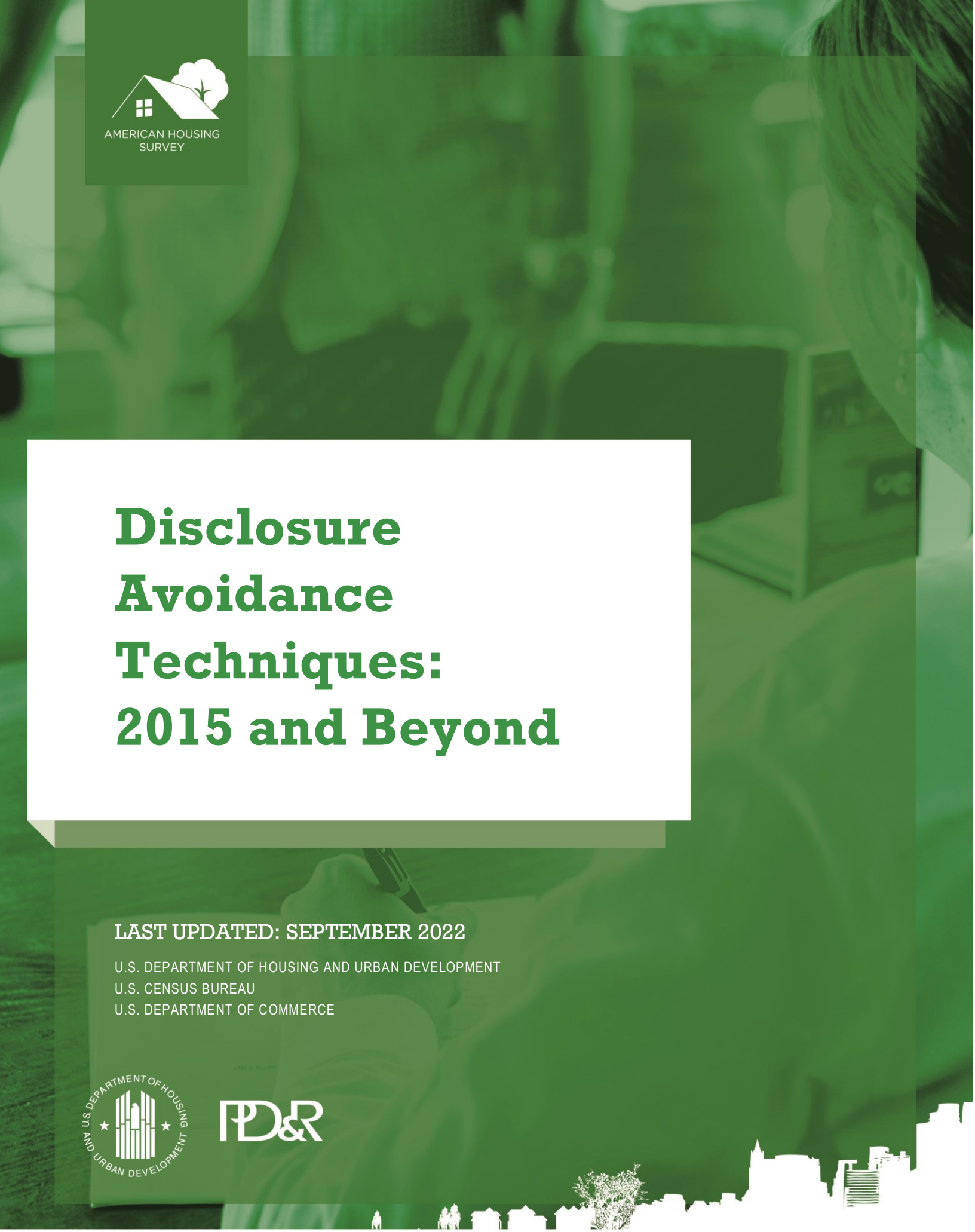

Disclosure Avoidance Techniques: 2015 and Beyond

LAST UPDATED: SEPTEMBER 2022

U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT
U.S. CENSUS BUREAU
U.S. DEPARTMENT OF COMMERCE

AMERICAN HOUSING SURVEY

Contents

AMERICAN HOUSING SURVEY

List of Exhibits

1. Overview

The purpose of this document is to provide a general overview of the various disclosure avoidance techniques that HUD and the Census Bureau applied to the American Housing Survey (AHS) summary table estimates and microdata for 2015 and beyond.

