Redesigning the American Housing Survey Sample: How and Why

By

Toni Warner

Aref Dajani

U.S. Census Bureau

Shawn Bucholtz

U.S. Department of Housing and Urban Development

U.S. Department of Housing and Urban Development | Office of Policy Development and Research

Redesigning the American Housing Survey Sample: How and Why

By:

Toni Warner
Aref Dajani
U.S. Census Bureau

Shawn Bucholtz
U.S. Department of Housing and Urban Development

March 2020

PREFACE

In preparation for redesigning the American Housing Survey (AHS) sample in 2015, the U.S. Department of Housing and Urban Development (HUD) and the U.S. Census Bureau, the survey's sponsor and implementer, respectively, evaluated several aspects of the AHS sample—including attrition rates, the representativeness of the longitudinal sample, and issues of changing geographic definitions. Two reports were prepared by HUD and the Census Bureau based on the AHS data between 1985 and 2009: (1) "Measuring Attrition in Long-Term Longitudinal Surveys," (Dajani et al., 2012) and (2) "The American Housing Survey 2015 Redesign: Impact and Analysis" (Warner and Dajani, 2011).

The two reports were written for internal HUD and Census Bureau discussions during the redesign process and, as such, contain references and allusions that would be unclear to non-AHS experts. The present document synthesizes the two reports to provide an overview of the decision process that HUD and the Census Bureau undertook before redesigning the AHS sample. Although it is faithful to the content and thrust of the original reports, this document aims to be more generally understandable by eliminating repetitive material and allusions that are not useful to most readers. In the present report, no additional analyses are conducted; this report merely synthesizes the findings of the two internal reports to provide AHS users with insight into the sample redesign decision process.

It should be noted that the two original reports were written in 2011, reflect the thinking at that time, and do not cover all aspects of the redesign. The two reports also do not always agree with each other; whenever appropriate, the present report alerts readers to differences in interpretation without taking sides or censoring either report.

All the analyses in the two reports use AHS data through 2009. This can confuse readers because by 2015, two more AHS surveys were available but were not used in the analyses cited in the two papers used to create this report. When updated data are readily available, the present report will note them in footnotes. The 2009 viewpoint also affects issues discussed in the reports. For example, the issue of budgetary attrition was very much in the minds of the HUD and the Census Bureau authors. Between the 2007 and 2009 AHS surveys, 4,991 cases were dropped from the sample to reduce survey costs; these units were added back to the sample in 2011, and the sample was further augmented with an additional 3,694 cases.

TABLE OF CONTENTS

LIST OF EXHIBITS

1. INTRODUCTION

The American Housing Survey (AHS), sponsored by the U.S. Department of Housing and Urban Development (HUD) and conducted by the U.S. Census Bureau, is a longitudinal survey of housing units that provides an ongoing picture of the U.S. housing stock. It also provides information on homeowners, renters, the condition of homes, the cost of living in them, amenities available, and types of heating/cooling and other equipment, along with data on renovation, maintenance, and repair. The survey is also one of the key sources of data on unoccupied homes as well as homes that are no longer part of the housing stock.

When first implemented in 1973, AHS was an annual survey with over 50,000 occupied and unoccupied housing units in its sample. At that time, it was called the "Annual Housing Survey." A new sample was drawn in 1974 that was in use until 1985; housing units in this sample were interviewed annually until 1981. In 1981, AHS became a biennial survey, and its name changed to the "American Housing Survey." Over 50,000 national housing units continued to be in the sample. Additional samples of 4,000 to 5,000 housing units were added for major metropolitan areas, although not all metropolitan areas were surveyed every other year.

In 1985, HUD and the Census Bureau conducted a major sample redesign. The new sample remained a longitudinal panel; that is, housing units once sampled remained in the sample for each iteration of the survey. An additional sample of newly constructed housing units was added to the AHS national (AHS-N) and metropolitan (AHS-MS) samples for each iteration to ensure that samples remained representative of the housing stock. The 1985 sample with supplements of newly constructed housing units for each iteration was in use for almost 30 years. In 2015, HUD and the Census Bureau redesigned the survey and drew a new sample. This report provides an overview of the decision-making process for redesigning the AHS sample in 2015, including how HUD and the Census Bureau arrived at the 2015 sample.

2. THE 1985 AMERICAN HOUSING SURVEY SAMPLE

In 1985, HUD and the Census Bureau implemented a longitudinal sample design for AHS in which selected housing units remained in a national sample over time. This evolved into a 28-year panel, with final interviews conducted in 2013.

The sample was supplemented by units that were selected to account for new construction and to oversample selected populations. Although most surveys conducted by the Census Bureau focus on the householder, AHS focused on the housing unit; when a householder in an AHS housing unit moved, the in-movers to the unit were interviewed during the next iteration of AHS. Out-movers from the housing units were not interviewed.

The sample frame for the 1985 sample included housing units that were identified during the 1970 and 1980 Decennial Censuses and was supplemented over time with those newly identified during the 1990 and 2000 Decennial Censuses. Housing units identified from new construction permit data were also included in the sampling frame.

This design served a wide variety of stakeholders, including those interested in:

- The **up-to-date status of the housing market**, who would look at 1 year of cross-sectional data.
- **Incremental change** (that is, 2 years apart for national and up to 5 years for metropolitan areas) in the housing market, who could link successive interviews through unique housing unit identifiers in a Public Use File (PUF) specific to its year of interview.
- **Change over a longer period**, who could link files from many years apart for housing units that had completed interviews in both the beginning and end of the period of interest.
- The **dynamics of activity in the housing market**, who could link all PUFs and select households that had completed every interview from the beginning and end of the period of interest.

2.1 Sample Frame, Sample Selection, and Interview Eligibility

In order to understand the issues of attrition, representativeness, and budgetary constraints that influenced the decisions for the new 2015 sample, it is necessary to understand how the AHS longitudinal

sampling frame was structured. The cumulative sample included all housing units (often referred to as cases) that have been visited since 1985. As of the 2009 AHS, 102,594 cases were part of the cumulative sample; that is, this is the total number of distinct cases when each PUF from 1985 through 2009 are combined.[1] These units were sampled at some point between 1985 and 2009.

For the 2009 sample, 62,163 of the total 102,594 housing units, or 61 percent, were selected; this number is the result of the removal of prior demolished units, oversamples that are no longer of interest,[2] and budgetary restrictions. Of those units, 59,555 housing units were selected for interview in 2009, with the resulting 3-percent loss due to recently demolished units or units under construction; of these, 53,488 completed interviews. Therefore, the unweighted nonresponse rate was (59,555 - 53,488) / 59,555, or 10.2 percent.

2.2 American Housing Survey Completed Interviews

AHS interviews can be completed with respondents for units:

1. That are **occupied**.

2. Where the **resident usually lives elsewhere** (such as second homes and seasonal homes).

3. That are **vacant** (that is, they are currently unoccupied but are not under construction of any kind and are in a condition that allows for being occupied).

Surveys often cannot be completed for some units in the sample for a variety of reasons.

2.3 American Housing Survey Non-Interviews

The reasons that an AHS interview cannot be completed are classified into three broad groups:

- **Type A:** These are housing units where an interview could not be completed because of nonresponse, such as refusal or because a respondent was not available. These units remain eligible to be interviewed in subsequent iterations.

