



Independent Statistics and Analysis

**U.S. Energy Information
Administration**

2020 RECS Square Footage Data Collection and Estimation Methodology

March 2023



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 U.S. Government. The views in this report should not be construed as representing those of the U.S. Department of Energy or other federal agencies.

Table of Contents

Overview	3
RECS square-footage concepts and definitions	3
2020 RECS household survey square-footage questions.....	4
Square-footage response editing and imputation.....	5
Editing	5
Imputation	6
Estimation methods.....	6
Total square footage (TOTSQFT_EN)	6
Heated and cooled square footage	7
Comparing 2020 RECS and previous RECS square-footage estimates.....	8
Appendix A. 2015 RECS square-footage analysis summary.....	9

Table of Figures

Figure 1. Square-footage questions, RECS 2020 4

Table of Tables

Table 1. Public microdata file respondent-reported square footage variable names and descriptions, 2020 RECS 5

Table 2. Respondent-reported square-footage editing data change rates by housing unit type, 2020 RECS 6

Table 3. Square footage variable imputation rates, 2020 RECS 6

Table 4. Square-footage adjustment scenarios for basements, 2020 RECS 7

Overview

The square footage, or size, of a home is important for understanding its energy use. The amount of energy used for major end uses, such as space heating and air conditioning, are strongly related to the size of the home. This document discusses the methods we used to collect square footage and related data as part of the *2020 Residential Energy Consumption Survey (RECS)* and how we estimated final square footage values for each responding household. This document also compares the 2020 methods and estimates to those we employed and produced in the 2015 and previous RECS studies.

We used different square footage estimation methods for the 2020 RECS than we used in previous rounds. These new estimates represent a break in the square footage data series, and we advise data users to avoid comparing square-footage averages or square-footage energy-use intensities with previous RECS estimates.

We conducted the 2020 RECS household survey using only self-administered (web-based and paper) questionnaires. These modes replaced computer-assisted personal interviews (CAPI) used partially or wholly for every previous RECS household survey, which required us to change how we collected and estimated square footage. When we used CAPI modes, interviewers measured the dimensions of the homes on-site, then we used those measurements to calculate the square footage of the homes. For the new self-administered modes, we relied on the respondent's estimate of the home's square footage and systematically adjusted those estimates based on their answers to related questions (for example, the square-footage *include* questions discussed in the next section of this report.)

This document discusses:

- RECS square-footage concepts and definitions
- Household survey square-footage questions for 2020 RECS
- Square-footage response editing and imputation
- Estimation methods
- Comparing 2020 RECS and previous RECS square-footage estimates

We collaborated with IMG-Crown and RTI International on the 2020 RECS household survey data collection, imputation, and square-footage analysis.

RECS square-footage concepts and definitions

For RECS, the square footage of a home is defined as the two-dimensional floor area of the home that is enclosed from the weather by exterior walls, which is also the area where residential energy-consuming activities occur. Although the data collection and estimation methods changed for the 2020 RECS, the RECS definition of square footage remained the same.

A home's square footage can be divided into four areas: the main living area of the home, attic, basement, and garage. The basement and all main living areas (the first floor, second floor, etc.) are always included in the RECS definition of square footage. Attics are only included if they are heated, cooled, or finished. Garages are only included if they are heated or cooled and directly attached to the home. All single-family, apartment, and mobile homes have a main living area. For RECS square-footage estimation, only single-family detached and attached homes can have a basement, attic, or garage.

A home’s heated and cooled square footage is the portion of the home’s total square footage that is heated or cooled by space-conditioning equipment. A home with no air-conditioning equipment or where the air-conditioning equipment is not used, for example, will have no cooled square footage.

Because of differences in definitions and methods, RECS square footage estimates may not be appropriate to compare with other data sources, which may consider only furnished or livable space.

2020 RECS household survey square-footage questions

Estimating total, heated, and cooled square footage for each 2020 RECS case began with a series of questions on the RECS Household Survey (Figure 1). We modified and added questions from previous RECS studies to improve response quality and support new estimation techniques required for the entirely self-administered modes for the 2020 RECS Household Survey. These questions replaced the interviewer measurements that formed the basis of previous rounds of RECS square-footage estimates.