Title 13, Section 9 of the United States Code (U.S.C.) requires the U.S. Census Bureau to keep confidential the information collected from the public under Title 13, the authority under which the AHS data is collected. Disclosure avoidance is the process for protecting the confidentiality of data, as required under Title 13 U.S.C. A disclosure of data occurs when someone can use published statistical information to identify an individual who has provided confidential information.

All AHS data products released to the public are first reviewed by the Census Bureau Disclosure Review Board (DRB) to ensure that no identifiable Title 13 data are or may be disclosed. If the DRB determines that the requested statistical product does or reasonably could result in such disclosure, then the data product will be modified prior to approval for release to the public. Increased prevalence of administrative records and disclosure research in recent years has led the U.S. Department of Housing and Urban Development (HUD) and the Census Bureau to take increasingly strict measures to protect the data from re-identification.

For more information on disclosure avoidance techniques for 2013 and earlier years of the AHS, see *Disclosure Avoidance Techniques: 1985-2013*.

2. Vulnerability to Disclosure

For each year of the AHS, HUD and the Census Bureau produce two microdata products that contain individual responses to survey questions: the internal use file (IUF) and the public use file (PUF). The IUF is only available for approved researchers and not released publicly. It contains the individual responses as provided by the respondent and detailed geographic information (e.g., census block, parcel number).

The PUF is released publicly on the AHS web site. Its purpose is to allow data users and the general public to conduct their own statistical analysis, including summary statistics and regression modeling while protecting the confidentiality of AHS respondents. The PUF is derived from the IUF. However, the PUF is altered in numerous ways to avoid the disclosure of a respondent's name or address. Generally, there are three types of disclosure we aim to avoid:

1. **Direct disclosure of a respondent's name or address:** Including a respondent's name or address would be a clear violation of confidentiality.
2. **Indirect disclosure of a respondent's address through disclosure of detailed spatial information:** Including precise spatial information such as census block, or inherently spatial information such as distance to water, could result in an indirect disclosure of the respondent's address.
3. **Re-identification of a respondent's name or address via a re-identification attack**: A re-identification attack occurs when an attacker matches an external data source with precise name or address information to the individual PUF responses using information common to both datasets. The AHS is vulnerable to this type of attack due to the large number of housing attributes included in the survey.

HUD and the Census Bureau release summary table estimates through the AHS Table Creator. The summary table estimates in the AHS Table Creator are derived from the IUF. The AHS Table Creator contains more geographic information than is available on the PUF, which allows users to create summary table estimates that cannot otherwise be created using the PUF.

The summary table estimates in the AHS Table Creator are themselves a disclosure risk from a database reconstruction attack. A database reconstruction attack occurs when an attacker is able to reconstruct individual IUF records using the summary table estimates. To guard against database reconstruction attacks, numerous disclosure avoidance techniques are applied to the summary table estimates in the AHS Table Creator.

It is important for AHS users to note that summary estimates derived using the PUF may not match summary estimates derived from the AHS Table Creator. This is because of the disclosure avoidance techniques applied to the PUF.

Section 3 of this document details the disclosure avoidance techniques applied to the summary table estimates. Section 4 details the disclosure avoidance techniques applied to the IUF to create the PUF.

3. Disclosure Avoidance Techniques Applied to Summary Table Estimates in the AHS Table Creator

Summary table estimates are available for 2015 and later via the AHS Table Creator. Table Creator allows for customized tables, which are presented in an easy to read and understand format. There are two types of disclosure avoidance techniques applied to summary table estimates: cell suppression and rounding.

3.1. Cell Suppression

Disclosure-based cell suppression has been added to AHS Table Creator per DRB guidelines. Suppressed cells are displayed with an ‘S.’ Suppression rules apply when an estimate is based on less than three unweighted AHS observations and when one of the two conditions below is present:

- When an AHS Table Creator estimate is based on a variable available only in the IUF (i.e., including numerous geographic indicators).
- When an AHS Table Creator estimate is based on a variable available from the PUF but is cross-tabulated with a column variable by-group that is based on a variable available only in the IUF.