- **Type B:** These are not eligible for interviewing because the units are not in the housing stock for that specific survey cycle (for example, because they are still under construction). These units may become eligible to be interviewed in subsequent cycles.

- **Type C:** These units are determined to have permanently left the housing stock due to demolition or other reasons and are, therefore, permanently ineligible to be interviewed.

The complete list of the reasons why an AHS interview could not be completed is provided in Appendix C.

2.4 Attrition in the 1985 Sample

The 1985 AHS longitudinal sample was in place for nearly three decades, and although the frequency of interviews for any one housing unit is only once every 2 years, it is plausible that a household occupying a housing unit may not want to participate in AHS during each survey cycle. **The rate at which housing units leave the sample permanently or temporarily is known as the *attrition rate*.**

Attrition rates are measured in a variety of ways, and attrition measurements can produce different results. For the purposes of assessing the AHS attrition rates, **three measures of attrition were calculated in this study's previous report, "Measuring Attrition in Long-Term Longitudinal Surveys.":**

1. Permanently ineligible attrition.

2. Construction/conversion/prohibited occupancy attrition.

3. Nonresponse attrition.

Attrition could also occur because of administrative reasons, such as budget cuts.

Measuring the historical attrition rate for the 1985 sample helps determine the degree to which attrition may be a challenge for future survey design; HUD and the Census Bureau analyzed the AHS data from 1985 to 2009 to evaluate the attrition in the survey sample.

[1] By 2013, the last year that units from the sample starting in 1985 were interviewed, the national AHS sample included 125,049 cases.

[2] These oversamples included 10,299 units located next to AHS regular units in 1985, 1989, and 1993; 6,736 units oversampled in six large metropolitan areas in 1995, 1999, and 2003 ("the Big Six"); and a 6,351 rural oversample from 1987 and 1991.

2.4.1 Permanently Ineligible Attrition (Prior or Current Determination of AHS Type C Non-Interview)

Housing units can be permanently removed from the housing stock for reasons such as demolition or relocation of the unit or disaster loss. These units are deemed permanently ineligible for the AHS sample. On average, AHS experiences permanent loss of housing units at a rate of 2.0 percent per survey year. There is nothing inherent in the design of AHS that causes this type of permanent attrition, however. AHS addresses permanent attrition of housing units by adding newly constructed units to the sample for each survey iteration. This helps ensure that the sample remains representative of the housing stock. It must also be noted that this type of attrition is important to AHS and to other Census Bureau statistics because it serves as a measure of housing stock loss, and it has been used in Census Bureau population estimate calculations.

2.4.2 Construction/Conversion/ Prohibited Occupancy Attrition (AHS Type B Non-Interview)

At the beginning of a survey cycle, newly constructed units are added to the AHS sample. The Census Bureau selects newly constructed units by determining where building permits have been issued. In some instances, the unit is not habitable because construction has not started or is not complete at the time of survey; even though these cases are eligible for the sample, they are deemed ineligible for an interview for the current survey year. They are eligible to be included in the sample in subsequent years, however.

In other instances, **a housing unit may have been converted to an institutional unit, or occupancy is either determined to be prohibited or not possible because the housing unit's interior is exposed to the elements**. Similar to incomplete newly constructed units, these housing units are deemed ineligible for an interview for the current survey but eligible for inclusion in the sample in subsequent years if they return to the housing stock.

These housing units are designated as Type B Non-Interviews. This type of attrition averages approximately 2.9 percent per survey year and is an important measure for understanding temporary changes to the housing stock, especially considering housing market fluctuations. This type of attrition is also unavoidable for housing unit surveys, but it is not likely correlated with the survey design.

Further analysis is required to determine whether the probability of a housing unit being Type B for a given survey is significantly increased if it was previously determined to be a Type B.

2.4.3 Nonresponse Attrition (AHS Type A Non-Interview)

In AHS, some housing units are designated as "Type A" Non-Interviews (referred to hereafter as nonresponse) for the survey year; they are, however, eligible to be included in the sample in subsequent years. As noted earlier, the reasons for Type A nonresponse are sub-classified in AHS and include:

- Housing unit could not be located.
- Householder could not be located.
- Householder refused to participate.
- Language barriers prevented interview from taking place.

Virtually all surveys suffer from some level of nonresponse. Often, the most desirable type of nonresponse is random across the population, meaning it is not correlated with survey questions or any particular sub-population. In cross-sectional surveys, random nonresponse means a smaller sample, which in turn reduces the statistical precision of the estimates but does not cause population estimates made from the sample to become biased. It can reasonably be expected that AHS will suffer from nonresponse during any particular survey.

2.4.3.1 Persistent Nonresponse Attrition

AHS is a longitudinal survey; therefore, **nonresponse in one survey year could be related to nonresponse in a previous survey year and thus be classified as persistent nonresponse attrition**. One reason this may occur is that a householder refuses to participate in the current and future surveys because the previous survey was too long or the respondent feels the survey is administered too frequently (that is, respondent fatigue). Another reason is that respondents may not be home during the times of the year when the enumerator visits to administer the survey; this could happen each survey year. An additional reason is because the householder refuses to participate in any survey. Regardless of the

reasons, if the nonrespondent remains in the housing unit and the unit remains in the survey, the impact of persistent nonresponse attrition is a reduction in the longitudinal sample size beyond what might be expected due to random nonresponse, where current nonresponse is unrelated to prior survey nonresponse.

There are equally valid reasons why nonresponse rates for housing units that remain in the survey might be lower than what would be expected by random chance. For instance, if the survey is well designed and relatively easy for the respondent to complete, their prior experience may have been positive, so they may be likely to complete the survey in a subsequent year. It is worth noting that households in the AHS sample units may change over time. As such, persistent nonresponse for a particular housing unit may end if a new householder moves into the unit.

Although the reasons for negative and positive influences on nonresponse rates are important, what matters most is the impact of nonresponse rates on the overall panel size. The following three subsections provide outcome measures for persistent nonresponse attrition. The outcome measures used are single-year nonresponse rates, consecutive-year nonresponse rates, and intermittent attrition for the panel.

2.4.3.2 Single-Year Nonresponse Rates as a Measure of Persistent Nonresponse Attrition

Most current users of AHS create cross-sectional estimates for a single year and are mainly concerned about the single-year nonresponse rate. Measuring the nonresponse rate for the entire sample and comparing it to the subset of cases that have been in the AHS-N sample since 1985 can help determine whether the length of time in the survey has an impact on nonresponse rates.

Exhibit 1: Single-Year Nonresponse Rates For All Interviewed Housing Units

Year	Unweighted (%)	Weighted (Standard Error) (%)
1985	4.2	3.9 (0.13)
1987	3.2	3.2 (0.13)
1989	4.2	4.1 (0.12)
1991	4.4	4.5 (0.11)
1993	4.1	4.3 (0.12)
1995	7.5	7.1 (0.13)

Exhibit 1: Single-Year Nonresponse Rates For All Interviewed Housing Units (cont.)

Year	Unweighted (%)	Weighted (Standard Error) (%)
1997	9.9	9.8 (0.17)
1999	9.9	9.2 (0.15)
2001	10.0	9.9 (0.16)
2003	9.2	8.7 (0.14)
2005	10.8	10.9 (0.17)
2007	12.4	12.4 (0.17)
2009	10.2	10.5 (0.16)

Exhibit 1 shows that nonresponse rates were stable at about 4 percent between 1985 and 1993, increased to 10 percent in 1997, and have remained stable since then. This pattern is consistent with other surveys that experienced nonresponse rate increases in the late 1990s (Atrostic et al., 2001). It is also important to note the unweighted and weighted nonresponse rates,[3] when different, are not vastly dissimilar, indicating that the patterns of nonresponse are not very different across strata.

