

**Office of Management and Budget**

**Statistical Policy Working Paper<sup>1</sup>**

*Federal Agency Approaches to Providing Statistical Information  
on Detailed Asian and Native Hawaiian and Other Pacific Islander Groups*

---

<sup>1</sup> This working paper is intended to inform interested parties and to encourage discussion of Federal statistical policy issues and standards. It does not represent any change in OMB policy. Any views expressed herein are those of the author, Brian A. Harris-Kojetin, Ph.D., Statistical and Science Policy, Office of Management and Budget.

## Executive Summary

The Office of Management and Budget (OMB) *Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity* (available at [http://www.whitehouse.gov/omb/fedreg\\_1997standards/](http://www.whitehouse.gov/omb/fedreg_1997standards/)) provide a minimum standard for maintaining, collecting, and presenting data on race and ethnicity for all Federal reporting purposes. These standards, which provide a common language to promote uniformity and comparability for data on race and ethnicity, provide minimum categories for Federal data on race and ethnicity: American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, and White. These standards do not limit the collection of data to these categories; however, any collection that uses more detail must be organized in such a way that the additional categories can be aggregated back into these minimum categories for data on race and ethnicity.

The Department of Health and Human Services (HHS) recently issued data standards for race and ethnicity that include additional granularity, with all categories aggregating to the OMB standard categories. The HHS standards provide additional granularity for the Asian (seven additional categories), Native Hawaiian or Other Pacific Islander (four additional categories), and Hispanic (four additional categories) subpopulations beyond the OMB minimum standard categories. Specifically, the new HHS standards expand the data category for Asian to distinguish Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, and Other Asian; and the Native Hawaiian or Other Pacific Islander category has been expanded to distinguish Native Hawaiian, Guamanian or Chamorro, Samoan and Other Pacific Islander.

Both the OMB and HHS standards explicitly acknowledge the need for adequate samples to provide information of acceptable statistical quality and to maintain confidentiality of respondents' information. Although some detailed information regarding racial and ethnic groups is currently available from the Decennial Census and the American Community Survey, confidentiality constraints and quality concerns often preclude Federal agencies from publishing results for small groups or subgroups on even large national surveys. To help address concerns about the limited availability of information on detailed race groups from Federal data collections, the Statistical and Science Policy Branch at OMB reached out to principal statistical agencies to examine ways agencies are thinking creatively to collect and report detailed race and ethnicity data from their surveys.

A common approach taken by statistical agencies to increase sample sizes has been to pool data across months or years of data collection. This approach is a key component of the American Community Survey design, and it is also being used in other major household surveys. Another approach to increase sample sizes for detailed racial and ethnic groups is to oversample members of those groups at a higher rate so that they make up a larger proportion of the sample than they do the overall population. Specific examples of each of these approaches are highlighted in this paper and included in the appendix.

Although obtaining adequate sample sizes to achieve acceptable statistical quality and maintain confidentiality has been a challenge, Federal statistical agencies have been able to provide some valuable information on detailed racial and ethnic groups through oversampling and pooling of

data across years. These reports can serve as useful examples for other agencies to consider applying in their large-scale household surveys to provide more information for these detailed racial and ethnic groups.

## **1. OMB Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity**

### *A. Background*

The Paperwork Reduction Act, PL 104-13, charges the OMB Director with developing and overseeing the implementation of government-wide policies, principles, standards, and guidelines concerning statistical collection procedures and methods. Statistical classifications, definitions, and data sources encourage uniformity in data collection, analysis, and dissemination. They are designed and managed to support the full range of research and analytical objectives in a specific subject matter area rather than the needs of a specific program or a specific study. The general criteria OMB has used for evaluating the standards have been relevancy, accuracy, currency, efficiency, minimization of burden, and stability (“continuity” and/or “comparability”).

For more than 35 years, OMB’s standards for data on race and ethnicity have provided a common language to promote uniformity and comparability for data on race and ethnicity for the population groups specified in the standards. They were developed in cooperation with Federal agencies to provide consistent data on race and ethnicity throughout the Federal Government.

