabilities living in noninstitutionalized households, including distributions by age, race, sex, victims' types of disabilities, and other victim characteristics. Data are for the nation only.[221]

Responsible Unit: Bureau of Justice Statistics (BJS), Department of Justice

Authorization: Title 42, United States Code, Section 3732 of the Justice Systems Improvement Act of 1979, authorizes BJS to collect statistics on victimization.

Frequency of Data Publication: Annual

Timing of Data Release: The 2009-2012 report was made available in February 2014.

Modes to Access Data: Publications & Products: Crime Against People with Disabilities, BJS (pdf, ascii, csv)

Methodology: The report series is based on data from the National Crime Victimization Survey (NCVS), an ongoing data collection conducted by the U.S. Census Bureau for the Bureau of Justice Statistics (BJS). The NCVS is a self-report survey in which interviewed persons are asked about the number and characteristics of victimizations experienced during the prior 6 months. The NCVS collects information on nonfatal personal crimes (rape or sexual assault, robbery, aggravated and simple assault, and personal larceny) and household property crimes (burglary, motor vehicle theft, and other theft) both reported and not reported to police. In addition to providing annual level and change estimates on criminal victimization, the NCVS is the primary source of information on the nature of criminal victimization incidents.

Survey respondents provide information about themselves (e.g., age, sex, race and Hispanic origin, marital status, education level, and income) and whether they experienced a victimization. Information is collected for each victimization incident about the offender (e.g., age, race and Hispanic origin, sex, and victim-offender relationship), characteristics of the crime (including time and place of occurrence, use of weapons, nature of injury, and economic consequences), whether the crime was reported to police, reasons the crime was or was not reported, and experiences with the criminal justice system.

The NCVS is administered to persons age 12 or older from a nationally representative sample of households in the United States. The NCVS defines a household as a group of members who all

---

[221] Text source: Crime Against Persons with Disabilities, 2009–2012 - Statistical Tables, BJS website.

reside at a sampled address. Persons are considered household members when the sampled address is their usual place of residence at the time of the interview and when they have no usual place of residence elsewhere. Once selected, households remain in the sample for 3 years, and eligible persons in these households are interviewed every 6 months either in-person or over the phone, for a total of seven interviews.

Generally, all first interviews are conducted in-person. New households rotate into the sample on an ongoing basis to replace outgoing households that have been in the sample for the 3-year period. The sample includes persons living in group quarters, such as dormitories, rooming houses, and religious group dwellings, and excludes persons living in military barracks and institutional settings, such as correctional or hospital facilities, and the homeless.[222]

<u>Application of ACS Data in Methodology</u>: In two ways, the ACS is utilized in the preparation of the data series on crime against persons with disabilities.

First, the NCVS adopts questions from the U.S. Census Bureau's American Community Survey (ACS) to measure the rate of victimization against people with disabilities. The NCVS does not identify persons in the general population with disabilities. In 2007, the ACS Subcommittee on Disability Questions developed the disability questions based on questions used in the 2000 Decennial Census and earlier versions of the ACS. The questions identify persons who may require assistance to maintain their independence, be at risk for discrimination, or lack opportunities available to the general population because of limitations related to a prolonged (6 months or longer) sensory, physical, mental, or emotional condition.

In 2008, the U.S. Census Bureau changed some of the disability questions on the ACS. The question about sensory disability was separated into two questions about blindness and deafness, and the questions about physical disability were asked only about serious difficulty walking or climbing stairs. Also, questions on employment disability and going outside of the home were eliminated in 2008. Census Bureau analysis of 2007 and 2008 ACS disability data revealed significant conceptual and measurement differences between the 2007 and 2008 disability questions. The Census Bureau concluded that data users should not compare the 2007 estimates of the population with disabilities and those of later years. Because the 2007 and 2008 NCVS disability questions mirrored the ACS, estimates of victimization of people with disabilities from the 2007 and 2008 NCVS should not be compared.

Second, ACS data are used to generate estimates of populations of persons with disabilities and without disabilities by various demographic characteristics, including sex, race, ethnicity, and age. These figures serve as the denominator in estimates of age-adjusted victimization rates for persons with and without disabilities, by type of disability, with the numerator being provided by the NCVS. For 2008 through 2010, the Census Bureau ran special tabulations for BJS to

---

[222] For more information on NCVS methodology, see the NCVS website, Bureau of Justice Statistics.

generate the population estimates. However, due to budgetary restrictions, for 2011 and 2012 BJS used the ACS Public Use Microdata Sample (PUMS) for this purpose.[223]

<u>ACS Questions Utilized (as suggested by the methodology)</u>: P3-6, P17-19

<u>Major Uses</u>: The data series on crimes against persons with disabilities guides the efforts of law enforcement agencies and organizations that advocate for persons with disabilities.

---

[223] Erika Harrell, <u>"Crime against Persons with Disabilities, 2009–2012 - Statistical Tables,"</u> BJS, February 2014, pp. 9-11.

XI. U.S. Department of Labor

A. Bureau of Labor Statistics

68. Consumer Price Index

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

69. *Current Population Survey – Labor Force*

Purpose of the Dataset: The Current Population Survey (CPS) is a monthly survey of households conducted by the Bureau of Census for the Bureau of Labor Statistics (BLS). It provides a comprehensive body of data on the labor force, employment, unemployment, persons not in the labor force, hours of work, earnings, and other demographic and labor force characteristics.[224]

The CPS serves as the foundation for a series of BLS-funded supplemental data collections on the U.S. labor force, including:

- Veterans (with the Department of Veterans Affairs)
- Displaced workers
- Contingent workers
- Job tenure and occupational mobility
- Unemployment insurance
- Disability (with the Department of Labor's Office of Disability Employment Policy)
- School enrollment (with Census and the National Center for Education Statistics)
- Fertility (with Census)
- Volunteers (with the Corporation for National and Community Service)
- Civic engagement (with the Corporation for National and Community Service)

Responsible Unit: Bureau of Labor Statistics (BLS), carried out by the Census Bureau.

Authorization: Title 13, United States Code, Section 182, and Title 29, United States Code, Sections 1-9

Frequency of Data Publication: BLS publishes data reports on the following topics based on the monthly CPS:

- Employment status (A,Q,M)[225]
- Characteristics of the employed (A,Q,M)
- Characteristics of the unemployed (A,Q,M)
- Veteran status (A,Q,M)
- Persons not in the labor force (A,Q)
- Multiple jobholders (A,Q)
- Weekly earnings (A,Q)
- Geographic Profile of Employment and Unemployment (A)
- Union affiliation (A)
- Minimum wage workers (A)
- Absences from work (A)
- Women in the labor force (A)

---

[224] Text source: "Labor Force Statistics from the Current Population Survey," BLS website.
[225] Annual, Quarterly, Monthly

- Women's earnings (A)
- Labor force characteristics by race and ethnicity (A)
- Labor force characteristics of persons with a disability (A)

Data from BLS-sponsored supplements appear on the following schedule:

- Annual – veterans, volunteers, school enrollment, civic engagement, disability, the working poor
- Biennial (even years) – displaced workers, job tenure and occupational mobility, fertility
- Once a decade – contingent workers

Timing of Data Release: The CPS Employment Situation for a given month is issued on the first Friday of the next month. Publication schedules for the other reports vary.

Modes to Access Data: Labor Force Statistics from the CPS, BLS

- By format
  - Databases
  - Tables
  - Publications
- by subject:
  - Demographics
  - Labor Force Characteristics
  - Earnings
  - Disability
- Geographic Profile of Employment and Unemployment
- Employment and Earnings Online

Methodology: The universe for the CPS is about 136 million households. From their Master Address File, the Census Bureau selects approximately 72,000 households each month. Of these, approximately 60,000 households are eligible for interviews. (Note: 'Eligible' can be simplistically defined as an occupied housing unit having at least one person in the civilian noninstitutional population.) The Census Bureau actually interviews about 55,000 households each month. This sample includes 10,000 eligible housing units from the monthly supplementary sample to improve state-level estimates of health insurance coverage for low-income children, also known as the CHIP expansion.

The CPS sample is a probability sample based on a stratified two-stage sampling scheme: selection of sample primary sampling units (PSUs) and selection of sample housing units within those PSUs. In general, the CPS sample is selected from lists of addresses obtained from the Master Address File (MAF) with updates from the United States Postal Service (USPS) twice a year.

In the first stage of sampling, PSUs are selected. These PSUs consist of counties or groups of contiguous counties in the United States, and are grouped into strata. The CPS is a state-based

design. Therefore, all PSUs and strata are defined within state boundaries and the sample is allocated among the states to produce state and national estimates with the required reliability, while keeping total sample size to a minimum.

The variables chosen for grouping PSUs in each state into strata reflect the primary interest of the CPS in maximizing the reliability of estimates of labor force characteristics. Basically, the same set of stratification variables, from the 2010 Decennial Census and the American Community Survey (ACS), are used for each state: unemployment statistics by gender; number of families maintained by a woman; and the proportion of occupied housing units with three or more people. In addition, the number of persons employed in selected industries and the average monthly wage for selected industries are used as stratification variables in some states. The industry-specific data are averages over the period 2000 through 2008 and are obtained from the Quarterly Census of Employment and Wages program of the BLS.

The 2010 sample design comprises three frames: unit, coverage improvement (CI) and group quarters (GQ). Within these sampling frames, housing units are sorted based on characteristics of the ACS and geography. Then, from each frame, a systematic sample of addresses within the sample PSUs is obtained.

Each sample is divided into eight approximately equal panels, called rotation groups. A rotation group is interviewed for four consecutive months, temporarily leaves the sample for eight months, and then returns for four more consecutive months before retiring permanently from the CPS (after a total of eight interviews).[226]

This survey's estimation procedure adjusts weighted sample results to agree with independently derived population estimates of the civilian noninstitutionalized population of the United States and each state (including the District of Columbia). These population estimates, used as controls for the CPS, are prepared monthly to agree with the most current set of population estimates that are released as part of the Census Bureau's population estimates and projections program.

The population controls for the nation are distributed by demographic characteristics in two ways:

- Age, sex, and race (White alone, Black alone, and all other groups combined)
- Age, sex, and Hispanic origin

The population controls for the states are distributed by race (Black alone and all other race groups combined), age (0-15, 16-44, and 45 and over), and sex. The independent estimates by age, sex, race, and Hispanic origin, and for states by selected age groups and broad race categories, are developed using the basic demographic accounting formula whereby the population from the 2010 Decennial Census data is updated using data on the components of

---

[226] Text source: Bureau of Labor Statistics, "Supporting Statement, Current Population Survey (OMB Control Number 1220-0100)," Part B, April 7, 2014.

population change (births, deaths, and net international migration) with net internal migration as an additional component in the state population estimates.[227]

Application of ACS Data in Methodology: The ACS is used in three ways to stratify the survey sample:

- Primary sampling units (PSUs) are designed in part on the basis of select ACS geographic characteristics—unemployment statistics by gender, female householder, and the number of people in a housing unit.
- The ACS provides the housing unit sampling frame within those PSUs.
- Sample housing units are selected from that frame on the basis of specific ACS characteristic microdata.[228]

ACS Questions Utilized (as suggested by the methodology): number of people, P3-6, P15, P29, P35-39

While it is possible that additional ACS questions could be used in sample stratification and estimation procedures for a particular topical supplement, the report authors have not found any examples so far.

Major Uses: The labor force information gathered through the survey is of paramount importance in keeping track of the economic health of the Nation. The survey is the official source of monthly data on total employment and unemployment, with the monthly Employment Situation report designated a Principal Federal Economic Indicator (PFEI). The CPS data are used monthly, in conjunction with data from other sources, to analyze the extent to which the various components of the American population are participating in the economic life of the Nation.

The labor force data gathered through the CPS are provided to users in the greatest detail possible, consistent with the demographic information obtained in the survey. In brief, the labor force data can be broken down by sex, age, race, ethnicity, marital status, family composition, educational level, disability status, and various other characteristics. Through such breakdowns, one can focus on the employment situation of specific population groups as well as on the general trends in employment and unemployment. Moreover, the survey yields data on the basic status and characteristics of people who have stopped looking for work because they believe no jobs are available – discouraged workers. Information of this type can be obtained only through demographically-oriented surveys such as the CPS.

In addition to being a source of much detailed data at the national level, the CPS provides information that is crucial in examining the employment situation at the sub-national level. For the 50 states, the District of Columbia, New York City, and the Los Angeles-Long Beach-Glendale

---

[227] Census Bureau, "Current Population Survey, December 2013, Food Security File," Technical Documentation, CPS-13, December 2013.

[228] Bureau of Labor Statistics, "Supporting Statement, Current Population Survey (OMB Control Number 1220-0100)," Part B, April 7, 2014.

metropolitan division and the respective balances of state, data are produced using estimating equations based on regression techniques. These models combine current and historical data from the CPS with data from other surveys.

The basic CPS data are also used as an important platform on which to base the data derived from the various supplemental questions administered in conjunction with the survey. By coupling the basic data from the monthly survey with the special data from supplements, one can gain valuable insights on the behavior of American workers and on the social and economic health of their families. Recent supplements have produced data on poverty, health insurance, volunteering, displaced workers, and veterans with service-connected disabilities, for example.

The monthly CPS data are first published by means of a news release, The Employment Situation, and, on the same day, are often discussed by the Commissioner of Labor Statistics before the Joint Economic Committee of the U.S. Congress. There is wide interest in this initial release among Government policy makers, legislators, economists, the media, and the general public. The data are subsequently published in much greater detail on the BLS website and in a variety of publications. Special analyses of data by BLS economists often are reported in the Monthly Labor Review, also published by BLS.

While the data from the CPS are used in conjunction with data from other surveys—notably the Current Employment Statistics (CES) survey—in assessing the economic health of the Nation, they are unique in various ways. They provide monthly, nationally representative measures of total employment that include self-employed workers, farm workers, and unpaid family workers. By contrast, establishment surveys are generally restricted to the nonagricultural wage and salary sector. The CPS provides timely and detailed data on all job seekers, and on all persons outside the labor force, while payroll-based surveys cannot, by definition, cover these sectors of the population, and other national surveys that include a few questions on employment (such as the American Community Survey) do not have sufficient space (or legal mandate) to collect information at this level of detail. Finally, the CPS data on employment, unemployment, and persons not in the labor force can be linked separately to the demographic characteristics of the many groups which make up the Nation's population, while data from other surveys are often devoid of demographic information.[229]

---

[229] Text source: Bureau of Labor Statistics, "Supporting Statement, Current Population Survey (OMB Control Number 1220-0100)," Part A, December 17, 2013.

70. *Local Area Unemployment Statistics*

Purpose of the Measure: The Local Area Unemployment Statistics (LAUS) program produces estimates of unemployment in states and local areas. These statistics are key indicators of local economic conditions, and are used by state and local governments for planning and budgetary purposes, and as determinants of the need for local employment and training services and programs. They also are used by economic forecasters, researchers, and bond and mortgage underwriters. In addition, local area unemployment estimates are used to determine the eligibility of an area for benefits in various federal assistance programs.

The LAUS program is a Federal-State cooperative program. The Bureau of Labor Statistics (BLS) develops concepts, definitions, and technical procedures and then works with state workforce agencies, who prepare labor force and unemployment estimates. Monthly and annual average estimates of employment and unemployment are prepared in state agencies for more than 7,300 unique geographic areas: states, the District of Columbia, and Puerto Rico; labor market areas, such as metropolitan and micropolitan areas; counties and equivalents; cities with a population of 25,000 or more; and all cities and towns in New England, regardless of population.[230]

Responsible Unit: Bureau of Labor Statistics (BLS), Department of Labor

Authorization: 29 U.S.C. § 1-8; 29 U.S.C. § 49l-2. The latter charges the Secretary of Labor with providing "statistics on employment and unemployment status of national, State, and local populations, including self-employed, part-time, and seasonal workers . . . ."

Frequency of Data Publication: Monthly and annual

Timing of Data Release: Three-four weeks after the end of the reference month, following this schedule.

Modes to Access Data: LAUS, BLS (xls, pdf, gif)

Methodology: In 1950, the Department of Labor's Bureau of Employment Security (now Employment and Training Administration) published a handbook, *Techniques for Estimating Unemployment*, so that comparable estimates of the unemployment rate could be produced for all states. This led to the formulation of the "handbook method" in the late 1950s. The handbook method is a series of computational steps designed to produce local employment and unemployment estimates, using available data at a much lower cost than a direct survey. The handbook method relies heavily on data derived from the unemployment insurance (UI) system.

In 1972, BLS assumed technical responsibility for the program and began to refine the concepts and methods used to estimate the labor force, employment, and unemployment at the subnational level. In 1973, a new system for developing labor force estimates was introduced, combining the handbook method with the concepts, definitions, and estimation controls from

---

[230] Text source: BLS, "Chapter 4, Measure of Unemployment in States and Areas," *BLS Handbook of Methods*.

the Current Population Survey. A major advantage of the CPS is that it applies uniformly across states, whereas UI data are affected by individual states' UI laws.[231]

Over the years, major improvements have been made to the UI database, an integral input to state and area estimation. The UI database project, conducted in 1976-78, standardized all UI claims data used in state and area labor force estimates, so that these data would be more consistent with the conceptual underpinnings of unemployment used in the CPS, as well as more comparable from state to state. The result of this project was the regular, automated development of data on UI claimants certifying to unemployment for the week including the 12th day of the month (the CPS reference week). These data are based on the claimants' state/county/city of residence and exclude those who had earnings from employment in the certification week.

In January 2005, a major program Redesign was implemented. Work on the Redesign began in Fiscal Year 2001, with a budget initiative to enhance the quality and quantity of LAUS program statistics. Major LAUS Redesign components include improvements to the current method of State and large area estimation including 'real time benchmarking', extending our 'best estimating' techniques to more areas, improving the methods used in all other areas through better techniques and input data, and updating the geography with 2000 Census-based areas.

For all states and the District of Columbia, the Los Angeles-Long Beach-Glendale, metropolitan division, New York City, and the respective balances of state, models based on a "signal-plus-noise" approach are used to develop employment and unemployment estimates. The model of the signal is a time series model of the true labor force which consists of three components: A variable coefficient regression, a flexible trend, and a flexible seasonal component. The regression techniques are based on historical and current relationships found within each state's economy as reflected in the different sources of data that are available for each state—the CPS, the Current Employment Statistics (CES) survey, and the UI system. The noise component of the models explicitly accounts for autocorrelation in the CPS sampling error and changes in the average magnitude of the error. In addition, the models can identify and remove the effects of outliers in the historical CPS series. While all the state models have important components in common, they differ somewhat from one another to better reflect individual state labor force characteristics.

The Redesign bivariate models incorporate a major change in the approach to benchmarking and the benchmarking process. Rather than continue with an annual average State benchmark applied retrospectively that reintroduces sampling error to the historical monthly estimates, the Redesign approach uses a reliable real-time monthly national benchmark for controlling current State model estimates of employment and unemployment. In this process, benchmarking is part of the monthly State model estimation process.

Under real-time benchmarking, a tiered approach to estimation is used. Model-based estimates are developed for the nine Census divisions that geographically exhaust the nation using

---

[231] Text source: *Ibid*.

univariate signal-plus-noise models. The Division models are similar to the State models, but do not use unemployment insurance claims or nonfarm payroll employment as variables. The division estimates are benchmarked to the national levels of employment and unemployment on a monthly basis. The benchmarked division model estimate is then used as the benchmark for the States within the division. The distribution of the monthly benchmark adjustment to the States is based on each State's monthly model estimate. In this manner, the monthly State employment and unemployment estimates will add to the national levels.

Monthly labor force estimates for two large substate areas—New York City and the Los Angeles-Long Beach-Glendale, CA metropolitan division and the respective balances of New York and California—are developed using bivariate signal-plus-noise models. We have also developed signal-plus-noise models for five additional substate areas and their State balances. The areas are: the Chicago-Naperville-Joliet, IL metropolitan division; the Cleveland-Elyria-Mentor, OH metropolitan area; the Detroit-Warren-Livonia, MI metropolitan area; the Miami-Miami Beach-Kendall, FL metropolitan division; and the Seattle-Bellevue-Everett, WA metropolitan division.

Area models, like the Division models, are univariate in design in that only the historical relationship of the inputs is considered—UI claims and CES inputs are not used each month in the estimation process. Area and balance of State models are controlled directly to the State totals, which are themselves controlled to the national CPS via the Census division models.

The LAUS Handbook method is an effort to estimate unemployment for an area, using available information without the expense of expanding a labor force survey like the CPS. The Handbook presents a series of estimating "building blocks," in which categories of unemployed workers are classified by their previous status. Two broad categories of unemployed persons are: (1) Those who were last employed in industries covered by state UI laws, and (2) those who either entered the labor force for the first time or reentered after a period of separation. Handbook inputs were updated using the 2000 Census and other improvements to Handbook estimation were implemented with January 2005 estimates.

Employment. The total employment estimate is based on data from several sources. The primary source for most metropolitan areas (MAs) is the Federal-state CES survey. The CES is designed to produce estimates of the total number of employees on payrolls in nonfarm industries for the particular area. In small labor market areas and the remainder of the MAs, the establishment employment data come from the Quarterly Report of Quarterly Census of Employment and Wages (ES-202 Report).

These "place-of-work" employment estimates must be adjusted to a place-of-residence basis, as in the CPS. Estimated adjustment factors have been developed using employment relationships which existed at the time of the most recent decennial census. The adjustment approach implemented in January 2005 is more dynamic than the previous one and incorporates commuting to nearby labor market areas. These factors are applied to the place-of-work employment estimates for the current period to obtain adjusted employment estimates, to which are added synthetically developed estimates for employment not

represented in the establishment series—agricultural workers, nonfarm self-employed and unpaid family workers, and private household workers.