When suppression rules apply, they apply to more than just a single estimate. They also apply to any other estimate that has a parent or child relationship to the suppressed estimate. Parent indicators are row (column) indicators that have rows (columns) indented under them. Child indicators are the indented rows (columns) that, when added together, sum up to the parent row (column). To do this, mutually exclusive indicators within each table stub were grouped according to parent/child relationships in order to identify which rows (columns) were “related” to one another. From this, related cells were flagged as requiring suppression to prevent multidimensional disclosure (by subtraction) of any other cells within the group where at least one of the cells had an unweighted count of less than 3.

Additionally, for all means and medians except interpolated medians (for example, Year Structure Built), when a mean or median cell count is less than 10, the cell is suppressed, and any replicated indicators are suppressed as well. Interpolated medians and means have a suppression threshold of 3.

3.2. Rounding

There are three types of summary table estimates found in the AHS Table Creator: housing unit counts (in thousands), means, and medians. All national housing unit count estimates, including margins of error, are rounded to the nearest thousand. All state and metro housing unit count estimates are rounded to the nearest hundred because they show one decimal place in Table Creator. All means and medians, including margins of error for means and medians, are rounded to four significant digits or fewer.

4. Disclosure Avoidance Techniques Applied to the IUF to Create the PUF

As noted in Section 2, IUFs contain individual responses to survey questions. HUD and the Census Bureau derived the PUF directly from the IUF. PUFs can be used to create custom tabulations, allowing users to delve further into the rich detail collected in the AHS. To help AHS users navigate the PUF, HUD and the Census Bureau created the AHS Codebook online, listing the variables in the PUF and numerous pieces of information about each variable.

Disclosure avoidance techniques are applied to numerous variables in the PUF. To help AHS PUF users understand when a disclosure avoidance technique has been applied to a PUF variable, the AHS Codebook includes the *DISCLOSURE* field that lists the specific disclosure technique applied to that variable.

The subsections below describe the disclosure avoidance techniques applied to the PUF.

4.1. Removal of Personally Identifiable Information Variables

Variables that directly identify a housing unit or person are withheld from the PUF. These variables include *NAME*, *ADDRESS*, *PHONE NUMBER*, *LATITUDE*, *LONGITUDE*, and *PARCEL NUMBER*.

4.2. Removal of Detailed Political or Census Geographic Variables, and Inherently Spatial Variables

The 2015 and later IUFs contain a full complement of political and census geographic variables, including census block, block group, tract, incorporated place, census-designated place, core-based statistical area (CBSA), urban area, county, county subdivision, state, census division, and census region. Additionally, numerous external variables have been added to the IUF based on a political or Census boundary, including the Economic Research Service's Rural-Urban Continuum Code and Rural-Urban Commuting Area and the U.S. Forest Service's Wildland-Urban Interface (WUI) areas.

The 2015 and later integrated national longitudinal sample PUFs include only two census geographic variables: *DIVISION* (census division) and *OMB13CBSA* (core-based statistical area). The 2015 and later metropolitan area longitudinal oversamples (hereafter referred to as metropolitan area samples) PUFs do not include *DIVISION*. Only certain PUF cases have a specific *OMB13CBSA* code, while others have a more general code. For more details on geography in the PUFs, see the document: *AHS PUF Geography 2015 and Beyond*.

There are several additional variables in the IUF that are inherently spatial, meaning they correspond to a known geographic area or they represent a spatial concept that is easily observed. Exhibit 4.2.1 lists the inherently spatial variables on the IUF that are withheld from the PUF and Exhibit 4.2.2 describes

changes to the public use versions of variables in the 2021 topical Wildfire Risk module due to the spatial nature of these variables.

Exhibit 4.2.1. IUF Inherently Spatial Variables Removed from the PUF

IL30PER	IL50PER	IL80PER	GEOAREA	DEGREE
FMR	AMI	NEARWATER	WATFRONT	NEARSFD
NEARSFA	NEARMH	NEARBUSIN	NEARFACT	

Exhibit 4.2.2. Geographic Restrictions in the 2021 Wildfire Risk Module

In 2021, the AHS collected data on housing characteristics related to wildfire risk such as roofing and siding materials, vegetation around the home, and how respondents would be alerted to wildfire emergencies.

