Overview

EIA makes available a public-use microdata file for each RECS survey cycle. The 2015 file is a valuable tool for users conducting detailed analysis of home energy use. This document provides some background on the RECS design, as well as useful tips and examples that will guide users through the use of the RECS microdata.

RECS sample design

The RECS sample was designed to estimate energy characteristics, consumption, and expenditures for the national stock of occupied housing units and the households that live in them. The 2015 RECS allows for separate estimation for Census regions and divisions. (The return to the traditional sample size for the 2015 RECS does not allow for state-level estimation, as was available for the expanded 2009 RECS.) To produce estimates for these geographies and the total U.S., the sample cases were properly weighted to represent the population, including the residences not in the sample. In a sense, a case's weight indicates the number of households that the particular case represents.

Base sampling weights, which are the reciprocal of the probability of being selected for the RECS sample, were first calculated for each sampled housing unit. The base weights were adjusted to account for survey nonresponse and ratio adjustments were used to ensure that the RECS weights add up to Census Bureau estimates of the number of occupied housing units for 2015. The variable **NWEIGHT** in the data file represents the *final sampling weight*, accounting for different probabilities of selection and rates of response, and being adjusted for the Census Bureau housing unit estimates. NWEIGHT is the number of households in the population that the observation represents. For example, if NWEIGHT for a household is 10,000, that household represents itself and 9,999 other non-sampled households. More details about the sample design can be found in the *RECS 2015 Technical Documentation – Summary*.

Sampling error

Estimates from a sample survey like RECS are not exact but are statistical estimates with some associated sampling error in each direction—the result of generating estimates based on a sample rather than a census of the entire population. Sampling error provides a measure of the accuracy of a particular estimate for a characteristic based on how common and variable it is in the population, given a particular sample size.

Standard errors are used in conjunction with survey estimates to measure sampling error, construct confidence intervals, or perform hypothesis tests. A relative standard error (RSE) is defined as the standard error (square root of the variance) of a survey estimate, divided by the survey estimate, and multiplied by 100. In other words, the RSE is the standard error relative to the survey estimate on a scale from zero to 100. The larger the RSE, the less precise the survey estimate is of the true value in the population. An RSE is shown for each estimate in the RECS tables.