Exhibit 2 presents the AHS Type A unweighted and weighted nonresponse rates for a subset of housing units: those selected for interview each year that were part of the original sample selected in 1985. A comparison of exhibits 1 and 2 reveals a similar pattern of nonresponse rate increases. This similarity suggests that housing units in the sample since 1985 are no more likely to have higher nonresponse rates than housing units added to the sample in later years. Combining this finding with the known increase in nonresponse rates in other surveys, it is reasonable to conclude that response fatigue does not appear to be a factor in nonresponse rates.

This imperfect measure of nonresponse attrition due to response fatigue does not measure the nonresponse rate for households that have been in the survey for multiple years. Although this type of analysis is certainly possible for AHS, it is beyond the scope of this analysis. What matters, however, is the impact of persistent nonresponse attrition, and the magnitude does not appear very high.

[3] The AHS data are weighted by geography and various housing and population characteristics to ensure representativeness. More information on the weighting is available at https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/identifying_samples.html.

Exhibit 2: Single-Year Nonresponse Rates For Units Sampled in 1985

Year	Unweighted (%)	Weighted (Standard Error) (%)
1985	4.2	3.9 (0.13)
1987	3.1	3.3 (0.13)
1989	4.3	4.1 (0.13)
1991	4.1	4.5 (0.11)
1993	3.9	4.2 (0.12)
1995	7.5	7.3 (0.15)
1997	9.6	9.7 (0.18)
1999	10.2	9.1 (0.16)
2001	9.6	10.0 (0.19)
2003	9.2	8.7 (0.16)
2005	10.7	10.8 (0.20)
2007	12.1	12.5 (0.21)
2009	10.7	10.6 (0.20)

2.4.3.3 Consecutive-Year Nonresponse Rates as Measure of Persistent Nonresponse Attrition

One of the popular uses of the longitudinal structure of AHS is to measure near-term changes in housing units by using data from consecutive surveys. One example involves measuring the income difference between households that were interviewed in 2007 and 2009; another would be using Components of Inventory Change (CINCH) analysis to determine housing stock changes (Eggers and Moumen, 2011). For near-term changes, it is important to have completed interviews for a pair of consecutive years, but it is not necessary to have completed interviews for all years.

Exhibit 3 presents the nonresponse rates for pairs of consecutive years; it shows that nonresponse rates for pairs of consecutive years have been between 15 and 18 percent since 1997. Most of the joint nonresponse rates can be attributed to nonresponse in one of the survey years. In addition, the pattern in Exhibit 3 mimics the pattern in Exhibit 2 wherein nonresponse rates began to increase in 1995 but have been steady since then, with a minor uptick in 2007.

Exhibit 3: Nonresponse Rates for Housing Units Selected for Interviews in Consecutive Years

Year	Unweighted (%)	Weighted (Standard Error) (%)
1985 & 1987	5.9	5.9 (0.20)
1987 & 1989	6.0	6.1 (0.20)
1989 & 1991	7.0	7.0 (0.20)
1991 & 1993	7.1	7.1 (0.13)
1993 & 1995	9.5	9.4 (0.17)
1995 & 1997	13.7	13.2 (0.18)
1997 & 1999	15.4	15.2 (0.20)
1999 & 2001	15.6	15.0 (0.18)
2001 & 2003	15.2	14.8 (0.20)
2003 & 2005	15.7	15.2 (0.18)
2005 & 2007	17.9	17.9 (0.21)
2007 & 2009	17.8	17.9 (0.20)

2.4.3.4 Intermittent Attrition for the Panel as a Measure of Persistent Nonresponse Attrition

A small number of AHS users require the longitudinal panel to have a completed interview for most or all years for which the housing unit has been in the sample. These AHS users are typically measuring spells of activity or performing event-history analysis, such as tracking "underwater" mortgage status[4] over time. An AHS user's ability to conduct longitudinal analysis with the full panel may be impacted by missing data in 1 or more years. Nonresponse (that is, missing interviews) for 1 or more years is known as intermittent attrition.

Exhibit 4 presents cumulative and average intermittent attrition rates for the 60,932 housing units that entered the sample in 1985 or later and remained in the sample through 2009, including 35,053 housing units that entered the sample in 1985.

[4] "Underwater mortgage status is where the loan due on a house is greater than its market value.

Exhibit 4: Single-Year Nonresponse Rates for Units Sampled in 1985

Number of Survey Years in Sample	One or More Type A Nonresponses (%)	Two or More Type A Nonresponses (%)	Three or More Type A Nonresponses (%)	Average Number of Type A Nonresponses
13 (since 1985)	47	26	16	1.2
12 (since 1987)	52	29	17	1.3
11 (since 1991)	50	28	16	1.2
10 (since 1993)	48	27	15	1.2
9 (since 1995)	44	25	13	1.0
8 (since 1997)	44	24	15	1.1
7 (since 1999)	42	20	9	0.8

Exhibit 4 reveals several important intermittent attrition rate outcomes for AHS:

- Fewer than one-half of the housing units (47 percent) that have been in the survey from 1985 to 2009 have nonresponses. Therefore, 53 percent of the housing units have zero nonresponses.

- Of the housing units that have been in the sample for the 13 surveys (1985 through 2009), only 16 percent have three or more Type A nonresponses. Therefore, 84 percent of housing units in the sample since 1985 have at least 11 completed interviews.

- The average number of Type A nonresponses—which could be expected to rise as the longitudinal sample grows longer—appears to level off at 1.2 to 1.3 times for respondents missing interviews after being in the sample for at least eight interviews.

2.4.4 Administratively Ineligible Attrition

There are administrative reasons why housing units may be deemed ineligible for the sample for a given year, including budgetary reductions and special interest samples. Like many federal programs, AHS has been subject to budgetary fluctuations throughout its history, and this has resulted in portions of the sample not being interviewed for a given year. For instance, 8 percent of the 2005 sample was not interviewed in 2007 and 2009 due to budgetary restrictions.

- **Additional housing units have been added to the AHS-N sample at various times for special purposes**, including oversampling metropolitan areas, improving manufactured housing coverage, and oversampling HUD-assisted renters. **Although these housing units have not been removed from the sample, many have been interviewed intermittently**.

- Administrative attrition is an unfortunate but common characteristic of many surveys, but it is not related to survey design. That being said, HUD and the Census Bureau can do a better job of identifying specific reasons for administrative attrition in the PUF so that analysts can better understand the historical nature of administrative attrition.[5]

2.4.5 Absolute Attrition Rate

Up to this point, this synthesis of the two papers on AHS sample attrition and the sample redesign for 2015 has relied heavily on Dajani et al 2012, but from this point on, the synthesis will rely more heavily on Warner and Dajani 2011. In discussing attrition, the two papers both use permanent attrition, but they define it differently. Exhibit 5 uses the term as defined in Warner & Dajani 2011 (page 6)[6]:

"**Permanent Attrition:** In this measure, attrition is defined as **units that were selected into the sample in 1985 and achieved a 'non-interview' status at some point**, remaining so through 2009. The two types of 'non-interview' statuses are:

[5] In 2011, the Census Bureau released a file called the Sample Case History File (better known by users as the "Where Did They Go" file) that gives the history of every case that has been in the AHS-N sample.

