

## **Appendix B.**

### **Sample Design and Weighting**

#### **ELIGIBLE UNIVERSE**

The universe of interest for the American Housing Survey National Sample (AHS-N) is the residential housing units in the United States that exist at the time the survey is conducted. This excludes group quarters and businesses. See also the definition of a Housing Unit in the definitions of Appendix A.

#### **SAMPLE SIZE**

The 2011 national data are from a sample of housing units interviewed between July and December 2011. The same basic sample of housing units is interviewed every two years until a new sample is selected. The U.S. Census Bureau updates the sample by adding newly constructed housing units and units discovered through coverage improvement efforts.

For the 2011 American Housing Survey-National (AHS-N) sample, approximately 72,400 sample housing units were selected for interview. This includes a supplemental sample of 2,708 housing units in the Los Angeles MSA.

In 2011, the Department of Housing and Urban Development (HUD) requested that the AHS-N sample size be increased from 56,200 to 64,750. Two methods were used to increase the sample size. First, the roughly 5,200 units that were reduced from the 2007 sample<sup>1</sup> were reinstated. Second, approximately 3,100 units that were selected during the 2000 redesign, but were never interviewed, were introduced as a supplement. Later, HUD requested an oversample of approximately 5,300 subsidized housing units.

About 2,700 of the 72,400 total units included for interview were found to be ineligible because the unit no longer existed or because the units did not meet the AHS-N definition of a housing unit.

Of the 69,700 eligible sample units, about 8,800 were classified (both occupied and vacant housing units) as "Type A" noninterviews because (a) no one was at home after repeated visitors or (b) the respondent refused to be interviewed.<sup>2</sup> This classification produced an unweighted overall response rate of 87 percent. The weighted overall response rate was 87 percent.

#### **SAMPLE SELECTION**

The Census Bureau has interviewed the current longitudinal sample of housing units since 1985. First, the United States was divided into areas made up of counties or groups of counties and

---

<sup>1</sup> Due to budgetary constraints, roughly 8 percent of units were taken out of the 2007 sample and were not interviewed until 2011.

<sup>2</sup> Previous versions of this document used to include units that were "unable to be located" in the counts for Type A noninterviews. This is incorrect.

independent cities known as primary sampling units (PSUs). A sample of these PSUs was selected. Then, a sample of housing units was selected within these PSUs.

**Selection of sample areas.** The sample for AHS is spread over 394 PSUs. These PSUs cover 878 counties and independent cities with coverage in all 50 states and the District of Columbia.

If there were over 100,000 housing units in a PSU at the time of selection, the PSU was known as a self-representing (SR) PSU because it was removed from the probability sampling operation; it was in sample with certainty. The sample from the PSU represents only that PSU. There are 170 SR PSUs.

The Census Bureau grouped the remaining PSUs and selected one PSU per group, proportional to the number of housing units in the PSU, to represent all PSUs in the group. These selected PSUs are referred to as non-self-representing (NSR) PSUs. The sample NSR PSUs for AHS are a subsample of the Current Population Survey's (CPS) sample areas based on the 1980 census.

**Selection of sample housing units.** The AHS sample consists of the following types of units in the sampled PSUs:

- Housing units selected from the 1980 census
- New construction in areas requiring building permits
- Housing units missed in the 1980 census
- Other housing units added since the 1980 census
- Housing units selected from the 2000 census

**Housing units selected from the 1980 census.** The Census Bureau picked a systematic sample so every unit had a 1 in 2,148 chance of being included in the AHS.

In areas where addresses were complete (at least 96 percent of units having a house number and street name) and permits were required for new construction, housing units that received 1980 census long-form questionnaires were sorted by the following items:

- PSU
- Central city, urbanized area, urban outside urbanized area, rural
- Owner, renter, vacant for rent, vacant for sale, other types of vacant housing units
- Number of rooms
- Value of housing unit or gross rent
- Manufactured/mobile home or not a manufactured/mobile home

In areas where addresses were not complete or permits were not required for new construction, land areas were sorted using a formula incorporating the following items:

- PSU
- Central city, urbanized area, urban outside urbanized area, rural
- Median value of housing unit

- Number of children under 6 years old
- Number of elderly people
- Number of owner-occupied housing units
- Number of manufactured/mobile homes
- Number of housing units lacking some plumbing
- Number of owner-occupied housing units whose value is below \$45,000
- Number of renter-occupied housing units with rent below \$200
- Number of Black and Hispanic people
- Number of 1-room housing units

**New construction in areas requiring building permits.** In areas that required building permits for new construction, the Census Bureau selected a sample of permits. These permits did not cover manufactured/mobile homes or conversion of buildings to residential use.

**Housing units missed in the 1980 census.** The Census Bureau conducted a special study which identified units at addresses missed or inadequately defined in the 1980 census. A sample of these identified units was selected.

**Other housing units added since the 1980 census.** If extra units were added in buildings or manufactured/mobile home parks where AHS already had sample units, a sample of these extra units was selected. To find when whole buildings were built (in addition to building permits, mentioned above) or were converted from nonresidential to residential use, the Census Bureau listed all residential buildings in a sample of areas around the country, found any additional buildings, and selected a sample of their units.

