



*Independent Statistics & Analysis*

U.S. Energy Information  
Administration

---

# Residential Energy Consumption Survey (RECS) 2015 Technical Documentation Summary

May 2017



This report was prepared by the U.S. Energy Information Administration (EIA), the statistical and analytical agency within the U.S. Department of Energy. By law, EIA's data, analyses, and forecasts are independent of approval by any other officer or employee of the United States Government. The views in this report therefore should not be construed as representing those of the U.S. Department of Energy or other federal agencies.

---

## Contents

---

|   |    |
|---|----|
| Figures.....                            | 3  |
| Overview and History of RECS .....      | 4  |
| Data Products.....                      | 6  |
| Coverage and Sample Design.....         | 7  |
| Sample selection.....                   | 7  |
| Household survey .....                  | 9  |
| Questionnaire design.....               | 9  |
| Data Collection.....                    | 10 |
| In-person interviews .....              | 10 |
| Web and paper questionnaires.....       | 10 |
| Response rate and nonresponse bias..... | 11 |
| Editing and data quality .....          | 11 |
| Item imputation.....                    | 12 |
| Weather and geographic data .....       | 13 |
| Square footage data.....                | 14 |
| Consumption and expenditure data .....  | 15 |
| Weighting and sampling error .....      | 16 |
| Relative standard error (RSE) .....     | 16 |
| Confidentiality of information .....    | 16 |

---

## Figures

---

|   |   |
|---|---|
| Figure 1. Timeline of EIA’s 2015 Residential Energy Consumption Survey..... | 4 |
| Figure 2. Graphic of 2015 RECS sample design.....                           | 7 |

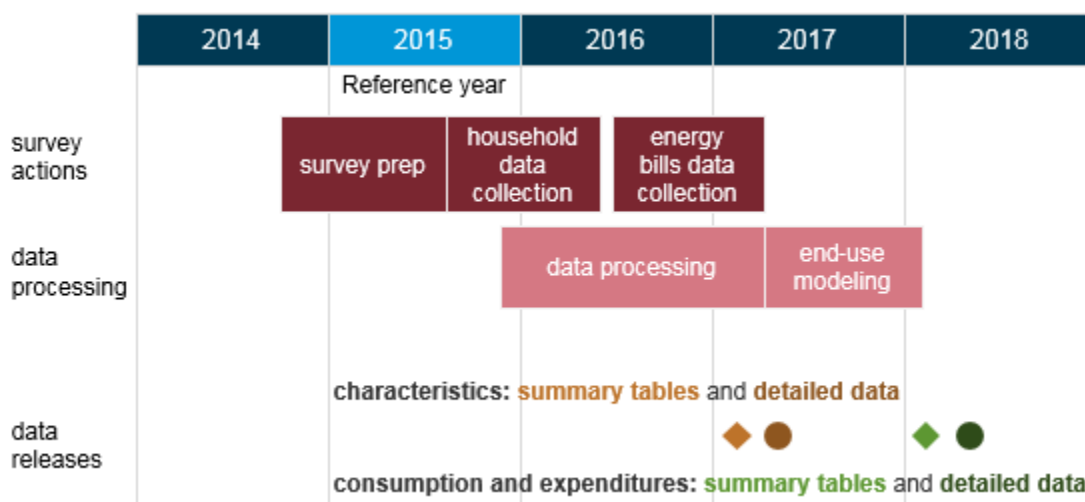
## Overview and History of RECS

The Residential Energy Consumption Survey (RECS) is a periodic study conducted by the U.S. Energy Information Administration (EIA) that provides detailed information about energy usage in U.S. homes. RECS is a multi-year effort (Figure 1) consisting of a household survey phase, data collection from household energy suppliers, and end-use consumption and expenditures estimation.

The household survey collects data on energy-related characteristics and usage patterns of a national representative sample of housing units. The Energy Supplier Survey (ESS) collects data on how much electricity, natural gas, propane/LPG, fuel oil, and kerosene were consumed in the sampled housing units during the reference year. It also collects data on actual dollar amounts spent on these energy sources.

EIA uses models (energy engineering-based models in the 2015 survey and non-linear statistical models in past RECS) to produce consumption and expenditures estimates for heating, cooling, refrigeration, and other end uses in all housing units occupied as a primary residence in the United States. Originally conducted by trained interviewers with paper and pencil, the 2015 study used a combination of computer-assisted personal interview (CAPI), web, and mail modes to collect data for the Household and Energy Supplier Surveys.