Unemployment. The estimate of unemployment is an aggregate of the estimates for each of the two building-block categories. The "covered" category further consists of two unemployed worker groups: (1) Those who are currently receiving UI benefits and (2) those who have exhausted their benefits. Only the number of those currently collecting benefits is obtained directly from an actual count of UI claimants for the reference week. The estimate of persons who have exhausted their benefits is based upon the number actually exhausting benefits in previous periods "survived" using a conditional probability approach based on CPS data.

The second category, "new entrants and reentrants into the labor force," cannot be estimated directly from UI statistics, because unemployment for these persons is not immediately preceded by the period of employment required to receive UI benefits. In addition, there is no uniform source of new entrants and reentrants data for States available at the LMA level; the only existing source available is from the CPS at the State level. Separate estimates for new entrants and for reentrants are derived from econometric models based on current and historical state entrants data from the CPS. These model estimates are then allocated to all Labor Market Areas (LMAs) based on the age population distribution of each LMA.

Substate adjustment for consistency and additivity. Each month, Handbook estimates are prepared for labor market areas that exhaust the entire state area. To obtain a labor force estimate for a given area, a "Handbook share" is computed for that area which is defined as the ratio of that area's Handbook estimates of employment and unemployment to the sum of the Handbook estimates of employment and unemployment for all LMAs in the state. These ratios are then multiplied by the current, statewide estimate for employment and unemployment to produce the final adjusted LMA estimates.

Once each year, labor force estimates are revised to reflect updated input data and new Census Bureau population controls. As part of this procedure, all of the state and substate models are reviewed, revised as necessary, and then reestimated; this reestimation is called "smoothing."

When new population controls are available from the Bureau of the Census, typically in January, CPS estimates for all states, the District of Columbia, New York City; the Chicago-Naperville-Joliet, IL metropolitan division; Cleveland-Elyria-Mentor, OH metropolitan area; Detroit-Warren-Livonia, MI metropolitan area; Los Angeles-Long Beach-Glendale, CA metropolitan division; Miami-Miami Beach-Kendall, FL metropolitan division; New Orleans-Metairie-Kenner, LA metropolitan area; and, the Seattle-Bellevue-Everett, WA metropolitan division are adjusted to these controls. Additionally, the time series regression models for the states and model-based areas are reestimated based on the latest input data.

Other substate estimates for previous years are also revised on an annual basis. The updates incorporate any changes in the inputs, such as revisions to establishment-based employment

estimates or claims data and updated historical relationships. The revised estimates are then readjusted to the latest statewide estimates of employment and unemployment.[232]

In 2015, the LAUS estimation methodology will revised. The 2015 LAUS Redesign includes improved time-series models for the census divisions, states, select substate areas, and the balances of those states; an improved real-time benchmarking procedure to the national Current Population Survey (CPS) estimates; an improved smoothed seasonal adjustment procedure; and improved treatment of outliers.

Non-modeled area estimation improvements include: updated Dynamic Residency Ratios (DRR); more accurate estimates for all-other employment; more accurate estimation of agricultural employment; and improved estimation of non-covered agricultural unemployment. Handbook estimation is now done at the county level instead of at the Labor Market Area (LMA) level, which better reflects local conditions. The Redesign also introduces estimation inputs from the American Community Survey (ACS) to replace inputs that were previously obtained from the decennial census long-form survey.[233]

Application of ACS Data in Methodology: Beginning in Fiscal Year 2015 LAUS plans to introduce methodological improvements with ACS as an input to its estimation techniques.

The LAUS program had been reliant on the long form data as the basis for developing substate estimates for self-employed, unpaid family workers, private household workers, and agricultural workers throughout the decade. These data elements represent employment that is either not covered by unemployment insurance compensation programs or not included in the payroll survey data CES, thus the Census long form had been the sole source for this type of information at the local level.

ACS data are issued on an annual basis and they do not represent a single point in time as did the decennial Census, which represented April 1 in the year that the Census was conducted. Instead ACS data are estimates that span 1 year, 3 years, or 5 years depending on the population level of each area. To ensure coverage of all LAUS geography, which includes areas with 25,000 population or more and all cities and towns in New England regardless of population, the 5-year estimates must be used. In addition to covering all LAUS geography, the 5-year estimates use the largest sample size and are the most statistically reliable of the ACS estimates. However, since they represent a 5-year span they cannot be directly used to develop current monthly estimates.

The most current source of the needed data inputs is the CPS which does not have the geographic detail of the ACS. The proposed methodology will utilize the strengths of the CPS and the ACS to develop monthly estimates of self-employed, unpaid family, and private

---

[232] Text source: "Local Area Unemployment Statistics: Estimation Methodology," BLS website.
[233] Text source: "Questions and Answers on the Local Area Unemployment Statistics (LAUS) Program 2015 Redesign," BLS website.

household workers (collectively known as ``all-other'' employment) and agricultural workers at the needed level of geography.

Enhanced procedures for developing other substate areas. Utilizing ACS data to replace the Census long form data facilitated the enhancement of some of the substate methodologies making up the building-block approach used to develop independent substate estimates. Revisions are proposed for:

- the methodology of adjusting place-of-work data to a place-of-residence basis,
- the estimation of what is known as ``all-other'' employment,
- the estimation of agricultural employment, and
- the estimation of agricultural unemployment not covered by unemployment insurance.

In addition, using the ACS, substate estimates will be developed at the county level rather than the labor market area level. A brief discussion of the new methodologies is below.

--Place-of-Work Residency Adjustment. The LAUS program uses the same labor force concepts as the CPS. Thus employment inputs from the CES and Quarterly Census of Employment and Wages (QCEW) programs, which are based on place-of-work, must be adjusted to reflect the worker's place of residence per the CPS. To accomplish this, Dynamic Residency Ratios (DRRs) are applied to CES and QCEW employment inputs for LAUS estimation. This methodology assumes that resident employment in an area is a function of the relationship between employed residents and jobs not only in that area, but in other areas within commuting distance. The procedure is more dynamic than the use of a single residency ratio insofar as job count changes in commuting areas can affect resident employment.

In the past, journey-to-work data from the decennial Census were incorporated into the DRRs. Journey-to-work data were not available from 2010 Census due to the discontinuation of the long form. For the LAUS 2015 redesign, DRRs will be computed using ACS journey-to-work data in the same manner that they are computed now with one major modification. Currently, an area must be the destination workplace of at least 100 resident commuters (50 in New England) to be considered a potential commuter area. BLS proposes replacing these criteria with a percentage threshold. In the new set of DRRs, commuter areas will be limited to those areas that are the work destination of no less than 10 percent of resident commuters. This will eliminate marginal commuter areas included in the previous methodology to account for potential future growth.

The previous threshold for DRR commutation areas reflected the ten-year span between Census journey-to-work data releases. The inclusion of a relatively high number of areas would accommodate any potential changes to commuting patterns over the ensuing decade. The new data source for DRRs, ACS journey-to-work data, is intended to be updated every five years. The increased frequency in the availability of commutation data will make the list of commutation areas more responsive to changing commuting patterns, reducing the need to include minor destinations which may grow in importance over time.

--Estimation of All-Other Employment. The current method uses Census 2000 data as the starting point for the self-employed, unpaid family, and private household workers (known as ``All-Other Employment'') and moves it forward through time by applying the relationship of all-other employment to the nonfarm wage and salary employment estimate at the time of the Census.

The new method uses the relationship of each area's share of ACS all-other employment to the State's total ACS all-other employment. This relationship is then used to allocate a monthly 5-year weighted average of each State's CPS estimate of all-other employment. A weighted average of the CPS estimate is used because, depending on the State's CPS sample size, the monthly estimate for this element may be volatile due to sampling error. This monthly 5-year weighted average consists of the current month's estimate averaged with the same month's estimate going back years, with more weight placed on the more current estimates. This technique borrows strength from prior estimates while preserving seasonal trends.

--Estimation of Agricultural Employment. The current method uses the Census 2000 data as the base and moves the estimate forward using a monthly change factor based on a State's membership in a multi-State agricultural region.

The new method for estimation of agricultural employment uses a similar approach as the all-other employment method. A monthly 5-year weighted average of each State's CPS estimate of agricultural employment is developed and allocated to substate areas using each areas' share of the State's total ACS agricultural employment. This method is State-specific and eliminates the need for an agricultural regional factor.

--Estimation of Non-covered Agricultural Unemployment. This is an optional procedure that is currently utilized by 19 States. The current procedure uses an indirect approach for the estimation of agricultural unemployment not covered by unemployment insurance. It assumes that there is unemployment associated with employment and that the unemployment rate in non-covered agriculture is related to the rate of unemployment in covered sectors of the economy. To estimate non-covered agricultural employment, the annual average of covered agricultural employment from the QCEW program is subtracted from the covered agricultural employment estimate that is developed each month (as described in the prior section). Seasonal factors derived from CPS agricultural data from 1977-1982 are applied to account for seasonality.

The new method replaces the annual average QCEW covered agricultural employment with a 1-year lagged monthly estimate of agricultural employment from the QCEW and eliminates the potentially outdated seasonal factors. Use of a 1-year lagged monthly estimate will incorporate seasonal trends into the estimate, simplifying the calculation and making it more responsive to long-term changes in seasonal patterns.

--New Procedure for Estimating Employment and Unemployment at the County Level. Labor market areas (LMAs) are independently estimated using a building block approach that incorporates the new methods discussed above and other methods still currently in use. The employment component is comprised of non-agricultural wage and salary employment, all-

other employment and agricultural employment. While the unemployment component is derived by summing the estimates of non-covered agricultural unemployment (if applicable), total unemployment insurance (UI) continued claims without earnings, unemployed exhaustees and unemployed entrants into the labor force.

The current procedure consists of first developing these independent substate estimates at the LMA level and then disaggregating them into counties and cities. With the exception of non-agricultural wage and salary employment, all inputs for estimating the components of employment and unemployment are readily available at the county level (Minor Civil Division (MCD) level in New England, MCDs being cities and towns). Aggregating these more geographically detailed data into LMAs is an unnecessary step that results in the distortion of these data when they are reallocated backed to the county level or MCD level, particularly for some of the unemployment components.

The new method proposes to first develop the independent substate estimates at county level and then sum them to their appropriate LMA. This approach will result in more accurate estimates and will allow better operational flexibility for future updates to the geographic definitions of LMAs as counties (MCDs in New England) are the basic component of LMA geographic definitions issued by the Office of Management (OMB), as well as for small labor market areas as defined by the BLS.

The current method estimates the labor force in LMAs, which are defined to comprise one or more counties (MCDs in New England). Employment and unemployment inputs are entered at the LMA level. In a multi-county LMA, county unemployment estimates are disaggregated from the LMA using the share of UI claims for the experienced unemployed, the share of the 16-19 population for unemployed new entrants, and the share of the 20+ population for unemployed re-entrants.

The new procedures discussed above for estimating the employment components of all-other employment and agricultural employment produce these estimates at the county level. The non-agricultural wage and salary employment component, which is provided by the CES and the QCEW programs, is generally available at the LMA level and must be allocated into the counties that comprise the LMA. This will be accomplished by using ACS non-agricultural wage and salary employment ratios derived from the most recent ACS five-year dataset to distribute the CES/QCEW LMA data to its component counties (and MCDs in New England). This step is not needed for single county LMAs.

All of the necessary inputs for estimating unemployment are already available at the county (and MCD) level. The new procedure results in more accurate county estimates by estimating the level of persons who remain unemployed after exhausting their eligibility for unemployment insurance benefits (known as exhaustees) at the county level and by avoiding the disaggregation of entrants from interstate LMAs.[234]

---

[234] Bureau of Labor Statistics, "Comment Request on the Local Area Unemployment Statistics Program," *Federal Register* Volume 79, Number 175, September 10, 2014, pp. 53787-53791.

(Indirectly, ACS data also contribute to the LAUS through their use to estimate international migration for the Census Bureau's Population Estimates, which provide population controls for the LAUS estimates.[235])

ACS Questions Utilized (as suggested by the methodology): P29-31, 41-46

Major Uses: The estimates are used in economic analysis by public agencies and private industry and for State and area allocations and eligibility determinations according to legal and administrative requirements. Implementation of policy and legislative prerogatives could not be accomplished as now written without collection of the data.

BLS has identified 26 federal programs that utilize LAUS data, collectively distributing $90.3 billion in FY2013.[236] The Workforce Investment Act of 1998 (PL 97-300) requires these data:

- to allocate funds under Title II (Adult Education and Literacy programs);
- to implement Title I, which establishes state workforce investment boards, state plans, and local workforce investment areas. The local workforce investment areas are to support a one-stop delivery system that provides core employment-related services and other employment and training services;
- to implement Title I, which provides that one of the standards for eligibility for services under the Act is that the ratio of participants being trained on the job in the public sector to participants in the private sector shall not exceed the ratio of civilian government employment to private sector employment in the area.[237]

Other requirements include the use of these data:

- to allocate funds to States to establish and maintain public employment offices as required by the Wagner-Peyser Act (29 USC 49 et seq.) and to require that public labor exchange activities be part of the one-stop system and integrated into the state workforce development plan;
- to indicate economic distress in areas to identify political jurisdictions eligible for Federal assistance under the Urban Development Action Grant Program (Section 119 of Title I of the Housing and Community Development Act (PL 93-383)). The standards include local unemployment as an indicator;
- by the Employment and Training Administration to identify Labor Surplus Areas designated under Executive Orders 10582 and 12073 in the implementation of amendments to the Small Business Act (PL 96-302), which provides that Federal procurement contracts be set aside for businesses located in labor surplus areas;

---

[235] For differences between ACS data on unemployment and LAUS data, see "ACS Questions and Answers," Local Area Unemployment Statistics, Bureau of Labor Statistics website.

[236] BLS, "Administrative Uses of Local Area Unemployment Statistics," November 21, 2013.

[237] It is expected that the recently passed Workforce Innovation and Opportunity Act of 2014, which replaces WIA, relies on LAUS data in a similar fashion.

- to allocate Department of Commerce funds for certain programs to areas of high unemployment as required by the Public Works and Economic Development Act of 1965 (42 USC 3121).[238]

---

[238] Text source: Bureau of Labor Statistics, "Supporting Statement- Part A, Local Area Unemployment Statistics," December 7, 2011.

71. *Employment Projections*

Purpose of the Measure: The Employment Projections program (EP) develops information about the labor market for the United States as a whole for 10 years in the future. It is the only comprehensive source of regularly produced occupational employment and wage rate information for the U.S. economy. Projections of employment and wage estimates by nonfarm industry are made for the full Standard Occupational Classification system, which includes over 800 detailed occupations. In addition to projecting employment for each detailed occupation, the program depicts the education, related work experience, and on-the-job training typically needed for occupations.

Responsible Unit: Office of Occupational Statistics and Employment Projections, Bureau of Labor Statistics (BLS), Department of Labor

Authorization: 29 U.S.C. § 1-8; 29 U.S.C. § 49, et. seq. 29 USC 49l-2 charges the Secretary of Labor with maintaining a nationwide workforce and labor market information data system that "includes statistical data that . . . enumerate, estimate, and project employment opportunities and conditions at national, State, and local levels in a timely manner, including statistics on industrial distribution of occupations, as well as current and projected employment opportunities, wages, benefits (where data is available), and skill trends by occupation and industry . . . ."

Frequency of Data Publication: Every other year.

Timing of Data Release: The most recent projections are for 2012-2022 and were released in December 2013.

Modes to Access Data:

- Employment Projections, BLS website, including databases and tables (html, xls)
- Occupational Outlook Handbook, BLS website

Methodology: Projecting employment in industry and occupational detail requires projections of the total economy and its sectors. BLS develops its projections in a series of six steps that examine: 1) the size and demographic composition of the labor force; 2) aggregate economic growth; 3) commodity final demand; 4) input–output; 5) industry output and employment; 6) occupational employment and openings. Each step, based on separate procedures and models and on related assumptions, goes through several iterations to ensure internal consistency as assumptions and results are reviewed and revised. Together, the six components provide the analytical framework needed to develop detailed employment projections. BLS analysts solve each component sequentially.[239]

Application of ACS Data in Methodology: EP includes information about education and training requirements for hundreds of occupations. BLS assigns occupations to a designation within

---

[239] BLS, "Chapter 13, Employment Projections, *BLS Handbook of Methods*.

three categories: typical entry-level education, work experience in a related occupation, and typical on-the-job training.

BLS economists assign occupations to categories based on analyses of qualitative and quantitative information. The ACS is a source of quantitative data on educational attainment; other data on educational and training requirements come from the Occupational Information Network and the National Center for Education Statistics. Education attainment data from the ACS are a useful analytical tool that presents the percent distribution of workers employed in an occupation, broken down by their highest level of education attained. The educational attainment data published by BLS are based on the Census Bureau's microdata files.[240]

Occupational statistics in the ACS are coded under the Census Bureau's 2010 Census occupation classification system. Both the BLS category system and the BLS Occupational Employment Statistics (OES) data—the source data for the National Employment Matrix, are coded based on the federal government's Standard Occupational Classification (SOC) system. Although the Census Bureau's system also is based on the SOC, it does not provide the same level of detail as the matrix. As a result, some detailed SOC occupations have the same educational attainment data, because they are combined in the Census Bureau's occupational classification system.

The EP educational attainment data use 2 years of ACS microdata from 2010—11. The data present the highest level of education attained by those 25 years and older in the current workforce. Two years of data are used to improve the reliability of the estimates.

The educational attainment distributions allow data users to better discern whether there are multiple education and training possibilities. For example, because 86.5 percent of speech-language pathologists have at least a master's degree in the 2010—11 ACS data, this is a clear indication that getting a master's degree is the typical way to become a speech-language pathologist. However, educational attainment data for other occupations may be more varied. For example, the 2010—11 ACS data show that 19.4 percent of chefs and head cooks have "less than a high school diploma"; 28.7 percent have a "high school diploma or equivalent"; 21.7 percent have "some college, no degree"; 16.2 percent have an "associate's degree"; and 12.4 percent have a "bachelor's degree." The educational attainment distribution for chefs and head cooks suggests that there is more than one way to enter this occupation. Data show the highest level of education the survey respondent has attained—not necessarily the level of education required for the occupation.

Like any sample survey, the ACS is a household sample survey and is subject to response and coding error, as well as sampling error.

The educational attainment data do not always closely reflect the education assignment in the category system. One major difference is that the category system reflects typical entry-level educational requirements, whereas ACS data report the level of education attained by workers already in the occupation. This can lead to cases in which the attainment data reflect a higher

---

[240] Text source: "Employment Projections- Measures of Education and Training," BLS website.

level of educational attainment than the category selection. In other cases, the category assignment reflects a higher level of education than the attainment data show.

In addition to respondent or coding errors, there are several reasons that the educational attainment data may not match the category assignment. Examples are: underemployment, individual choice, and the trend of "upskilling," in which the educational attainment of workers continues to rise over time. Also, because of changing entry requirements, individuals entering an occupation may need a higher level of formal education than for those persons who are already working in it.

(Labor force projections are based on expectations of the future size and composition of the population, as well as on the trends in labor force participation rates of different age, gender, race, and ethnic groups. The Population Projections of the U.S. Census Bureau is used by the BLS to develop these labor force projections. [241] The ACS contributes to EP indirectly through the Census Bureau's Population Projections, providing the data on international migration.)

ACS Questions Utilized (as suggested by the methodology): P11, P45-46

Major Uses: EP serves a variety of users who need information about expected patterns of economic growth and the effects these patterns could have on employment. Data users include individuals seeking career guidance and organizations and individuals offering career guidance resources. In addition, policymakers, community planners, and educational authorities, who need information for long-term policy planning purposes, make use of BLS employment projections, as do states in preparing state and local area projections.

Projections are released online biennially in December or January and in both the Monthly Labor Review and the Occupational Outlook Quarterly. The Review typically includes an overview article and an article on each of the major components of the projections: the labor force, the aggregate economy, industry output and employment, and occupational employment and job openings. The Quarterly publishes articles related to career preparation, such as occupational profiles, jobseeking information, and understanding wage and benefits data. Part of each projection study is the release of the Occupational Outlook Handbook (OOH), which contains extensive information about hundreds of occupations. In addition to presenting outlook data for each occupation, this publication includes information on the nature of the work, education and training requirements, working conditions, and wages. The OOH is used as a primary source of information for people choosing a career and is available in many career centers of high schools and colleges, as well as in libraries.

---

[241] Text source: "Employment Projections- Projections Methodology," BLS website.

72. *Standard Occupational Classification System*

Nature and Purpose of the Classification System: The Standard Occupational Classification (SOC) system classifies all occupations in the economy, including private, public, and military occupations, in order to provide a means to compare occupational data produced for statistical purposes across agencies. It is designed to reflect the current occupational work structure in the U.S. and to cover all occupations in which work is performed for pay or profit.[242]

The 2010 SOC groups occupations using a tiered system with four levels, shown below with the number of categories at each level, with 840 detailed occupations organized into broad occupations, minor occupation groups, and major occupation groups.

The SOC coding system uses a 6-digit code, with the first two digits indicating the major occupation group. The third digit indicates the minor occupation group, the fourth and fifth digits indicate the broad occupation, and finally the sixth digit indicates the detailed occupation.

Responsible Unit: Interagency SOC Policy Committee. The Bureau of Labor Statistics (BLS) chairs the SOC Policy Committee and provides staff resources to carry out the work of the committee.

Authorization: Budget and Accounting Procedures Act of 1950 (31 U.S.C. 1104(d)) and the Paperwork Reduction Act of 1995 (44 U.S.C. 3504(e)). 31 U.S.C. 1104(d) notes that the "President shall develop programs and prescribe regulations to improve the compilation, analysis, publication, and dissemination of statistical information by executive agencies."

