Independent Statistics and Analysis
U.S. Energy Information Administration

Supporting Statement B for Residential Energy Consumption Survey

Part B: Collections of Information Employing Statistical Methods

OMB No. 1905-0092

Form EIA-457A *Household Survey*

Form EIA-457D *Energy Supplier Survey: Household Propane Usage*

Form EIA-457E *Energy Supplier Survey: Household Electricity Usage*

Form EIA-457F *Energy Supplier Survey: Household Natural Gas Usage*

Form EIA-457G *Energy Supplier Survey: Household Fuel Oil/Kerosene Usage*

July 2024

Independent Statistics & Analysis
www.eia.gov

U.S. Department of Energy
Washington, DC 20585

Table of Content

S

B.1. Respondent Universe..... 1
B.2. Statistical Methods..... 1
B.3. Maximizing Response Rates..... 7
B.4. Test Procedures and Form Consultations..... 9
B.5. Statistical Consultations..... 10

B.1. Respondent Universe

The respondent universe for the 2024 Residential Energy Consumption Survey (RECS) will be the population of all housing units (HUs) occupied as a primary residence in the 50 states and the District of Columbia. The definition of an occupied housing unit is the same as that used by the U.S. Census Bureau, which excludes military barracks, college dormitories, nursing homes, prisons, vacant homes, and seasonal homes. According to 2022 American Community Survey (ACS) estimates, the U.S. has approximately 144 million total housing units. Approximately 130 million of these are occupied, primary housing units, which is the target population for RECS.

B.2. Statistical Methods

The 2024 RECS sample is designed to meet the precision requirements for key household energy consumption metrics and to produce estimates of household energy characteristics, consumption, and expenditures for various geographic levels such as national, state, and select Metropolitan Statistical Area (MSA) levels. The 2020 RECS was the first in program history to target estimates for all 50 states and DC. The 2024 RECS will be the first to target select, large MSAs. The MSAs listed below were selected based on demographic and geographic diversity and, due to sampling constraints, selected from MSAs that are entirely within a single state.

- Atlanta-Sandy Springs – Roswell, GA
- Baltimore-Columbia-Towson, MD
- Birmingham, AL
- Detroit-Warren-Dearborn, MI
- Houston-Pasadena-The Woodlands, TX
- Miami-Fort Lauderdale-West Palm Beach, FL
- Oklahoma City, OK
- Phoenix-Mesa-Chandler, AZ
- Riverside-San Bernardino-Ontario, CA
- Rochester, NY

Target Number of Completes and Precision Requirements

EIA resources allow for up to 20,000 completed Household Survey responses and associated data collected from the follow-on *Energy Supplier Survey* operation. Based on the 2020 RECS results, 20,000 cases will support precise estimates of key energy-use metrics for various geographic subpopulations, as well as publishable estimates for key demographic subpopulations and estimates for emerging energy topics.

The 2024 RECS sample design is similar to the design used for the 2020 RECS and will be optimized to meet precision requirements for key household energy consumption metrics, including total average site energy consumption and average consumption for individual fuels used in the residential sector. The precision requirements are based on relative standard errors (RSEs), which is defined as standard error of an estimate divided by the estimate itself (for RECS, mean is the estimate), expressed as a percentage. Table 1 displays the precision requirements, by geographic domain, for the 2024 RECS sample design.

Table 1. Relative Standard Error (RSE) Precision Requirements

Geography	Outcome Variable	RSE of Mean
United States	Average total site energy consumption	1%
	Average site electricity consumption	1%
	Average natural gas consumption	1%
	Average propane consumption	4%
	Average fuel oil consumption	4%
All Census Regions	Average total site energy consumption	2%
	Average site electricity consumption	2%
	Average natural gas consumption	2%
Northeast Census Region only	Average fuel oil consumption[1]	5%
All Census Divisions [2]	Average total site energy consumption	3%
	Average site electricity consumption	3%
	Average natural gas consumption	3%
50 States and DC	Average total site energy consumption	4%
10 Metropolitan Statistical Areas*	Average total site energy consumption	5%

* To simplify the sample allocation process, as described below, we will apply an additional requirement of 5% RSE for Average total site energy consumption for the remainder of the areas within the states that have selected MSAs.