**Figure 1. Timeline of EIA’s 2015 Residential Energy Consumption Survey**



The scope and purpose of RECS differ slightly from similar EIA products that report *residential* energy data. RECS samples homes occupied as a primary residence, which excludes secondary homes, vacant units, military barracks, and common areas in apartment buildings. As a result, RECS estimates do not represent sector-level estimates, but they are best suited for comparison across different characteristics of homes within the residential sector. Estimates for residential wood consumption are not included in RECS total site consumption estimates, but are shown in separate tables.

---

EIA instituted the following survey design revisions, content changes, and variable updates for the 2015 RECS:

- **A return to the traditional number of respondents after an increase in the previous RECS.** The total number of responding households is 5,686 in the 2015 RECS compared to 12,083 in the 2009 RECS. Due to the smaller sample size, no state level estimates are available for the 2015 RECS.
- **A new sampling development methodology.** The Census Bureau's Public Use Microdata Areas (PUMAs) were used instead of counties (used in previous RECS surveys) in the first stage of the multi-stage area probability design. This change reduced unequal weighting effects and also reduced the sampling variability in the first stage by using energy data from the Census Bureau on PUMAs. In the second stage of sampling, Census Block Groups were used instead of Census Blocks (used in the previous RECS) to reduce clustering and possible intra-class correlations.
- **The introduction of new survey modes for the household survey.** Prior to the 2015 RECS, all iterations of the study were conducted entirely through in-person interviews with trained interviewers at the sampled households. For the 2015 RECS household survey, 5,686 questionnaires were completed using a combination of three modes: in-person computer-assisted personal interviews (CAPI), paper questionnaires sent through the mail, and web questionnaires accessed by a URL and password sent through the mail.
- **New questions on emerging technologies and usage behavior.** The 2015 RECS asked questions about smartphones and tablets, household behavior related to heating and cooling, detailed usage of cooking appliances, smart meters and smart thermostats, LED bulbs, lighting controls, and television peripherals.

The RECS is authorized under the Federal Energy Administration Act of 1974 (Public Law 93-275) as amended, and the Energy Policy Act of 1992.

---

## Data Products

---

EIA has released a variety of RECS products across survey cycles tailored to a wide range of data users. These releases include summary-level tables of energy-related household characteristics, detailed tabulations of energy consumption intensities across key variables, and public-use microdata for customized analysis of home energy use.

The 2015 RECS releases include articles highlighting key findings, standard tables, and a microdata file. All current and historical products are available on the [RECS homepage](#).

RECS data are also used as critical inputs for EIA sector level forecasts, including the [Annual Energy Outlook](#), as well as for energy program analysis.

## Coverage and Sample Design

The 2015 RECS used a multistage area probability sample design, where the population was divided into successively smaller statistically selected areas starting from large geographic areas and ending with individual housing units. The population for the 2015 RECS sample design included all housing units occupied as primary residences in the 50 states and the District of Columbia. Because EIA benchmarks to occupied housing totals from the U.S. Census Bureau's American Community Survey (ACS), the RECS uses the Census Bureau definition of a [housing unit](#). The RECS includes single-family homes, units in multifamily buildings, and mobile homes. The RECS excludes vacant, seasonal, vacation homes, and group quarters such as prisons, military barracks, dormitories, and nursing homes. Through the sampling methodology described below, all 118.2 million occupied primary households in the United States are represented by the 2015 RECS sample.