In order to reduce respondent burden, the AHS limited Wildfire Risk questions to certain geographic areas. Housing units had to be in one of the two following geographies to be eligible for the wildfire risk questions:

1. The unit was located in one of the following Census divisions: South Atlantic, West South Central, Mountain, or Pacific. **OR**
2. The unit was located in a Wildland Urban Interface (WUI)[1] area.

Because WUI areas cannot be included in the PUF for disclosure reasons, the public use version of the Wildfire Risk variables only shows values for records located within the four Census divisions.

4.3. Removal of Housing Unit and Household Characteristic Variables

Public property tax records and other administrative data are potential sources of detailed housing unit and household characteristic information that could be used by an attacker in a re-identification attack. To guard against such an attack, numerous AHS IUF housing unit characteristics are removed from the PUF. Exhibit 4.3.1 lists these variables.

Exhibit 4.3.1. IUF Housing Unit and Housing Characteristic Variables Removed from the PUF

NUNITS	DENS	RECROOMS	LIVING	FAMROOMS
AGERES	COOP	SUBDIV	TPARK	GATED
VACANCY2	WCN	HUDSAMP	MOVM	

4.4. Removal of Mortgage and Financial Variables

Public deed and mortgage records are potential sources of detailed mortgage and financial information that could be used by an attacker in a re-identification attack. To guard against such an attack, numerous AHS IUF variables were withheld in the 2015 and beyond AHS PUFs. Exhibit 4.4.1 lists the mortgage and financial variables that are withheld. The variables in the table below may not have been collected in all

[1] More information on Wildland Urban Interface boundaries is available at https://www.usfa.fema.gov/wui/.

survey years. For example, some of the variables that were collected but withheld from the 2015-2019 PUFs were no longer collected after the mortgage section of the AHS was redesigned in 2021.

Exhibit 4.4.1. IUF Mortgage Variables Removed from the PUF

MORTSTAT	REFILWPAY	INTPM	PMIAMT	HELOCBAL[2]
MORTTYPE[3]	PRICE	REFIINCPER	ADJPM	MORTGOV
MORTDOC	MORTTERM	FORSALE	REFIEXTLN	PTCHAM
LENMOD	PRIPMT	BALLOONAMT	MORTSUB	UNPBALAMT
REFILWINT	MINPM	PMIPMT	OTHAMT	HELOCADD[2]
YEARBUY	REFILWPER	RATEPM	REFIOTH	HELOCLIM[4]
MORTYEAR	DWNPAYSRC	REFICSH[2]	MORTSRC	OTHPMT
PMTFREQ[5]	MORTARM	LOTVAL	PTCHYR	OTRPM
FXDPM	INSPMT	MORTADDTN[4]	REFICSHAMT[4]	

4.5. Removal of Eviction Variables

The 2017 AHS included a topical module on eviction. Public eviction records are a potential source of information that could be used by an attacker in a re-identification attack. To guard against such an attack, all IUF variables from the eviction module have been removed from the PUF.

Exhibit 4.5.1. IUF Eviction Variables Removed from the PUF

EVIC	EVICCOURT	EVICFEAR	EVICFORCL	EVICNORNT
EVICRECRD	EVICORDER	EVICCONDM	EVICPAID	EVICLNDLD
EVICBEHND	EVICPRECT	EVICRAISE	EVICKIDS	EVICNOFIX
EVICPREV	EVICPRECT	EVICDANGR		

4.6. Topcoding and Bottomcoding

Topcoding is a disclosure limitation technique that involves limiting the maximum value of a variable allowed on the file to prevent disclosure of units with extreme values in a distribution (e.g., outliers).

Top and bottom coded variables were edited up or down to a point determined by the topcoding rules. For these years, topcodes are calculated at the CBSA level for the metropolitan area samples or nationally for all other cases.

To preserve confidentiality, it is the policy of the DRB that there must be at least three cases included in the calculation of a mean at each geographic level. It is not unusual in the AHS PUF, particularly in the AHS metropolitan area sample PUF, for a variable's universe of cases to be so small that there is not a

[2] From 2015-2019, HELOCADD, HELOCBAL, and REFICSH were withheld from the PUF. Beginning in 2021, these variables were released on the PUF.

[3] From 2015-2019, MORTTYPE was withheld from the PUF. Beginning in 2021, MORTTYPE was released on the PUF with collapsed categories. The uncollapsed categories are available on the IUF as MORTTYPE_IUF.