Housing Unit Frame

The 2024 RECS will use an address-based sampling frame that starts with a vendor address list derived from the U.S. Postal Service's (USPS) Computerized Delivery Sequence (CDS) file. The frame that is derived from this file will include all mailable residential addresses in the 50 states and DC. Excluded from the frame will be non-active (No-Stat) addresses, P.O. Box addresses that correspond to households that also have a street address for mail delivery, and drop-point addresses that have five or

[1] A precision requirement for fuel oil consumption is only imposed for the Northeast as it is a highly regional fuel and used less frequently in other parts of the country.

[2] The divisions are the same as the standard Census divisions except that the Mountain division is split into Mountain South (AZ, NM, and NV) and Mountain North (CO, ID, MT, WY, and UT).

more units in a multi-family building. The frame will include P.O. Box addresses that are known as "Only Way to Get Mail (OWGM)" and drop-point addresses that have fewer than five units. In total, the USPS-based frame will cover about 99% of all occupied, primary housing units in the U.S.

Considerations for Only Way to Get Mail (OWGM) and Drop Point Addresses

A P.O. Box is a secured, uniquely numbered box located at a U.S. Post Office or a qualified business where a USPS customer can get mail delivery. Only the OWGM P.O. Box addresses will be included in the frame, because they are the only option for the corresponding households to get mail. These P.O. Boxes account for about 1% of addresses nationally.

For OWGM addresses, we will mail an invitation to the P.O. Box, but attempt to collect the physical address for responding households via a special question on the Household Survey form. Having a physical address for OWGM households will allow us to better link critical weather station data to household respondent data and to include these households in the Energy Supplier Survey's data collection.

Drop points are single-mail receptacles shared by multiple housing units and there is a variable on the CDS file that indicates which addresses are drop points. Drop points are found primarily in multifamily buildings with only 2 to 4 units, and only these units are included in the 2024 RECS frame. Drop points with 5 or more units are about 25% of the overall drop points, and are likely to be institutional units that are out of scope for RECS. For these reasons, drop points with 5 or more units are excluded from the 2024 RECS frame.

Drop points create operational problems such that specific households cannot be targeted for mailing reminders and incentive payments. To eliminate the need for a special data collection protocol for sampled drop point housing units, EIA and its contractor will statistically substitute any selected drop point address with a non-drop point multi-family (in most cases physical) neighbor Under this substitution approach, if a drop point unit was selected, a substitute unit in a non-drop point multi-unit building with the same number of units will be randomly sampled instead. This method was chosen because most drop points are expected to be similar in building structure to their non-drop point neighbors.[3, 4]

Determining Targeted Completed Cases

To meet the RSE requirements presented in Table 1 above, the 2024 RECS will use a stratified sample design by geographic area, which consists of 61 strata. The 61 strata include 40 states, DC, 10 MSAs from the remaining 10 states, and the remaining part of the 10 states that have MSAs. For example, GA will be divided into two strata, one for the Atlanta-Sandy Springs – Roswell MSA, and the other one for

[3] Lewis, T., McMichael, J., & Looby, C. (2023). "Evaluating substitution as a strategy for handling U.S. Postal Service drop points in self-administered address-based sampling frame surveys." Sociological Methodology, 53(1), 158-175. https://doi.org/10.1177/00811750221147525

[4] Amaya, A., LeClere, F., Fiorio, L., & English, N. 2014. "Improving the utility of the DSF address-based frame through ancillary information." Field Methods, 26(1), 70–86.

the rest of GA. To simplify the process of the sample allocation, we will also set 5% RSE targets for the remaining part of the 10 states that have MSAs.