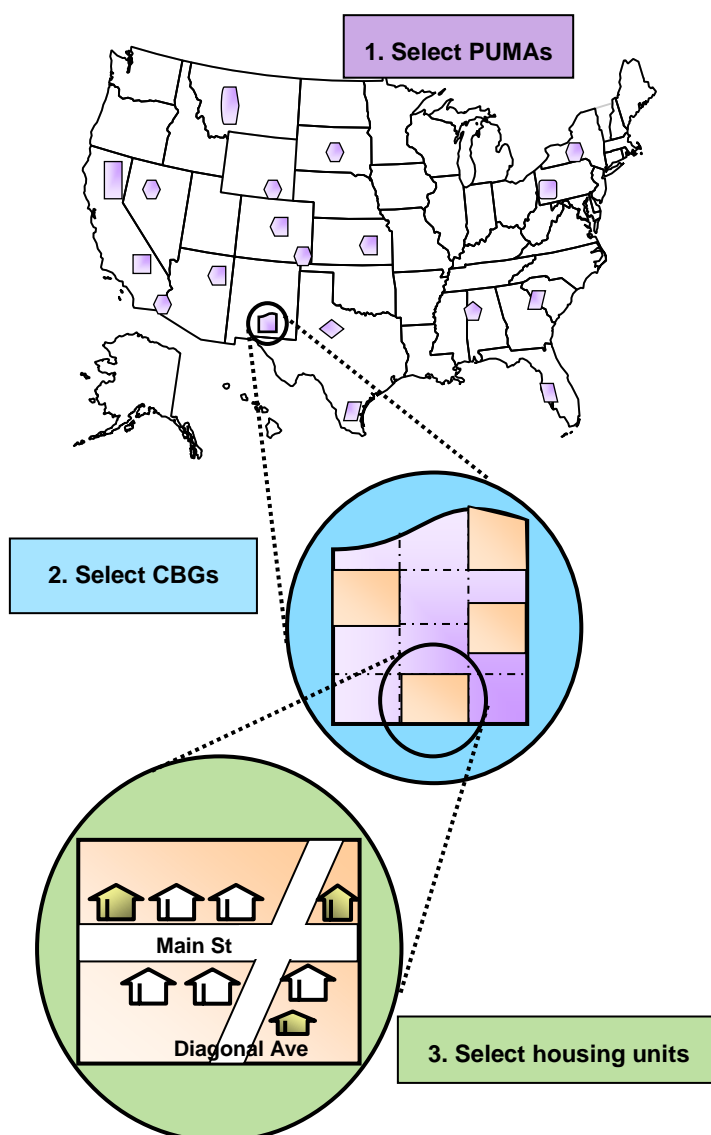
### Sample selection

The first step in the RECS sample design was to divide the entire United States into geographic areas. This subdivision allowed geographic control of the sample allocated to the Census Divisions. Precision constraints for estimates of average energy consumption per household were then applied at the national, Census Region, and Census Division levels to develop the sample allocation for the geographic domains. The 2015 RECS design has 19 geographic domains.

Sample selection began by randomly choosing Public Use Microdata Areas (PUMAs), which are geographic areas designated by the Census Bureau that contain at least 100,000 people and are based on census tracts and counties. A total of 200 PUMAs were selected in this first stage of sampling in the RECS multistage area probability design. In the second stage, the selected PUMAs were divided into Census block groups (CBGs), which are geographic areas designated by the Census Bureau that contain between 600 and 3,000 people. A sample of 800 CBGs, 4 per PUMA, was selected.

In the final stage of the design, households were randomly sampled from a list of

Figure 2. Graphic of 2015 RECS sample design





---

households in the selected CBGs. In most CBGs, the list of households was created from the U.S. Postal Service's Delivery Sequence File (DSF). In about 30% of CBGs where the DSF coverage of households was expected to be low, in-person field listing methods created the household list or supplemented the DSF<sup>1</sup>. A total of 12,753 households were ultimately sampled the 2015 RECS.

---

<sup>1</sup> McMichael, J. P., Shook-Sa, B. E., Ridenhour, J. L., & Harter, R. (2013). The CHUM: A frame supplementation procedure for address-based sampling. In *2013 Research Conference Papers* Washington, DC. Available at <http://www.fcsn.gov/events/papers2013.html>

---

## Household survey

---

### Questionnaire design

Each of the 2015 RECS household survey interviews was conducted using one of three modes: an in-person interview, a web questionnaire, or a paper questionnaire. The in-person interviews were conducted by trained professional interviewers who use a laptop with a specialized software called Blaise. The web address for the online questionnaire and the paper questionnaire were mailed to selected households and completed without the help of an interviewer.