Although there are statutory requirements that specify many different needs for racial and ethnic data by the Federal Government, development of the data standards stemmed in large measure from new responsibilities to enforce civil rights laws. Data were needed to monitor equal access in housing, education, employment, and other areas, for populations that historically had experienced discrimination and differential treatment because of their race or ethnicity. Given the myriad of laws and Federal uses, there was and continues to be a clear need for the collection and use of compatible, non-duplicated, exchangeable racial and ethnic data by Federal agencies. OMB is in a unique position to coordinate these measurement activities across agencies.

The OMB Standards provide a minimum categorization for maintaining, collecting, and presenting data on race and ethnicity for all Federal reporting purposes. The standards stipulate *how* agencies must collect data on race and ethnicity **if** they are collecting this information—the standards do not require agencies to gather data on race and ethnicity. Moreover, the standards recognize that because of reliability and confidentiality constraints, some data that are collected may not necessarily be published by the agency.

### *B. Categories and Definitions*

The standards have five categories for data on race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. There are two categories for data on ethnicity: “Hispanic or Latino,” and “Not Hispanic or Latino.”

These are defined as follows:

*American Indian or Alaska Native.* A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

*Asian.* A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

*Black or African American.* A person having origins in any of the black racial groups of Africa. Terms such as “Haitian” or “Negro” can be used in addition to “Black or African American.”

*Hispanic or Latino.* A person of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race. The term, “Spanish origin,” can be used in addition to “Hispanic or Latino.”

*Native Hawaiian or Other Pacific Islander.* A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

*White.* A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

The standards require that respondents be permitted to select one or more of the race categories, so that they may identify with any and all categories. Furthermore, the standards should not be construed to limit the collection of data to the categories described above. The collection of greater detail is encouraged; however, any collection that uses more detail must be organized in such a way that the additional categories can be aggregated into these minimum categories for data on race and ethnicity.

A core principle of the standards is that self-identification is the preferred means of obtaining information about an individual’s race and ethnicity, except in instances where self-reporting is not practicable or feasible (e.g., when identification is done by funeral personnel, observer identification may be used).

It is important to note that the categories in this classification are social-political constructs and should not be interpreted as being scientific or anthropological in nature. They are not to be used as determinants of eligibility for participation in any Federal program.

### *C. Use of the OMB Standards for Record Keeping and Reporting*

The OMB *Standards* are required to be used as a minimum for all federally-sponsored statistical data collections that include data on race and/or ethnicity, except when the collection effort focuses on a specific racial or ethnic group. The standards are used not only in the Decennial Census (which provides the data on the size of universe, which often serves as the “denominator”

for many measures), but also in household surveys, on administrative forms (e.g., school registration and mortgage lending applications), and in medical and other research.

Federal agencies should produce and publish as much detailed information on race and ethnicity as possible. However, Federal agencies must not present data on detailed racial and ethnic categories when the collection involves a sample of such size that the data on the smaller categories would be unreliable, or doing so would compromise data quality or confidentiality standards.

## **2. Department of Health and Human Services—New Standards for Race and Ethnicity Data**

The Department of Health and Human Services (HHS) recently finalized new standards for collecting and reporting data on race, ethnicity, sex, primary language and disability status in accordance with section 4302 of the Patient Protection and Affordable Care Act, PL 111-148. Section 4302 requires the Secretary of HHS to establish data collection standards for race, ethnicity, sex, primary language, and disability status. The law requires that, once established, these data collection standards be used, to the extent practicable, in all national population health surveys.

As described below, the new HHS data standards include additional granularity for race and ethnicity, but all of the categories roll up to the OMB standards. Because additional granularity in the race and ethnicity categories is important for documenting and tracking health disparities, large Federal surveys such as the National Health Interview Survey (NHIS), the Current Population Survey (CPS), and the American Community Survey (ACS) have implemented a more granular strategy, particularly for Hispanic and Asian subpopulations. (These and other Federal surveys are discussed in further detail in the following sections.)