Figure 1. Square-footage questions, RECS 2020

16. What is the square footage of your home?
 , square feet → *Go to #18*
 Don't know

17. We understand that you might not know how many square feet your home is. A home’s size is closely related to its energy use, so we would like you to provide an estimate. Please tell us which category best describes the square footage of your home.

- Less than 600 square feet
- 600 to 799 square feet
- 800 to 999 square feet
- 1,000 to 1,499 square feet
- 1,500 to 1,999 square feet
- 2,000 to 2,499 square feet
- 2,500 to 2,999 square feet
- 3,000 square feet or more
- Don't know *(Select this option only if you are unable to provide a guess) → Go to #20*

18. What sources of information did you use to determine your home’s square footage? Please select all that apply.

- I remember from when I bought or leased my home
- I checked an Internet site (for example: Zillow)
- I checked my property tax documents
- I checked my lease or mortgage documents
- I asked someone else who knows
- I made an estimate
- Other (please specify):

19. Are any of the following areas included in the square footage you provided?

	Yes	No	Don't Know
a. Basement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Attic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Attached garage	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Data source: U.S. Energy Information Administration, 2020 Residential Energy Consumption Survey (RECS) paper questionnaire

New for the 2020 RECS questionnaire was a follow-up, square-footage range question (Question 17 in Figure 1) for respondents who did not provide an answer to the numeric question (Question 16 in Figure 1).

Table 1 shows the corresponding public microdata file variable names and descriptions for each of the square-footage questions in Figure 1. Question 18 responses were not released in the public file. An additional variable, TOTSQFT_EN, is derived from Questions 16, 17, and 19 and represents the official

square-footage estimate for each RECS case. We further discuss estimation for TOTSQFT_EN below in the *Estimation methods* section.

Table 1. Public microdata file respondent-reported square footage variable names and descriptions, 2020 RECS

Question number	Variable	Description
16	SQFTEST	Respondent-reported square footage
17	SQFTRANGE	Respondent-reported square-footage range
19	SQFTINCB	Respondent-reported square-footage includes basement
19	SQFTINCA	Respondent-reported square-footage includes attic
19	SQFTINCG	Respondent-reported square-footage includes attached garage

Data source: U.S. Energy Information Administration, 2020 Residential Energy Consumption Survey (RECS)

In addition to these specific square-footage questions, items such as the type of housing unit, the number of stories, and whether the basement is heated or cooled are also important for RECS total, heated, and cooled square-footage estimation.

Square-footage response editing and imputation

Editing

After we collected the 2020 RECS household survey data, we implemented a series of quality control checks and response-correction procedures to improve square-footage data quality. These checks included:

- Reviewing high and low outlier responses
- Reviewing inconsistent square footage and related characteristics responses (for example, comparing the number of rooms to the respondent square-footage estimate)
- Detecting inadvertent data entries or typos

Some corrections were implemented automatically via a batch-edit process, but other failed edits required analyst review. In some cases, analysts compared square-footage responses with related variables or external sources. Corrections, where necessary, generally involved deleting respondent-provided data and setting the square-footage response to *missing* and then assigning an imputed value. A total of 3.54% of cases required a change to any respondent square-footage variable.

Table 2. Respondent-reported square-footage editing data change rates by housing unit type, 2020 RECS

	Edits to SQFTEST	Edits to SQFTRANGE	Edits to any include variables	Edits to any square-footage variables
Mobile homes	4.51%	2.67%	0.00%	6.77%
Single family detached	2.54%	1.08%	0.02%	3.33%
Single family attached	2.17%	2.34%	0.23%	4.22%
Apartments 2-4 units	1.56%	2.15%	0.00%	3.32%
Apartments 5 or more units	1.19%	2.46%	0.00%	3.03%
Total US	2.38%	1.52%	0.04%	3.54%

Data source: U.S. Energy Information Administration, 2020 Residential Energy Consumption Survey (RECS)

Imputation

We used a hot-deck imputation method for missing (due to nonresponse or deletion during editing) or *don't know* numeric square-footage responses. In this method, we matched a recipient household that had a missing value for a variable being imputed to a similar donor household that had a response. If provided, we used the recipient's square-footage range response to select a donor and imputed the numeric response. For example, if the recipient answered the range question as *1,000 to 1,499 square feet* then the imputed numeric estimate value would be within the range of 1,000 to 1,499 square feet.