All household survey questionnaires covered the same topical areas, with questions about the type and number of energy-consuming devices, usage patterns, structural characteristics of the home, household demographics, and energy supplier information. The questionnaire topics were as follows:

- Housing unit characteristics
- Appliances
- Electronics
- Space heating
- Air conditioning
- Water heating
- Lighting
- Energy programs
- Energy bills
- Energy suppliers
- Household characteristics
- Energy assistance
- Housing unit measurement (for in-person interview only)
- Scanning of sample energy bills (for in-person interview only)

While all three questionnaire modes covered the same topical areas, the in-person questionnaire generally collected additional detail beyond the most critical content that was collected by the web or paper questionnaires. For example, the Appliances section of all questionnaires collected information about the number of refrigerators in each home, followed by detail about up to two refrigerators including the size, age, and door arrangement of each. The in-person questionnaire also collected the same detail about the third refrigerator, for the nearly 5% of households that have three or more refrigerators. The published estimates in RECS housing characteristics tables and the variables in the public use microdata file reflect items collected from respondents across all three modes. Items only collected in the in-person interviews are excluded from published data because they were not weighted to represent the population.

Each time EIA conducts the RECS, we review the content from the previous round and revise the questionnaire as appropriate. The content revisions typically include adding or dropping questions to keep up with current household technologies. For 2015, the questionnaire design process also included soliciting data user input, prioritizing the potential changes, and then pretesting most of the new or substantially revised questions on a small number of people. New questions for 2015 cover

---

smartphones and tablets, household behavior related to heating and cooling, more detailed information on the usage of cooking appliances, and smart meters and smart thermostats. Some of the notable revisions from the 2009 RECS were to the lighting section (to include LED bulbs and lighting controls) and television peripheral section (to include counts of devices such as cable and satellite boxes and DVRs). The 2015 RECS questionnaires are available on the [EIA website](#).

## Data Collection

The RECS household survey was conducted on a voluntary basis with respondents. All questionnaires were completed between August 2015 and April 2016. A total of 5,686 questionnaires were completed: 2,417 in-person interviews (42.5%), 2,122 web questionnaires (37.3%), and 1,147 paper questionnaires (20.2%). The 2015 RECS is the first iteration of the study to use web and paper questionnaire modes. Prior to the 2015 RECS, EIA conducted a series of web and paper pilot tests to assess the feasibility of using these modes. The results of these pilot tests, as well as an overall assessment of 2015 RECS data quality, will be available in the coming months.

### *In-person interviews*

About 100 trained interviewers worked on the 2015 RECS, including bilingual interviewers who conducted interviews in Spanish when necessary. They were provided with an employee badge, 2015 RECS business cards, and 2015 RECS brochures that they carried when visiting households. To increase legitimacy and gain cooperation, households selected for in-person interviews were mailed an advance letter and postcard before an interviewer visited their home. Interviewers made several visits at different times of the day and on different days of the week in an effort to maximize the response rate. If contacted, all sampled households received a monetary, unconditional incentive as a token of appreciation. On average, in-person interviews were completed in 67 minutes, including the scanning and housing unit measurement.

When an in-person interview respondent was a renter or did not pay some or all of the household energy bills directly, a follow-up survey was attempted with a rental agent or landlord to get more information about the housing unit. The Rental Agent Survey asked about the equipment and fuels used by the household for space heating, water heating, air conditioning, and cooking, as well as the method of bill payment. Interviews with rental agents were conducted either in-person or by phone soon after the in-person household interview was completed. On average, the Rental Agent Survey took 10 minutes to complete. The household survey initiated 851 Rental Agent Survey cases, of which 524 (62%) were successfully completed. The data from the Rental Agent Survey were used to edit household survey responses; see the **Editing and data quality** section for more detail.

### *Web and paper questionnaires*

To increase legitimacy and gain cooperation, households selected to receive web and paper questionnaires received a sequence of mailings. These mailings included an advance postcard, invitation letters (which included the web address and paper questionnaire), and reminder postcards. Materials were available in both English and Spanish, typically with the Spanish text printed below or on the back. The entire sequence of six mailings was spaced over 2.5 months. All households in the web and paper phase of collection received a monetary, unconditional incentive as a token of appreciation; responding households received an additional monetary incentive. Once a household's web or paper questionnaire was completed, they were sent a thank you letter.

On average, web questionnaires were completed in 34 minutes. Eighty-one percent of web respondents used a desktop or laptop computer, while the remaining 19% used a smartphone or tablet. The

---

completion time of the 20-page paper questionnaire is unknown, but it is likely similar to the web questionnaire because it included the same content.