HHS's new data standards for race and ethnicity are a slightly modified version of the ACS and Decennial Census questions, which have been tested and structured to increase response rates, validity, and reliability. These new standards provide additional granularity for Hispanic (four additional categories), Asian (seven additional categories), and Native Hawaiian or Other Pacific Islander (four additional categories) subpopulations beyond the OMB minimum standard categories. Specifically, the new standards expand the data category for Asian to distinguish Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, and Other Asian; and the Native Hawaiian or Other Pacific Islander category has been expanded to distinguish Native Hawaiian, Guamanian or Chamorro, Samoan, and Other Pacific Islander. These more detailed race and ethnicity categories, which are modeled on the ACS and recent Decennial Census questions, roll up to the OMB standard five categories. As with the OMB standards, respondents are also instructed to select all categories that apply (i.e., they may select more than one racial category).

In its implementation guidance, HHS states that the minimum data standards on race and ethnicity must be used in all population health surveys conducted or sponsored by HHS in which respondents either self-report information or a knowledgeable proxy provides information about the person or responds for all persons in a household. Like the OMB standards, the HHS data standards represent minimum standards and are not intended to limit the collection of needed

data. Agencies desiring more granularity are permitted and encouraged to collect additional data on race and ethnicity as needed as long as the minimum standard is included, the data can be aggregated up to the minimum standard, and the sample size supports the resulting estimates.

In a similar manner to the OMB standards, HHS implementation guidance makes clear that only those demographic categories for race and ethnicity with adequate sample sizes to provide statistically reliable data should be reported. Information on the validity and reliability of the data should be included, whenever possible, to enable readers to judge the credibility of the findings. Furthermore, HHS guidance notes that when consultation with the agency statistician determines that the race and ethnicity data for any particular demographic group are considered statistically unreliable, then the data for such demographic groups should not be reported separately unless accompanied by the appropriate caveats.

### **3. Reporting of Detailed Race and Ethnicity Data in the Decennial Census and American Community Survey<sup>3</sup>**

The remainder of this paper provides some examples of products from Federal statistical surveys that depict information on detailed Asian and Native Hawaiian and Other Pacific Islander groups. This section provides a brief description of the products that are currently available from the Decennial Census and the American Community Survey, while the next section describes approaches used in major Federal household surveys.

#### *A. 2010 Decennial Census Products*

The Decennial Census is described in Article I, Section 2 of the Constitution. It calls for an actual enumeration of the people every ten years, to be used for apportionment of seats in the House of Representatives among the States. The first official census was conducted in 1790 and has been conducted every ten years, generally on April 1 in years ending in a zero. Since 1975, the Census Bureau has also had responsibility to produce small-area population data needed to redraw state legislative and congressional districts. Other important uses of Census data include the distribution of funds for government programs, such as Medicaid, and planning locations for schools, roads, and other public facilities.

A copy of the race question for the Decennial Census appears in Appendix A. As noted above, the following detailed Asian groups are collected: Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, and Other Asian. The following Native Hawaiian or Other Pacific Islander groups are also collected: Native Hawaiian, Guamanian or Chamorro, Samoan and Other Pacific Islander.

The Decennial Census program produces a variety of data products to provide detailed information on the U.S. population, including counts and statistical estimates of different characteristics for a wide range of racial and ethnic groups. Most Census data are available for many levels of geography, including states, counties, cities, ZIP codes, census tracts, and blocks.

---

<sup>3</sup> Note: The descriptions of the Decennial Census and American Community Survey are drawn from the Census Bureau website and other Census Bureau publications and materials.