[6] As Warner & Dajani go on to say, "This measure of attrition provides a conservative estimate. In each case, these sample units are not replaced with new samples but are followed from time period to time period. This is a burden on the field representatives but not on the respondents. Although there is interest in knowing how large this group is, no information is gathered from these units after this attrition occurs."

- Type A–refusals and housing units that field representatives are unable to locate.

- Type B–housing units under construction or an empty mobile home pad."

Exhibit 5: AHS-N Absolute Attrition Rates, 1985–2009

Year	Permanent (%)	Survey Year-Specific (%)	Intermittent (%)
1985	0.3	5.5	5.5
1987	0.4	8.1	8.1
1989	0.5	9.6	11.1
1991	0.6	10.3	14.1
1993	0.7	10.2	16.3
1995	0.9	13.1	20.6
1997	1.0	15.6	25.7
1999	1.3	15.3	29.8
2001	1.5	16.0	33.9
2003	1.8	15.3	36.9
2005	2.4	17.6	40.8
2007	3.4	20.0	45.2
2009	6.2	19.3	48.4

The other two terms used in Exhibit 5 are defined similarly as to how they are used in the Attrition paper. The Impact and Analysis paper calculates these rates based on the cases in the original 1985 sample, excluding cases added after 1985 to represent new units in the sample.

AHS absolute attrition rate analysis indicates that permanent loss to the 1985 sample from its inception to 2009 has been 6.2 percent—this attrition rate remained under 1 percent until 1997 (12 years into the sample), crested at 3 percent in 2007 (22 years into the sample), and then almost doubled to over 6 percent in 2009. Perhaps more importantly, nearly two-thirds of the attrition occurred after 2005, suggesting that after 20 years, attrition could be expected to increase at a higher rate.

The survey-year-specific attrition rates are considerably higher than the permanent attrition rate, both in absolute and rate-of-change terms. The absolute value exhibited periods of stability, however, from 1989 to 1993, from 1997 to 2003, and from 2007 to 2009; between these periods of stability are large increases.

The intermittent attrition rates are much larger than the permanent and survey-year-specific attrition rates. In absolute terms, nearly one-half of the sample selected in 1985 could not be surveyed in at least 1 year between 1985 and 2009. The rate of change in attrition was consistent between 1985 and 2009, increasing by nearly 4 percent each additional survey year.

This attrition rate analysis evaluated the types of attrition unique to AHS and focused on the types of attrition that may be, or likely are, related to survey design.[7] The findings of this analysis suggest that while attrition is not a serious problem, it is still a challenge for AHS, as single-year Type A nonresponse rates have stabilized at around 10 percent.

3. CHALLENGES PRESENTED BY THE 1985 SAMPLE

Over time, several challenges emerged in continuing to use this sample. Key among these challenges is the increasing attrition rate. In addition, changes in the definition of geographic areas, such as the definition for metropolitan areas, began to present increasing challenges to following data disclosure rules and thus preserving the privacy and anonymity of respondents. It also became increasingly difficult to ensure the representativeness of the sample. Finally, budgetary variations often led to variations in sample sizes that could impact the power of the data.

3.1 Challenge #1: Attrition Rates

AHS attrition rate analysis indicates that permanent loss to the sample (6.2 percent over 24 years, as shown in Exhibit 5 for the original 1985 sample) has been low over the 24 years since the 1985 design. Perhaps more importantly, nearly two-thirds of the attrition occurred after 2005, suggesting that after 20 years, attrition could be expected to increase at a higher rate.[8]

[7] The "*Measuring Attrition in Long-Term Longitudinal Surveys*" report argues in several places that observed attrition is not unique to AHS.

[8] This sentence suggests that, if HUD and the Census Bureau were to choose a sample design using the same sample for 20 years or more, one could expect respondent fatigue to increase Type A nonresponse as time goes on. The Measuring Attrition report casts doubt on the respondent fatigue argument. This difference is mentioned because the following discussion of design options is taken from the Impact and Analysis report and mentions respondent fatigue frequently.

3.2 Challenge #2: Geographic Indicators and Their Impact on Disclosure

The 1985 AHS sample was of sufficient size given the disclosure rules used during the period studied to permit sub-national estimates, including for census regions; urban and rural areas; Office of Management and Budget (OMB)-defined metropolitan areas;[9] and, in some cases, HUD-defined sub-areas within metropolitan areas.[10] The Census Bureau included various geographic identifiers in the AHS-N and AHS-MS PUFs.

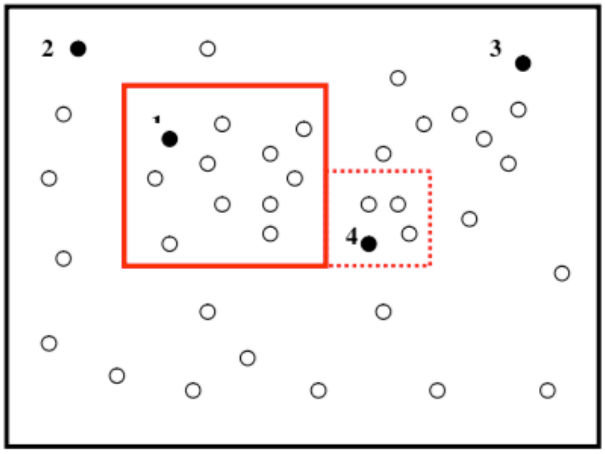
The metropolitan areas identified on the AHS-N PUF for the 1985 AHS sample are defined using the 1983 OMB definitions and are based on 1980 geography, and they have not been updated in AHS since 1985. This is one of the biggest issues for AHS-N users when they are using public use data.[11] Other data sources offer estimates for metropolitan areas using updated definitions and geography, such as the OMB's 2003 definitions. This makes it especially hard to link and compare AHS to other survey data.

Census data are released under Title 13 of the U.S. Code, which prohibits wrongful disclosure of information on individuals. To make the identification of individuals highly unlikely, the Census Bureau does not identify geographic areas with fewer than 100,000 people in PUF (Hawala, 2001). Because there are already several geographic identifiers on the AHS-N PUFs, each new geographic identifier added has the potential to violate disclosure standards, which is one of the reasons why AHS-N PUF metropolitan area definitions have not been updated since 1985.

Exhibit 6 provides a visual representation of a disclosure issue that would arise if a new geographic identifier was added to the PUF. The black box represents a state, and each circle represents an AHS sample point. Within the state, there is a 1983 metropolitan area represented by the red box. The metropolitan area has 150,000 people, so AHS sample points within the red solid box included an identifier for the metropolitan area on the 1985 AHS PUF.

Points outside the box were identified as not being in a metropolitan area. There is no disclosure violation under this scenario because the metropolitan area includes 150,000 people.

Exhibit 6: An Illustration of the Impact of Changing Geography on Disclosure



Point	Metro 83	Metro 03
1	Yes	Yes
2	No	No
3	No	No
4	No	Yes

Suppose, however, that the metropolitan area was expanded in 2003 to include the area denoted by the red dashed line. This area includes 25,000 people. If the 2003 AHS were to include a new geographic identifier for the 2003 metropolitan areas (in addition to the identifier for the 1983 metropolitan areas) in PUF, sample point #4 could be identified geographically as being in an area with fewer than 100,000 people.[12] This identification on the 2003 AHS PUF would then violate Census Bureau disclosure rules.