Frequency of Data Publication: Historically, each decennial census year, most recently in 2010. Breaking with tradition, the next SOC will be released in 2018.

Timing of Data Release: The 2018 revision date is intended to minimize disruption to data producers and users by promoting simultaneous adoption of revised occupational and industry classification systems for those data series that use both. Given the multiple interdependent programs that rely on the SOC, the interagency SOC Policy Committee has determined that the timing of the SOC revision should follow the North American Industry Classification System (NAICS) revisions, which occur for years ending in 2 and 7. Thus, the next such year for a SOC revision is 2018, following the 2017 NAICS revision. This date has the additional benefit of coinciding with the beginning year of the American Community Survey's next 5-year set of surveys. The Office of Management and Budget intends to consider revisions of the SOC every 10 years from 2018.

For the 2018 revision, the SOC Policy Committee began planning in early 2012, and expects to solicit public input through the first Federal Register notice in spring 2014. This process is expected to lead to publication of the proposed revision for public comment through the

---

[242] Text source: Office of Management and Budget, "Standard Occupational Classification (SOC)—Revision for 2018; Notice," *Federal Register*, Vol. 79, No. 99, May 22, 2014, p. 29620.

second Federal Register notice in spring 2015 and publication of the final 2018 SOC structure in spring 2017.[243]

Modes to Access Classification System: Standard Occupational Classification Materials, BLS

Methodology: The SOC is developed on the basis of nine classification principles. The SOC Policy Committee is considering what, if any, revisions to make in the 2010 SOC Classification Principles for the 2018 version.

Principle 9 states that "The U.S. Bureau of Labor Statistics and the U.S. Census Bureau are charged with collecting and reporting data on total U.S. employment across the full spectrum of SOC major groups. Thus, for a detailed occupation to be included in the SOC, either the Bureau of Labor Statistics or the Census Bureau must be able to collect and report data on that occupation."

Classification Principle 9 deals with "collectability" – that is, whether data can actually be collected on the occupation. For a detailed occupation to be included in the SOC, either BLS or the Census Bureau must be able to collect and report data on the occupation. BLS and the Census Bureau are responsible for producing data across the entire range of occupations in the U.S. labor market, and conduct comprehensive household and business surveys that collect occupational data.

Collectability is partly a function of the size of the occupation – it must be large enough to be detected in sample household or business surveys. In evaluating collectability, however, the SOC Policy Committee does not use a specific employment size cut-off. Small occupations that are concentrated in certain industries or geographic areas may be collectable, while occupations of similar or larger employment that are spread throughout the economy may not be collectable.

Size is not the only consideration in collectability. Collectablity is also related to the type of data collection used, specifically the comprehensive household and business surveys in BLS and the Census Bureau.

In general, household surveys can collect a more limited set of information for identifying the occupation of individuals than is possible in business surveys. For example, the Current Population Survey (CPS) and the American Community Survey (ACS) collect the individual's job title and a very brief description of the person's most important activities or duties. In most household surveys, coders are not able to recontact the respondent for clarification. Since less information is available for assigning classification codes, household surveys generally provide less occupational detail than business surveys. Thus, occupational categories with fine distinctions from one another may not be collectable in household surveys.

Household surveys such as the CPS and ACS are the main sources of demographic information on workers by occupation, especially educational attainment, gender, age, and race/ethnicity.

---

[243] Text source: BLS, "Revising the Standard Occupational Classification," March 2014.

In addition, the CPS and ACS are the main sources of occupational data for parts of the workforce not covered by business surveys, namely the self-employed, unpaid family workers, and workers in private households and most agriculture industries. Occupations that are primarily comprised of these types of workers or mainly found in private households or agriculture therefore must be collectable on household surveys.

Business surveys collect data on occupations directly from employers. These surveys rely on the employer for information about the workers' duties, and often coders may recontact the employer to obtain clarifications. Thus, it is often possible to obtain detailed information about the work performed, allowing occupational coding to more detail than possible in most household surveys. Business surveys provide data on employment, wages, and benefits by occupation, and sometimes on other characteristics of the job or worker. In general, however, business surveys do not provide demographic information, nor do they include the self-employed, unpaid family workers, workers in private households, or workers in most agriculture industries.

The SOC Policy Committee indicates it seeks to collect the following information for any occupation considered for inclusion in the 2018 SOC:

- Nature of the work performed
- Differentiation of work from other detailed SOC occupations
- Job titles
- Numbers of jobs or workers in the occupation
- Types of employers
- Education and training
- Licensing
- Tools and technologies
- Professional or trade associations and unions

Application of ACS Data in Methodology: The ACS is essential to the construction of a detailed useful SOC in that it is the primary vehicle for "collectability," as follows:

- the classification process must rely first and foremost on household surveys, that is, the CPS and the ACS;
- the ACS is far larger than the CPS and so is able to:
  - generate much more information about occupational titles, including emerging occupations;
  - the number of workers in the occupation; and
  - the distribution of highest level of educational attainment within the occupation;
- only the ACS asks a question about job duties (P46); and
- the ACS can provide substantial demographic detail for an occupation.

ACS Questions Utilized (as suggested by the methodology):

- In the construction of SOC detailed occupations: P45-46

- To demonstrate the collectability of information about an occupation: P3-48

Major Uses of the Classification: The Standard Occupational Classification (SOC) is one of several standard classification systems established by OMB to ensure coordination of Federal statistical activities. All Federal agencies that publish occupational data for statistical purposes are required to use the SOC to increase data comparability (and thus, data utility) across Federal programs.

Federal occupational data programs that rely on the SOC include:

- The U.S. Census Bureau publishes occupational data annually, collected through the American Community Survey (ACS).
- The Department of Defense publishes data that cross-reference military occupational codes of the Army, Navy, Air Force, Marine Corps, and Coast Guard with civilian equivalent occupations.
- The National Center for Education Statistics (NCES) publishes data collected through the School and Staffing Survey (SASS) on the employment of elementary and secondary teachers, principals, and other school staff, as well as detailed information on their education, training, and background characteristics. NCES publishes detailed data on postsecondary instructors and professors collected through the Integrated Postsecondary Education Data System (IPEDS). In addition, NCES conducts various longitudinal studies that follow high school and college students into their working years and uses the SOC to classify their occupations.
- Biennially, the Bureau of Labor Statistics' Employment Projections (EP) Program publishes the Occupational Outlook Handbook and Career Guide to Industries. In addition, EP publishes the Occupational Outlook Quarterly.
- The Current Population Survey (CPS), a joint program of the Census Bureau and the Bureau of Labor Statistics, uses the Census occupational classification system, which is derived from the Standard Occupational Classification.
- The Bureau of Labor Statistics' National Compensation Survey (NCS) program provides comprehensive measures of occupational wages; employment cost trends; and benefit incidence and detailed plan provisions. Detailed occupational earnings are available for selected metropolitan and nonmetropolitan areas, nine Census divisions, and on a National basis. Employment cost trends and information on the incidence and detailed provision of employee benefit plans are published for major occupational groups.
- The Bureau of Labor Statistics' Occupational Employment Statistics (OES) program produces cross-industry occupational employment and wage estimates for the Nation, all States, the District of Columbia, Guam, Puerto Rico, the U.S. Virgin Islands, metropolitan areas, metropolitan divisions, and nonmetropolitan areas. OES also publishes national industry-specific occupational employment and wage estimates for sectors and 3-, 4-, and selected 5-digit North American Industry Classification System (NAICS) industries.

- The Employment and Training Administration's (ETA) Occupational Information Network (O*NET) system is a comprehensive database of occupational competency profiles. ETA sponsors the development, updating, and dissemination of O*NET information through a grant with the North Carolina Employment Security Commission. The O*NET system is based on the Standard Occupational Classification (SOC) system and also provides information on additional detailed occupations within a SOC category in selected instances. The O*NET Content Model of occupational descriptors is the foundation for a series of survey questionnaires that go out to incumbent workers in various occupations which form the basis for the O*NET occupational competency profiles. The O*NET system is the successor to the Dictionary of Occupational Titles, which was last published by the Department of Labor in 1991.

- The Equal Employment Opportunity Commission (EEOC) uses SOC occupational classifications, and equivalent Census occupational classifications, to create broader categories as part of the Commission's data survey and enforcement programs. Under the survey program, employer workforce information is collected periodically from private-sector firms on the Employer Information Report (EEO-1), and public sector employers on the State and Local Government Report (EEO-4).

- The National Science Foundation (NSF) Division of Science and Resources Statistics (SRS) Web site provides access to the Scientists and Engineers Statistical Data System (SESTAT), a comprehensive and integrated system of information about the employment, educational and demographic characteristics of scientists and engineers in the United States.[244]

Information about occupations—employment levels, trends, pay and benefits, demographic characteristics, skills required, and many other items—is widely used by individuals, businesses, researchers, educators, and public policy-makers. In addition to Federal agencies, users of the occupational data include industrial and labor relations practitioners, students considering career training, job seekers, career and employment counselors, educational institutions, and employers wishing to set salary scales or to locate to a new facility.[245]

---

[244] BLS, "2010 SOC User Guide FAQs and Acknowledgements," February 2010.
[245] Text source: BLS, "Revising the Standard Occupational Classification," March 2014.

B. Office of Disability Employment Policy and Office of Federal Contracting Compliance Programs

73. *Disability Employment Tabulation*

Purpose of the Measure: The Disability Employment Tabulation shows the disability status and diversity of the labor force and population 16 and over. The Tabulation has similar content to that found in the Equal Employment Tabulation, and provides in-depth labor force characteristics of people with a disability, with more details on occupation, education, and earnings. Tables may include employment status, detailed occupation, occupation groups, citizenship, educational attainment, age, race, sex and earnings.[246]

Responsible Unit: Census Bureau for the Department of Labor's Office of Disability Employment Policy and Office of Federal Contracting Compliance Programs

Authorization: Disability Employment Policy Office (29 USC 557b) and Office of Federal Contracting Compliance Programs (Section 503 of the Rehabilitation Act of 1973, as amended).

Frequency of Data Publication: Only release thus far was on March 14, 2013. The sponsors have yet to determine if there will be another tabulation and when it would be released.

Timing of Data Release: A special tabulation of three-year ACS data for 2008-2010 was released in 2013.

Modes to Access Data: Disability Employment Tabulation, Census Bureau (html, csv, pdf, ftp)

Methodology/Application of ACS Data in Methodology: The Disability Employment Tabulation 2008-2010, a 49-table tabulation which shows the disability status and diversity of the labor force and population 16 and over, is based on ACS three-year data 2008-2010. Similar to the Equal Employment Tabulation, the Disability Employment Tabulation provides "in-depth labor force characteristics of people with a disability, with more details on occupation, education, and earnings." The tabulations include estimates for the "U.S. total, states, the District of Columbia and Puerto Rico, metro/micro areas, EEO County Sets, and PUMS."[247]

ACS Questions Utilized (as suggested by the methodology): P3-8, P11, P17-19, P29, P36-37, P45-48

Major Uses: The primary purpose of the special tabulation is to increase the capacity of the DOL sponsoring agencies to achieve their respective missions.

The Disability Employment Tabulation allows examination of labor force diversity by disability status. It is the source for detailed occupational statistics by disability status in the workforce for local areas, and is the first product from the Census Bureau providing in-depth labor force characteristics of individuals with a disability, with more detail on occupation, education, and earnings. The tabulation provides reliable and accurate data to further education, research, and

---

[246] Text source: "Disability Employment Tabulation: Main," Census Bureau website.

[247] Text source: *Ibid*.

policy initiatives that improve employment opportunities and outcomes for individuals with disabilities.[248]

---

[248] Census Bureau, "American FactFinder (AFF) and The DOL Disability Employment Tabulation 2008-2010: Tutorial," presentation, June 2013, slide 4.

C. Employment and Training Administration

74. Data on Economically Disadvantaged Youth and Adults

Nature and Purpose of the Measure: Estimates of the number of economically disadvantaged youth and adults are prepared for use in the within-state allocation formulas for Workforce Investment Act (WIA) Youth and Adult programs starting in Program Year (PY) 2013. Data are available for the nation, state, county, place, subdivision, and areas of indigenous peoples (American Indian, Alaskan Native, Hawaiian Homeland).

Responsible Unit: Employment and Training Administration (ETA), Department of Labor

Authorization: The Workforce Investment Act (WIA) requires states to use data on economically disadvantaged Youth and Adults in their within-state allocation formulas. (WIA was recently amended and renamed the Workforce Innovation and Opportunity Act)

Frequency of Data Publication: ETA anticipates updating the data every five years.

Timing of Data Release: Data for the 2006-2010 period were released in March 2013.

Modes of Data Access: "Updated Data for Persons Defined as Economically Disadvantaged Youth and Adults," ETA website (xls)

Methodology/Application of ACS in the Methodology: An economically disadvantaged Youth is defined as "an individual who is age 16 through 21 who received an income, or is a member of a family that received a total family income that, in relation to family size, does not exceed the higher of the poverty line, or 70 percent of the Lower Living Standard Income Level (LLSIL)." The definition of a disadvantaged Adult is similar; but, per WIA Section 132(b)(1)(B)(v)(I), the age restriction is 22 to 72. WIA requires college students and members of the Armed Forces to be excluded from the number of disadvantaged Youth and Adults to the extent practicable. The LLSIL is an income level determined annually by the Secretary based on the most recent lower living family budget issued by the Secretary.

The data sets are special tabulations of American Community Survey data obtained by ETA from the Census Bureau based on the five-year moving average using data collected between January 1, 2006 and December 31, 2010. The data sets matching the WIA definition of economically disadvantaged Youth and Adults are intended to be used in the within-state allocation formulas for WIA Youth and Adult programs starting in Program Year (PY) 2013.

ACS Questions Utilized (as suggested by the methodology): number in household, P2, P4, P10, P26, P29, P35-37, P42, P47-48[249]

Major Uses: WIA grantees use the estimates to allocate funds within-state. Researchers may use these datasets as well.

---

[249] While labor force status is not mentioned in the description of the methodology, a data table on the topic is offered.

75. Lower Living Standard Income Level

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology: Use of CPI

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

76. Occupational Information Network (O*NET)

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology: Disability questions (see p. 62 of Part A of O*NET ICR)

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

XII.   U.S. Department of Treasury

A.   Community Development Financial Institutions Fund

77. CDFI Information Mapping System

Purpose of the Data Series:

Responsible Unit:

Authorization:

Frequency of Data Publication:

Timing of Data Release:

Modes to Access Data:

Methodology:

Application of ACS Data in Methodology:

ACS Questions Utilized (as suggested by the methodology):

Major Uses:

XIII. U.S. Department of Transportation

A. Bureau of Transportation Statistics

78. *National Household Travel Survey Transferability Statistics*

Nature and Purpose of Dataset: National Household Travel Survey (NHTS) Transferability Statistics provides estimates of average weekday household person trips, vehicle trips, person miles traveled, and vehicle miles traveled, for each census tract in the United States.

The estimates were made using national travel data collected in the 2009 National Household Travel Survey (NHTS), sponsored by the Federal Highway Administration, and applied to individual census tracts using the 2005 to 2009 American Community Survey (ACS) estimates from the Bureau of Census.[250]

Responsible Unit: Bureau of Transportation Statistics (BTS), Department of Transportation

Authorization: 49 U.S.C. § 111, which authorizes the Secretary of Transportation to collect, compile, analyze and publish a comprehensive set of transportation statistics for transportation decision-making and planning purposes.

Frequency of Data Publication: NHTS Transferability Statistics were created in 1995, 2001, and 2009. A new NHTS will be conducted in 2015-2016.

Timing of Data Release: NA

Modes to Access Data: 2009 NHTS Transferability Statistics data download (SAS, tab-delimited)

Methodology/Application of ACS Data in Methodology: For the 2009 NHTS Transferability Statistics, the NHTS national sample was divided up into six regions, and within each region further divided up into urban, suburban and rural areas. This was done to reduce geographical differences and improve the accuracy and reliability of the estimates.

The transferability analysis requires all U.S. Census tracts to be labeled in terms of their urbanicity (urban, suburban, rural). If a complete urban-rural continuum dataset could be obtained for Census block groups, a dataset for tracts could be created by calculating the mean or median value assigned to block groups contained by a given tract and assigning that value to the tract. Since a complete block group dataset cannot be obtained, a dataset cannot be created for tracts in this manner. For this reason, a new urbanicity measure was created and calculated for all tracts. This measure is based on the population density of a Census tract (converted to a centile score) and on whether the Census tract is in an urban area or urban region/division. The 2010 Census tract and urban boundaries were used in combination with 2010 Census information on population and land area.

The selection of explanatory variables to be used in the analysis relied partially on previous work in the 2001 NHTS Transferability Study. In addition other NHTS household variables were

---

[250] "About NHTS Transferability Statistics," BTS website.

examined for potential inclusion. This examination also included the requirement that comparable data be available in the Census ACS public data tables at the census tract level. This became a significant constraint in developing the life-cycle household variables. The NHTS defined life-cycle variables do not have equivalent counterparts in the ACS data tables. As a result, alternative life-cycle variables were constructed that could be used with the available ACS data tables. The final set of explanatory variables used in the analysis includes:

1. Household income. This variable was converted from the household income categories in the NHTS data to a point estimate, using the mid-point of each category range. For the last category, household income above $100,000 , more detailed Census household income tables were used to derive a weighted average of $147,500 for that category. The natural log of household income was also used in some cases to reflect the non-linear relationship sometimes observed between higher household income and trips taken.   Household income is the best available proxy for household wealth, which is assumed to the primary driver of discretionary travel expenditure.
2. Count of household vehicles
3. Number of members in household
4. Homeowner (yes or no)
5. Number of workers in household
6. Life-Cycle, 1 or more children in household, less than 18 years old
7. Life-Cycle, 1 person household, less than 65 years old
8. Life-Cycle, 2 or more person household, all less than 65 years old
9. Life-Cycle, 2 or more person household, at least one 65 or more years old

The relationship between each dependent variable (person trips, vehicle trips, person miles, and vehicle miles) and the explanatory variables was estimated using multiple linear regression.[251]

A SAS database has been created with those estimates, along with a data dictionary.[252] The 1995 and 2001 NHTS Transferability Statistics used decennial census long-form data.

ACS Questions Utilized (as suggested by the methodology): number of household members, P4, P29, P47-48, H12, H17

Major Uses of the Dataset: NHTS Transferability Statistics are a valuable resource for state and local transportation planners as it allows them to model transportation flows without needing to conduct an independent travel survey. While individual census tract estimates may have limited accuracy in some cases, they can be very beneficial to local governments, and other interested customers, who often do not have the budget and/or time for conducting their own surveys. Using these estimates can make economic sense for those agencies, even if the results are less accurate than if they conducted their own survey. Additionally NHTS has the advantage of using questions standardized across the geographic sample (with only small variations for the

[251] BTS, "NHTS Transferability Technical Report," Draft 4, April 3, 2014.
[252] "About NHTS Transferability Statistics," BTS website.

add-ons), which would not be possible when comparing local surveys with differing methodologies. [253]

---

[253] BTS, *op. cit.*

B. Federal Highway Administration

79. *National Household Travel Survey*

Nature and Purpose of the Dataset: The National Household Travel Survey (NHTS) collects data on local travel by the American public. The survey gathers trip-related data such as mode of transportation, duration, distance and purpose of trip. It also gathers demographic, geographic, and economic data for analysis purposes. NHTS survey data are collected from a sample of U.S. households and expanded to provide national estimates of trips and miles by travel mode, trip purpose, and a host of household attributes.[254]

Responsible Unit: Federal Highway Administration (FHwA), Department of Transportation (DOT). The Federal Transit Administration, AAA, and the Public Policy Institute of AARP also contributed funding to the national sample of the 2009 NHTS. Most of the 2009 NHTS interviews were sponsored by 20 Add-on partners who are state and metropolitan transportation planning organizations that fund additional samples to use as a household travel survey for their respective jurisdictions.[255]

Authorization: Title 23 United States Code, Section 502 authorizes the DOT to carry out advanced research and transportation research to measure the performance of the surface transportation systems in the United States, including the efficiency, energy use, air quality, congestion, and safety of the highway and intermodal transportation systems. The DOT is charged with the overall responsibility to obtain current information on national patterns of travel, which establishes a data base to better understand travel behavior, evaluate the use of transportation facilities, and gauge the impact of the department's policies and programs.[256]

Frequency of Data Publication: The NHTS was conducted in 2001 and 2009. Prior to that, DOT conducted the Nationwide Personal Transportation Survey (NPTS) in 1969, 1977, 1983, 1990, and 1995 and the American Travel Survey (ATS) in 1977 and 1995. A new NHTS will be conducted in 2015. Data collection is tentatively planned to start in autumn of 2015 and will continue for a full year.[257]

Timing of Data Release: First public data release in January 2010, enhanced release in the fall 2010.

Modes to Access Data: National Household Travel Survey Online Analysis Tools and Data Center, Oak Ridge National Laboratory for FhWA (SAS, ascii, html, xls, and csv)

Methodology: The NHTS was conducted as a telephone survey, using Computer-Assisted Telephone Interviewing (CATI) technology. The sample was a list-assisted random digit dialing (RDD) telephone number sample. There were 150,147 households in the landline sample.

---

[254] Text source: "National Household Travel Survey," Bureau of Transportation Statistics website.

[255] Text source: FHwA, "2009 National Household Travel Survey: User's Guide," Version 2, October 2011.

[256] Text source: FHwA, "Supporting Statement: National Household Travel Survey, Part A," October 22, 2007, p.1.

[257] "National Household Travel Survey," Oak Ridge National Laboratory website.