### **Housing Units Selected From Census 2000**

The following adjustments were made to the AHS-N sample in 2011 by adding certain types of units selected from Census 2000:

In 2005, a new sample of manufactured/mobile homes was selected from Census 2000 in an attempt to improve coverage of manufactured/mobile homes built between 1980 and 2000. These units were selected at the same rate as the 1980-based units. One-half of this sample was included in the 2005 interviewing and, as a result, one-half of the 1980-based sample was not included. The units interviewed in 2005 were also interviewed in 2007, 2009, and 2011.

In 2005, a sample of assisted living units was selected from Census 2000 in an attempt to improve coverage of the elderly. The Census Bureau purchased address lists of assisted living facilities from two vendors and supplemented these lists with facility addresses from various assisted living facility internet sites. These addresses were matched to Census 2000 to identify Census 2000 housing units in assisted living facilities from which the resulting sample of these units was drawn. Because the universe was so small, a double sample of units was selected. Before interviewing, the resulting sample was screened to confirm eligibility. There were 921 units in the initial sample with 486 remaining after screening. The units interviewed in 2005 were also interviewed in 2007, 2009, and 2011.

### Subsidized Housing Units

A sample of subsidized housing units was selected from address lists provided by HUD in an attempt to improve coverage of housing units receiving rent subsidies. The Census Bureau received address lists from HUD in 2010. These lists include the Public and Indian Housing Information Center (PIC), Tenant Rental Assistance Certification System (TRACS), and Home Investment Partnership Program (HOME) address lists. The lists were matched to the sample PSUs to enable the Census Bureau to select housing units receiving HUD subsidies. There were 5,259 units in the initial sample with 5,064 eligible for interview.

### SUPPLEMENTAL METROPOLITAN SAMPLE

In 2011, the Census Bureau supplemented the national sample in the Los Angeles metropolitan area. This metropolitan area was based on the 2003 OMB definition for the Los Angeles-Long Beach-Glendale Metropolitan Division.

Table 1 provides the size of the supplemental sample in the Los Angeles metropolitan area.

**Table 1. 2011 Supplemental Sample Size for the Los Angeles AHS-National-Based Metropolitan Area**

Metropolitan Area	Basic Sample	Supplemental Sample	Total Sample Size
Los Angeles, CA	1,884	2,708	4,592

**Interview Activity.** Table 2 summarizes the interview activity for the Los Angeles AHS-National metropolitan area. The table provides the response rate, number of eligible units (comprised of completed interviews and noninterviews), and the number of units visited but ineligible for interview.

**Table 2. Interview Activity of the Los Angeles AHS-National-Based Metropolitan Area**

Metropolitan Area	Unweighted Response Rate <sup>3</sup> (percent)	Weighted Response Rate <sup>4</sup> (percent)	Eligible Units			Ineligible Units <sup>5</sup>
			Total	Interviewed	Not Interviewed <sup>6</sup>	
Los Angeles, CA	85.1%	84.7%	4,482	3,813	669	110

3 The unweighted response rate is computed by dividing the unweighted number of interviews by the unweighted total number of cases eligible for interview and multiplying by 100.

4 The weighted response rate is computed by dividing the weighted number of interviews by the weighted total number of cases eligible for interview and multiplying by 100.

5 Sampled units were determined to be errors on the frame or structures that did not meet the definition of a housing unit.

6 Sample units were visited but occupants were not at home after repeated visits or were unavailable for some other reason.

## ESTIMATION FOR AHS-NATIONAL

Each housing unit in the AHS sample represents itself and over 2,000 other units. The exact number it represents is its "weight." The weight is calculated in six steps. The purpose of these steps is to minimize both sampling errors and errors from incomplete data. The result of these steps is also to force consistency with some major categories of data in other Census Bureau surveys, so figures on these categories do not actually depend on the AHS sample, but on the other surveys.

1. **Basic weight.** The Census Bureau assigned each unit a weight to reflect its probability of selection. With rare exceptions, the AHS-National weight is 2,148. One exception is the special living sample which is assigned a weight of 1,074.
2. **Sample adjustment.** An adjustment was made to the units to account for the introduction of the 2000-design cases, the addition of the supplemental sample in the Los Angeles metro area, and for the addition of the oversample of subsidized housing units. This adjustment was made to ensure the additional sample would not inflate the national housing unit estimates.
3. **Noninterview adjustment.** An adjustment was made for refusals and occupied units where no one was home. The calculations for this adjustment do not include units the Census Bureau could not locate. The earlier weight was multiplied by the following factor:

$$\frac{\text{Interviewed units} + \text{Units not interviewed}}{\text{Interviewed units}}$$

It was assumed the units missed are similar in some ways to the units interviewed for AHS.