The consequence of the Census Bureau disclosure rule is that updated geography is not typically included in the AHS PUFs. In limited circumstances, such as the AHS-MS, some updated geography is provided.

The number of disclosure issues created by the inclusion of geographic identifiers on PUF is directly related to the number of geographic identifiers included. Every new geographic identifier included on

[9] https://www.bls.gov/bls/omb-bulletin-17-01-revised-delineations-of-metropolitan-statistical-areas.pdf

[10] These sub-metropolitan areas were defined by HUD and use a combination of Census Urban Area designations and expert opinion.

[11] The metropolitan areas included in the current AHS-MS PUFs were defined using either the 2003 OMB definitions (based on 1990 geography) or HUD-specific definitions that may not coincide with OMB.

[12] An in-depth explanation of adjustments is provided in Appendix B of the American Housing Survey for the United States: 2009, available at http://www.census.gov/prod/2011pubs/h150-09.pdf.

PUF, whether it is a new type of geography or a new vintage of an existing geography, must be intersected with each existing geographic identifier to ensure that small areas cannot be identified.[13]

One possibility for the 2015 sample design is to reduce the number of geographic identifiers on PUF and replace them with other geographic identifiers of interest to the AHS user community, including identifying counties or groups of counties and adding new metro area definitions when they become available.

3.3 Challenge #3: Maintaining a Representative Sample

A key to any successful survey design is ensuring that the survey sample is representative of the population under study. This is especially true for AHS' longitudinal survey design, in which the sample was drawn many years ago and sample replacement is small. If key conditions change in the housing universe and new samples are not added at a rate or type that ensures that the total sample is representative of the universe, then conclusions drawn from the sample may be biased.

The Census Bureau addresses this challenge in two ways. First, as mentioned previously, the 1985 AHS-N and ASH-MS samples are supplemented with new housing units over time to account for new construction. Although the supplemental samples are small, adding newly constructed housing may help ensure that the AHS sample reflects the housing universe.

Second, and more importantly, the Census Bureau performs several weighting adjustments, including non-interview, primary sampling unit, new construction, and demographic adjustments. Each of these adjustments are based in part on measures from other Census Bureau surveys, including the decennial census, current population survey, housing vacancy survey, survey of construction, and manufactured homes survey. The aggregate impact of the weighting adjustments should be a sample that is representative of the current housing universe.

To assess the impact of the weighting adjustments, comparisons can be made between variables that are present in both AHS and other surveys, assuming that the variables are derived from questions that are the same or substantially similar. Schwartz (2011) conducted such an analysis, comparing estimates from the 2007 AHS and the 2007 American Community Survey (ACS) for a selected set of variables that were derived from the same or similar questions. The 2007 ACS includes a much larger and more current sample than AHS, as the ACS sample is not longitudinal but rather is drawn freshly for each iteration. Although there are important differences between the two surveys, especially in the manner in which occupancy is measured, they are similar enough to draw general conclusions about how AHS compares with a larger and more recent survey.

Schwartz took great care to detail why differences in AHS and ACS may exist, including how particular questions are asked and how choice sets are formulated. It was beyond the scope of Schwartz's study, however, to parse the reasons behind those differences into categories, such as differences in questions, answer sets, survey timing, or sample bias. It was also beyond the scope of Schwartz's study to draw definitive conclusions about sample bias in AHS.

An examination of the results, in the context of Schwartz's explanation about each of questions and their corresponding answer sets, may lead the researcher to conclude that there is no systematic sample bias in AHS. This conclusion may be supported by three pieces of evidence from Schwartz's analysis:

- Estimates from five variables (Units in Structure, Number of Bedrooms, Home Value, Year Householder Moved into Home, and Age of Householder) show an inconsistency of statistically significant differences between AHS and ACS among the categories of answers. For example, Units in Structure, Number of Bedrooms, and Home Value are statistically significantly different in the two surveys, but the Year Household Moved into Home only shows small differences, with the median being the same in both surveys, and the Age of Householder is not statistically significantly different across the two surveys. Furthermore, there is no definitive pattern in the differences. These five variables likely represent the most straightforward questions and, therefore, are the least likely to be impacted by differences in how or when the questions were asked.

- Estimates from the Year Built variable show statistically significant differences in the AHS

[13] The Census Bureau has identified the sub-county variables METRO3 and ZONE as causing the most disclosure-related problems. This is due to the multiple types of geographies used to build each of these variables, including new vintages for each geography. For instance, METRO3 is composed of metropolitan areas, Census Bureau urban areas, and Census Bureau central cities.

and ACS estimates for all categories. As Schwartz indicated, there appears to be a slight bias in AHS toward older homes. The differences between categories are very small, however, and the difference in this variable's median values for AHS (1973) and ACS (1974) is also small.

- Estimates from the Rooms in Structure variable show statistically significant differences in the AHS and ACS estimates for nearly all categories. As Schwartz indicated, there appears to be a slight bias in ACS toward very small and very large homes, but the differences are small. Furthermore, the differences may be attributable to minor differences in what AHS and ACS consider to be a "room" for the purposes of the question.

3.4 Challenge #4: Achieving an Appropriate Sample Size in an Uncertain Budget Environment

The challenge of achieving an appropriate sample size is more often a matter of budget than of sample design or survey administration. Nevertheless, in a complex survey like AHS, it is important to establish goals for the survey, determine the sample size required to meet those goals, and evaluate survey outcomes to ensure that goals are being achieved. Moreover, the sample size is an important component when considering options for redesigning the 2015 AHS. An appropriate sample design is one that is flexible enough to maintain usefulness when budget challenges require sample size changes to be made.

A general goal in the AHS sample design is to produce estimates for 5-percent sub-populations at the national level that have high reliability and precision. HUD has defined high reliability and precision to be a percent coefficient of variation (%CV) of 5 percent.[14]

There are many variables within AHS that produce sub-populations of approximately 5 percent, and calculating the %CV for each of these variables is burdensome. Therefore, four variables that produce sub-populations were chosen, and %CV was calculated at the national and regional levels. The four variables are:

- Percent for rent/sale.
- Percent seasonal.
- Number of buildings with 5 to 9 units.
- Number of buildings with 10 to 19 units.

Exhibit 7 presents the results of the %CV calculations for various sample sizes. By design, %CV values decrease as the sample size grows. At the 1985 national sample size (approximately 65,000), the %CV values for the four selected sub-populations are near 5 percent. Therefore, it can reasonably be concluded that the 1985 national sample size meets the goal of the survey design.

[14] %CV is the ratio in percentage terms of the standard deviation to the mean of a variable.