During the administration of the household recruitment interview, the CATI program automatically assigned a travel date to each household. The interviewer revealed the travel date to the household respondent during the interview.

The use of travel diaries on household travel surveys has been shown to improve the accuracy of trip reporting. Therefore, all household members in households who completed a household recruitment interview were sent diaries for their travel day. Of the over 300,000 household members who completed person interviews, 72 percent reported having filled out the travel diary.

The 2009 NHTS interviews were conducted from March 17, 2008 through May 7, 2009. The first assigned Travel Day was March 28, 2008 and the last assigned travel date was April 30, 2009.

The survey must be conducted over at least a 12-month period so that seasonal variations in travel are represented. As in 1995 and 2001, the 2009 NHTS took 14 months, rather than 12 to complete. This was because interviewers were trained in waves and it took a few months to train all the interviewers needed for the study. The survey weights adjust for the monthly differences in number of interviews completed.

Travel day dates were assigned to all seven days of the week, including holidays. The intent was to represent travel across an entire year.

Interviews were conducted with households in all 50 States and the District of Columbia. This survey was designed as a list-assisted random digit dialing survey, to yield an equal probability sample of households with landline telephones. The national sample was increased in 20 Add-on areas: Phoenix AZ, Tucson AZ, California, Florida, Georgia, Iowa, Cedar Rapids IA, Indiana, North Carolina, Piedmont Region NC, Omaha NE, New York, South Carolina, South Dakota, Tennessee, Texas, Virginia, Vermont, Chittenden County VT, and Wisconsin. The target sample size was 25,000 completed households for the national sample.

Data collection for the national landline sample of the 2009 NHTS, as well as the 20 Add-on samples, was conducted utilizing staff at Westat's Telephone Research Centers (TRCs) and their vast team of at home interviewers (AHI) located across the nation. The centers used were located in Frederick, MD, Sarasota, FL, Sacramento and Merced, CA, and Rockville, MD. Westat is a social science research firm headquartered in Rockville, Maryland.

A total of 196,619 households were recruited for the 2009 NHTS. The final dataset contains information on the 150,147 useable households. Although the definition of a useable household required only 50 percent of adults to complete a person interview, Westat completed person interviews with or by proxy for 93 percent of the eligible adult household members. The total number of household members enumerated in these completed households was 351,178.[258]

---

[258] Text source: FHwA, "2009 National Household Travel Survey: User's Guide," Version 2, October 2011.

NHTS developed weights to reflect the selection probabilities and adjustments to account for nonresponse, undercoverage, and multiple telephones in a household. The overall steps in the weighting process were as follows:

- Construction of base weights—the base weights are the reciprocals of the telephone frame sampling rates within each sample group (the sample group was a particular telephone number sample for a particular study1 at a particular point in time, with five sample groups being taken over the course of the year for most of the study areas).
- Construction of jackknife replicate base weights—the replicate weights are designed to allow the user to easily produce valid jackknife variance estimators based on the sample design;
- Household-level nonresponse adjustments (done within each sample group and study area separately);
- Composition of the household nonresponse-adjusted base weights by domain (the domains are defined by the quarter in which the travel date occurred and by final geography as determined by geocoding each household's address);
- Household-level raking and trimming (using the composite weights);
- Person-level nonresponse adjustments (for nonresponding persons within useable households);
- Person-level weights (with person-level raking and trimming).[259]

There are several different weights, and it is important that the appropriate weight is used for a particular estimate. There are household weights, person weights and travel day weights.

- Household weights (WTHHFIN) are used whenever one is tabulating an estimate at the household level as opposed to the person level, such as number of households by household vehicle ownership and distribution of households by number of household drivers.
- Vehicle weights (WTHHFIN) are the same as the household weight since the vehicle is considered a household attribute. Use the vehicle weight for items such as vehicles by vehicle type or by vehicle age.
- Travel day weights (WTTRDFIN) are used for estimates involving numbers of trips or miles of travel, for example, number of vehicle trips by trip purpose. Only trips in privately-owned vehicles that are reported by the driver should be counted in estimating vehicle trips. For example, if a person reports being a passenger in a vehicle driven by another household member, that trip would not be counted.
- Person weights (WTPERFIN) are generally used for estimates of non-household and non-travel day items of interest, for example workers by gender, drivers by annual miles estimated, etc.[260]

---

[259] Westat, "Task N: Weighting Report," November 17, 2010, pp. 4-5.
[260] Text source: FHwA, "2009 National Household Travel Survey: User's Guide," Version 2, October 2011.

<u>Application of ACS Data in Methodology</u>: The ACS was used in the development of the 2009 NHTS data in three ways:

- Coverage adjustments to adjust all selected characteristics simultaneously until all weight sums are within a small tolerance level (e.g., 0.001 percent) of the population estimates for each characteristic simultaneously. Characteristics such as race, ethnicity, Census Division, MSA size, household size, number of household vehicles, worker status were used in developing coverage adjustments.[261]
- Non-coverage adjustments on the basis of ACS journey-to-work data.[262]
- Validation of NHTS data with ACS journey-to-work data.[263]

Regarding the use of the ACS data as controls: The estimation procedure took composite weights and adjusted them to be consistent with independent controls for the domains based on various demographic categories, in a process called raking. This reduces the variances for any characteristics which are correlated to these demographic characteristics (as the independent controls have much lower variability). The source for these controls was the 2008 American Community Survey (ACS), and the controls consisted of attributes such as: geography, race, ethnicity and number of household vehicles.[264]

The control totals for all dimensions were derived from one-year 2008 ACS estimates where possible, using 2006-08 three year estimates where necessary to impute distributions for areas for which 2008 one-year ACS estimates were not available.

Geographic raking cells were always defined at the county level (either single counties or sets of counties). For the larger counties, the ACS had estimates of control totals that are based on the one year of data alone (the most recent year). This was the ideal situation. For many other counties, the ACS had an estimate, but this estimate is a moving average based on the three most recent years. ACS moving averages were used only in such a way that they did not contradict the one-year estimates for any given area. The three-year estimates were only used to "fill in" whatever was missing from the one-year estimates (only percentages were utilized). Finally, there were counties for which there was no estimate.

For a geographic raking cell to have a one-year ACS estimate, it was necessary that all counties within the raking cell had one-year ACS estimates. The raking cell estimate was the summation of the one-year ACS county estimates. If all counties within a raking cell had three-year ACS estimates, but only some had one-year ACS estimates, then a three-year estimate was computed for these domains, by taking the summation of the county-level three-year estimates. There was no mixing of one-year and three-year estimates in any summations. If a

---

[261] *Ibid.* See Table 1-3, p. 1-7.
[262] FHwA, "Supporting Statement: National Household Travel Survey, Part B," October 22, 2007, p.7.
[263] *Ibid*, p. 12.
[264] Text source: FHwA, "2009 National Household Travel Survey: User's Guide," Version 2, October 2011, p. 5-4.

particular raking cell had at least one county with no ACS estimate at all (one- or three-year), then no estimate was generated for the raking cell.[265]

ACS Questions Utilized (as suggested by the methodology): P3-8, P11, P17-19, P29-37, P45-48, H1, H10, H12, H17

Major Uses: The NHTS is a tool in the urban transportation planning process; it provides data on personal travel behavior, trends in travel over time, trip generation rates, national data to use as a benchmark in reviewing local data, and data for various other planning and modeling applications.

The transportation research community, including academics, consultants and government, use the NHTS extensively to examine:

- travel behavior at the individual and household level;
- the characteristics of travel, such as trip chaining, use of the various modes, amount and purpose of travel by time of day and day of week, vehicle occupancy, and a host of other attributes;
- the relationship between demographics and travel; and
- the public's perceptions of the transportation system.

People in various fields outside of transportation use the NHTS data to connect the role of transportation with other aspects of our lives. Medical researchers use the data to determine crash exposure rates of drivers and passengers, including the elderly, who have heightened morbidity and mortality rates. Safety specialists study the accident risk of school-age children, particularly when they are traveling on their own by walking or biking. Social service agencies need to know more about how low-income households currently meet their travel needs.[266]

The NHTS is designed primarily to obtain data needed for performance measurement, policy analyses, program planning, and program management by agencies within the Department of Transportation (DOT): the Office of the Secretary (OST), Federal Highway Administration (FHWA), the Federal Transit Administration (FTA), the National Highway Traffic Safety Administration (NHTSA), and the Bureau of Transportation Statistics (BTS).

The NHTS provides data not available from any other source, and has a widely disparate user community. Other Federal agencies depend on the NHTS data, for instance the Department of Energy (DOE) uses the data to estimate residential transportation-related energy use, the Environmental Protection Agency (EPA) uses the data as an input to the standard air-quality estimation model, and the Centers for Disease Control (CDC) use NHTS to assess the connection between transportation choice, land use, and health. State and local transportation agencies and transit authorities use the data for benchmarking local statistics and to examine trends in use. The data is also widely used by a range of researchers and public and private agencies, for example the Transportation Research Board, Highway Users Federation for Safety and Mobility,

---

[265] Westat, *op., cit*.
[266] "Introduction to the 2009 NHTS," Oak Ridge National Laboratory website.

American Association of State Highway and Transportation Officials, the Insurance Institute for Highway Safety, the travel and tourism industries, National Transportation Safety Board (NTSB), University Transportation Centers (UTCs) and the broader transportation research community.

DOT uses the data to estimate the amount and nature of passenger travel in the U.S., the relationship between socio-economic characteristics and travel patterns, and trends in passenger travel. These data are critical to develop policies related to congestion, mobility, safety, finance, and for forecasting future demand.

Travel behavior data informs on complex and inter-related policies and performance measures. The growing array of travel data needs are a direct result of a growing and changing U.S. population. This creates an even greater need for key measures of travel behavior to support safety, mobility, and congestion strategies as volumes increase, the traveling population ages, the mobility needs of new immigrants require special consideration for alternative modes, and as safety and air quality issues continue to trouble many urban areas. In addition, infrastructure improvements are advocated to safely encourage walking and biking to improve American's health, new land-use strategies are being implemented to reduce vehicle travel and increase transit use, and new vehicle types are being marketed to reduce energy consumption and greenhouse gas emissions.

OST, FHWA, FTA, NHTSA and BTS use NHTS data to address a number of issues in the Department's Strategic Plan. Specific applications are outlined below.

Safety. Ensuring the safety of the American public when traveling has been a long standing mission of the Department of Transportation. The NHTS is the only source of data available on the level of use of the transportation system by mode of travel and demographic group. Specifically, NHTS data is used by the Department, Administration, and transportation organizations for:

- Calculating exposure rates by age, gender, and vehicle type to compute risk for crash and fatalities. These data are used to assess impacts of demographic shifts and vehicle technology on the safety of the American public. The risk assessment tool, TrafficSTATS, developed by the AAA Traffic Safety Foundation is one example of this use (http://www.aaafoundation.org).
- Development of educational campaigns that reach target audiences. For example, understanding who is traveling at high accident times allows educational campaigns to target traveling market segments, including details on the age and sex of the driver, the number of people in the vehicle, the purpose of travel, and vehicle characteristics which can impact collision severity, such as vehicle age, type, etc.
- Analysis of the incidence of walk and bike trips, characteristics of those making these trips and the trips themselves, such as time of day, and trip purpose to establish baseline measures of exposure and address high risk areas (e.g. rural roads) and demographic groups (e.g. new immigrants).
- Information on rare modes, such as motorcycle use, to help understand the growth in motorcycle accidents and fatalities in recent years.

- Evaluating new safety initiatives such as Safe Routes to School to understand travel to school to help monitor special programs related to increasing the safety of children walking and biking to school.

Congestion. Reducing the level of congestion in U.S. cities is one of the top priorities of the DOT. The NHTS plays an important role in understanding the travel behavior that contributes to the congestion issue. Travel demand is generated by the choices that people make to carry out their daily activities. The steady increase in travel demand over the past fifty years (especially vehicle travel) has created high levels of congestion in our urban areas. Congestion is no longer only a weekday work commute issue. In fact, approximately half of all travel during peak commute times is for non-work purposes. The NHTS is the only source of information on non-work travel. The NHTS supports policy and planning by DOT, the Administration, and Congress in measuring in the demand side of congestion including:

- Vehicle occupancy during congested times to measure changes in carpool rates,
- Trip purpose distribution for peak and off-peak travel, including non-work,
- Mode share for all trips by time of day and day of week, and
- Trends in time, distance, and speed for work and non-work travel.
- Homeland security issues? E.g., Capacity & congestion of designated evacuation routes

Mobility. Mobility issues are particularly acute for the elderly, new immigrants, and the poor. Issues include access to and use of alternative means of transportation, the range of daily mobility, and the relative cost of transportation for the household.   In particular, the NHTS provides data that supports an examination of:

- Mobility of people of color and language minorities, including whether their reduced mobility is due to race/Hispanic origin, income, residential location, or other barriers to access and mobility.
- Women's travel issues, particularly the travel behavior of working mothers who continue to retain primary responsibility for family and household needs, and elderly women who may be isolated when they give up driving
- Teen travel behavior, especially in the pre-driving age and through the later teens as graduated licensing becomes more common. Teen travel behavior has impacts on safety, household trip generation, and future transportation service needs and demand.
- Transit use is analyzed in terms of the socioeconomic characteristics of users, availability and access to transit, and trip characteristics including wait time, trip purpose, length, travel party size, and time of day.

Economic Issues. Financing options for the highway system and the quantification of the cost of transportation issues such as congestion on the U.S. economy are top issues in the DOT, Congress, and Administration. National data on passenger travel, as collected in the NHTS, informs on key aspects of economic issues related to cost and finance of the surface transportation system. These include:

- Evaluation of highway finance options,
- Measures of efficiency of surface transportation modes, such as travel time data and trends in travel time,
- The impact of user fees based on estimates of the socio-economics of the traveler, the purpose of travel during peak periods, and the other relevant characteristics of peak period travel,
- Characteristics of travel to work, with particular emphasis on the NHTS data serving as a bridge for state and metropolitan planners to use Decennial Census data in their travel models, and
- Data on vehicle ownership, vehicle characteristics and amount of travel are used in the FHWA revenue forecasting model to forecast Highway Trust Fund receipts for use by DOT and the Department of Treasury

Private Vehicle Fleet Characteristics. The NHTS is the only national source of data on the composition of the household vehicle fleet, particularly vehicle type and age, and how it has changed over time. In addition, the Energy Information Agency partners with DOT to append data on fuel efficiency, gas cost (at the household's location in the month of data collection), and annual fuel use. This data is critical for determining both trends in fuel use and understating the changing types and levels of emissions. Important changes in the character of the vehicle fleet have been tracked using the NHTS data including:

- Increased ownership and use of SUVs and the impacts of that trend on fuel use and vehicle emissions,
- The relative cost of travel and fuel usage by the type of autos owned by the household and the household's location,
- Increased ownership and use of hybrid vehicles
- Changes in the overall fuel efficiency of the residential vehicle fleet, and
- Changes in the overall age of the residential vehicle fleet and availability of safety features to key demographic groups.

Local Level Planning and Policy. In addition to DOT policy issues and Strategic Plan goals, a key function of NHTS data is in the planning processes of States and MPOs (Metropolitan Planning Organizations).   NHTS data is used to supplement, or even substitute for, local data on key variables needed in the policy and planning process. States and MPOs have used NHTS data as inputs into the travel demand forecasting, safety planning, and air quality analyses that are mandated by Congress.   Since many large urban areas collect local data, NHTS is most useful to very small MPOs with limited resources and/or to States who are piecing together data from a number of urban areas or require some estimates of rural travel behavior to fill in State data gaps.[267]

---

[267] Text source: FHwA, "Supporting Statement: National Household Travel Survey, Part A," October 22, 2007, pp. 2-5.

Oak Ridge National Laboratory annually publishes an extensive compendium of uses of NHTS data.[268]

C. Federal Highway Administration, Bureau of Transportation Statistics, and Federal Transit Administration

80. *Census Transportation Planning Products*[269]

Nature and Purpose of the Dataset: The Census Transportation Planning Products (CTPP) is a set of special tabulations designed by transportation planners using large sample surveys conducted by the Census Bureau. From 1970 to 2000, the CTPP and its predecessor used data from the decennial census long form.[270] With the advent of the American Community Survey (ACS), the CTPP has been redesigned as a special tabulation of multi-year ACS data.

The CTPP data product based on 2006 – 2010 5-year ACS Data consists of almost 200 residence based tables, 115 workplace based tables and 39 flow tables (home to work) for over 325,000 geographies.

As a special tabulation, the CTPP data products are different from standard ACS data products in several key ways. For instance:

- CTPP 2006-2010 data product has journey to work flow data while ACS standard data product does not provide flow data;
- CTPP 2006-2010 data product has workplace tables down to small geographies including Traffic Analysis Zone (TAZ) and Census Tract, but the workplace tables in standard ACS only go to county level;
- The CTPP has many more transportation specific tables. The CTPP 2006-2010 has seven tables with commute mode in a cross-tabulation, while ACS standard products have only two tables.
- CTPP 2006-2010 data product has customized geographies including Traffic Analysis Zone and Transportation Analysis District.[271]

Responsible Unit: Social, Economic and Housing Statistics Division, Census Bureau for agencies in the Department of Transportation (Federal Highway Administration, Bureau of Transportation Statistics, and Federal Transit Administration) and the American Association of State Highway and Transportation Officials (AASHTO), in cooperation with the Census Subcommittee of the Transportation Research Board.

A CTPP Oversight Board was created to provide direction and oversight to the overall CTPP program. Voting members of the CTPP Oversight Board members include representatives from states and MPOs. Ex-officio members include representatives of U.S. DOT agencies, the CB, and a number of census data users from academia and the consultant community. The CTPP Oversight Board reports to the AASHTO Standing Committee on Planning (SCOP).

---

[269] An alternative CTPP website is hosted by the American Association of State Highway and Transportation Officials.

[270] When the CTPP was based on the decennial census, the second "P" stood for "Package."

[271] Text source: "CTPP 2006-2010 Frequently Asked Questions," Federal Highway Administration website.

Authorization: 49 U.S.C. § 111, which authorizes the Secretary of Transportation to collect, compile, analyze and publish a comprehensive set of transportation statistics for transportation decision-making and planning purposes.

Frequency of Data Publication: A CTPP was prepared after the 1990 and 2000 decennial censuses. At present, a number of CTPP-branded products are available on the web; while some are clearly Census Bureau special tabulations, it is not entirely clear if others are. ACS-based CTTP products include:

- One-year data (2005, 2007) transportation profiles
- Three-year data (2005-2007, 2006-2008, 2008-2010, 2010-2012) profiles for states, counties, and places with more than 20,000 people outlining key population, residence, workplace, journey-to-work, and flow-data characteristics.
- Five-year data (2006-2010) at a national level as well as at smaller state, county, place, census tract, block groups (where possible), and Transportation Analysis Zones (TAZ) geographic levels.

The next special tabulation is envisioned to be based on the American Community Survey 2012–2016 data set. By using 2012-2016 ACS, the CTPP will take advantage of the increased sample size of the ACS, fully implemented in 2012, and the increased computer opt-in responses.

Timing of Data Release: The Census Bureau delivered the 2006-2008 CTPP in February 2010. The 2006-2010 CTPP was released in October 2013.

Modes to Access Data: ACS-based CTPP data products are available at FHwA and AASHTO.

Methodology/Application of ACS Data in Methodology: CTPP products are prepared by the Census Bureau according to the specifications developed by the client group. Tables are organized into three groups: residence-based tables (188), workplace-based tables (116), and worker home-to-work tables (40).[272]

ACS Questions Utilized (as suggested by the methodology): P2-10, P14, P20, P29, P31-37, P40-48, H1, H8, H12, H17

Major Uses: Census data on demographic characteristics, home and work locations and journey to work travel flows are key inputs to a variety of state, regional and local transportation policy and planning efforts. They also support corridor and project studies, environmental analyses and emergency operations management.

The CTPP program provides data to support over 20 Federal planning requirements. Included are data on residences, work places, journey to work characteristics and flows. The CTPP is a national source of data that provides journey-to-work information, including data on mode

---

[272] A complete list of available tables can be found at CTPP tables-5YR List public.xlsx.

splits, travel times, and travel distances. The CTPP represents a significant source of data for supporting a wide variety of transportation planning tasks, such as:

- Policy studies
- Travel demand modeling
- Congestion management
- Emergency preparedness
- Corridor and project studies
- Transit new start and service planning
- Environmental justice studies
- Air quality conformity
- Environmental justice reviews
- Trends analyses

The CTPP is the only source of journey to work data for many small and medium MPOs that lack the resources to do more comprehensive travel surveys. CTPP data, particularly on mode use and travel times, becomes even more important as transportation agencies become more involved in performance management and in understanding relationships between transportation, climate change, sustainability, reliability, and accessibility.[273]

---

[273] Text source: "Census Transportation Planning Products Program 2008–2012, Mid-Program Report," October 2010, available on the AASHTO website. Author not specified.

XIV. U.S. Department of Veterans Affairs

A. National Center for Veterans Analysis and Statistics

81. *National Survey of Veterans*

Nature and Purpose of the Dataset: The National Survey of Veterans (NSV, officially "The National Survey of Veterans, Active Duty Service Members, Activated National Guard and Reserve Members, Family Members and Survivors") collects data on the demographic characteristics of the Veteran population of the United States, as well as data on the delivery and usage of Veteran services and benefits. Data are for the nation as a whole.