**Census Briefs.** The Census Briefs provide analysis of 2010 Census population and housing topics, and include graphs and tables. They focus on the most important aspects of the topics, as well as explore the geographic distribution of the subject matter. This series is a basic analytic tool that is useful for introducing the public to 2010 Census population and housing data. The Census Bureau released in March 2012 a brief on *The Asian Population: 2010*. This report (provided in Appendix A and available at <http://www.census.gov/prod/cen2010/briefs/c2010br-11.pdf>) includes substantial information on detailed Asian groups, including population size and some information on geographic differences.

Additionally, in May 2012, the Census Bureau released a brief on *The Native Hawaiian and Other Pacific Islander Population: 2010*. This report (provided in Appendix A and available at <http://www.census.gov/prod/cen2010/briefs/c2010br-12.pdf>) includes substantial information on detailed Native Hawaiian and Other Pacific Islander groups, including population size and some information on geographic differences.

The *Census 2010 Summary File 1 product*, released in summer 2011, provides important information on detailed Asian groups and detailed Native Hawaiian and Other Pacific Islander groups in a series of tables on age, sex, households, families, relationship to householder, housing units, and group quarters. Most tables are shown down to the block or census tract level.

More specifically, Summary File 1 (SF-1) provides data for a number of detailed Asian groups that approached or exceeded a threshold of 7,000 at the national level in 2000. For 2010, the Census Bureau produced three related SF-1 tables to provide details on the following Asian groups: Asian Indian, Bangladeshi, Bhutanese, Burmese, Cambodian, Chinese (except Taiwanese), Filipino, Hmong, Indonesian, Japanese, Korean, Laotian, Malaysian, Nepalese, Pakistani, Sri Lankan, Taiwanese, Thai, and Vietnamese. Specifically, the following tables are available through American Fact Finder on the Census Bureau website:

- PCT5<sup>4</sup> presents data for people who reported only one Asian ethnic group, such as Korean.
- PCT6 presents data for people who reported one or more Asian ethnic groups, such as Korean and Japanese.
- PCT7 presents data for people who reported Asian ethnic group(s), in any combination of ethnic group(s) or with other race group(s), such as Korean *and* White.

Summary File 1 also provides data for a number of detailed Native Hawaiian and Other Pacific Islander groups that approached or exceeded a threshold of 7,000 at the national level in 2000. For 2010, the Census Bureau produced three related SF-1 tables to provide details on the following Native Hawaiian and Other Pacific Islander groups: the Polynesian population, including data on Native Hawaiians, Samoans, Tongans, and Other Polynesians; the Micronesian population, including data on Guamanian or Chamorro, Marshallese, and Other Micronesians; and the Melanesian population, including data on Fijians, and Other Melanesians. Specifically:

- PCT8 presents data for people who reported only one NHPI ethnic group, such as Samoan.

---

<sup>4</sup> PCT5 is the table designation in American Fact Finder. These are included here to make it easier to quickly bring up the referenced tables.

- PCT9 presents data for people who reported one or more NHPI ethnic groups, such as Samoan and Fijian.
- PCT10 presents data for people who reported NHPI ethnic group(s), in any combination of NHPI ethnic group(s) or with other race group(s), such as Samoan *and* White.

Samples of these Summary File 1 tables are available in Appendix A. Many additional tables can be created using American Fact Finder on the Census web site at the following link: <http://factfinder2.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t>

The *Census 2010 Summary File 2 product*. Detailed demographic information from the 2010 Census Summary File 2 was released at the end of April 2012, adding a new layer of detail to the population and housing topics released last year from the 2010 Census. Information, such as age, relationship, and homeownership, previously available only for an area's entire population, is now available for many detailed racial and ethnic groups in communities across the country, at different levels of geography. Each Summary File 2 table is presented for up to 331 detailed race and Hispanic origin groups, down to the census tract level, including information for the U.S., regions, metropolitan statistical areas, and other geographic areas that may cross state boundaries. These include iteration groups for the total population, 75 major race groups, 114 American Indian and Alaska Native groups, 47 Asian groups, 43 Native Hawaiian and Other Pacific Islander groups, and 51 Hispanic or Latino origin groups. Descriptions of these groups are provided in Appendix A.