Exhibit 7: Percent CVs for Selected AHS Sub-Populations

	For Rent/Sale				Seasonal			
	32,000	65,000	100,000	200,000	32,000	65,000	100,000	200,000
National	7.6	5.3	4.4	3.1	7.3	5.2	4.2	3.0
Census Regions								
Northeast	18.1	12.8	10.5	7.4	14.9	10.5	8.6	6.1
Midwest	15.1	10.7	8.7	6.2	16.6	11.7	9.6	6.8
South	11.5	8.1	6.7	4.7	11.7	8.3	6.8	4.8
West	18.0	12.8	10.4	7.4	16.0	11.3	9.2	6.5
	Buildings with 5 to 9 Units				Buildings with 10 to 19 Units			
	32,000	65,000	100,000	200,000	32,000	65,000	100,000	200,000
National	6.2	4.4	3.6	2.5	6.6	4.6	3.8	2.7
Census Regions								
Northeast	13.5	9.6	7.8	5.5	15.2	10.8	8.8	6.2
Midwest	14.2	10.1	8.2	5.8	14.8	10.4	8.5	6.0
South	10.2	7.2	5.9	4.2	10.2	7.2	5.9	4.2
West	12.3	8.7	7.1	5.0	13.8	9.7	8.0	5.6

It is worth noting, however, that the 1985 AHS-N sample size does not produce Census region-level estimates that meet the 5-percent %CV criteria. In fact, the AHS-N sample size would need to be increased three-fold to 200,000 before Census region-level estimates approach a %CV of 5 percent.

4. SAMPLE DESIGN OPTIONS CONSIDERED FOR THE 2015 AMERICAN HOUSING SURVEY SAMPLE REDESIGN

Several criteria were considered when evaluating which sample design best suits current and future stakeholders of AHS, including the ability to calculate cross-sectional and longitudinal estimates, the change from the current design, adaptability to changing geography and definitions, respondent fatigue, loss of housing units due to attrition of units, and an uncertain budget environment.

In general, purely longitudinal sample designs were deemed best at calculating long-term longitudinal estimates and involved the least amount of change from the 1985 AHS design. Purely cross-sectional (that is, one response) sample designs were deemed best at calculating cross-sectional estimates, were adaptable to changing definitions and conditions, did not suffer from attrition, and were robust to updating geography.

Rotating panels were thought to be best at calculating short-term longitudinal estimates. They also offered reasonable adaptability to changing definitions and conditions, suffered minimal attrition, and updated geography reasonably well.

Because of the wide variety of users of AHS data, and in order to consider challenges raised by the 1985 survey sample design, four sample design options were considered. Reliability and precision can be assessed by referring to exhibit 9 in chapter 5 for the predicted percent coefficients of variation for differing samples sizes.

4.1 Option #1–Full Single Panel

Total sample size of 205,000, with all units allocated in one single longitudinal panel (panel 00) for a panel lifespan of up to 30 years. The metropolitan oversampling design is included in appendix A.

Full Single Panel

Year	Panel
2015	00
2017	00
2019	00
...	00
2039	00
2041	00
2043	00

4.2 Option #2–Full Dual Panel

Total sample size of 205,000 units divided equally into two samples of 102,500 units. The two samples would be interviewed on alternating schedules, with 4 years between each interview (that is, Sample 1 interviewed in 2015, 2019, 2023, and so on; and Sample 2 interviewed in 2017, 2021, 2025, and so on). The metropolitan oversampling design is included in Appendix A.

Full Dual Panel

Year	Panel	Panel
2015	01	
2017		02
2019	01	
...		02
2039	01	
2041		02
2043	01	

4.3 Option #3–Hybrid Single and Rotating Panels

Total sample size of 205,000 units, with 100,000 units in the longitudinal panel and 105,000 units divided equally among five rotating samples of 21,000 units each. The metropolitan oversampling design is included in Appendix B.

4.4 Option #4–Rotating Panels

Total sample size of 205,000 units divided equally into five rotating panels of 41,000 units each. The metropolitan oversampling design is included in Appendix A.

Hybrid Single and Rotating Panels

Year	Panels										
2015	00	01	02	03	04	05					
2017	00		02	03	04	05	06				
2019	00			03	04	05	06	07			
2021	00				04	05	06	07	08		
2023	00					05	06	07	08	09	
2025	00						06	07	08	09	10
2027	00							07	08	09	10	11

Rotating Panels

Year	Panels										
2015	01	02	03	04	05						
2017		02	03	04	05	06					
2019			03	04	05	06	07				
2021				04	05	06	07	08			
2023					05	06	07	08	09		
2025						06	07	08	09	10	
2027							07	08	09	10	11

5. SAMPLE DESIGN OPTIONS DISCUSSED

HUD and the Census Bureau circulated the sample design options to a variety of data users and other stakeholders within the housing industry and convened a conference to discuss the options with data users and other stakeholders. Among the issues considered before selecting a sampling method for the 2015 redesigned sample were attrition rates, geographic identifiers, maintaining representativeness, and achieving appropriate sample sizes in an uncertain budget environment.

5.1 Attrition Rates

The evaluation of attrition rates in AHS showed that attrition rates increased considerably after 2005, or 20 years into the sample. Depending on your view of what is an appropriate attrition metric and acceptable attrition rate, you may find the attrition rate acceptable at a time prior to 20 years.[15]

Sample design options 1, 2, and the 00 panel of option 3 assume that housing units remain in the sample for up to 30 years. It may be reasonably expected that the trajectory of attrition rates in these two options will be similar to the trajectory of attrition rates found in the 1985 AHS sample, although that assumption cannot be confirmed until a future time.

Option 2 may result in a smaller attrition rate because housing units are only sampled once every 4 years. It should also be noted that options 1 and 2 could have a sample period of less than 30 years.

Sample design option 4 and the rotating panels in option 3 are in the sample for only 10 years. The attrition rate analysis shows the attrition rates after 10 years to be between 1 and 20 percent, depending on how attrition is measured. Attrition rates below 20 percent may be acceptable to most AHS users.

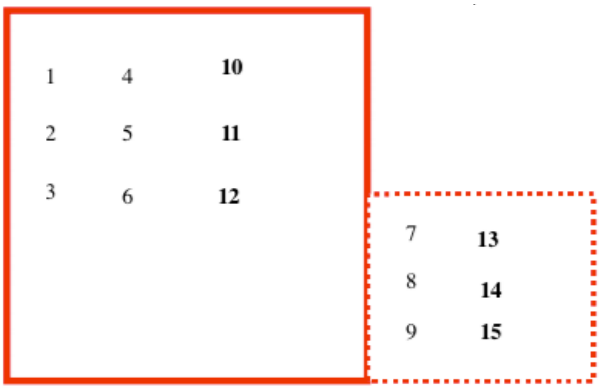
[15] Atrostic et al. (2001) noted that nonresponse has been increasing in federal surveys since the 1990s. Additionally, measurement of AHS attrition shows that several AHS attrition types are not related to the design of the survey (but instead "come with the territory") and that although attrition has been challenging, it has not been a particularly serious problem in AHS.

5.2 Geographic Identifiers

Unfortunately, each of the sample design options will suffer from disclosure issues caused by the inclusion of geographic identifiers. One strategy for dealing with disclosure issues, however, is to use a technique known as "salting," which is the deliberate introduction or removal of sample units in order to introduce inaccurate results. In the case of AHS, the salting technique could be used to remove observations that result in disclosure, as well as a random selection of observations that do not result in disclosure, and replacing them with an equivalent number of observations within and outside of the new geography.

Exhibit 8 shows how salting may be used to mitigate disclosure issues. In this example, observations 1 through 9 are part of the original sample, with 1 through 6 being inside a metropolitan area (denoted by the solid red box) and observations 7 through 9 being outside of the metropolitan area. Suppose that the 2010 metropolitan area (denoted by the solid red box) is expanded for 2020 by adding a new county (denoted by the dashed red box). Observations 1 through 3 are kept in the sample, observations 4 through 6 are randomly chosen to be deleted and replaced with observations 10 through 12, and observations 7 through 9 are deleted and replaced with observations 13 through 15. In this example, it is impossible to determine which observations are in the new part of the metropolitan area.