Responsible Unit: National Center for Veterans Analysis and Statistics, Department of Veterans Affairs

Authorization: While past NSVs have been conducted under the general authorization of U.S. Code Title 38, Section 527 which requires the VA Secretary to gather data for the purposes of planning and evaluating VA programs, the 2010 NSV also included the requirement, at the direction of P.L. 108-454, Section 805, to assess beneficiary awareness of VA benefits and services. The Public Law also expanded the survey populations in the 2010 NSV to include in addition to Veterans, other beneficiary groups: Active Duty Service members; demobilized National Guard and Reserve members; Family members and Surviving spouses.[274]

Frequency of Data Publication: The first NSV studies were conducted in the late 1970s, and subsequent NSVs were conducted in 1987, 1993, 2001, and most recently in 2010.

Timing of Data Release: The 2010 NSV report was released in October 2010.

Modes to Access Data: 2010 NSV report and spreadsheet (pdf, xls)

Methodology: The 2010 National Survey of Veterans, Active Duty Service members, demobilized National Guard and Reserve members, family members and survivors (NSV), targets six distinct populations. Procedures were designed to sample the following populations:

- Veterans
- Demobilized National Guard and Reserve members
- Active Duty Service members, including currently activated National Guard and Reserve members (excluding those in zones of conflict, primarily Afghanistan and Iraq)
- Veteran spouses
- Spouses of active military (including spouses of currently activated National Guard and Reserve members)

---

[274] Text source: Westat, "National Survey of Veterans, Active Duty Service Members, Demobilized National Guard and Reserve Members, Family Members, and Surviving Spouses," Final Report, Deliverable 27 to the Department of Veterans Affairs, October 18, 2010, p. viii.

- Surviving spouses

The design for the 2010 NSV split the six target populations into two overall sampling groups. The first was an address-based group that included Veterans, Veteran spouses, and Surviving spouses. The second group was list-based and was comprised of active duty military members, including currently activated National Guard and Reserve members, demobilized National Guard and Reserve members, and Active Duty spouses.

Sections in the 2010 NSV instruments included:

- Military Background
- Familiarity with Veteran Benefits
- Disability
- Vocational Rehabilitation
- Health Status
- Health Care
- Health Insurance
- Education and Training
- Employment
- Life Insurance
- Home Loans
- Burial Benefits
- Burial Plans
- Internet Use
- Income
- Demographics
- Your Veteran Spouse
- Disability (Veteran Spouse)
- Educational Assistance

Data collection began October 16, 2009, and ended March 19, 2010. We used two data collection methods or approaches, one for the address-based sample group and one for the list-based sample group, to collect the 2010 NSV data. A total of 10,972 surveys were completed across the various survey populations of interest. There were 8,710 completed surveys received from Veterans. The response rate for the household screening survey was 32.3 percent; the response rate for the Veteran Survey was 66.7 percent, for an overall response rate of 21.5 percent. Survey data were weighted to represent the entire non-institutionalized Veteran population. Weights incorporated the probability of selection, survey nonresponse, and were post-stratified to known population totals.[275]

Application of ACS Data in Methodology: The ACS was used in the NSV in two ways. One, sampling weights were developed to account for the selection probability of each sample unit

---

[275] Text source: *Ibid,* Chapter 2: Study Methodology.

and to compensate for unit nonresponse. Ratio adjustments were applied such that the sums of the sampling weights match known population targets.

For the address-based samples, the Veteran sampling weights were adjusted by using data from the VA's VetPop model and the 2008 American Community Survey. As part of this procedure, the ACS was used to derive the proportions of Veterans living in group quarters and then exclude such Veterans from the control totals. Estimates for Veterans' spouses were also calculated from the 2008 American Community Survey, which asks the Veteran status of all individual members of the sampled households.[276]

Second, the 2010 NSV report compares its results with the 2001 NSV, VA's 2007 Employment History Survey, and the Census Bureau's 2008 American Community Survey. The intention of these comparisons is to provide the reader with some sense of change or trends over time. The comparisons between the ACS and the NSV provide a quality check of the NSV data in terms of the weighted, demographic distribution of the 2010 NSV. [277]

ACS Questions Utilized (as suggested by the methodology): P2-6, P11, P16, P26-29, P35-37, P47-48, H17

Major Uses: The NSV is of great importance to VA because it is often the only source of information on Veterans who are not currently using VA benefits and services. The NSV is designed to help VA improve services for beneficiaries and their families. It will collect information that is not available in VA administrative files.

The NSV provides VA, Congress, stakeholders, and the public more accurate descriptions and assessments of the characteristics of the Veteran population to evaluate existing programs and policies, to establish baseline measures before planning and implementing new programs and policies, and to monitor progress of programs and policies and their impacts on the population. The NSV will provide information to support VA policy, planning, and quality improvement decisions.

Because of the infrequent administration of the NSV, the data from any single administration are the only source of information on Veterans who are not currently receiving benefits or services from the VA. The NSV is currently the only tool for reaching this population, and the NSV is the only existing tool that will allow for direct comparisons between Veterans using VA services and those who are not. The NSV is also the only existing tool that can evaluate VA outreach efforts and assess benefit awareness among Veterans. Since the data from the NSV are so important, and because the NSV is conducted infrequently, it is important that we minimize the total survey error associated with the study design.

VA uses the collected survey data to:

- Understand the level of awareness the Veterans have on VA services and benefits;
- Determine current and future use of VA services and benefits;

---

[276] *Ibid*, "Appendix B National Survey of Veterans Detailed Description of Weighting Procedures."
[277] Text source: *Ibid*, p. 47.

- Determine best methods to disseminate information on VA benefits and services; and
- Craft VA policies to plan and make quality improvement decisions.

VA's overall goal in using the data will be to monitor and improve Veterans' services and benefits for the health and well-being of the Veterans.[278]

Secretary Shinseki was quoted as saying about the 2010 NSV: "By hearing directly from Veterans and their family members, we gain valuable information to help us serve them better."[279]

---

[278] Text source: Department of Veterans Affairs, "Supporting Statement, Information Collection for The National Survey of Veterans, Active Duty Service Members, Activated National Guard and Reserve Members, Family Members and Survivors (NSV), Part A," July 27, 2009.

[279] Office of Public and Intergovernmental Affairs, Department of Veterans Affairs, "VA to Survey Veteran Households," December 3, 2009.

B. Office of the Actuary

82. *Veteran Population Projection Model*

Purpose of the Measure: The Veteran Population Projection Model (VetPop) provides the latest official Veteran population projection from the Department of Veterans Affairs (VA). VetPop is an actuarial projection model developed by the Office of the Actuary (OACT) for Veteran population projection. VetPop provides living Veteran counts by key demographic characteristics such as age, gender, period of service, and race/ethnicity at various geographic levels.

Responsible Unit: Office of the Actuary, Department of Veterans Affairs

Authorization: Statutory requirements for Veterans data come from the following sources: Veterans Benefits Program (38 U.S.C. 317), Veterans Employment and Training Program (38 U.S.C. 4103 et seq.), Veterans Needs (Reports to Congress, 38 U.S.C. 542, 4107(c)), Veterans Outreach Program (Disabled, 38 U.S.C. 4103A(a)(1) and (b)(2)), Veterans Population (State Projections, 38 U.S.C. 8131(l), Veterans Program (Evaluation, 38 U.S.C. 527), Veterans State Estimates of Nursing Home Care (38 U.S.C. 8134(a)(1)). Among the various provisions is authorization within 38 USC 8131 which states that "The veteran population of each State shall be determined on the basis of the latest figures certified by the Department of Commerce."

Frequency of Data Publication: As of 2014, the VA has made seven projections of the Veterans population.

Timing of Data Release: Normally, every two years. VetPop2014 was released October 1, 2014 and covers 2013-2043.

Modes to Access Data: Veterans Population Project Model 2014, National Center for Veterans Analysis and Statistics (xls)

Methodology: VetPop2014 is the 7th generation of the OACT Veteran Population Projection Model with significant improvements in data, methodology, and modeling processes. Similar to the previous Veteran Population Projection Model 2011 (VetPop2011), it is a bottom-up model which projects future Veteran population at the county level as a starting point. The county level projections are then aggregated to provide Veteran information at larger geographic units such as congressional districts, states, and at the national level. The VetPop2014 actuarial model uses both Veteran record-level data and survey data from a wide variety of sources including VA, Department of Defense (DoD), U.S. Census Bureau, Department of Treasury's Internal Revenue Service (IRS), and the Social Security Administration (SSA). These data sources enabled OACT to develop the VetPop2014 Model using advanced actuarial and predictive modeling methods for three critical modules -- the Separation Module, the Mortality Module, and the Migration Module.[280]

---

[280] Office of the Actuary, Department of Veterans Affairs, "Veteran Population Projection Model 2014," October 2014.

Application of ACS Data in Methodology: ACS data on certain demographic characteristics of veterans is used to model population projections. These characteristics include service history, disability rating, age, gender, and migration. In particular, VetPop2014 Migration Module developed the county Veteran migration models for various age and gender cohorts using historical longitudinal data from VA, IRS, and ACS.[281]

ACS Questions Utilized (as suggested by the methodology): P3-6, P15, P26-28

Major Uses: These statistics are important for understanding the Veterans population of the U.S. and for planning policies and programs accordingly. Due to the latest data enhancements, the VetPop2014 projected more living Veterans in the future compared to VetPop2011. Overall, while the male Veteran population steadily decreases, woman and minority Veteran population are projected to increase over the next 30 years. Another noticeable trend for the Veteran population is the projected higher growths in the Southern and Western regions.[282]

---

[281] *Ibid.*
[282] *Ibid.*

XV.   Independent Agencies

A. **Environmental Protection Agency**

83. *America's Children and the Environment*

Nature and Purpose of the Measure: *America's Children and the Environment* (ACE) is the Environmental Protection Agency's (EPA's) report presenting data on children's environmental health. ACE brings together information from a variety of sources to provide national indicators in the following areas:

- Environments and Contaminants (conditions in the environment, such as levels of air pollution)
- Biomonitoring (contaminants measured in the bodies of children and women of child-bearing age, such as children's blood lead levels)
- Health (rates at which selected health outcomes occur among U.S. children, such as the annual percentage of children who currently have asthma)

Accompanying each indicator is text discussing the relevance of the issue to children's environmental health and describing the data used in preparing the indicator. Wherever possible, the indicators are based on data sources that are updated in a consistent manner, so that indicator values may be compared over time.[283]

Responsible Unit: Office of Policy, EPA

Authorization: Under Executive Order 13045, EPA and other federal agencies are directed to "make it a high priority to identify and assess environmental health risks and safety risks that may disproportionately affect children."

Frequency of Data Publication: The first ACE report was published in December 2000, the second in February 2003, and the third in January 2013.

Timing of Data Release: NA

Modes to Access Data: America's Children and the Environment, Third Edition (ACE3), EPA website (pdf)

Methodology: In choosing indicators for ACE3, EPA considered a variety of factors, including public interest, magnitude of prevalence and/or trend in prevalence, extent of exposure, severity of health outcome, past EPA actions to address the issue, and research findings indicating or suggesting that an environmental exposure may contribute to a children's health outcome. ACE3 includes topics for which scientific evidence is sufficient to conclude there is a causal relationship between exposure and health effects, as well as topics for which there is

---

[283] Text source: "America's Children and the Environment, Third Edition (ACE3)—Frequently Asked Questions," EPA webpage.

less extensive scientific evidence. Inclusion of a topic in ACE3, therefore, does not imply that a cause-effect determination has been made.

The selection of topics involved generating a list of children's environmental health issues of potential interest, evaluating availability of suitable databases relevant to those topics, and considering indicators that might be derived from those databases. EPA obtained input from members of EPA's Children's Health Protection Advisory Committee (CHPAC) on each stage of this process, including input on the ultimate selection of topics and indicators presented in ACE3. Independent external peer reviewers provided their opinions to EPA regarding the suitability of the indicators and other information provided for each topic. EPA revised the report based on the peer review comments, and comments received from the public.[284]

Regarding the presence of hazardous air pollutants: EPA's National Air Toxics Assessment (NATA) provides estimated concentrations of 181 HAPs in ambient air for the year 2005. NATA is the most comprehensive resource on potential human exposure to and risk of adverse health effects from HAPs in the United States. Monitoring data are insufficient to characterize HAP concentrations across the country because of the limited number of monitors, and because concentrations of many HAPs may vary considerably within a metropolitan area or region.

Under NATA, EPA develops modeled estimates of ambient concentrations of HAPs using estimated emissions data from major, area, onroad mobile, and non-road mobile sources. These emissions data are collected and updated periodically, and are maintained in an emissions inventory. The most recent assessment developed estimated ambient concentrations of 179 air toxics for the year 2005. A computer model provided estimates for every census tract in the United States. The modeled estimates generally are consistent with the limited set of ambient air toxics monitoring data, although at many locations the model estimates for some HAPs are lower than measured concentrations.[285]

Application of ACS Data in Methodology: The ACS provides the denominator for several indicators of the presence of hazardous air pollutants:

- Percentage of children ages 0 to 17 years living in census tracts where estimated hazardous air pollutant concentrations were greater than health benchmarks in 2005
- Percentage of schoolchildren attending schools in census tracts where estimated hazardous air pollutant concentrations were greater than health benchmarks in 2005
- Percentage of children ages 0 to 17 years living in census tracts where the cancer risk from estimated hazardous air pollutant concentrations was at least one in 10,000 in 2005, by race/ethnicity and family income

---

[284] Text source: *Ibid.*

[285] Text source: EPA, *America's Children and the Environment*, January 2013, p. 52.

- Percentage of children ages 0 to 17 years living in census tracts where the non-cancer risk from estimated hazardous air pollutant concentrations exceeded health benchmarks in 2005, by race/ethnicity and family income[286]

ACS Questions Utilized (as suggested by the methodology): number in household, P4-6, P10, P47-48

Major Uses: America's Children and the Environment (ACE) has three principal objectives:

- First, it compiles data from a variety of sources to present concrete, quantifiable indicators for key factors relevant to the environment and children's health in the United States.
- Second, it can inform discussions among policymakers and the public about how to improve data on children's health and the environment.
- Third, it includes indicators that can be used by policymakers and the public to track trends in children's environmental health, and ultimately to help identify and evaluate ways to minimize environmental impacts on children.[287]

---

[286] *Ibid*, pp. A-7, B-7-8.

[287] Text source: Text source: "America's Children and the Environment, Third Edition (ACE3)—Frequently Asked Questions," EPA webpage.

84. EJView

Nature and Purpose of the Measure: EJView, formerly known as the Environmental Justice Geographic Assessment Tool, is a mapping tool that allows users to create maps and generate detailed reports based on the geographic areas and data sets they choose. EJView includes data from multiple factors that may affect human and environmental health within a community or region, including: demographic, health, environmental, and facility-level data.[288]

Responsible Unit: National Geospatial Program, Environmental Protection Agency (EPA)

Authorization: 42 U.S. Code § 4332, National Environmental Policy Act of 1969, and Executive Order (E.O.) 12898 - Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations - was issued by President William J. Clinton in 1994. Its purpose is to focus federal attention on the environmental and human health effects of federal actions on minority and low-income populations with the goal of achieving environmental protection for all communities.[289]

Frequency of Data Publication: EJView was released in May 2012. Prior versions were offered in 2008 and 2010.

Timing of Data Release: NA

Modes to Access Data: EJView, EPA (mapping tool)

Methodology: EJView uses an area-weighted method of population estimation. Population and housing statistics are created by overlaying the specified study area (buffered point, user-digitized polygon or map window) with the appropriate Census summary level geography (block, block group, tract, or county). For each Census polygon the respective population values are adjusted proportionally (area-weighted) based on the percentage of the polygon that lies within the study area. For example, given 3 block groups intersecting the study area with the following characteristics:

- 100% contained with a population of 15;
- 50% contained with a population of 12;
- 20% contained with a population of 10

The total estimated population is (1 * 15) + (0.5 * 12) + (0.2 * 10) = 23

This method assumes an even distribution of population over each Census geographic unit. The accuracy of the method depends on how even the population distribution actually is and on the size of the geographic units. This is the best population estimation method available without having additional information about the population distribution.[290]

---

[288] Text source: "About EJView," EPA webpage.
[289] Text source: "Summary of Executive Order 12898 - Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations," EPA webpage.
[290] *Ibid.*

Categories of EJView data elements include community-based EPA grants, sites reporting to EPA, water monitoring stations, places (for schools, hospitals, and worship places), nonattainment areas, health, demographics, and water and boundaries features.[291]

Application of ACS Data in Methodology: Demographic data by census tract is available from the 2006-2010 ACS, including:

- Population Density (People/sq mi) - Persons Per Square Mile is computed by dividing the total population count by the land area in square miles. Available by Blockgroup, by Tract, and by County.
- Per Capita Income - Per Capita Income is computed by dividing the collective income for all persons 15 years and over by the total population count within the area. Available by Blockgroup, by Tract, and by County.
- Below Poverty (%) - Percent Below Poverty is computed by dividing the sum of persons living below the poverty level by the number of persons for whom poverty status is determined. Available by Tract, or by County.
- Education < 12G (%) - The percentage of adults (18 years and older) who have not completed high school. Available by Blockgroup, by Tract, and by County..
- HS Diploma Only (%) - The percentage of adults (18 years and older) who have a high school diploma only. Available by Blockgroup, by Tract, and by County.
- College Degree (%) - The percentage of adults (18 years and older) who have a Bachelors degree or above. Available by Blockgroup, by Tract, and by County.
- Age < 18 years (%) - The percentage of population under 18 years old. Available by Blockgroup, by Tract, and by County.
- Homes pre-1950 (%) - The percentage of homes built before 1950. Available by Blockgroup, by Tract, and by County..
- Speak English < Well (%) - Percent Speak English Less Than Well is computed by dividing the sum of population speaking English less than well with the sum of population. Available by Blockgroup, by Tract, and by County.
- Female (%) - Percent Female is computed by dividing the female population by the total population. Available by Blockgroup, by Tract, and by County.
- Rental Units(%) - Percent Rental Units is computed by dividing the number of rented housing units by the total number of occupied housing units. Available by Blockgroup, by Tract, and by County.
- Minority (%) - Percent Minority includes all races except non-Hispanic white persons. Available by Blockgroup, by Tract, and by County.[292]

ACS Questions Utilized (as suggested by the methodology): P3-6, P11, P14, P47-48, H2, H17

Major Uses: In 1994, Executive Order 12898 (EO 12898) mandated that each federal agency make achieving environmental justice part of its mission by identifying and addressing disproportionately high and adverse human health or environmental effects of its programs,

---

[291] See "EJView—Description of Map Features," EPA webpage.
[292] Text source: *Ibid*.

policies, and activities on minority, low-income, tribal, and indigenous populations. EJView facilitates implementation of E.O. 12898 by enabling users to determine the extent to which environment and human health are fairly protected for all people regardless of race, color, national origin, or income, at the community level.[293]

---

[293] EPA, Plan EJ 2014, September 2011.

85. EnviroAtlas

Nature and Purpose of the Dataset: EnviroAtlas is a collection of interactive tools and resources that allows users to explore the many benefits people receive from nature, often referred to as ecosystem services. Though critically important to human well-being, ecosystem services are often overlooked.

EnviroAtlas brings together data on ecology and environmental resources, community infrastructure, populations, economics, and public health to create a more complete picture of the relationships between nature, people, and well-being. Using a landscape approach, EnviroAtlas can help determine the ecological conditions that might be likely in an area, as well as the underlying reasons for those conditions. A landscape approach looks across connected geographic areas and can provide a wealth of information about a location within the context of its surroundings. Its "5,000-foot" view of an area enables an understanding and visualization of important ecological values and patterns of change that may not be evident when looking at smaller, local land areas.[294]

The current version of EnviroAtlas contains an interactive map with over 250 data layers related to ecosystem services, an eco-health relationship browser, analysis tools, and other resources.

At the community level, information currently is available for six communities (Durham, NC; Milwaukee, WI; Phoeniz, AZ; Pittsburgh, PA; Portland, ME; and Tampa, FL).

Responsible Unit: EnviroAtlas is a collaborative project developed by EPA, in cooperation with the US Geological Survey (USGS), the US Department of Agriculture's Natural Resources Conservation Service (NRCS) and Forest Service, and Landscope America.

Authorization: 42 U.S. Code § 4332, National Environmental Policy Act of 1969,

Frequency of Data Publication: In June 2013, EPA released a password-protected beta version of EnviroAtlas. After incorporating feedback, the current version of EnviroAtlas (Ver 1.0) became available online April 2014.

Development plans for EnviroAtlas extend out to 2017, with regular updates taking place as they become available. By 2017, EnviroAtlas will grow to include over 50 communities in the United States.[295] Each community area boundary is based on selected block groups within the 2010 US Census Urban Area boundary.

Timing of Data Release: The EnviroAtlas is being updated and expanded on a rolling basis.[296]

Modes to Access Data: EnviroAtlas is a dynamic effort to provide data and the ability to analyze relationships between ecosystems, the services and benefits they provide, the communities

---

[294] Text source: "EnviroAtlas—About EnviroAtlas," EPA webpage.
[295] A map of the first 23 communities is available.
[296] See "What's the status of EnviroAtlas?," EPA webpage.

that depend on them, and the factors that affect these services. The data can be used in the Interactive Map, downloaded, or accessed through web services.

- EnviroAtlas Interactive Map
- EnviroAtlas web services provide access to data layers through ESRI's ArcGIS Server and Open Geospatial Consortium (OGC) service formats. The services are published in the EPA's GeoPlatform Hosting Environment and include metadata that is discoverable through the EPA's Environmental Data Gateway (EDG) and Data.gov. Each EnviroAtlas web service can be imported into applications (such as ArcGIS Desktop, ArcGIS.com, Bing Maps, or OpenLayers) or into the EPA's GeoPlatform Online mapping interface.
- EnviroAtlas national and community data are available to download as geodatabases.