[4] From 2015-2019 HELOCLIM, MORTADDTN, and REFICSHAMT were withheld from the PUF. Beginning in 2021 these variables were released on the PUF with noise injection.

[5] From 2015-2019, PMTFREQ was withheld from the PUF. Beginning in 2021, PMTFREQ was released on the PUF with collapsed categories. The uncollapsed categories are available on the IUF as PMTFREQ_IUF.

minimum of three cases greater than or equal to the topcode predetermined or calculated for that variable. In these instances, the value of the topcode is lowered until there are at least three cases that can be included in the calculation of the mean. In the rare instances where there are not three eligible cases in the universe for a variable, all applicable values are set to a not-reported code.

4.7. Rounding

Rounding limits the number of unique values in the data and protects against rare-event situations. Exhibit 4.7.1 lists the rounding rules applied to the AHS. Exhibit 4.7.2 lists the IUF variables subject to rounding in the PUF. All variables that are created from or edited against variables that are rounded are re-calculated and/or re-edited following rounding.

Exhibit 4.7.1. Rounding Rules for AHS Variables

Unrounded	Rounded
1–7	4
8–999	Round to the nearest 10
1,000–49,999	Round to the nearest 100
50,000 or more	Round to the nearest 1,000

Exhibit 4.7.2. IUF Variables Rounded in the PUF

RENT	HCAMT	ELECAMT	GASAMT	OILAMT
OTHERAMT	TRASHAMT	WATERAMT	MAINTAMT	POOLAMT
PARKING	TOLL	TAXI	FERRY	TRANAMT
WAGP	SEMP	INTP	SSP	SSIP
PAP	RETP	JOBCOST		

4.8. Collapsing

Collapsing of categorical variables into more generalized categories was done to protect against rare-event situations. Collapsing was done by reducing the number of categories in the PUF variables. The following IUF variables are collapsed in the PUF:

Exhibit 4.8.1. IUF Variables Collapsed on the PUF

YRBUILT	HUDSUB	HHRACE	NATVTY
UNITFLOORS	MHWIDE	RACE	HHNATVTY
MORTTYPE	PMTFREQ	PETSCAT	PETSDOG

4.9. Recoding

Recoding of numeric variables into categorical variables reduces the number of unique values. Recoding is also done when a PUF variable is categorical but created from two or more IUF variables. The following IUF variables are collapsed or recoded for disclosure purposes in the PUF:

Exhibit 4.9.1. IUF Variables Recoded on the PUF

UNITSIZE	LOTSIZE	STORIES
DIST	MISCPMT (using INSPMPT and PMIPMT)	

4.10. Perturbation

Some AHS PUF variables that represent a year or number of years are perturbed, or slightly altered in a non-random way, to protect against rare-event situations. Variables that are perturbed may cause additional variables to be re-edited and recoded to preserve confidentiality and consistency with other demographic variables. The following variables are perturbed: *AGE*, *HHAGE*, *MOVE*, *HHINUSYR*, and *INUSYR*.

4.11. Noise Injection

To allow for the release of some mortgage and financial characteristics, multiplicative noise is injected into the real values. Appendix A contains detailed information on the noise injection process. Table 4.11.1 lists the IUF variables injected with noise in the PUF. The variables in the table below may not have been collected in all survey years. For example, some of the variables that were noise injected on the 2015-2019 PUFs were no longer collected after the mortgage section of the AHS was redesigned in 2021.

Exhibit 4.11.1. IUF Variables Injected with Noise on the PUF

PMTAMT	INTRATE	LOTAMT	MAINTAMT	AMMORT
PROTAXAMT	MARKETVAL	MORTPURCH	TOTHCAMT[6]	MORTAMT[6]
INSURAMT	HOAAMT	TOTBALAMT	PMTONLY	HELOCLIM[7]
MORTADDTN[7]	REFICSHAMT[7]			

In the AHS Table Creator, mortgage and financial tables are derived from non-noise injected IUF data. This means that estimates shown in the AHS Table Creator for these variables will be different from those derived from the PUF.

Appendix A

Additional Noise Injection Information

This appendix provides additional information on the noise injection process for continuous mortgage and financial variables within both the *HOUSEHOLD* and *MORTGAGE* tables. This information will help researchers account for the additional errors caused by noise injection.