Exhibit 8: An Illustration of the Impact of Changing Geography on Disclosure



Sample Point	Metro 2010	Metro 2020
1, 2, 3	Yes	Yes
4, 5, 6	Yes	Deleted
7, 8, 9	No	Deleted
10, 11, 12	—	Yes
13, 14, 15	—	Yes

It is important to note the salting technique may not work well when more than two geographic identifiers are present. For example, in exhibit 8, Metro 2010 and Metro 2020 are identified. Suppose that a third geographic identifier, county, was included. If the 2010 metropolitan area was composed of one county, and the county being added for the 2020 metropolitan area was part of another group of counties, then the mere inclusion of the county identifier on the PUF would produce a disclosure violation.

To understand this, consider Exhibit 9, which is a modified version of the table from Exhibit 8 with an additional column for a county identifier. In this example, observations 13 through 15 are exactly identified as being in the new metro area and as part of county group 45. Anyone with a list of changes to the metropolitan areas between 2010 and 2020 could immediately determine that exact county where observations 13 through 15 reside because it has been disclosed that observations 13 through 15 belong to county group 45 and to a metropolitan area. This is a disclosure violation.

Exhibit 9: Illustration of Disclosure with Three Geographic Identifiers

Sample Point	Metro 2010	Metro 2020	County
1, 2, 3	Yes	Yes	County 1
4, 5, 6	Yes	Deleted	County 1
7, 8, 9	No	Deleted	County Group 45
10, 11, 12	—	Yes	County 1
13, 14,15	—	Yes	County Group 45

Although salting can mitigate the disclosure problems of changing geography, it also has the side effect of reducing the longitudinal sample. In the above example, only observations 1 through 3 would be present in succeeding surveys. Observations 10 through 15 could not be used for longitudinal analysis of changes since the previous survey, although they could be used for such analysis in future surveys.

None of the sample designs will alleviate disclosure issues. Disclosure issues are only alleviated if geographies selected at the beginning of the sample remain fixed in vintage and no new geography types are added.

5.3 Maintaining a Representative Sample

The Census Bureau took great care to ensure that the 1985 AHS sample was representative of the changing housing universe. Their strategy included selecting new construction into the sample as well as a robust weighting scheme that benchmarks certain variables to other data sources. Furthermore, it can reasonably be concluded that analysis comparing the 2007 ACS to the 2007 AHS does not reveal bias caused by an out-of-date AHS sample.

The sampling strategies for the various design options are very similar, with the rotating panel design being slightly superior because the sample is replaced at a greater rate and with benchmarks to current estimates. In addition, the Census Bureau will continue to use a weighting strategy benchmarked to the most recent data sources available to ensure that the sample continues to be representative of the housing universe.

5.3.1 Selecting New Sample for the Full Panel Designs (Options 1, 2, and the 00 Panel for Option 3)

The initial sample will be selected from the Master Address File (MAF) the first time, and new growth will be added.[16] The initial sample will be benchmarked to the 2010 Decennial Census and will be based on the following stratifications:

- Geography.
- Tenure.
- Type of structure.
- Number of rooms.[17]

New growth units added to the sample will be stratified only on geographic variables (that is, county, tract, block, and ZIP Code). It is expected that geographic stratification is sufficient to be representative of the types of housing units built since the last new growth sample was selected.

[16] In addition, it was initially planned that the subsidized housing sample will be refreshed every 2 years to include the most current units.

[17] The MAF does not identify number of rooms. This measure will be based on a proxy measure using average home values and rents determined from the 5-year ACS data.

5.3.2 Selecting Sample in a Rotating Panel Design (Options 3 and 4)

The initial sample will be selected in the same manner as the initial sample in the single-panel design (benchmarked to the 2010 Decennial Census). The differences in sampling occur during subsequent sample years.

In subsequent years, the newly selected sample will be benchmarked to the 2010 Census in terms of the type of housing unit (that is, manufactured home, one-unit structures, two- or more unit structures) and tenure (that is, owned, rented, vacant). These would then be benchmarked to the most current ACS data available in terms of home value and rent.

5.4 Achieving an Appropriate Sample Size in an Uncertain Budget Environment

The discussion of the AHS sample design options included consideration of how those options fared in response to an uncertain budget environment. In this context, the rotating panel design is superior to a single panel.

In the full-panel designs (single or dual), the sample size is chosen upfront. In the best case, the initial survey starts with a small sample size, then subsequently increases if budgets permit additional samples to be added. Therefore, if budgets were to decrease in subsequent years, the sample would have to be reduced. In fact, this has been the history of the 1985 AHS-MS samples, for which variations in budgets have resulted in both reductions and expansions to the number of AHS-MS surveys conducted during a cycle. The net result has been an inconsistent pattern of AHS-MS surveys and limited longitudinality.

In the rotating-panel designs, budget variations could be accommodated simply through reductions to the sample size for the upcoming panel. Consider option 4 and that for 2015, each panel costs $10 million for 40,000 observations (for a total of 200,000 observations). If the 2015 budget permitted $50 million in the initial year, all five panels could be completed. But suppose that in the next survey cycle (2017), the budget was reduced by $5 million. As previously planned in the panel design, panels 2, 3, 4, and 5 could be completed for $40 million. Panel 6, the new panel, could be scaled back to include only 20,000 observations and $5 million, for a total survey cost of $45 million. If budgets were to return to 2015 levels for 2019, panel 7 could then be increased to 60,000 observations, bringing the total observations back to 200,000. In all these cases, the longitudinality of the panels is maintained for 10 years. Furthermore, the decrease in observations for 2017 is shared across the entire survey, as opposed to specific metro areas.

6. THE 2015 REDESIGNED AMERICAN HOUSING SURVEY SAMPLE

The editor, as part of synthesizing the two reports, added this section to describe the redesign chosen by HUD and the Census Bureau. It is intentionally brief and contains links to recent documents that fully describe the redesign.

After due consideration, HUD opted for **Option 1–Full Single Panel**. For 2015, HUD and the Census Bureau selected an entirely new sample for AHS. The 2015 AHS sample is composed of an integrated national sample and independent metropolitan area samples. The national sample is described as integrated because it incorporates a few different types of samples, including:

- A representative sample of the nation.
- Representative oversamples of each of the 15 largest metropolitan areas.
- A representative oversample of HUD-assisted housing units.[18]

HUD and the Census Bureau intend to survey the entire integrated national sample once every 2 years. As such, it is a longitudinal panel with a 2-year survey cycle.

The independent metropolitan area samples include representative samples of 10 selected metropolitan areas. For 2015, the 10 selected metropolitan areas represent one-half of what HUD and the Census Bureau refer to as the "Next 20" group of metropolitan areas (the second half will be included in the 2017 AHS). The Next 20 group is a subset of metropolitan areas ranging from the 16th to 50th largest, by population. HUD and the Census Bureau intend to survey each member of the Next 20 group of metropolitan areas once every

[18] Includes units in the public housing, project-based rental assistance, and Housing Choice Voucher programs.

4 years; as such, the Next 20 group of independent metropolitan area samples is a longitudinal panel with a 4-year survey cycle.