Table 3. Square footage variable imputation rates, 2020 RECS

Variable	Cases imputed	Eligible cases	Imputation rate
SQFTEST	6,566	18,496	35.5%
SQFTRANGE	1,557	18,496	8.4%
SQFTINCB	1,102	6,713	16.4%
SQFTINCA	1,796	9,132	19.7%
SQFTINCG	1,901	8,820	21.6%

Data source: U.S. Energy Information Administration, 2020 Residential Energy Consumption Survey (RECS)

More than one-third of cases (35.5%) required imputation for the numeric square-footage question (SQFTEST). We imputed most of these missing responses within the range the respondent provided in the range question (SQFTRANGE). Only 8.4% of respondents reported neither the numeric square footage nor the range. For these cases, we imputed the square-footage range first, then the numeric square-footage value, based on the corresponding square-footage range.

Estimation methods

Total square footage (TOTSQFT_EN)

To estimate the total square footage (TOTSQFT_EN) for each 2020 RECS household, we started with the reported (or imputed) numeric value (SQFTEST) and adjusted, if applicable, that value based on

responses to other questions. We only made adjustments for single-family homes (TYPEHUQ = 2 or 3) because only those housing types could include basement, attic, or garage spaces. Mobile home (TYPEHUQ = 1) and apartment (TYPEHUQ = 4 or 5) values were not adjusted, and the final square-footage estimates for those cases is the respondent-reported or imputed value.

For single-family homes, we adjusted the numeric values based on responses to questions about what spaces were included in the square footage estimate, such as basements, attics, or attached garages (for example, SQFTINCB.) For basements, if a respondent reported that a basement within the home was present and *was not* included in their estimate, then we added an estimated area for the basement to the final square footage estimate. If the respondent reported that the basement was present and *was* included in the square footage estimate, then we did not adjust the square footage for that area. [Table 4](#) shows examples for various basement adjustment scenarios.

Table 4. Square-footage adjustment scenarios for basements, 2020 RECS

Type of home	Number of stories or levels	Has basement	SQFTINCB (SQFTTEST includes basement)	SQFTTEST (respondent estimate)	Basement square footage added	TOTSQFT_EN (total square footage)
Single-Family	One	No	NA	1,000	NA	1,000
Single-Family	One	Yes	Yes	1,000	0	1,000
Single-Family	One	Yes	No	1,000	1,000	2,000
Single-Family	Two	Yes	Yes	1,500	0	1,500
Single-Family	Two	Yes	No	1,500	750	2,250
Apartment	One	NA	NA	750	NA	750
Mobile home	One	NA	NA	1,000	NA	1,000

Data source: U.S. Energy Information Administration, 2020 *Residential Energy Consumption Survey* (RECS)

We used similar adjustment methods for attic and garage areas. For these areas, however, we may have subtracted some space from the respondent estimate if those spaces were not finished (attics) or conditioned. For example, if a respondent reported that their square-footage estimate included an unfinished attic, then we subtracted an estimated value for the attic space from the numeric square-footage value.

Heated and cooled square footage

Once we determined the total square-footage estimate, we calculated a heated and cooled square footage for each case. For respondents that did not have or use heating or air-conditioning equipment, heated or cooled square footage was estimated as 0. For respondents who did heat or cool their homes, we calculated the corresponding heated and cooled square footage values using the total square-footage estimate, equipment type responses, and responses to questions about particular heated or cooled spaces in the home, such as basements (for example, BASEHEAT). A home with an unheated basement, for example, has less heated square footage (TOTHSQFT) than total square footage (TOTSQFT_EN). A multi-room apartment that uses a window air-conditioning unit to cool only one room has less cooled square footage (TOTCSQFT) than total square footage (TOTSQFT_EN).

