2025 American Housing Survey Sample Design and Weighting

The 2025 American Housing Survey (AHS) is a longitudinal survey where sample housing units will be 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 2025 AHS, 183,500 sample housing units will be interviewed. 175,000 of these are for the longitudinal panel starting in 2025 and 8,500 comprise a bridge sample selected from responding housing units in the 2023 AHS to allow for break-in-series analyses.

AHS-N SAMPLE DESIGN

The AHS-N employs a multiple stage sample design. First, the U.S. was divided into areas consisting of counties or groups of counties and independent cities known as primary sampling units (PSUs). The sample for AHS includes just over 500 PSUs. These PSUs cover approximately 880 counties and independent cities with coverage in all 50 states and the District of Columbia.

The counties of the top 60 Core Base Statistical Areas (CBSAs), according to the 2020 projected OMB definitions, were included in the first-stage sample design with certainty. These PSUs are known as a self-representing PSUs because they represent themselves and no other PSUs.

The Census Bureau grouped the remaining PSUs into first-stage strata and selected one PSU per strata, proportional to the number of housing units in the PSU, to represent all PSUs in the strata. The selected PSUs are referred to as nonself-representing PSUs because they represent themselves and all of the other PSUs in the same strata. No first-stage strata crossed state boundaries in the new 2025 AHS sample design.

Allocation of sample housing units. The new AHS-N sample allocation scheme attempted to ensure high quality estimates of 20 CBSAs and 17 states, giving higher priority to the CBSAs when the two goals conflicted. To that end, the 2025 National Sample of 175,000 was first allocated proportionally to every state. Then 20 CBSAs were given enough sample to ensure that they had at least 3,000 and no more than 7,000 sample units. From here, some states "donated" sample to others: large states donated because they will produce good estimates with less than their proportionally allocated sample size, and small states donated because they will not produce good estimates without a much larger overall sample size. This pool of donor sample was then allocated to states that were deemed to viably produce reliable state-level estimates.

The AHS-N sample includes an oversample of 8,469 HUs in two HUD programs: public housing and private-project based (or multifamily) units.

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

- All valid housing units on the Census Bureau Master Address File as of January, 2024

The first stage consists of selecting one PSU from each strata based on probabilities proportional to their size as determined by the 2020 Census. The second and third stage sample selections are each systematic random samples from the same ordered list. A combination of geographic variables, tenure, and the identification of housing units participating in HUD programs was used to sort the list of HUs prior to selecting the sample. Two stages of systematic sampling were needed to first oversample the HUD HUs and then sample down to the desired level of 175,000 nationally. The bridge sample was selected systematically nationwide, sorted on the original sorting order of the 2015 design.

AHS-N ESTIMATION

Each housing unit in the AHS sample represents itself and a varying number of other units. The exact number it represents is its "weight." The weight was calculated in six steps and the purpose of these steps is to minimize both sampling errors and errors from incomplete data. The description that follows is largely consistent with the weighting methodology for the prior cohort interviewed from 2015 to 2023, although newer methods for noninterview adjustments will be employed. Research is continuing to improve weighting methodology writ large.

The final weight is the product of the weights and factors described below and was calculated in the order listed below. The calculation of each factor used all of the previously calculated factors.

1. *Base weight.* Each unit is assigned a weight to reflect its probability of selection. The basic weight accounts for the sample selection of PSUs in the first stage and housing units in the second stage. The base weight also accounts for any reductions or expansions of the sample.

2. *First-stage ratio adjustment.* The weights are adjusted for differences between the number of housing units for the complete set of nonself-representing (NSR) PSUs and the estimated number of housing units nonself-representing (NSR) PSUs where only the sample PSUs contribute to the estimate. This can be represented as:

$$\frac{\text{Number of housing units in all PSUs of the nonself-representing PSUs}}{\text{Estimated number housing units in all PSUs of the nonself-representing PSUs}}$$

   This adjustment is done separately for groups that have yet to be determined.

3. *Noninterview adjustment.* An adjustment will be made for noninterviews of occupied units. The calculations for this adjustment will not include units the Census Bureau could not locate. The earlier weight will be multiplied by the following factor:

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

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

This adjustment is done separately within cells that are related to the propensity to respond and the variable of interest. These cells will be constructed using estimated propensities to respond based on a logistic regression model with independent variables yet to be determined by planned research.

4.   *Adjustment to Independent Housing Unit totals.*   Independent totals of housing units will be used to adjust for known deficiencies by multiplying the earlier weight by the following factor:

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

This adjustment will be completed separately for groups that have yet to be determined by planned research.

The known totals will be based on the Census Bureau's Survey of Construction and Survey of Manufactured Home Placements. Note that final AHS figures for the categories above are not really based on the AHS sample findings, but on the independent sources.

5.   *Adjustment to Independent Population Totals.* Comparability of population estimates among the surveys will be ensured by multiplying the earlier weight by the following factor:

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

This adjustment will be completed separately for groups that have yet to be determined by planned research.

Repetitions. The new construction and demographic adjustments will be repeated to help match both sets of independent estimates simultaneously.

*Small groups***.** In each step of weighting, many items were cross-classified; so some groups may have few cases. When a group is too small (less than 30 cases) or the adjustment factor is too extreme (greater than 2.0 for the noninterview adjustment or outside a range of 0.5 to 2.0 for the demographic adjustment) the Census Bureau combined the group with one or more other groups that are similar in most respects. These criteria will undergo some further research as well to fine tune their optimality.

AHS VARIANCE ESTIMATION

Replicate weights will be calculated for both the AHS-N and will be used to calculate both the estimates provided in publications and be made available to the public. The AHS-N replicate weights will use Balanced Repeated Replication estimate the variation due to the selection of PSUs in the first-stage and Successive Difference Replication (SDR) to estimate the variation due to the selection of HUs within each of the self-representing PSUs.