The following national-level highlights on median age provide an example of the characteristics available for different racial and ethnic groups. Among detailed Asian groups, the Hmong alone or in any combination population was the youngest Asian group with a median age of 20.5; Okinawan alone was the oldest with a median age of 55.4. Among detailed Native Hawaiian and Other Pacific Islander groups, Marshallese alone was the youngest with a median age of 20.3; Tokelauan alone was the oldest with a median age of 41.4.

The Summary File 2 tables can be found on the Census Bureau's American FactFinder website at <http://factfinder2.census.gov> by using the "Population Groups" filter to select the specific race or ethnic groups of interest.

#### *B. American Community Survey Products*

While the main function of the Decennial Census is to provide *counts* of people for the purpose of congressional apportionment and legislative redistricting, the primary purpose of the American Community Survey (ACS) is to measure the changing social and economic *characteristics* of the U.S. population. For the 2000 Census, additional questions were asked of a sample of persons and housing units (generally 1 in 6 households) on topics such as income, education, place of birth, and more. Information on those topics is now available as part of the ACS. The ACS collects and produces population and housing information every year instead of once every ten years. Collecting data every year provides more up-to-date information throughout the decade about the U.S. population at the local community level.



About 3.5 million housing unit addresses are selected annually across every county in the Nation. The annual ACS sample is smaller than that of the Census 2000 long-form sample, which included about 18 million housing units. As a result, the ACS needs to combine population or housing data from multiple years to produce reliable numbers for small counties, neighborhoods, and other local areas. To provide information for communities each year, the ACS provides one, three, and five-year estimates<sup>5</sup>.

The ACS uses the same race question as the Decennial Census. Data on the characteristics of detailed Asian groups and detailed Native Hawaiian and Other Pacific Islander groups are available in the ACS in products such as the Selected Population Profiles, which provide detailed characteristics such as education, income, poverty, health insurance, commuting, etc. In a similar manner to the Decennial Census products, tables can be produced at varying levels of geography, from the entire Nation to states, counties or census tracts; however, multiple years of ACS data may be needed for less populous areas. For example, ACS one-year estimates cover only areas with a population of at least 65,000, and three-year estimates are available only for areas with a population of at least 20,000. The five-year estimates cover all levels of geography.

The Census Bureau in May 2012 released new detailed estimates about the social, economic and housing characteristics of hundreds of race, tribal, Hispanic and ancestry groups at numerous geographic levels. This is the first time this level of statistical detail has been available for groups since the 2000 Census. The new products, based on the 2006-2010 American Community Survey, are generally comparable to estimates generated from the 2000 Census “long form.” Estimates are available for groups that meet a size threshold and for geographic areas that meet a completed survey response threshold. Up to 300 tables are included on topics such as educational attainment, fertility, nativity, citizenship, income, poverty and homeownership. They are contained in two data products:

- *The 2006-2010 American Community Survey Selected Population Tables.* Characteristics estimates are available for the 392 race, Hispanic origin, ancestry and tribal groups that had a national population of at least 7,000. Estimates are published for an individual group in a particular geographic area if it had at least 50 group members during the five-year survey period. Census tracts are the lowest geographic level available to groups meeting the two thresholds.
- *The 2006-2010 American Community Survey American Indian and Alaska Native Tables.* Characteristics estimates are available for 950 American Indian and Alaska Native groups that had a national population of at least 100. Estimates are published for an individual tribe in a particular geographic area if that group had at least 50 group members during the five-year survey period. Fewer geographic types are available, but the release includes Alaska Native Regional Corporations and American Indian and Alaska Native Areas and Hawaiian Home Lands.

---

<sup>5</sup> Five-year estimates from the ACS were originally intended to have similar characteristics as the long-form. The current sample size of the ACS is projected to yield responses from about 11.75 million housing units over five years, due to the subsampling of nonrespondents.