Methodology: Indicators and information are organized in three sections on the Interactive Map:

- Ecosystem Services and Biodiversity—seven ecosystem services benefit categories.
  - Clean air
  - Clean and plentiful water
  - Natural hazard mitigation
  - Climate stabilization
  - Recreation, culture, and aesthetics
  - Food, fuel, and materials
  - Biodiversity conservation
- Demographics and the Built Environment
- Supplemental Maps—organized into four sections: Boundaries, Biophysical Vector Data, Biophysical Raster Data, and EPA Waters Data.[297]

Indicators are selected and developed for their ability to provide some information about the provision of a particular ecosystem service. Many of the indicators do not describe an ecosystem service by themselves but provide one piece of information of a complex puzzle. Some indicators get much closer to a true ecosystem service than others. The Community data, for example, gets much closer to ecosystem services values than does the national data. This is because models have been developed that can be applied to the Community but not the nation because we do not have all of the necessary data. As ecosystem services science continues to evolve, EnviroAtlas will provide data that gets closer to ecosystem services. In the future, tools will be added to allow the user to take a group of indicators and combine them into an index. Taken collectively, a group of indicators can get much closer to an ecosystem services value than can an individual indicator.

The development of indicators relies on the availability of nationally and locally available data sets which provide inputs to models and calculations. EnviroAtlas supports the development of

---

[297] Text source: "EnviroAtlas Data—Data Organization," EPA webpage.

some of these data sets. Land cover for example is a critical data set that is necessary for the computation of many of the EnviroAtlas data layers. EnviroAtlas supports the development of the National Land Cover Dataset (NLCD), a 30 m resolution product, and is responsible for the development of the high resolution land cover, a 1 m resolution product, being used for the selected Communities. For investigating changes over time, it is important to have land cover data available for multiple time periods. NLCD is produced every 5 years. Other data sets such as stream hydrography, soils, demographics, topography, and economic data are all used in combination with land cover in various models to produce our indicators.[298]

EnviroAtlas communities are selected based on geography, ongoing local research, and the ability to leverage other EPA projects. In general, the community boundaries represent the census block groups within the census urban area for the main community in the area.[299]

Currently, users can upload shapefiles directly into the EnviroAtlas interactive map viewer. We are also developing the functionality to open web services from other sites directly into the map.

Application of ACS Data in Methodology:

- National Demographics: The demographic data for the nation is summarized by census tract, and uses the 2006-2010 American Community Survey data and the 2010 US Census tract boundaries.

  - Total Population
  - Population Density (per square mile)
  - Percent Minority
  - Percent Below Poverty Level
  - Percent Age Less Than 5 Years Old
  - Percent Age Less Than 18 Years Old
  - Percent Population Age Greater Than 64 Years Old
  - Percent Housing Units Built Before 1950
  - Percent 25 Years And Over With Less Than A High School Degree
  - Percent 25 Years And Over With A High School Degree
  - Percent Linguistically Isolated Households
  - Population of American Indian and Alaskan Native
  - Population of American Indian and Alaskan Native Below Poverty

- Community Demographics: For selected communities included in the community component of EnviroAtlas, demographic data is available in block group summaries also uses data from the 2006-2010 American Community Survey data and as well as the 2010 US Census Block Group boundaries.[300]

---

[298] Text source: "EnviroAtlas Data—Data Development," EPA webpage.
[299] Text source: "EnviroAtlas—Frequently Asked Questions," EPA webpage.
[300] Text source: "EnviroAtlas Data—Data Organization," EPA webpage.

- o   Total Population
- o   Population under 1 year old (pdf)
- o   Percent population under 1 year old
- o   Population under 13 years old
- o   Percent population under 13 years old
- o   Population over 70 years old
- o   Percent population over 70 years old
- o   Population other than White, non-Hispanic
- o   Percent population other than White, non-Hispanic
- o   Population with income below twice the US poverty level
- o   Percent population with income below twice the US poverty level[301]

ACS Questions Utilized (as suggested by the methodology): number in household, P4-6, P11, P14, P47-48, H2

Major Uses: Using EnviroAtlas, many types of users can access, view, and analyze diverse information to better understand how various decisions can affect an array of ecological and human health outcomes. The information provided in EnviroAtlas can be used to educate, inform policy and planning decisions, and to support future research in environmental management, planning, public health, and sustainability. The long-term goal for EnviroAtlas is to build decision-support tools into the interactive map that will illustrate the relative relationships between benefits received from ecosystems under existing and alternative conditions.[302]

---

[301] Text source: "EnviroAtlas—Data Fact Sheets," EPA webpage.
[302] Text source: "EnviroAtlas—About EnviroAtlas," EPA webpage.

86. *Residential Construction Trends in America's Metropolitan Regions*

Nature and Purpose of the Measure: Across the United States, many neighborhoods are experiencing dramatic transformations. Parking lots, underused commercial properties, and former industrial sites are being replaced by condominiums, apartments, townhouses, and small-lot single-family homes. These examples of residential infill—building new homes in previously developed areas—can help to create new housing choices, make neighborhoods livelier, increase the tax base, protect rural landscapes, reduce infrastructure costs, and conserve natural resources. Infill can also provide significant environmental benefits when compared with conventional suburban development. Developing more compactly in a location surrounded by existing development means that residents can drive less if they choose, reducing air pollution, and that less paved surface is needed for roads and parking lots, reducing the amount of polluted stormwater runoff flowing into waterways (see Environmental Benefits of Smart Growth for more information).

While examples of successful infill housing projects abound, big questions still remain: Do such examples add up to a fundamental shift in the geography of residential construction? Is infill housing construction on the rise? In which metropolitan regions is the shift to infill most significant? EPA explored these questions in a series of reports, *Residential Construction Trends in America's Metropolitan Regions*. The 2012 report compares the location of new homes to data about preexisting land cover to determine where infill development was taking place in 209 U.S. metropolitan regions between 2000 and 2009. [303]

Responsible Unit: Office of Sustainable Communities, EPA

Authorization: National Environmental Policy Act of 1969, Sections 102 and 4332

Frequency of Data Publication: released in 2009, 2010, and 2012.

Timing of Data Release: NA

Modes to Access Data: Residential Construction Trends in America's Metropolitan Regions, EPA webpage (pdf)

Methodology: This study combines data about the location of new housing construction with data about land cover and protected land uses in order to calculate the total percentage of all new home construction that is infill—or built in previously developed areas. It compares infill housing construction trends in all metropolitan regions with populations greater than 200,000 in the conterminous United States (excluding Alaska and Hawaii). There were three main steps in this study. (1) Determine the location of new home construction. (2) Determine the location of infill areas. (3) Calculate for each metropolitan region the percentage of all new housing construction that is infill.

---

[303] Text source: EPA, *Residential Construction Trends in America's Metropolitan Regions: 2012 Edition*, December 2012, p. iii.

Step 1: Determine the location of new home construction. Data on the precise geographic location of new home construction are not nationally available. However, the 2009 American Community Survey (ACS) 5-Year Estimates provide housing unit counts for each census block group, broken down by the time period of construction. This study obtained data regarding homes built during two different time periods, 2000 to 2004 and 2005 to 2009, for all census block groups within the 209 metropolitan areas included in the study.

Step 2: Determine the location of infill areas. This study analyzed land cover data to classify the development status of census block groups in 2001. All new homes built in block groups classified as Fully Developed are considered to be infill. This study defines infill housing as new homes built in census block groups that were fully developed in 2001.

Step 3: Calculate the percentage of all new housing construction that is infill. The final step in this analysis is calculating—for each metropolitan region—the percentage of all new housing units that are infill (i.e., built in block groups that are classified as Fully Developed). All other housing construction, by definition, is not infill. This definition relies on one key assumption: if a block group is Not Fully Developed, then all additional housing units will be built in undeveloped land areas that are available for new development.

This study measured infill housing construction in 209 metropolitan regions across the contiguous United States. In summary, each calculation involved the following three steps:

- Total homes: Sum of housing units built in all block groups located inside the metropolitan statistical area.
- Infill homes: Sum of housing units built in Fully Developed block groups located inside the MSA.
- Percentage of all homes that are infill: Infill homes divided by total homes.

This study calculated these metrics for both time periods (2000 to 2004 and 2005 to 2009) as well as the entire study period (2000 to 2009).[304]

Application of ACS Data in Methodology: Five-year ACS data are used to identify new home construction.

ACS Questions Utilized (as suggested by the methodology): H1-2

Major Uses: The report is used by national and local organizations to assess the impact of new home construction on land cover and ascertain the implications of the findings for smart growth policies and programs.

---

[304] Text source: *Ibid*, pp. 24-32.

87. [Smart Location Database](#)

Nature and Purpose of the Measure: The Environmental Protection Agency's (EPA) Smart Location Database (SLD) was developed to address the growing demand for data products and tools that consistently compare the location efficiency of various places. The SLD summarizes several demographic, employment, and built environment variables for every Census block group (CBG) in the United States. The attributes measured serve as indicators of the commonly cited "D" variables that have been shown in the transportation research literature to be related to travel behavior. The Ds include concepts such as residential and employment density, land use diversity, design of the built environment, access to destinations, and distance to transit.[305]

During the past two decades, the planning profession has seen an explosion of interest in the roles that land use and urban design play in shaping the transportation habits, health, and livelihood of urban and suburban residents. Researchers in the fields of transportation planning and public health have begun to isolate and measure the relationships between the built environment in which we live and work and our propensity to choose walking, transit, or driving to meet our everyday transportation needs. These studies tend to focus on neighborhood characteristics such as the density of development, mixing of land uses, connectivity of street networks, availability of transit, and accessibility to destinations via car, transit, or foot.

These kinds of studies enable planners and community advocates to quantify the potential benefits of local land use decisions such as encouraging compact and mixed-use development, allowing for more jobs and housing to be in walkable and transit-rich neighborhoods, and reducing the amount of new low-density development occurring at the outer suburban fringe.

Developing data that summarize built environment characteristics unfortunately can be expensive and time consuming. Moreover, each time a new community wants to conduct a planning study, the same general kinds of data must be identified, gathered, and processed. We wondered, therefore, if an economy of scale could be achieved by developing data about the built environment at the block group scale for the entire United States. These data would necessarily rely on national sources or widely used data standards.

Therefore, the results could be inferior to locally derived metrics that rely on detailed land use data available only at the local scale. We hypothesized, however, that a nationwide study could produce data that are sufficient for many local and regional studies that would not otherwise move forward because they lack resources. We also hypothesized that making nationally consistent data freely available could spur the development of third party planning analysis tools that significantly reduce barriers to entry for communities seeking to analyze the potential effects of land use decisions.

The U.S. Environmental Protection Agency's (EPA's) Smart Location Database (SLD) includes more than 90 variables summarizing conditions for every census block group in the United

---

[305] Text source: Kevin Ramsey and Alexander Bell, ["Smart Location Database: Version 2.0 User Guide,"](#) March 14, 2014, p. 2.

States. It is broken into 10 topic areas: administrative, area, demographics, employment, density, diversity (of land use), design, transit service, destination accessibility via automobile and via transit, and regional summaries.[306]

Responsible Unit: Office of Sustainable Communities, EPA

Authorization: Clean Water Act (CWA), Section 104(b)(3) and Clean Air Act (CAA), Section 104(b)(3)

Frequency of Data Publication: The SLD was first released in early 2012. Version 2.0 was released in March 2014. EPA hopes to update this database regularly—at least every decennial census. Such plans are contingent on the continued availability of funding, however.

Timing of Data Release: NA

Modes to Access Data: The SLD is a free resource available to the public for download, web service, or viewing online. Options are described below:

- Download: The SLD can be downloaded as a single file geodatabase at EPA's Environmental Dataset Gateway. Users who only wish to download data for a single state, metro region, or locality can use EPA's Clip and Ship tool.
- Web service: The SLD is available as an Esri mapping service, REST, SOAP, WMS, and KML. See the SLD web service for details.
- Viewing online: Several variables from the SLD are available for viewing online.[307]

Methodology: The Smart Location Database (SLD) summarizes over 90 demographic, built environment, transit service, and destination accessibility attributes for every census block group in the United States. The SLD reflects housing, population, and employment conditions in 2010, street network conditions in 2011, and transit-service conditions in late 2012.

Data sources used to develop the SLD include several Census datasets (TIGER/Line, 2010 summary file 1, American Community Survey, and Longitudinal Employer-Household Dynamics), NAVTEQ highway/streets and parks data, Protected Areas Database of the United States (PAD-US), fixed-guideway transit station locations from the TOD Database, and local transit service data shared in the General Transit Feed Specification (GTFS).[308]

Planned updates to the SLD will summarize these environment and accessibility characteristics in the form of a Smart Location Index that represents a block group's relative location efficiency

---

[306] Kevin Ramsey and Alexander Bell, "The Smart Location Database: A Nationwide Data Resource Characterizing the Built Environment and Destination Accessibility at the Neighborhood Scale," *Cityscape: A Journal of Policy Development and Research*, Volume 16, Number 2, 2014, pp. 146-147.

[307] Text source: Kevin Ramsey and Alexander Bell, "Smart Location Database: Version 2.0 User Guide," March 14, 2014, p. 3.

[308] Text source: *Ibid*, p. 9.

when compared to other block groups within the same metropolitan region. EPA also plans to periodically update attributes and add new attributes to reflect latest available data. [309]

Application of ACS Data in Methodology: ACS data were used to calculate the following data elements:

- Number of households in CBG that own zero automobiles
- Percent of zero-car households in CBG
- Number of households in CBG that own one automobile
- Percent of one-car households in CBG
- Number of households in CBG that own two or more automobiles
- Percent of two-plus-car households in CBG

ACS Questions Utilized (as suggested by the methodology): H12

Major Uses: The Smart Location Database (SLD) enables the consistent comparison of different locations across the U.S. in terms of their land use characteristics, infrastructure, and accessibility to nearby destinations. SLD attributes were selected for their utility in travel demand modeling studies. However they may also be relevant to land use scenario planning, community development, and public health studies, particularly when obtaining local data is not possible or practical.

- Developing indicators of location efficiency—EPA is using the Smart Location Database to develop simple indicators of location efficiency. For instance, EPA is working with the U.S. General Services Administration to develop a Smart Location Index that scores census block groups based on their built environment and accessibility characteristics. Block groups that are associated with reduced vehicle miles traveled receive higher scores compared to other block groups within the same metropolitan region. Advanced users could create similar composite indicators to compare walk ability, compact design, or other built environment characteristics. EPA hopes to include additional indicators of location efficiency in forthcoming updates to the Smart Location Database.
- Scenario planning and travel demand modeling—Planners can use the Smart Location Database as baseline data in scenario planning, sketch planning, and travel demand studies when more detailed or consistent local data are unavailable. Analysts can also use elasticities found in the research literature to adjust outputs of travel or activity models that are otherwise insensitive to variation in the built environment.
- Conducting nationwide research studies and developing tools—Building on previous research, EPA is conducting a nationwide modeling study to predict employee commute travel (e.g., average trip distances, mode share, vehicle miles traveled, etc.) based on workplace neighborhood characteristics. If successful, this study will

---

[309] Text source: "Smart Location Database of Census Block Groups in the State of Florida – 2013," Federal Geographic Data Committee webpage.

produce equations for refining travel demand models. This study and others like it also make it possible to create simple online tools to help more communities analyze the potential outcomes of proposed land use development.

- Comparing urban form among metropolitan regions—Researchers can use the Smart Location Database in nationwide studies that compare metropolitan regions based on urban form characteristics. For instance, analysts could determine the percentage of residents that live in walkable or transit-rich neighborhoods. EPA's 2012 study Residential Construction Trends in America's Metropolitan Regions used the Smart Location Database in conjunction with data from the National Land Cover Database and the American Community Survey to measure and compare infill housing development.

- Modeling impervious surface growth—EPA analyzed variables in a previous version of the Smart Location Database and the National Land Cover Database to create a model and simple spreadsheet tool for estimating new impervious surface growth associated with land use development scenarios. This model is sensitive not only to density of development but also to its relative centrality within the surrounding metropolitan region. For details, see EPA's Impervious Surface Growth Model.[310]

---

[310] *Ibid*.

88. *[Human Well-Being Index](#)*

<u>Nature and Purpose of the Measure</u>: The U.S. Environmental Protection Agency (EPA) has developed a holistic approach for characterizing human well-being at multiple scales: A Human Well-being Index (HWBI). The HWBI uses a substantial suite of measures to evaluate the influence of social, economic and environmental service flows on components of human well-being in an integrated fashion based on eight domains of well-being (Figure 1) applicable to communities of all scales whether national, regional or local.

Figure 1 represents the conceptual model underlying the HWBI and depicts the relationships of natural and built capital, goods and services (including social and economic inequities), the domains of well-being as well as its sub-elements, and the value system of the entity being examined (i.e., relative importance values). When tracked over time, the index has the potential to serve as a measure of sustainability as a function of human well-being and may be linked to alternative decisions that change the ecological, economic, and social states of defined populations. The metrics and methodologies for constructing multiple-scale measures have been developed and, using this approach, a U.S. HWBI is calculated along with the functional relationships to selected economic, societal and environmental services flows. The HWBI is distinct from other well-being indices in that the approach is scalable and calculated domain and HWBI values are responsive to changes in select economic, social and ecological services, making it well suited as an informative endpoint in the sustainability decision-making process.

The HWBI is designed to inform and empower decision makers to equitably weigh and integrate human health, socio-economic, environmental, and ecological factors to foster sustainability in the built and natural environments; thereby, helping local decision makers understand the effects on sustainability of alternative policies and actions.

In 2011, the Sustainable and Healthy Communities Research program (SHC) in EPA's Office of Research and Development (ORD) coined the term TRIO for Total Resource Impacts and Outcomes. The concept of TRIO encompasses any number of holistic community decision-making approaches that address all three pillars of sustainability—economic, societal and environmental.

A TRIO approach would evaluate both tangible costs (e.g., capital investment, tax revenue, permitting) and less tangible "costs" (e.g., community service disruption, loss of natural services) that may impact quality of life. This expanded assessment process would be accomplished across all three pillars of sustainability in parallel to help identify not only the expected attributes of the decision but also the unintended consequences of all decision options. SHC determined the need to adopt or develop an approach or index of human well-being that fully embraced the TRIO aspects of the developing research program. The Human Well-Being Index (HWBI) is the culmination of this research effort.

<u>Responsible Unit</u>: [Sustainable and Healthy Communities Research Program](#) (SHC), Office of Research and Development (ORD), Environmental Protection Agency (EPA)

Figure 1. Conceptual framework for evaluating the influence of service flows on well-being endpoints for the construction of a Human Well-Being Index.
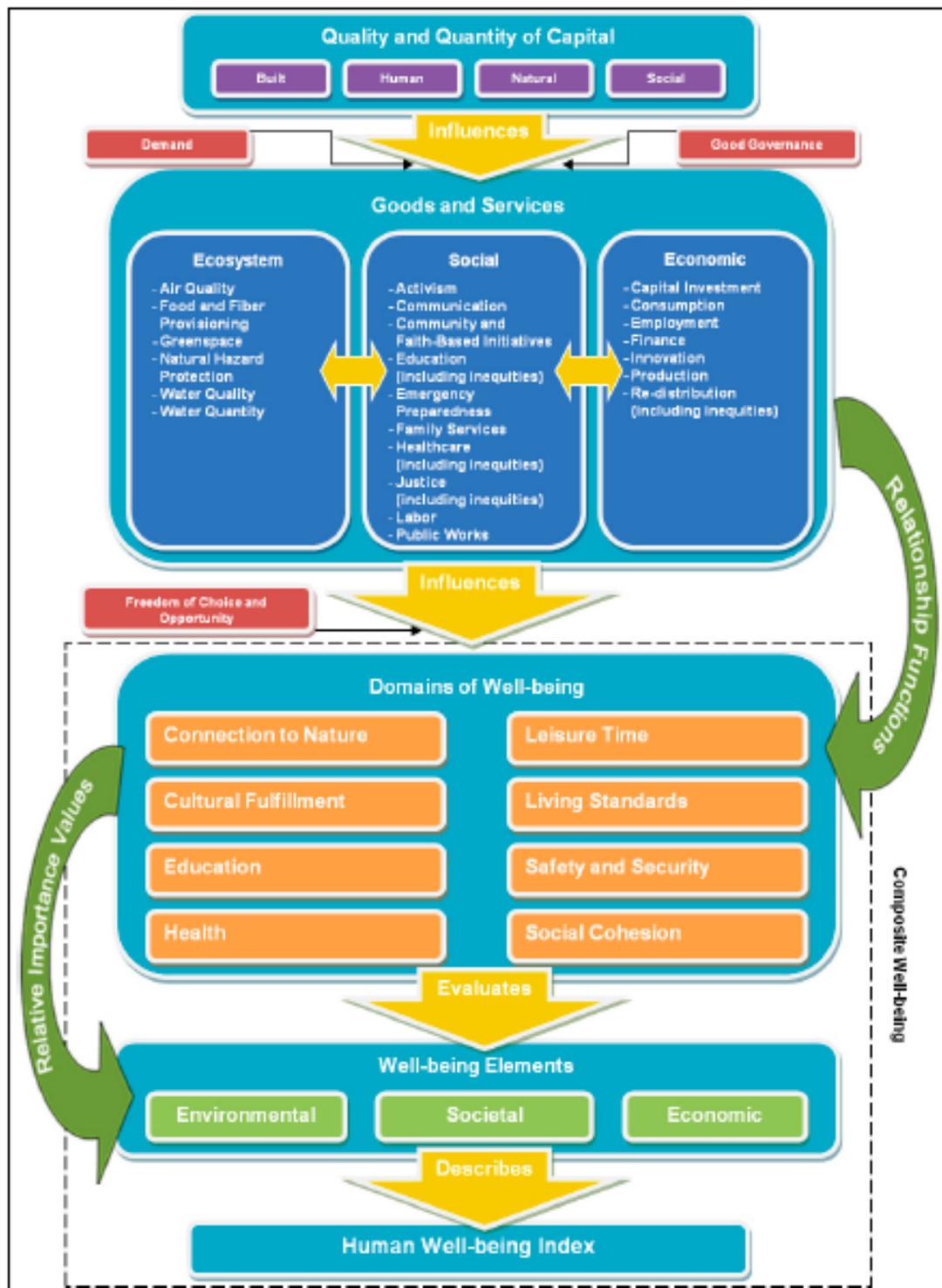
<u>Authorization</u>: 42 U.S. Code § 4332, National Environmental Policy Act

<u>Frequency of Data Publication</u>: The first HWBI was published in Fall 2014.[311]

<u>Timing of Data Release</u>: NA

<u>Modes to Access Data</u>: J. Kevin Summers, Lisa M. Smith, Linda C. Harwell, Jason L. Case, Christina M. Wade, Kendra R. Straub, and Heather M. Smith, "An Index of Human Well-Being for the U.S.: A TRIO Approach," Sustainability 2014, 6, 3915-3935 (pdf)

<u>Methodology</u>: The HWBI is an index based on a combination of objective data, subjective data collected through surveys, available data from other well-being surveys, and combined at the smallest spatial scale generally available (most often county level data). The conceptual model depicted in Figure 1 shows the interplay of goods and services and their influence on the eight domains of well-being used in HWBI, the use of relative importance values to describe the community value structures, and the combination of all information into the three well-being elements and their subsequent combination into a single value representing well-being. The index utilizes 8 domains, 25 indicators and 79 metrics (Table 2).