A.1. Noise Injection Process

Noise factors were created using random, independent pulls from a Laplace distribution with a mean of 1 and a beta value of $k*(1/\sqrt{N})$, with N being the number of responses in geographic area and k being a proportional multiplicative factor determined in conjunction with the DRB. Original values are then multiplied by the noise factors following data edits.

[6] Calculated using the noise injected values of their components.

[7] HELOCLIM, MORTADDTN, and REFICSHAMT were withheld from the 2015-2019 PUFs, but were released on the 2021 PUF with noise.

The inclusion of N in the Laplace distribution's scale parameter ensures that geographic areas with fewer observations (and thus more vulnerable) of these mortgage/financial-characteristic variables will have more noise injected. N is calculated using the square root of the count of responses for that variable in CBSA. Records in the integrated national longitudinal sample that were not in one of the released metropolitan area samples were grouped together as their own category. As such, records in the independent metropolitan area sample and the Top 15 group of metropolitan area longitudinal oversamples[8] file have more noise injected as compared to records in the integrated national longitudinal sample.

Noise factors are independent across variables and years. Noise is applied at the variable level in a single year for all observations. Thus, the mean noise factor for a variable in any given year is 1. This limits the impact of noise injection at the aggregate level and allows means and correlations to be maintained. However, the more a variable is sliced, the higher the risk that the mean noise factor will deviate from 1. See the impact of this below.

A.2. Level of Noise Injected

During the planning phase, several tests were conducted to measure the impact of various levels of noise on data quality and fitness of use. AHS staff worked closely with DRB to set the level of noise (expressed as k in the Laplace distribution's scale parameter).

A.3. Impact of Noise Injection

This section describes the error created via the noise injection process in a range of research designs. These error rates do not incorporate margins of error. Estimates for smaller groups will have increasingly larger margins of error, before noise injection. The sampling error + noise injection error should be considered before using noise-injected data.

A.3.1. Cross-Sectional Analysis

For the integrated national longitudinal sample, excluding the Top 15 group of metropolitan area longitudinal oversamples, the impact of noise injection in cross-sectional analysis is minimal. For the independent metropolitan area longitudinal oversamples and Top 15 metropolitan areas within the integrated national longitudinal sample, the impact of noise is greater. The more variables and levels are used to filter the mean, the higher the error.

For estimates at the metro area level, Census does not recommend using noise-injected data for analysis that cross-tabulates more than three variables because the error rates will be high. For example, the error caused by noise injection of the mean of *MARKETVAL* in Memphis by *BEDROOMS*, *FIRSTHOME*, and *BATHROOMS* is, on average, 4 percent. If your analysis requires this level of granularity, use of the IUF will be required.

Table Creator can be used to understand the impact of noise injection. Table Creator estimates of mortgage and financial indicators are derived from the IUF, which does not have noise injected. A table in

[8] The Top 15 group of metropolitan area longitudinal oversamples use the 2013 Office of Management and Budget's core based statistical area definitions as of February 2013.

Table Creator that is close to the population of interest can be compared to a similar table derived from the PUF. The difference between the two tables is a measure of the error created by noise injection for this subpopulation. In the example below, the error of filtering house value by race within a metropolitan area is about 1 percent overall, while the value groupings see larger error due to the large number of groupings. Reducing the number of value groupings will reduce the impact of noise injection.

Exhibit A.3.1. 2015 Memphis, TN-MS-AR CBSA – Value, Purchase Price, and Source of Down Payment–Owner-occupied Units

Characteristics	Householder Race – Black alone	
	Generated from Table Creator (thousands)	Generated from PUF (thousands)
Total	**111.8**	**111.8**
Value		
Less than $10,000	1.5	1.5
$10,000 to $19,999	1.7	2.7
$20,000 to $29,999	2.8	2.8
$30,000 to $39,999	7.3	7.1
$40,000 to $59,999	13.8	16.9
$60,000 to $79,999	15.5	14.9
$80,000 to $99,999	17.5	18.2
$100,000 to $119,999	11.5	10
$120,000 to $149,999	13.5	14.8
$150,000 to $199,999	14.2	10.8
$200,000 to $299,999	9.1	9.3
$300,000 to $399,999	1.7	1
$400,000 to $499,999	.	0.2
$500,000 to $749,999	1.5	1.3
$750,000 or more	0.2	0.2
Median (dollars)	90,000	90,547

A.3.2. LOTAMT and MORTPURCH

LOTAMT and *MORTPURCH* are impacted by noise injection more than other noise-injected variables. Beginning in 2021, *HELOCLIM*, *MORTADDTN*, and *REFICSHAMT* were also added to the PUF with noise and are similarly impacted.