A sample report from the ACS is available in Appendix B. Many additional tables can be created using American Fact Finder on the Census web site at the following link: <http://factfinder2.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t>

#### **4. Approaches for Providing More Detailed Reporting of Small Racial and Ethnic Groups in Major Federal Household Surveys**

A recent GAO report (2006) noted that there were 235 OMB-approved statistical or research surveys that collect information from households or individuals. This section focuses on a few major Federal household surveys to illustrate their standard practices for reporting data on race and ethnicity and describes ways in which they provide more detailed information on small racial and ethnic groups. The surveys discussed below are all ongoing household surveys that have relatively large samples and represent a significant investment of resources. Along with the Decennial Census and the ACS, these surveys represent much of the demographic statistical infrastructure of the Federal statistical system and provide a large core of data that informs policy makers, businesses, and households as well as directing the allocation of billions of dollars in Federal programs every year.

To help address concerns about the limited availability of information on detailed racial and ethnic groups from Federal data collections, the Statistical and Science Policy Branch at OMB reached out to principal statistical agencies that conduct major Federal household surveys to examine ways agencies are thinking creatively, in light of research and resource limitations, to collect and report detailed race and ethnicity data in their surveys. The following examples illustrate some current practices and may serve as useful exemplars to other agencies when they conduct surveys that would be large enough to produce estimates for more detailed racial and ethnic groups.

##### *A. Current Population Survey (CPS)*

The CPS provides information on the U.S. labor force and is the source of the unemployment rate released each month, usually on the first Friday of the month. The CPS is sponsored by the Bureau of Labor Statistics (BLS), and the data are collected by the Census Bureau using personal or telephone interviews. The sample size of the CPS is about 60,000 households per month. Interviewers ask household respondents to provide the race and ethnicity for everyone living in the household. The CPS collects respondents' identification with the following Asian race groups: Asian Indian, Chinese, Filipino, Japanese, Korean, and Vietnamese, as well as the remaining category Other Asian, if the respondent selects the Asian category from the minimum list of OMB race categories.

BLS publishes monthly labor force information on the overall Asian category, but not for Native Hawaiian and Other Pacific Islanders or for any of the detailed Asian subgroups. An example table is shown in Appendix C.

In addition, BLS also publishes annual averages, aggregating data across all twelve months from the CPS in order to provide more detailed tables by racial groups. In this publication

(<http://www.bls.gov/cps/cpsrace2010.pdf>), BLS is able to provide more detailed breakdowns for Asians by sex and age (see example table in Appendix C). Beginning with the report based on 2011 annual averages, BLS expects to be able to publish some estimates for Native Hawaiians and Other Pacific Islanders.

Similarly, the Department of Labor has recently released estimates based on annual averages from the Current Population Survey for seven specific groups of Asians including Asian Indians, Chinese, Filipinos, Japanese, Koreans, Vietnamese, and Other Asians. The data cover employment and unemployment for each of these groups for 2010. This report also includes labor force participation for 2010 and unemployment rates for 2003-2010 for Native Hawaiians and Other Pacific Islanders (available at

<http://www.dol.gov/sec/media/reports/AsianLaborForce/AsianLaborForce.pdf>).

BLS recently published an article primarily using three years of pooled data from the CPS (2008 to 2010) in order to provide more extensive information on employment by industry and occupation for those seven detailed Asian groups (included in Appendix C and available at <http://www.bls.gov/opub/mlr/2011/11/art1full.pdf>). The article also provides labor force participation rates, unemployment rates, and alternative measures of labor underutilization.

In addition, although BLS has data on these seven detailed categories going back to 2003, these detailed categories are not included on the CPS' PUMS file available to the public. BLS and the Census Bureau are examining a possible recode of these specific race categories so that these data might be included in forthcoming PUMS releases.