Comparing 2020 RECS and previous RECS square-footage estimates

For most of the history of the RECS program, we used interviewer measurements as the basis for square-footage estimates. Except for the 1997 RECS, all RECS studies relied on trained interviewers to take physical measurements of the dimensions of each respondent's home. For the 1997 RECS, we discontinued interviewer measurements as a cost-saving measure and relied on respondent estimates of heated square footage alone. For the 2015 RECS, we used interviewer measurements for a portion of the overall responding sample and imputed square footage for cases where measurements were not taken. Although the RECS collected respondents' numeric estimates of square footage in previous rounds, we primarily used the respondent data as a data quality check for the interviewer measurements. Appendix A includes a 2015 analysis comparing interviewer-measured and respondent-reported square footage values.

The quality of the 2020 estimates is sufficient to reflect the energy-consuming space in homes, as well as the heated and cooled spaces. In particular, the adjustment method we used for single-family homes produced square-footage estimates that more closely reflect our definition of energy-consuming space in homes than if we had used the respondent-reported numeric values alone. Updates to the questionnaire and new editing methods also resulted in improved quality in the respondent-reported values. However, because of the new collection and estimation methods used for square-footage data in the 2020 RECS, we advise data users to avoid comparing 2020 RECS square-footage estimates with previous RECS studies.

Appendix A. 2015 RECS square-footage analysis summary

This Appendix includes key findings from square-footage analysis we conducted prior to the 2020 RECS, as well as how we used the analysis to develop the new estimation methods for the 2020 RECS. More detailed information about this analysis is available on the 2015 RECS [Methodology](#) page. The analysis revealed six main points:

- In 2015, respondent-reported numeric square-footage values averaged 371 square feet less than measured square-footage estimates. In 31 of the 32 comparisons made across various subgroups, the average respondent estimate was lower than the measured square-footage estimate.
- Interviewer-measured square-footage estimates were not free of errors. In the 2015 RECS, for example, some physical measurements were incorrect.
- Square footage estimates by 2015 respondents living in single-family homes (both attached and detached) had the largest differences from the interviewer-measured estimates. The average respondent-reported home size was 312 to 556 square feet (depending on housing type and age of home) less than the interviewer-measured estimates.
- Respondents to the 2015 RECS who lived in apartments, on average, reported square-footage values that more closely matched measured square footage.
- Much (although not all) of the difference between 2015 respondent-reported and measured square footage is attributable to the difference between what we include in the definition of a home's square footage and the definition respondents used. Respondents used a variety of methods, including consulting official documents, online sources, and their own estimates of home size. For single-family homes, it appears that many 2015 respondents did not include basements, attics, or attached garages in their square-footage estimates.
- Respondent-reported square footage values for the 2015 RECS had higher nonresponse rates than other items. Because interviewers are trained to follow a certain data collection protocol, the measured square-footage estimates did not have similar item nonresponse rates.

We also analyzed alternative sources and approaches for estimating square footage.

- In 2015, three administrative data sources that have home square footage were evaluated to assess the feasibility to replace or augment square-footage estimates for RECS. These sources were Zillow.com, Acxiom, and CoreLogic. The match rates between RECS cases and these administrative sources ranged from approximately 40% to 60%. The match rates varied by housing unit type, with higher match rates for single-family homes and lower rates for apartments and mobile homes.
- For those cases that were matched, all three administrative sources underestimated the square footage, according to the RECS definition.

We considered conducting a subsample of in-person interviews to continue to support interviewer-measured square-footage estimates for the 2020 RECS. We did not pursue this approach due to cost and operational constraints. Instead, we focused on four ways to address the issues discovered in the 2015 analysis:

- Improve respondent-reported values by modifying the questionnaire, adding a *don't know* response option for the numeric question and a follow-up range question if the initial response was *don't know*.
- Retain and explore using a series of questions to determine what spaces the respondent included in their square-footage response
- Implement data quality checks to remove erroneous responses and responses that were inconsistent with related housing characteristics variables
- Implement an adjustment method that modified, if necessary, the respondent estimate to produce an alternative measure that more closely matched the RECS square-footage definition