## Response rate and nonresponse bias

The overall unweighted response rate for the 2015 RECS when accounting for ineligible cases was 50.8% (a total of 12,753 selected cases with 5,686 completes and 1,114 ineligible cases). The response rate was calculated using the American Association for Public Opinion Research (AAPOR) formula 3 (AAPOR, 2016):

$$RR = I/(I+R+E)$$

where I is the number of complete interviews<sup>2</sup>, R is the number of refusal and eligible incompletes, and E is the number of eligible cases estimated from cases with unknown eligibility.

The response rate for 2015 RECS was lower than the 79% response rate in 2009 RECS, which was expected because a significant number of the 2015 cases were self administered.

The lower response rate prompted a nonresponse bias study. The study compared response rates by sample sub-categories, and compared respondent and nonrespondent estimates for key sample frame variables. Differences in any comparisons could indicate potential nonresponse bias. In addition, demographic variables were compared to the American Community Survey (ACS) estimates to assess potential differences between the types of households responding to the RECS and the general household population of the United States.

The nonresponse bias study reached the following conclusions:

- Unit response rates differed slightly by some sample variables such as housing type, Census region, and urban/rural classification.
- Although statistical tests resulted in statistically significant differences for some variables, the differences were small and not of practical importance.
- Most comparisons with ACS were not statistically different; for variables that were significantly different, the differences were small.
- The potential for nonresponse bias in RECS was reduced by applying weighting adjustments, which are described in the Weighting and Sampling Error section.

Based on the nonresponse bias study, EIA concluded that the final weighted 2015 RECS estimates are generally not statistically or practically different from the population.

## Editing and data quality

EIA employed many strategies to ensure data quality in the 2015 RECS household survey. Checks and edits were built into the data collection software to reduce missing or invalid responses. Resources such as show cards, which have pictures and response options, were used during the in-person interviews. Pictures were used as a guide for some questions for respondents answering the web or mail questionnaire.

---

<sup>2</sup> Completed interviews include interviews where the respondent did not answer all questions in the survey. The respondent must have answered at least 7 out of 10 key RECS questions in order for the interview to be considered complete. Partially completed interviews that did not meet that definition were defined as eligible incompletes.

---

All completed interviews went through a validation process to ensure that the correct sampled households were interviewed, key questions were answered, and the in-person interviews were not falsified by the interviewer. All interviewer and respondent comments were reviewed for indications of data quality issues.

After the validation process, the data were thoroughly reviewed for outliers, inconsistent responses, square footage issues, and write-in responses when *other* was chosen as an answer to a question. If the review indicated that a response was incorrect, it was either changed to a different valid response using deductive reasoning or changed to indicate that it was missing and then imputed. Data collected from the Rental Agent Survey were used to verify the household responses; in most cases, Rental Agent Survey responses were used where they contradicted the household responses.

Quality control was performed at all stages of data processing to ensure the processing was done consistently and correctly.

## Item imputation

Item nonresponse occurs when respondents do not know or refuse to answer a question in the survey or when a response is determined to be invalid and removed during editing. Item imputation is the process of filling in the missing responses using a statistical model to produce a complete dataset and to reduce the bias associated with item nonresponse.

The hot-deck imputation method was used for the 2015 RECS. In this method, a *recipient* case that has a missing value for the item being imputed is matched with a similar *donor* case that has a response. The donor's value for that item is used to replace the missing value for the recipient case. After imputation, final editing reviews ensured questionnaire skip patterns were performed correctly.

Implementations of the hot-deck method differ on both how to group recipients and potential donors, and how to select the donor for a given recipient. The 2015 RECS used the Predictive Mean Neighborhood (PMN)<sup>3</sup> hot-deck method for household square footage and the Cyclical Tree-Based Hot Deck (CTB)<sup>4</sup> hot-deck method for the remainder of the household survey variables. The PMN method uses regression models to group recipients and potential donors, and chooses a donor with a statistical nearest neighbor approach. The CTB method uses classification trees to group recipients and potential donors, and uses a weighted sequential hot deck imputation procedure<sup>5</sup> in which weights are used to match chosen donors to recipients.

---

<sup>3</sup> Singh, A., Grau, E., & Folsom, R. (2004). Imputation and unbiased estimation: Use of centered predictive mean neighborhoods method. In *Proceedings of the 2004 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Toronto, Ontario, Canada* (pp. 4351-4358). Alexandria, VA: American Statistical Association. [Available as a PDF at <http://www.amstat.org/sections/srms/proceedings/>]

<sup>4</sup> Creel, D. V., & Krotki, K. (2006). Creating imputation classes using classification tree methodology. In *Proceedings of the Survey Research Methods Section, American Statistical Association, Joint Statistical Meeting 2006*, pp. 2884–2887.