LOTAMT, HELOCLIM, MORTADDTN, and REFICSHAMT data are uncommon and are more at-risk for disclosure; thus, more noise had to be injected. When producing estimates for independent metropolitan area samples and the Top 15 group of metropolitan area longitudinal oversamples, we do not recommend using these variables in an analysis that requires cross-tabulating them by more than one variable. For other uses, we recommend binning these variables into broad categories to absorb the impact of noise injection.

MORTPURCH is impacted by editing. There are many cases of *MORTPURCH* grouped around (and at) 100, and noise injection would often increase these values to above 100. Noise-injected values above 100 were then edited and rounded down to 100. Thus, mean noise is no longer 1 for *MORTPURCH*. When producing estimates for independent metropolitan area samples and the Top 15 group of metropolitan area longitudinal oversamples, we recommend binning noise-injected *MORTPURCH* data or

truncating/binning the upper end of the data to absorb the impact of post-noise editing. Beginning in 2021 *MORTADDTN has* a similar structure to *MORTPURCH*. However, the distribution of *MORTADDTN* includes a smaller percentage of records in this top category, so the impact is less severe.

A.3.3. True zeros

Original values of 0 (zero) are not impacted by the multiplicative noise injection. Thus, the more 0s a noise-injected variable has, the more the mean noise will deviate from 1. We recommend calculating means using non-0 values (for example, the mean of *INSURAMT* for those who paid at least $1 in insurance).

A.3.4. Longitudinal analysis

For cases that are in the integrated national longitudinal sample but are not part of the Top 15 metropolitan areas, the impact of noise injection for longitudinal analysis will be minimal. For the Top 15 metropolitan areas, which are interviewed every survey cycle, noise injection will have a greater impact. Ultimately, noise injection will hide true changes in means of non-0 values across years when the noise injected is, on average, larger than the fluctuation across time.

For the Top 15 metropolitan areas, longitudinal analysis should not be used for analysis cross-tabulating more than two variables (e.g., change in the mean of *MARKETVAL* in Memphis by *BEDROOMS* and *FIRSTHOME* from 2015 to 2017). If your analysis requires this level of granularity, use of the IUF will be required.

Longitudinal analysis using numeric noise-injected versions of *LOTAMT, HELOCLIM, MORTADDTN, REFICSHAMT* and *MORTPURCH* should be avoided. We recommend binning these variables into broad categories for longitudinal analysis.

A.3.5. Regression and correlation analysis

We conducted tests on the impact of the noise-injection process on regression results. To do this, we compared coefficients of the same variable in different regression models using mean absolute Z test scores.[9] We compared coefficients in regressions where the noise-infused variables were the independent *or* the dependent variable and when they were both. All regressions contained three control variables to better reflect use-case scenarios.

For both the integrated national longitudinal sample and independent metropolitan area longitudinal oversamples, all coefficient comparisons had Z-scores under 1.28 (90th percentile). This means that the coefficients were statistically equivalent. This is true even when data were limited to a single metropolitan area and/or if all 0 values were included and/or excluded in the regressions.

Care should be taken when measuring marginal effects in regressions. The random noise injected into variables will cause weak relationships to be hidden and strong relationships to appear as less strong.

[9] The formula used is Z score = $(\beta 1 - \beta 2) / \sqrt{((SE\beta 1)^2 + (SE\beta 2)^2)}$ following: Clogg, Clifford C., Eva Petkova, and Adamantios Haritou. Statistical Methods for Comparing Regression Coefficients Between Models. *American Journal of Sociology* 100: 1261-1293.

We have not tested a full range of more complex models (such as interaction effects with noise-injected variables).

A.3.6. Confirming the impact of noise on other types of analysis

We recognize that noise-injected data might be used in other types of analysis that will require different statistics on the impact of noise. If you have questions regarding how noise injection might impact your analysis, please contact the Census Bureau's American Housing Survey Branch at 1-888-518-7365 (toll free) or email ahsn@census.gov.

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