Objective and subjective data were collected from various publically accessible sources and organized hierarchically by spatial and temporal resolution (e.g., national, regional, state, and county by year) for the years 2000–2010 to populate these metrics. When multiple spatial scales existed for a metric, the finest scale (e.g., county versus state) was selected for processing. Data source determination was primarily driven by temporal and spatial coverage of offered data. To the extent possible, factors such as data reliability and credibility, historic data continuity, and future data accessibility were consider in the data selection process.

Data gaps caused by spatial and temporal disparities found among data sources were filled using a carry-forward substitution imputation technique using cross-year county or within year state or regional data. A secondary imputation was accomplished in an effort to calculate imputed values for counties exhibiting similar characteristics. From the spatially and temporally complete data set, county groupings were created using a combination of the Rural-Urban Continuum Code (RUCC) classifications and the Gini Index for Household Income Inequality (HII) quintiles (derived from the American Community Survey). The RUCC-HII permutations generated county data groupings that generally reflected the relative spatial relationship of a county to the nearest large urban center and its measured income dispersion. Within-year median values were calculated for each RUCC-HII banding. Missing values in the original aggregate of metric data were substituted with the resulting RUCC-HII banding values.

---

[311] J. Kevin Summers, Lisa M. Smith, Linda C. Harwell, Jason L. Case, Christina M. Wade, Kendra R. Straub, and Heather M. Smith, "An Index of Human Well-Being for the U.S.: A TRIO Approach," Sustainability 2014, 6, 3915-3935.

Table 2. Indicators and metrics associated with each of the eight domains used to characterize human well-being and calculate the index.

Domain	Indicator	Metrics
Connection to Nature	Biophilia	Connection to Life
		Spiritual Fulfillment
Cultural Fulfillment	Activity Participation	Performing Arts Attendance
		Rate of Congregational Adherence
Education	Basic Educational Knowledge and Skills of Youth	Mathematics Skills
		Reading Skills
		Science Skills
	Participation and Attainment	Adult Literacy
		High School Completion
		Participation
		Post-Secondary Attainment
	Social, Emotional and Developmental Aspects	Bullying
		Contextual Factors
		Physical Health
		Social Relationships and Emotional Well-being
Health	Healthcare	Population with a Regular Family Doctor
		Satisfaction with Healthcare
	Life Expectancy and Mortality	Asthma Mortality
		Cancer Mortality
		Diabetes Mortality
		Heart Disease Mortality
		Infant Mortality
		Life Expectancy
		Suicide Mortality
	Lifestyle and Behavior	Alcohol Consumption
		Healthy Behaviors Index
		Teen Pregnancy
		Teen Smoking Rate
	Personal Well-being	Happiness
		Life Satisfaction
		Perceived Health
	Physical and Mental Health Conditions	Adult Asthma Prevalence
		Cancer Prevalence
		Childhood Asthma Prevalence
		Coronary Heart Disease Prevalence
		Depression Prevalence
		Diabetes Prevalence
		Heart Attack Prevalence
		Obesity Prevalence
		Stroke Prevalence

Domain	Indicator	Metrics
Leisure Time	Activity Participation	Average Nights on Vacation
		Physical Activity
	Time Spent	Leisure Activities
	Working Age Adults	Adults who Provide Care to Seniors
		Adults Working Long Hours
		Adults Working Standard Hours
Living Standards	Basic Necessities	Food Security
		Housing Affordability
	Income	Incidence of Low Income
		Median Household Income
		Persistence of Low Income
	Wealth	Median Home Value
		Mortgage Debt
	Work	Job Quality
		Job Satisfaction
Safety and Security	Actual Safety	Accidental Morbidity and Mortality
		Loss of Human Life
		Property Crime
		Violent Crime
	Perceived Safety	Community Safety
	Risk	Social Vulnerability Index
Social Cohesion	Attitude toward Others and the Community	Belonging to Community
		City Satisfaction
		Discrimination
		Helping Others
		Trust
	Democratic Engagement	Interest in Politics
		Registered Voters
		Satisfaction with Democracy
		Trust in Government
		Voice in Government Decisions
		Voter Turnout
	Family Bonding	Exceeded Screen Time Guidelines
		Frequency of Meals at Home
		Parent-child Reading Activities
	Social Engagement	Participation in Group Activities
		Participation in Organized, Extracurricular Activities
		Volunteering
	Social Support	Close Friends and Family
		Emotional Support

Box-and-whisker analyses were completed for each fully enumerated HWBI metric. Extreme lower and upper outlier measures were set to minimum and maximum values, respectively. The maximum values were calculated to be three times the 75% percentile for each metric and the minimum values were calculated as minus three times the 25% percentile. Any outliers of this three times far-fence technique were set to the metric value closest to the fence. All data were

standardized on a scale from 0.1 to 0.9 following the Organization for Economic Co-Operation and Development's (OECD) Better Life Index approach with minor modification to account for the difference in scale. The resulting HWBI metric data set included both imputed and non-imputed standardized data for 3143 counties of the U.S that represented greater than 2.7 million data points, collectively.

The HWBI was derived from indicator scores calculated as the population weighted average of the standardized metric values. Indicator scores were averaged for each domain score. Finally, a scaled geometric mean was calculated across domain scores to produce the final the HWBI. Higher HWBI scores indicate greater levels of well-being. Both the hierarchical organization of the metric data and the step-wise calculation process provided the means for examining well-being and its constituents at multiple scales from the national level down to individual counties.[312]

The metrics chosen here for index development reflect measures of the human condition as opposed to the quality and quantity of goods and services supporting society. The metrics describing service flows will be used to model well-being as an endpoint measure in a predictive modeling framework. Therefore, the HWBI described in this document is an "ends" measure separated from the "means". By doing so, we can ultimately develop alternate scenarios for decision support tools for managers and policy makers. Information quantifying the delivery of social and economic services, and ongoing research within SHCRP seeking to measure ecosystem functions and quantify goods and services provisioning will provide information for model input. Modeling efforts will involve Relative Importance Values (RIVs). RIVs will also be used to link each service (ecosystem, social and economic) to each well-being domain by establishing their subjective importance.

Data collected by the following institutions and organizations was used in our index (* most used data sources):

- Centers for Disease Control and Prevention (CDC) *
- U.S. Census Bureau *
- General Social Survey (GSS)
- Gallup, Inc. (Gallup Brain, Gallup Healthways)
- Bureau of Labor Statistics (BLS)
- Bureau of Economic Analysis (BEA)
- U.S. Department of Health and Human Services (HHS)
- Federal Bureau of Investigation (FBI)
- American National Election Study (ANES)
- National Center for Education Statistics (NCES)
- Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS)
- National Oceanic and Atmospheric Administration (NOAA)
- Association of Religion Data Archives (ARDA)

---

[312] *Ibid*.

- University of South Carolina Hazards and Vulnerability Research Institute (HVRI)

These data sources were chosen based on the following criteria:

- Availability and access: The data were publicly available and easy to understand, access and extract.
- Reliability and data credibility: The sources collected data in a manner that was vetted by the professional community and had metadata available for review.
- Spatial preference: County-level data were the lowest geospatial level preferred, and could be rolled into larger scales as needed. In the absence of county-level data, or when it was not feasible to pull county-level data (i.e., data only available from local governmental sites; lack of compiled data from a single source), state, regional, and national-level data were used.
- Coverage: The data were available for a large portion of the United States.
- Chronological history and the likelihood that the data will continue to be collected: Data had a good history of collection or consistent collection. The goal was to initially create a time series beginning with the year 2000 and continuing through 2010; however, if the data were not available from a single data source for all years, other sources containing similar measurements were used to complete the time series.
- Subjective and objective data: Both subjective and objective data were included.[313]

Application of ACS Data in Methodology: Data from the American Community Survey are used in the construction of the HWBI in two ways:

- Construction of indicators for high school completion by age, postsecondary attainment by age, median value of owner-occupied housing units, owner-occupied housing units without a second mortgage or home equity loan, and monthly owner costs as a percentage of household income.[314]
- Imputation of missing data for counties on the basis, in part, of five-year household income inequality by county as measured in an ACS data brief.[315]

ACS Questions Utilized (as suggested by the methodology): P4, P11, P47-48, H17, H19, H22-23

Major Uses: The index offers a necessary measure of the influence of policies and services (environmental, economic and social) on aspects of social welfare and overall human well-being. These integrated concepts of the interactions of social, economic and environmental drivers allow a better understanding of the human condition and its collective relationship to

---

[313] Text source: U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory (NHEERL), Gulf Ecology Division, "Indicators and Methods for Constructing a U.S. Human Well-being Index (HWBI) for Ecosystem Services Research," March 2012.

[314] *Ibid*. This 2012 list may have been revised by the time of the 2014 article (footnotes 215 and 216), which does not identify the data source of each metric.

[315] Adam Bee, "Household Income Inequality within U.S. Counties: 2006–2010," Census Bureau, ACSBR/10-18, February 2012.

service flows and, thus, will permit decision makers to examine the impact of specific decision alternatives on the well-being of their constituencies. Coupling this type of decision scenario testing with specific targets of social equity and intergenerational equity should also permit selected decisions to create more sustainable conditions for communities.

As stated in earlier publications, the primary reason for the development of the HWBI is to include explicit connections between human well-being and environmental drivers and services. Increasingly, communities across the U.S. are examining the management of growth through sustainable development. The HWBI approach allows the U.S., states, counties and communities to assess the impact of decisions on the sustained well-being of their constituencies (e.g., effects of economic decisions, both intended and unintended, on social and environmental well-being).

Additionally, the HWBI allows these governmental entities to assess not only the direct impacts of decisions (e.g., effects of economic decisions on jobs) but also to assess the indirect impacts (unintended consequences) of these decisions (e.g., economic decisions on social and environmental issues). Many earlier indices focused on the point where human well-being and environmental conditions intersected rather than how they related. The HWBI represents a critical advancement in this area by emphasizing the symbiotic relationships between nature, humans and economies. Our development of the HWBI provides a significant step forward in a community's (or larger spatial entity's) ability to assess the short- and long-term impacts of potential decision alternatives on the well-being of their constituencies.[316]

---

[316] Summers et al., *op. cit.*

B. Equal Employment Opportunity Commission

89. *Equal Employment Opportunity Tabulation*

Nature and Purpose of the Dataset: The Equal Employment Opportunity (EEO) Tabulation serves as the primary external benchmark for comparing the race, ethnicity, and sex composition of an organization's internal workforce, and the analogous external labor market, within a specified geography and job category.[317]

The most recent tabulation was developed from 5-year American Community Survey data from 2006-2010. In November 2012, the EEO Tabulation 2006-2010 (5-year ACS data), a tabulation comparable to the 2000 Special EEO Tabulation was published, containing 488 detailed Census occupation categories based on the 2010 Standard Occupational Classification (SOC). This tabulation included estimates and percentages of the labor force for race and ethnicity by sex for all counties and for places of 50,000 or more, covering nearly 6,500 geographic entities (the nation, states, metro/micro areas, counties, and places). Characteristics included sex, race, and Hispanic origin, cross-tabulated by occupation, industry, age, educational attainment, and earnings. Two new measures were added: citizenship and unemployment status. Margins of error based on a 90-percent confidence level were pre-calculated and published. In addition, disclosure avoidance and rounding rules were applied.[318]

The EEO Tabulation is about one-and-a-half times larger than the 5-year American Community Survey release.

Responsible Unit: The Census Bureau produces this tabulation for four sponsoring Federal agencies: Equal Employment Opportunity Commission, Department of Justice, the Office of Federal Contract Compliance Programs (Department of Labor), and Office of Personnel Management.

Authorization: Age Discrimination in Employment Act (29 U.S.C. 623(a) - (d), 633(a); 29 CFR 1625.7(d)) and Civil Rights Act of 1964 (42 U.S.C. 2000e-2; 29 CFR 1607.4(B)). Also, see Executive Order 11246, as amended Sep. 24, 1965.

Frequency of Data Publication: An EEO Tabulation was released after each decennial census between 1970 and 2000. One tabulation has been produced on the basis of the 2006-2010 ACS. The sponsors of the tabulation have not yet determined the time of the next release.

Timing of Data Release: The ACS-based tabulation for 2006-2010 was released in November 2012.

---

[317] Text source: "Equal Employment Opportunity (EEO) Tabulation: Main," Census Bureau website.

[318] Text source: "About the Equal Employment Opportunity Tabulation," Census Bureau website.

Modes to Access Data: Census Bureau websites:[319]

- [EEO Tabulation](#) (xls, pdf).
- [American FactFinder](#) (html, xls, rtf, pdf)
- [EEO Parent Directory](#) (ftp site)

Methodology/Application of ACS Data in Methodology: The EEO Tabulation 2006-2010 (5-year ACS data) contains information similar to comparable tabulations from the 1970, 1980, 1990 and 2000 censuses. Characteristics include sex, race, and Hispanic origin, cross-tabulated by citizenship, occupation, industry, age, educational attainment, earnings, and unemployment status. The 2010 Census occupation categories used in this file are based on the 2010 Standard Occupational Classification (SOC) categories. Tables are tabulated for at least one of three geographic types: residence, worksite, and residence-to-worksite commuting flows. These are tabulated for one or more of the following geographic summary levels: U.S. Total, all states, the District of Columbia, and Puerto Rico, metro and micro areas, counties, places, and EEO county sets. All tables also include estimates, percentages, and margins of error.[320]

ACS Questions Utilized (as suggested by the methodology): P3-8, P11, P29-30, P35-38, P43-47

Major Uses: The EEO Tabulation serves as the primary external benchmark for comparing the race, ethnicity, and sex composition of an organization's internal workforce, and the analogous external labor market, within a specified geography and job category. It also is used by organizations to develop and update their affirmative action plans.[321]

---

[319] For additional information, see Census Bureau, "American FactFinder (AFF) and The Equal Employment Opportunity Tabulation 2006-2010: Tutorial," presentation, October 2014.
[320] Text source: "Equal Employment Opportunity (EEO) Tabulation: Main," Census Bureau website.
[321] *Ibid.*

C. Federal Financial Institutions Examination Council

90. *FFIEC Online Census Data System*

Nature and Purpose of the Measure: The Federal Financial Institutions Examination Council (FFIEC) is a formal interagency body empowered to prescribe uniform principles, standards, and report forms for the federal examination of financial institutions by the Board of Governors of the Federal Reserve System (FRB), the Federal Deposit Insurance Corporation (FDIC), the National Credit Union Administration (NCUA), the Office of the Comptroller of the Currency (OCC), and the Consumer Financial Protection Bureau (CFPB) and to make recommendations to promote uniformity in the supervision of financial institutions.

The FFIEC Online Census Data System (formerly named FFIEC Census Reports) is an online tool that can be used to access FFIEC census data by Metropolitan Statistical Area (MSA) / Metropolitan Division, county, and census tract. The system also provides data for non-MSA areas, counties, and census tracts. There are four report options – Census Demographic Data Summary, Census Income Data Summary, Census Population Data Summary, and Census Housing Data Summary.[322]

The FFIEC Census File now uses the 2006-2010 American Community Survey (ACS) for the majority of its demographic fields, with selected 2010 Census Summary File 1 tables used for population related fields.

The FFIEC publishes census data as part of the Home Mortgage Disclosure Act (HMDA) aggregate and disclosure data and as a separate census data file, both of which are available on the FFIEC website. The census data are used to provide context to HMDA and Community Reinvestment Act (CRA) data.[323]

Responsible Unit: Federal Financial Institutions Examination Council

Authorization: Home Mortgage Disclosure Act (PL 94-200), Section 310 (a); 12 USC 2809(a).This section notes that the "Federal Financial Institutions Examination Council shall compile each year, for each primary metropolitan statistical area, metropolitan statistical area, or consolidated metropolitan statistical area that is not comprised of designated primary metropolitan statistical areas, aggregate data by census tract for all depository institutions which are required to disclose data under section 304 or which are exempt pursuant to section 306(b). The Council shall also produce tables indicating, for each primary metropolitan statistical area, metropolitan statistical area, or consolidated metropolitan statistical area that is not comprised of designated primary metropolitan statistical areas, aggregate lending patterns for various categories of census tracts grouped according to location, age of housing stock, income level, and racial characteristics."

---

[322] Text source: "FFIEC Online Census Data System," FFIEC website.
[323] Text source: "Guidance on the Use of 2010 Census Data in Fair Lending Examinations," Federal Reserve website

Frequency of Data Publication: Updated annually. FFIEC relied on Census 2000 date through 2011 and shifted to reliance on the ACS in 2012.[324]

Timing of Data Release: Not applicable

Modes to Access Data: From the FFIEC website:

- FFIEC Online Census Data System (html, pdf)
- FFIEC Census Windows Application (xls, pdf, text)

Methodology/Application of ACS Data in Methodology: The FFIEC Online Census Data System relies on the ACS in two ways.

First, its baseline data come from the five-year ACS (2006-2010 at present). Historically, the census data "base" file was developed by the FFIEC every 10 years and was updated annually. Annual updates were made to reflect changes to MSA/MD boundaries announced by the Office of Management and Budget (OMB), median family income estimates and CRA distressed/underserved tracts as announced by the federal banking regulatory agencies.[325] Because of the changes in the collection of the 2010 census data, the FFIEC made corresponding changes to its data files. In October 2011, the FFIEC announced the incorporation of 2010 ACS data into the FFIEC-published census data file.

Currently, the FFIEC uses the 2010 ACS five-year estimate data to create a base file. The FFIEC will update the census information in the base file every five years. Implementation of the new data for consumer compliance and CRA examination purposes was in 2012, and the data will be utilized in the same manner that decennial data has been used in the past. In addition to the tract income data, the new base file will include updated race and ethnicity data. Those data will also be updated every five years going forward.

Although the ACS data are published annually as a rolling average, the FFIEC decided against publishing and using annual census data. Financial institutions and agency examination staff use census data when planning their activities and evaluating performance. The data used to determine low- and moderate-income census tracts for CRA examination purposes could change slightly from year to year, creating additional burden for financial institutions and agency examination staff, and increasing the anticipated time, cost, and complexity involved in using annual data. The FFIEC believes the modest changes in the annual data would provide only a limited benefit. Although the 2010 census data have resulted in many new census tracts as well as redefinition of some existing tracts, our studies have shown that the data generally remain the same for the majority of census tracts from year to year or over a multi-year time period. The five-year approach is seen as a significant improvement over the current 10-year data updates and will minimize the confusion and errors that might occur if the data were updated more frequently. Additionally, updating the data every five years should be sufficient

---

[324] "FFIEC Census Information Sheets," FFIEC website.
[325] Text source: "Guidance on the Use of 2010 Census Data in Fair Lending Examinations," Federal Reserve website.

to provide an accurate assessment of an institution's lending in low- and moderate-income geographies. [326]

The FFIEC's second use of ACS data is to annually update its estimate of median family income (MFI) for each metropolitan area.[327] The FFIEC annually adjusts the MFI of each census tract on the basis of the percent increase in its annual MFI estimate for each metropolitan area. The metro MFI estimate is taken from HUD's calculation for all metro areas with boundaries unchanged by OMB in 2013 and based on FFIEC's use of ACS estimates for all metro areas with boundaries that were changed by OMB in 2013.[328] (For FY2014, HUD has not adopted the new Metropolitan Statistical Area/Metropolitan Division (MSA/MD) delineations released by OMB).

ACS Questions Utilized (as suggested by the methodology): P5-6, P47-48, H1-2, H17[329]

Major Uses: Financial institutions and agency examination staff use the census data system to ensure compliance with HMDA and CRA requirements.

---

[326] Text source: "Press Release: FFIEC Announces the Use of American Community Survey Data In Its Census Data Files," FFIEC, October 19, 2011.

[327] See "FFIEC Median Family Income Report," FFIEC website.

[328] FFIEC, "Census File For Use With HMDA/CRA 2010 Demographics," May 2014.

[329] See "2010 Based Census Information (Years 2012 and Forward)," FFIEC website.