The integrated national sample was first interviewed between April 29, 2015, and September 11, 2015. The independent metropolitan area samples were first interviewed at the same time, except for Phoenix, which was interviewed between July 27, 2015, and October 23, 2015.

For more information on the redesigned integrated sample—and specifically for more details on the metropolitan sample—see **https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/metro_oversamp_hist_2015.html**.

APPENDIX A. METROPOLITAN OVERSAMPLING STRATEGIES FOR PROPOSED FULL PANEL SAMPLE DESIGNS[19]

Starting in 2015, there will no longer be separate American Housing Survey national (AHS-N) and metropolitan area (AHS-MS) surveys. Instead, a representative national sample will be surveyed during each survey, and it will be supplemented with additional "oversampling" cases in each of 60 metropolitan areas. Metropolitan oversample cases in 30 of the 60 metropolitan areas (Group A) will be surveyed in 2015 and in every other survey (2019, 2023, and so on). Metropolitan oversample cases in the other 30 of the 60 metropolitan areas (Group B) will be surveyed in 2017 and in every other year (2021, 2015, and so on).

The following Exhibits show how the oversample will be incorporated into each potential sample design.

Exhibit A-1: Full Single Panel With Metropolitan Oversampling

Year	Panel
2015	National – 00 Metropolitan Oversample – Group A
2017	National – 00 Metropolitan Oversample – Group B
2019	National – 00 Metropolitan Oversample – Group A
...	
2039	National – 00 Metropolitan Oversample – Group A
2041	National – 00 Metropolitan Oversample – Group B
2043	National – 00 Metropolitan Oversample – Group A

Exhibit A-2: Full Dual Panel With Metropolitan Oversampling

Year	Panel	Panel
2015	National – 01 Metropolitan Oversample – Group 1	
2017		National – 02 Metropolitan Oversample – Group B
2019	National – 01 Metropolitan Oversample – Group A	
...		National – 02 Metropolitan Oversample – Group B
2039	National – 01 Metropolitan Oversample – Group A	
2041		National – 02 Metropolitan Oversample – Group B
2043	National – 01 Metropolitan Oversample – Group A	

[19] Appendices A and B are taken from the Impact and Analysis report and explain the thinking in 2011 on how to link the metropolitan AHS to the design options for the nation that AHS considered in the report.

APPENDIX B. METROPOLITAN OVERSAMPLING STRATEGIES FOR A ROTATING PANEL DESIGN

Exhibit B-1: Metropolitan Oversampling in Rotating Panel Design

Year	Panel 1	Panel 2	Panel 3	Panel 4	Panel 5	Panel 6	Panel 7	Panel 8	Panel 9
2015	National-1 MetOvr Grp A-1	National-2 MetOvr Grp A-2	National-3 MetOvr Grp A-3	National-4 MetOvr Grp A-4	National-5 MetOvr Grp A-5				
2017		National-2 MetOvr Grp B-2	National-3 MetOvr Grp B-3	National-4 MetOvr Grp B-4	National-5 MetOvr Grp B-5	National-6 MetOvr Grp B-6			
2019			National-3 MetOvr Grp A-3	National-4 MetOvr Grp A-4	National-5 MetOvr Grp A-5	National-6 MetOvr Grp A-6	National-7 MetOvr Grp A-7		
2021				National-4 MetOvr Grp B-4	National-5 MetOvr Grp B-5	National-6 MetOvr Grp B-6	National-7 MetOvr Grp B-7	National-8 MetOvr Grp B-8	
2023					National-5 MetOvr Grp A-5	National-6 MetOvr Grp A-6	National-7 MetOvr Grp A-7	National-8 MetOvr Grp A-8	National-9 MetOvr Grp A-9

Grp A = Group A. Grp B = Group B. MetOvr = metropolitan oversample.

MetOvr Grp A = first 30 metropolitan oversample cases. MetOvr Grp B = other 30 metropolitan oversample cases.

MetOvr Grp A denotes that the metropolitan group within this panel is not surveyed in this year

Assumptions:

- Each metropolitan oversample group would be interviewed five times over 20 years.
- Total sample size is 325,000 cases, comprising:
    - A target national sample size of 85,000 cases, surveyed every year.
    - A target metropolitan oversample sample size of 240,000 cases, with only one-half (120,000 cases) surveyed each survey year.

Therefore, only 205,000 of the 325,000 total sample cases are surveyed in each year.

- Each panel has a total sample size of 65,000 cases, comprising:
    - National (17,000 cases).
    - Metropolitan oversample for group A (24,000 cases).
    - Metropolitan oversample for group B (24,000 cases).

Therefore, only 41,000 of the 65,000 total panel sample cases are surveyed in each year.

APPENDIX C. REASONS FOR NONINTERVIEW, WITH CODES

Type A (eligible for subsequent interviews)

1. No one home.
2. Temporarily absent.
3. Refused.
4. Unable to locate.
5. Language problem.
6. Other, occupied.

Type B (eligible for subsequent interviews)

10. Permit granted; construction not started.
11. Under construction, not ready.
12. Permanent or temporary business or commercial storage.
13. Unoccupied site for mobile home or tent.
14. Group quarters or converted to institutional unit.
15. Occupancy prohibited.
16. Interior exposed to the elements.
17. Not classified above; specify.

Type C (ineligible for subsequent interviews)

30. Demolished or disaster loss.
31. House or mobile home moved.
32. Unit eliminated in structural conversion.
33. Merged not in current sample.
36. Permit abandoned.
37. Not classified above.
38. Unit eliminated in subsampling.
39. Unit deleted in prelisting subsampling.
40. Unit already had a chance of selection.
41. Sample reduction for the current survey year.
42. Big Six metro supplement.

REFERENCES

Atrostic, B.K., Nancy Bates, Geraldine Burt, and Adriana Silberstein. 2001. "Nonresponse in US Government Household Surveys: Consistent Measures, Recent Trends, and New Insights," Journal of Official Statistics 17(2), 209–226.

Dajani, Aref N., Shawn J. Bucholtz, and Toni M. Warner. 2012. "Measuring Attrition in Long-Term Longitudinal Surveys," Working paper. https://www.huduser.gov/portal/datasets/ahs/Final_JSM_Attrition_Proceedings_Paper.pdf?#.

Eggers, Frederick J., and Fouad Moumen. 2011. Components of Inventory Change: 2007-2009. Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development & Research, Washington, DC. https://www.huduser.gov/portal/datasets/cinch/cinch09/cinch07-09.pdf.

Hawala, Sam. 2001. Enhancing the '100,000' Rule: On the Variation of the Percent of Uniques in a Microdata Sample and the Geographic Area Size Identified on the File. Proceedings of the Annual Meeting of the American Statistical Association. Atlanta, GA, August 5–9. Washington, DC: U.S. Census Bureau, Statistical Research Division.

Schwartz, Mary B. 2011. 2007 American Community Survey: A Comparison to Selected Housing and Financial Characteristics for the United States from the 2007 American Housing Survey. Washington, DC: U.S. Census Bureau, Housing and Household Economic Statistics Division.

Warner, Toni, and Aref Dajani. 2011. The American Housing Survey 2015 Redesign: Impact and Analysis. Working paper. http://www.huduser.org/portal/datasets/ahs/AHSRedesignWhitepaper.pdf?#.

U.S. Department of Housing and Urban Development
Office of Policy Development and Research
Washington, DC 20410-6000

March 2020