<sup>5</sup> Cox, B. G. (1980). The weighted sequential hot deck imputation procedure. In *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 721-726.

---

About two hundred variables in the 2015 RECS were imputed, but imputation rates for each item were generally low. The average imputation rate of all 216 variables was 3.7%. A total of 81% of the variables had imputation rates less than 5%, and 94% of the variables had imputation rates less than 10%.

## Weather and geographic data

EIA gathers weather and certain geographic indicators from other government agencies to complete the characteristics profile of sampled housing units. Daily heating degree days (HDD) and cooling degree days (CDD) for 2015 are available from the National Climate Data Center (NCDC) for each weather station in the United States. Each sampled RECS housing unit was matched to a local weather station in the area (most commonly the closest one to the house). The housing unit was then assigned the corresponding annualized HDD and CDD values. Thirty-year HDD and CDD averages were also used from the NCDC data. [Building America climate regions](#), which are defined using heating degree days, average temperatures, and precipitation data, were also matched to each housing unit.

EIA also matched two official Census Bureau geographic identifiers to each sampled housing unit, the Urban/Rural and Metropolitan/Micropolitan Statistical Area identifiers.

---

## Square footage data

---

A discussion of the household square footage data methodology will be available when the data are released in summer 2017.

---

## Consumption and expenditure data

---

A discussion of the consumption and expenditure data collected from the Energy Supplier Survey (ESS) and the end-use modeling and estimation processes will be available when the data are released in early 2018.



---

## Weighting and sampling error

---

The 2015 RECS used a multi-stage area probability design to select a sample of households that represents the population. To produce population estimates, the sample cases were weighted to represent all households, including those not in the sample. Base sampling weights, which are the reciprocal of the probability of selection for the RECS sample, were first calculated for each sampled household.

Similar to previous RECS, the 2015 RECS final analysis weights were calculated by applying eligibility, unit nonresponse, and poststratification adjustments to the base weights. The eligibility adjustment was calculated differently for the in-person and web/mail cases. Unlike the in-person cases where eligibility was determined by field interviewers, the eligibility of the web/mail cases was determined by a propensity model based on survey responses and contact mailing status. The Generalized Exponential Model (GEM)<sup>6</sup> calibration method was used for the nonresponse and poststratification adjustments. The weights were then post-stratified to the 2015 American Community Survey (ACS), which estimated a total of 118.2 million occupied housing units in the United States. The variables used for poststratification included Census Division, housing unit type, age of housing unit, and major heating fuel type.

The final analysis weight for each responding household is the number of households in the population that the observation represents. For example, if the analysis weight for a household is 5,000, that household represents itself and 4,999 other non-sampled households.

### Relative standard error (RSE)

Estimates from a sample survey like RECS are not exact but they are statistical estimates with some associated sampling error—the result of generating estimates based on a sample rather than conducting a census of the entire population. The standard error provides a measure of the precision of a particular statistic for a characteristic based on how variable it is in the population and a given sample size. Standard errors are used with survey statistics to measure sampling error, construct confidence intervals, or perform hypothesis tests. For the 2015 RECS, standard errors were estimated using Fay’s modified balanced repeated replication method<sup>7</sup> with a Fay factor of .5.

The relative standard error (RSE) measures how large the standard error is relative to the corresponding statistic; the larger the RSE, the less precise the survey statistic. The RSE is calculated as (standard error/statistic) x 100.

### Confidentiality of information

The Confidential Information Protection and Statistical Efficiency Act (CIPSEA) protects the privacy of respondents of federal surveys, including RECS. Any information collected that might permit the identification of respondents or their households is kept confidential and used only for statistical purposes. The household records EIA places on the public-use data file do not have name or address

---

<sup>6</sup> Folsom, R. E., & Singh, A. C. (2000). The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification. In *Proceedings of the American Statistical Association, Survey Research Methods Section*, pp. 598-603. Alexandria, VA: American Statistical Association.

<sup>7</sup> Judkins, D. R. (1990). Fay’s method for variance estimation. *Journal of Official Statistics*, 6, pp. 223–239.

---

information, and disclosure protection measures are taken to mask the data so that a sampled housing unit or its individual occupants cannot be identified by the public.