D. **National Science Board**

91. *Science and Engineering Indicators*

Nature and Purpose of the Measure: The *Science and Engineering Indicators* series is designed to provide a broad base of quantitative information about U.S. science, engineering, and technology for use by policymakers, researchers, and the general public. Each volume contains analyses of key aspects of the scope, quality, and vitality of the Nation's science and engineering enterprise in the context of global science and technology.

The report presents information on science, technology, engineering, and mathematics education at all levels; the scientific and engineering workforce; U.S. and international research and development performance; U.S. competitiveness in high technology; and public attitudes and understanding of science and engineering. A chapter on state-level science and engineering enables state comparisons on selected indicators. An Overview chapter synthesizes selected key themes emerging from the report.[330]

Responsible Unit: National Science Board, with the National Center for Science and Engineering Statistics (NCSES), National Science Foundation

Authorization: 42 U.S.C. § 1863 (j) (I), "The Board shall render to the President and the Congress no later than January 15 of each even numbered year, a report on indicators of the state of science and engineering in the United States."

Frequency of Data Publication: Biennial, in even-numbered years.

Timing of Data Release: NA

Modes to Access Data: *Science and Engineering Indicators*, National Science Board (html, xls)

Methodology: The National Science Board has selected 41 S&E indicators for inclusion in this digest. These indicators have been grouped into seven themes. Although each stands alone, collectively these seven themes are a snapshot of U.S. R&D capacity and outputs in the context of global trends affecting them. Exploration of areas that indicate capacity for innovation is a thread common to many of the themes presented here. As economies worldwide grow increasingly knowledge-intensive and interdependent, capacity for innovation becomes ever more critical.

Three themes provide a worldwide view, picturing R&D spending, research outputs, and STEM education. Four others share a domestic focus, providing information on U.S. R&D funding and performance, federal R&D support, the U.S. S&E workforce, and public research universities.

---

[330] Text source: National Science Board, "Letter of Transmittal," *Science and Engineering Indicators 2014*, NSB 14-01, February 2014.

Indicators may vary in successive volumes of the Science and Engineering Indicators series as different S&E policy issues emerge.[331]

Four criteria guide the selection of data for SEI:

- Representativeness. Data should represent national or international populations of interest.
- Relevance. Data should include indicators central to the functioning of the science and technology enterprise.
- Timeliness. Data that are not part of a time series should be timely, i.e., substantial and unmeasured changes in the population under study should not have occurred since the data were collected.
- Statistical and methodological quality. Survey methods used to acquire data should provide sufficient assurance that statements based on statistical analysis of the data are valid and reliable.[332]

<u>Application of ACS Data in Methodology</u>: National and state ACS data are used in the following measures:

- National
  - Listing of science and engineering occupations (ACS 2011)
  - Racial and ethnic distribution of employed individuals in S&E occupations, and of S&E degree holders, college graduates, and U.S. residents: 2010
  - Foreign-born workers in S&E occupations, by education level: Selected years, 2000–11
  - Employed in S&E occupations, by education level, 2011
- State
  - High school graduates or higher among individuals 25–44 years old, by state: 2001, 2006, and 2011
  - Postsecondary degree holders among individuals 25–44 years old, by state: 2001, 2006, and 2011
  - Bachelor's degree holders among individuals 25–44 years old, by state: 2001, 2006, and 2011
  - Bachelor's degree holders potentially in the workforce, by state: 2001, 2006, and 2011

<u>ACS Questions Utilized (as suggested by the methodology)</u>: P4-6, P7, P11, P45-46

<u>Major Uses</u>: By selecting a set of indicators, the Board seeks to contribute to the assessment of the state of U.S. science and engineering and to highlight issues of current opportunity or

---

[331] Text source: National Science Board, "Preface and Introduction," *Science and Engineering Indicators 2014 Digest*, NSB 14-02, February 2014.

[332] Text source: National Science Board, "Letter of Transmittal," *Science and Engineering Indicators 2014*, NSB 14-01, February 2014.

concern. These measures address an emerging set of trends of particular interest to planners and policymakers at all levels whose decisions affect our national S&E enterprise.[333]

---

[333] Text source: National Science Board, "Preface and Introduction," *Science and Engineering Indicators 2014 Digest*, NSB 14-02, February 2014.

E. National Science Foundation

92. *National Survey of College Graduates*

Purpose of the Measure: The National Survey of College Graduates (NSCG) is a longitudinal biennial survey conducted since the 1970s that provides data on the nation's college graduates, with particular focus on those in the science and engineering workforce. The survey samples individuals who are living in the United States during the survey reference week, have at least a bachelor's degree, and are under the age of 76. This survey is a unique source for examining various characteristics of college-educated individuals, including occupation, work activities, salary, the relationship of degree field and occupation, and demographic information.[334]

The NSCG is one of three principal surveys that provide data for the National Science Foundation's (NSF) Scientists and Engineers Statistical Data System (SESTAT). The purpose of the SESTAT database is to provide information on the entire U.S. population of scientists and engineers with at least a bachelor's degree. Historically, SESTAT has been produced by combining data from the Survey of Doctorate Recipients (SDR; representing persons in the general U.S. population who have earned a doctorate in science, engineering, or health (SEH) from a U.S. institution), the National Survey of Recent College Graduates (NSRCG; representing persons with a recently earned bachelor's or master's degree in SEH from a U.S. institution) and the NSCG. The NSCG constitutes the bulk of the records in the SESTAT database; accounting for approximately 60% of the records in the SESTAT system and slightly over 91% of the population estimate in 2010.[335]

Responsible Unit: National Center for Science and Engineering Statistics (NCSES), National Science Foundation

Authorization: Congress has mandated the National Science Foundation (NSF) "to provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources and to provide a source of information for policy formulation by other agencies of the Federal Government" (NSF Act of 1950, as amended; 42 U.S.C. 1862). A critical component of this mission is information on the science and engineering (S&E) workforce. NSF thus reports information for the United States based on data collected on the number, characteristics, and employment of the S&E workforce. Additionally, the Science and Engineering Equal Opportunities Act (Public Law 96-516) gave a further mandate to NSF to ensure that obtaining information on women, minority group members, and persons with disabilities in the S&E workforce were important considerations in data collection and analysis.[336]

---

[334] Text source: "NSCG - About the Survey," NSF website.

[335] Text source: NSF, "2013 National Survey of College Graduates Supporting Statement, Part A," September 20, 2012.

[336] Ronald S. Fecso, Mary J. Frase, and Nirmala Kannankutty, "Using the American Community Survey as the Sampling Frame for the National Survey of College Graduates," Working Paper, NCSES 12-201, August 2012, p. 4.

Frequency of Data Publication: The most recent NSCG round was carried out in 2013. The 1993, 2003, and 2010 cycles of the NSCG provided coverage of the nation's college-educated population as of the survey reference date. In addition to the 1993, 2003, and 2010 survey cycles, the NSCG was conducted biennially or triennially in the periods 1990–99 and 2000–09. For these within-decade iterations of the NSCG, the survey focused on the science and engineering (S&E) workforce component of the college-educated population.[337]

Timing of Data Release: The first data release from the 2010 NSCG was published in January 2013.

Modes to Access Data: NCSES, NSF:

- [NSCG Data](#) (SAS, ASCII)
- [Scientists and Engineers Statistical Data System](#) (SAS, ASCII)

Methodology: The U.S. Census Bureau conducted the NSCG for NCSES. Initial data collection used a self-administered mail survey and, beginning in the 2010 survey cycle, a self-administered Web survey. Non-respondents to the initial data collection were followed up using computer-assisted telephone interviewing (CATI). For each survey cycle, the data collection effort typically lasted 6 to 8 months.

In all survey cycles, the NSCG used a stratification sampling design to select its sample from the eligible sampling frame. Within the sampling strata, the NSCG used probability proportional to size (PPS) or systematic random sampling techniques to select the NSCG sample.

In the 1990s and 2000s, the NSCG sample was stratified by four sampling variables: 1) Demographic group, which was a composite variable that captured disability status, ethnicity, race, citizenship status at birth, and likelihood of a U.S.-earned degree; 2) Highest degree type; 3) Occupation; and 4) Sex.

Stratification variables were revised in the 2010 NSCG to take advantage of the additional information available on its new sampling frame, the American Community Survey and stratified by three sampling variables: 1) Demographic group; 2) Highest degree type; and 3) Occupation/bachelor's degree field, which was a composite variable that captured both occupation field and bachelor's degree field.

Every sample case in the NSCG has a sample weight that reflects the portion of the population the case represents. This sample weight reflects weighting adjustments that were conducted to account for the sample selection, nonresponse, trimming procedures to eliminate extreme weights, and raking procedures to ensure the sampling weights agree with sampling frame estimates for certain key sampling variables. The sample weights enable data users to derive survey estimates that reflect the NSCG target population.[338]

---

[337] Text source: "NSCG - About the Survey - Overview," NSF website.
[338] Text source: *Ibid*.

<u>Application of ACS Data in Methodology</u>: The 2010 NSCG incorporated a dual frame sample design where it selected a portion of its sample from the 2009 American Community Survey respondents who indicated they had a bachelor's degree or higher in any field of study. The remaining portion of the 2010 NSCG sample was selected from respondents to the 2008 NSCG.

Stratification variables were revised in the 2010 NSCG to take advantage of the additional information available on its new sampling frame, the American Community Survey and stratified by three sampling variables.

- Demographic group: A composite variable that captured disability status, ethnicity, race, citizenship status at birth, and likelihood of a U.S.-earned degree.
- Highest degree type
- Occupation/bachelor's degree field: A composite variable that captured both occupation field and bachelor's degree field.[339]

<u>ACS Questions Utilized (as suggested by the methodology)</u>: P4-6, P11-12, P45-46

<u>Major Uses</u>: The longitudinal data from the NSCG provides valuable information on careers, training, and educational development of the Nation's highly educated science and engineering population. These data enable government agencies to assess the scientific and engineering resources available in the U.S. to business, industry, and academia, and to provide a basis for the formulation of the Nation's science and engineering policies. Educational institutions use the NSCG data in establishing and modifying scientific and technical curricula, while various industries use the information to develop recruitment and remuneration policies.

Researchers, policymakers and other users of the data use information from the SESTAT database to answer questions about the number, employment, education, and characteristics of the S&E workforce. Because it provides up-to-date and nationally representative data, researchers and policymakers use the database to address questions on topics such as the role of foreign-born or foreign-degreed scientists and engineers, the transition from higher education to the workforce, the role and importance of postdocs, diversity in both education and employment, the implications of an aging cohort of scientists and engineers as baby boomers reach retirement age, and information on long-term trends in the S&E workforce.

Data from NSF's SESTAT component surveys are used in policy discussions of the executive and legislative branches of Government, the National Science Board, NSF management, the National Academy of Sciences, professional associations, and other private and public organizations. Some recent specific examples of the use of the SESTAT data are: the American Institutes of Research used the SESTAT data in the evaluation of the NSF's Alliance for Graduate Education and the Professoriate Program; the Commission on Professionals in Science and Technology regularly publishes data from SESTAT in their STEM Trends publications; the General Accounting Office used the SESTAT data to issue a report on education and disability.

---

[339] Text source: *Ibid*. Rationale, benefits, and issues regarding use of the ACS as the sampling frame for the NSCG are discussed in detail in Fecso, Frase, and Kannankutty, *op. cit.*

The Federal Reserve Bank of St. Louis used the SESTAT data to examine the pathway from Community College to a Bachelor's Degree and Beyond; and many Ph.D. students use the SESTAT workforce data in dissertations.[340]

---

[340] Text source: NSF, "2013 National Survey of College Graduates Supporting Statement, Part A," September 20, 2012.

F. Social Security Administration

93. *Financial Outlook for Old-Age, Survivors, and Disability Insurance Trust Funds*

<u>Nature and Purpose of the Measure</u>: The Old-Age, Survivors, and Disability Insurance (OASDI) program makes monthly income available to insured workers and their families at retirement, death, or disability. The Social Security Act established the Board of Trustees to oversee the financial operations of the OASDI Trust Funds. The Board is composed of six members, four by virtue of their positions in the Federal Government and two appointed by the President and confirmed by the Senate. The Social Security Act requires that the Board, among other duties, report annually to the Congress on the actuarial status and financial operations of the OASDI Trust Funds.[341]

<u>Responsible Unit</u>: Office of the Chief Actuary, Social Security Administration

<u>Authorization</u>: 42 USC 401(c)

<u>Frequency of Data Publication</u>: Annual

<u>Timing of Data Release</u>: Mid-year (the 2014 report to Congress was published July 28, 2014)

<u>Modes of Data Access</u>: [“Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds,”](#) Social Security Administration (pdf)

<u>Methodology</u>: The future income and cost of the OASDI program will depend on many demographic, economic, and program-specific factors. Trust fund income will depend on how these factors affect the size and composition of the working population as well as the level and distribution of earnings. Similarly, program cost will depend on how these factors affect the size and composition of the beneficiary population as well as the general level of benefits.

The Trustees make basic assumptions for several of these factors based on analysis of historical trends, historical conditions, and expected future conditions. These factors include fertility, mortality, immigration, marriage, divorce, productivity, inflation, average earnings, unemployment, real interest rate, and disability incidence and termination. Other factors depend on these basic assumptions. These other, often interdependent, factors include total population, life expectancy, labor force participation, gross domestic product, and program-specific factors. Each year the Trustees reexamine these assumptions and methods in light of new information and make appropriate revisions. The Trustees selected the assumptions for this report by the end of January 2014.

---

[341] Text source: The Board of Trustees, Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds, [“The 2014 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds,”](#) July 28, 2014, p. 1.

Future levels of these factors and their interrelationships are inherently uncertain. To address these uncertainties, this report uses three sets of assumptions, designated as intermediate (alternative II), low-cost (alternative I), and high-cost (alternative III). The intermediate set represents the Trustees' best estimate of the future course of the population and the economy.

With regard to the net effect on the actuarial status of the OASDI program, the low-cost set is more optimistic and the high-cost set is more pessimistic. The low-cost and high-cost sets of assumptions reflect significant potential changes in the interrelationships among factors, as well as changes in the values for individual factors.

While it is unlikely that all of the factors and interactions will differ in the specified directions from the intermediate values, many combinations of individual differences in the factors could have a similar overall effect. Outcomes with overall long-range cost as low as the low-cost scenario or as high as the high-cost scenario are very unlikely. This report also includes sensitivity analysis, where factors are changed one at a time, and a stochastic projection, which provides a probability distribution of possible future outcomes that is centered around the intermediate assumptions.

Readers should interpret with care the estimates based on the three sets of alternative assumptions. These estimates are not specific predictions of the future financial status of the OASDI program, but rather a reasonable range of future income and cost under a variety of plausible demographic and economic conditions.

The Trustees assume that values for each of the demographic, economic, and program-specific factors change toward long-range ultimate values from recent levels or trends within the next 25 years. For extrapolations beyond the 75-year long-range period, the ultimate levels or trends reached by the end of the 75-year period remain unchanged. The assumed ultimate values represent average annual experience or growth rates. Actual future values will exhibit fluctuations or cyclical patterns, as in the past.[342]

The diagrams below show the flow of the long-term OASDI projections methodology.[343]

Application of ACS in the Methodology: The American Community Survey (ACS) is one of multiple data sources used in Process 1 – Demography. In 2014, the following sets of ACS data were used in the demographic projections model:

- Estimates of the population by marital status, sex, and age from the ACS public use microdata samples (PUMS) for years 2000 – 2011.
- From ACS PUMS, number of existing marriages from 2000 – 2011 by age group of husband crossed with age group of wife.

---

[342] *Ibid*, pp. 75-76.
[343] Office of the Chief Actuary, Social Security Administration, "Long-Range OASDI Projection Methodology: Intermediate Assumptions of the 2014 Trustees Report," June 2014.

Chart 1:

Overview of Long-Range OASDI Projection Methodology

Process 1:
Demography

Process 2:
Economics

Process 3:
Beneficiaries

Process 4:
Trust Fund Operations and Actuarial Status

Chart 2:

Demography – Process 1

Trustees ultimate assumptions
Fertility
Mortality
Legal Immigration
Other Immigration

1.1   Fertility
Inputs: Historical U.S. births and female resident population
Outputs: Historical and projected central birth rates

1.2   Mortality
Inputs: Historical number of U.S. deaths by cause and U.S. resident population; medicare deaths and enrollments
Outputs: Historical and projected death probabilities

1.3   Legal Immigration
Inputs: Historical legal U.S. immigration levels
Outputs: Historical and projected annual legal immigration and emigration levels

1.4   Historical Population
Inputs: Historical U.S. population, undercounts, marital status data, other immigrant population data, and estimates of population in other components of Social Security area.
Outputs: Historical Social Security area population by age, sex, and marital status (including starting population)
*Historical other immigrant population by age and sex*

1.5   Other Immigration
Inputs: Historical levels of the foreign-born population
Outputs: Projected annual other immigration and emigration levels

1.6   Marriage
Inputs: Historical number of marriages, remarriage data, and consistent population (detailed data for a subset of the U.S. population)
Outputs: Historical and projected central marriage rates

1.7   Divorce
Inputs: Historical number of divorces and consistent population (detailed data for a subset of the U.S. population)
Outputs: Historical and projected central divorce rates

1.8   Projected Population
Inputs: Historical U.S. family data
Outputs: Projected data – Social Security area population by age, sex, and marital status; other immigrant population by age, sex and marital status; children by age of parent and average family size; projected deferred action for childhood arrivals (DACA) population by age and sex

Economics, Beneficiaries, and Trust Fund Operation and Actuarial Status
Note: 1.2, 1.3, 1.4, and 1.5 also provide outputs used by these processes

281

- Estimates of new marriages by age-group-of-husband crossed with age-group-of-wife from the ACS PUMS occurring, on average, end of years 2007 – 2010.
- The initial existing same-sex marriage grid as of 2013 is based on the number of same-sex marriages from 2004 – 2011 from the vital statistics offices in various states, and the same-sex marriage grid by age-group-of-husband crossed with age-group-of-wife from the ACS PUMS occurring, on average, at the end of years 2007 – 2010.
- From the ACS, foreign-born new arrivals by ACS year (2000-2011), entry year (1900-2011), age (0-100) and sex.

Since eligibility for auxiliary benefits is dependent on marital status, the Social Security area population is disaggregated by marital status. The four marital states are defined as single (having never been married), married, widowed, and divorced.

The distribution of the number of existing marriages are available for Census years 1940 – 2000 from decennial Census PUMS. These data are aggregated by age group of husband crossed with age group of wife. Additional tabulations from the ACS from 2000 – 2011 are incorporated to adjust these marital prevalence grids for changes since 2000. The grids are transformed from age grouped numbers to single year of age figures from ages 14 to 100+ for husband and wife using the two dimensional H.S. Beers method of interpolation.

Percentages of single, married, widowed, and divorced persons are calculated by taking the estimate for each marital status category and dividing them by the total number of people for each age group and sex based on either the CPS, more detailed Census numbers, if available, or ACS. Then, for each sex, if one age group has a higher or lower percentage than the surrounding age groups, an average of the surrounding groups replaces the original value. After verifying the percentages are close to the original data (and adjusted if needed), these percentages are multiplied by the total populations for each age, sex, and year to get a preliminary population for each age, sex, and marital status.

To keep the marriage prevalence grids and the marital status percentages smooth and consistent, several algorithms are used. First, the married population is adjusted so that the number of married males equals the number of married females. Then, the number of married persons for each age and sex is set equal to the marginal total of the associated year's marital prevalence grid. Finally, the other marital statuses population totals are adjusted to keep the total number of people in all marital statuses the same as calculated before splitting into marital statuses.

The initial existing same-sex marriage grid as of 2013 is based on the number of same-sex marriages from 2004 – 2011 from the vital statistics offices in various states, and the same-sex marriage grid by age-group-of-husband crossed with age-group- of-wife from the American Community Survey (ACS) public use microdata samples (PUMS) occurring, on average, at the end of years 2007 – 2010.

The NCHS stopped collecting data on the annual number of new marriages in the MRA in 1989. Less detailed data on new marriages from a subset of the MRA were obtained for the years 1989-1995. In 2008, the ACS added a question asking if a person was married in the last 12 months. Using this question along with ages of spouses, new marriages grids by age-group-of-husband crossed with age-group-of wife were developed. For the years in between the NCHS and ACS data, the marriage grids were linearly interpolated. These NCHS and ACS data are used to adjust the more detailed age-of-husband crossed with age-of-wife rates from the earlier years.

The ACS provides the number of foreign-born new arrivals, which is then used to separate the historical net other than legal immigration into those entering and those leaving. There are several other key inputs that go into this calculation, including an estimated undercount factor. This factor accounts for (1) differences between the foreign-born data from the ACS and the component pieces obtained from DHS, (2) differences between the ACS (Public Use Microdata Sample) and Census' total population, and (3) the foreign-born residing in Puerto Rico. The estimated other than legal immigration is calculated by taking the Beers'd foreign born from the ACS (after applying the undercount factors) and subtracting the legal new arrivals. The estimated historical other than legal emigration is then calculated as the difference between the net other than legal immigration and the estimated historical other than legal immigration. A series of steps are then taken to smooth the two categories. Assumptions split the historical other immigrants into those who arrive or depart the Social Security area as a never-authorized immigrant, nonimmigrant, and visa-overstayer immigrant.[344]

ACS Questions Utilized (as suggested by the methodology): P2-4, P7-9, P15, P20-21

Major Uses: The financial outlook of the OASDI Trust Funds guides the Board of Trustees' recommendations to Congress regarding legislative actions required to maintain the long-term solvency of the Trust Funds.

---

[344] Text source: *Ibid.*