#### *B. National Health Interview Survey (NHIS)*

The NHIS is the principal source of information on the health of the U.S. population. Survey results have been instrumental in providing data to track health status, health care access, and progress toward achieving national health objectives. The NHIS is conducted annually by the U.S. Census Bureau for the Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS). In 2010, household interviews were completed for about 90,000 persons living in over 34,000 households. For the race question, interviewers ask household respondents to provide the race and ethnicity for everyone living in the household and directly collect respondent's identification with the following detailed Asian race groups: Asian Indian, Chinese, Filipino, Japanese, Korean, and Vietnamese, as well as the remaining category Other Asian. The following Native Hawaiian or Other Pacific Islander groups are also collected: Native Hawaiian, Guamanian, Samoan, and Other Pacific Islander.

NCHS frequently pools data across years to generate sufficient sample size within small population groups. For the smallest subgroups, many years of data need to be pooled. NCHS has published a number of reports on different topics that have adopted this strategy for smaller racial/ethnic groups. For example, NCHS published the report: *Utilization of Dental Care Services by Asians and Native Hawaiian or Other Pacific Islanders: United States, 1997–2000* (included in Appendix C and available at <http://www.cdc.gov/nchs/data/ad/ad336.pdf>) by pooling four years of NHIS data.

NCHS has also taken another approach to enable greater reporting of detailed racial groups. Since 2006, the NHIS has been oversampling Asian American households. In addition, the Asian adults age 65 and older have a greater probability of being selected as the sample adult in that portion of the survey. Other changes were implemented in the 2011 NHIS to increase the overall sample size to enable monitoring of the changes due to the Affordable Care Act, and it is expected that these changes may result in a small increase in the Asian sample in the NHIS, above what is already expected as a result of the oversample.

*C. National Health and Nutrition Examination Survey (NHANES)*

NHANES is a program of studies designed to assess the health and nutritional status of adults and children in the United States; it is also sponsored by NCHS. The survey is unique in that it combines interviews and physical examinations. NHANES examines a nationally representative sample of about 5,000 people each year. These individuals are located in counties across the country, 15 of which are visited each year. The NHANES interview includes demographic, socioeconomic, dietary, and health-related questions. The examination component consists of medical, dental, and physiological measurements, as well as laboratory tests administered by highly trained medical personnel. Findings from this survey are used to determine the prevalence of major diseases and risk factors for diseases. Information is also used to assess nutritional status and its association with health promotion and disease prevention. NHANES findings are in addition the basis for national standards for such measurements as height, weight, and blood pressure.

The current NHANES panel (2011-2014) oversamples Asians and is accompanied by an extensive outreach program. HHS is reaching out to community organizations to spread the word about NHANES and encourage participation, as well as to identify sources for local interpreters and local media contacts. This oversample should provide extensive and detailed information about the health and nutrition of Asian Americans. Specifically, by Fall 2013, estimates of prevalence among Asian Americans of undiagnosed conditions such as hypertension, high cholesterol, and diabetes will be available.

*D. Additional efforts to provide more detailed reporting of small racial and ethnic groups*

Federal statistical agencies provided additional information about some efforts in other major survey programs to provide more detailed reporting on small racial and ethnic groups; however, these efforts may still not result in publication of estimates on detailed Asian groups. For example, the National Center for Education Statistics (NCES) sponsors the National Household Education Survey (NHES), which provides descriptive data on the educational activities of the U.S. population covering learning at all ages, from early childhood to school age through adulthood. In 2007, the NHES oversampled Asian and Pacific Islander children and obtained a large enough sample size for group analysis of Asian and Pacific Islander children (as well as some separate analyses of Asians), but the sample still was not large enough for the analysis of subgroups (even ones where the expected prevalence would be high such as Chinese). NCES had the same experience with oversampling in one of its longitudinal studies, the Early Childhood Longitudinal Study: Kindergarten Class of 2010-11. However, the Early Childhood Longitudinal Study Birth Cohort used information from birth certificates to identify children of

Chinese descent and oversampled those children. The resulting sample allowed analyses of such children independently of children of other Asian descent.