This adjustment was done separately for groups defined by cross-classifying the following data items if prior year data for the indicated items are available:

- Nine Census divisions
- 1990 central city, suburb, or nonmetropolitan
- 1990 urban or rural
- Manufactured/mobile home or not a manufactured/mobile home
- In a special living facility or not in a special living facility (only if the housing unit is not a manufactured/mobile home)
- Owner/for sale or renter/for rent
- Number of units in structure<sup>7</sup>

---

<sup>7</sup> If known from a previous survey; otherwise, the Census Bureau substituted whether or not units were drawn from building permits for these items.

- Number of rooms<sup>7</sup>
- Occupied, vacant year round, or seasonal/migratory vacant<sup>7</sup>

For seasonal/migratory vacants and year round vacants other than those for rent or for sale, units were cross-classified only by census region and 2000 central city/suburb/nonmetropolitan.

4. **PSU adjustment.** The Census Bureau adjusted for differences that existed in 1990 between the number of 1990 census housing units estimated from the AHS sample of non-self-representing PSUs and the 1990 census counts outside the self-representing PSUs. The earlier weight was multiplied by the following factor:

$$\frac{\text{1990 census housing units in all areas that could have been chosen as non-self-representing PSUs}}{\text{1990 census housing units estimated from the AHS sample of non-self-representing PSUs}}$$

This adjustment was done separately for groups defined by cross-classifying:

- Four Census regions
- Owner, renter, or vacant
- 1990 central city, suburb, or nonmetropolitan
- 1990 urban or rural
- Hispanic or non-Hispanic householder (only in South and West regions)
- Black or non-Black householder (only in South region)

5. **New construction adjustment.** The Census Bureau adjusted for known deficiencies in sampling new construction by multiplying the earlier weight by the following factor:

$$\frac{\text{Independent estimate}}{\text{AHS sample estimate}}$$

This adjustment was done separately for groups defined by cross-classifying:

- Four Census Regions
- Manufactured/mobile home or not a manufactured/mobile home
- Number of units in structure
- Year built (five-year categories after 1980 as shown in the publication)

Independent estimates were based on the Census Bureau's Survey of Construction and Survey of Manufactured Home Placements. Independent estimates were not available for public housing built after 2010. These units did not get a new construction adjustment.

Note that final AHS figures for the categories above were not based on the AHS sample findings, but on the independent sources.

Units built before 1980 were adjusted such that, when combined with the units that received the new construction adjustment, their estimates conformed to externally-provided housing unit totals.

6. **Demographic adjustment.** Comparability among the surveys was ensured by multiplying the earlier weight by the following factor:

$$\frac{\text{Independent estimate}}{\text{AHS sample estimate}}$$

This adjustment was done in two steps for occupied units. First, the factors were computed and applied for the Hispanic or non-Hispanic groups defined by cross-classifying:

- Nine Census divisions
- Owner or renter
- Hispanic or non-Hispanic householder
- Husband-wife, other male householder, or other female householder
- Age of householder

Next, the demographic adjustment was repeated with the same cells, except classified by the Black or non-Black groups, rather than the Hispanic or non-Hispanic groups.

Vacant for sale, vacant for rent, other year round vacant, and seasonal/migratory vacant units were cross-classified only by the four census regions and 2000 central city, suburb, or nonmetropolitan area.

The percentage of occupied and vacant units was based on the AHS itself. The distributions within occupied and vacant units were from the Census Bureau's Current Population Survey for occupied units and from the Housing Vacancy Survey for vacant units. The total number of all housing units in the United States was based on the 2010 Census adjusted to account for new and lost units. Note that final AHS figures for the categories above were not based on the AHS sample findings, but on the independent sources.

**Repetitions.** The new construction and demographic adjustments were repeated to help match both sets of independent estimates simultaneously. These adjustments were repeated until every cell's factor is between 0.98 and 1.02 or the change in each factor from one repetition to the next is less than 0.015.

**Small cells.** In each step of weighting, many items were cross-classified; so some cells may have few cases. When a cell was too small (less than 20 cases for the noninterview adjustment or for the demographic adjustment) or the adjustment factor was too extreme (greater than 2 for

the noninterview adjustment or the demographic adjustment), the Census Bureau combined the cell with one or more other cells that are similar in most respects. Cells for the PSU adjustment or the new construction adjustment were not combined.

## REVISION HISTORY

The 2011 AHS-N was issued in September 2013. We refer to the weights, sample, and methodology used to calculate those weights as the *original* weights, sample, and methodology.

In 2016, the Census Bureau updated the 2011 weighting methodology by applying the same methodology used in 2013. We refer to the weights, sample, and methodology used to calculate those weights as the *revised* weights, sample, and methodology and this appendix describes the revised weights and methodology. We now review how the methodology changed from the original to the revised.

**Removed the AHS-Metropolitan Sample.** The original sample used combined the national sample with the Metropolitan sample. The revised sample only uses the National sample.

**Revised method for combining the HUD supplement with the National sample.** We improved the matching and general methodology for combining the HUD supplement with the National sample.

**Revised the New Construction Adjustment.** We changed the level of the adjustments to the Region from the Census Division. Previously, we used the regional totals to make synthetic estimates at the division level. The revised methodology uses the regional totals directly and avoids untested assumptions about the method of distributing the region totals to division totals.