NCES also reported that the National Assessment of Educational Progress (NAEP), which is also referred to as the Nation's Report Card, oversamples schools that have a higher proportion of students of various racial and ethnic groups (e.g., African American, Hispanic, and Native American and Alaskan Native). Although NAEP data have not been pooled, NCES reports no technical impediments to pooling, though further study would be required.

The National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation reported that most of its surveys do not have the sample size that would allow reporting race categories in finer granularity than the five standard OMB race categories. The Center did provide a report as an example of how it is possible to report race combinations using the Survey of Doctorate Recipients (SDR), (available at <http://www.nsf.gov/statistics/infbrief/nsf12304/nsf12304.pdf>). The report suggests that there is more flexibility in reporting race data by the maximum number of respondents identifying with each race. The focus of the report was on multiple race reporting and did not report on detailed racial or ethnic groups.

## **5. Summary and Conclusions**

The OMB Standards for data on race and ethnicity permit and even encourage collection and reporting of more detailed racial and ethnic subgroups. Recent HHS standards also stress the collection of detailed racial and ethnic groups. However, both standards explicitly acknowledge the need for adequate samples to provide information of acceptable statistical quality and to maintain confidentiality of respondents' information. These reliability and confidentiality constraints often preclude Federal agencies from publishing results from small groups or subgroups on even large national surveys.

Despite the relatively large sample sizes of the major household surveys described in the previous sections, their ability to provide estimates for detailed race groups is quite limited. Statistical agencies carefully review their statistical products prior to publication to ensure that the data meet specified criteria for reliability as well as protecting confidentiality of respondents. Because all of the major Federal household surveys (with the exception of the Decennial Census and the ACS) are voluntary, it is considered critical that Federal statistical agencies be able to promise respondents that the information they provide will be kept confidential and used only for statistical purposes. To secure high levels of cooperation from respondents, even on mandatory collections, this is essential. Agencies are legally bound by that pledge of confidentiality and must institute appropriate safeguards and statistical methods to ensure that identifiable information is not disclosed in any statistical products.

A common approach taken by statistical agencies to increase sample sizes and thereby enable publication of more detailed racial and ethnic groups has been to pool data across months or years of data collection. In fact, this approach is a key component of the ACS design, and it is also being used in other major household surveys, such as the CPS and NHIS.

Another approach frequently used by statistical agencies to increase sample sizes and thereby enable publication of data for various racial and ethnic groups, including the more detailed groups, is to oversample members of those groups. While oversampling is a key aspect of the design of many surveys, it is important to note that there are limitations to this approach. If the sampling frame the agency is using does not contain information on the detailed race subgroups, then the agency cannot directly oversample the people it needs, making it much more difficult or expensive to obtain an appropriate sample. Furthermore, if a great deal of oversampling is needed relative to the size of the overall sample, it directly impacts the efficiency of the sample for making broader national estimates and estimates for other groups.

Federal statistical agencies have been able to provide some valuable information on detailed racial and ethnic groups through oversampling and pooling of data across years. These reports can serve as useful examples for other agencies to consider applying in their large-scale household surveys to provide more information for these detailed racial and ethnic groups.

### **Contents of Appendices**

#### Appendix A. 2010 Census Products

1. Copy of 2010 Decennial Census form that includes the race question
2. Census Brief on the Asian Population
3. Census Brief on the Native Hawaiian and Other Pacific Islander Population
4. Example reports from the SF-1 file from the 2010 Census
5. Detailed racial and ethnic groups that will be included in the SF-2 file from the 2010 Census

#### Appendix B. American Community Survey Products

1. Example reports from the ACS

#### Appendix C. Products from other major household surveys

1. Example of monthly BLS publication table showing employment by race
2. Example table from the 2010 CPS race report
3. BLS Article from the *Monthly Labor Review* providing more detailed estimates on Asian groups.
4. Report from NCHS pooling data from the NHIS to provide estimates for detailed Asian groups.