The sample allocation optimization process will rely on a "bottom-up" approach. First, we will calculate the sample sizes needed to meet the RSE requirements for the average total energy consumption at the strata level (e.g., 4% for Idaho, 5% for the Atlanta MSAs, 5% for the remainder of GA outside the Atlanta MSA) based on the weighted means and standard deviations calculated from the 2020 RECS. These stratum sample sizes will then be used to compute the expected RSEs (standard error/mean) at the division level. If an expected RSE does not meet the required RSE (e.g., 3% in the Pacific division), then we will determine the minimum sample size needed to meet the division requirement using the division-level means and standard deviations calculated from the 2020 RECS. This minimum sample size will be allocated using the Neyman allocation method to the corresponding strata in the division. The updated stratum sample size will then be the larger of the sample sizes needed to meet both the stratum requirements and the division requirements. The same allocation process will be used to verify that the resulting stratum sample sizes satisfy the regional (2%) and national (1%) requirements. We will perform the optimal allocation, as needed, to arrive at the minimum needed to meet all the requirements for average total energy consumption.

Sample sizes sufficient for average total energy consumption requirements are not necessarily sufficient to meet RSE requirements for average electricity and average natural gas consumption. Of these, natural gas generally requires larger sample sizes in some strata, because not all homes use natural gas, and its use can vary greatly by geography. For natural gas, we will first check to see whether the division requirements are met using the minimum sample size for meeting average total energy consumption precision requirements in the first step. We will consider the proportion of HUs in each stratum that use natural gas (from the 2020 RECS). If division requirements are not met, we increase the stratum sample sizes to meet the division RSE requirements. Then, we increase sample sizes for regional natural gas requirements, if necessary, using the same approach as that for total average energy consumption. After adjusting for natural gas precision requirements, the RSE requirements including those for electricity, propane, and fuel oils will be checked and verified until no further adjustments are required.

As a result of this process, the minimum sample size needed to meet all RSE precision requirements is 14,379, as shown in Table 2. However, EIA resources allow for a sample of up to 20,000 completed cases, which would better meet the needs of data users at state and federal levels across a broader array of outcome variables. The difference between 20,000 and the minimum sample of 14,379 will be allocated to strata in proportion to the number of occupied HUs in each stratum per the 2022 American Community Survey (ACS) 5-year estimates of occupied HU counts.[5] Table 2 shows the allocation with the sample expanded proportionately to 19,998 respondents.

[5] The 2022 ACS 1-year estimates could not be used because occupied HU counts were not published for all counties in the 10 MSAs of interest.

Table 2. Allocation of Targeted Completed Cases, by Stratum

Stratum	Min. Sample for Precision Requirements	Sample Inflated Proportionatel y to 20,000	Stratum	Min. Sample for Precision Requirements	Sample Inflated Proportionatel y to 20,000
AK	246	258	MO	207	317
AL MSA	111	131	MS	206	256
AL non-MSA	185	251	MT	297	317
AR	208	260	NC	211	395
AZ MSA	264	345	ND	283	297
AZ non-MSA	126	167	NE	220	255
CA MSA	96	159	NH	253	277
CA non-MSA	487	1019	NJ	386	540
CO	227	329	NM	204	240
CT	284	347	NV	237	289
DC	461	475	NY MSA	145	165
DE	164	181	NY non-MSA	726	1046
FL MSA	145	247	OH	191	405
FL non-MSA	139	410	OK MSA	119	144
GA MSA	155	256	OK non-MSA	145	188
GA non-MSA	102	178	OR	266	341
HI	252	274	PA	491	723
IA	140	198	RI	218	237
ID	207	237	SC	201	291
IL	226	448	SD	187	203
IN	193	312	TN	486	607
KS	150	201	TX MSA	141	254
KY	308	387	TX non-MSA	557	913
LA	244	323	UT	184	232
MA	430	553	VA	254	401
MD MSA	147	196	VT	188	200
MD non-MSA	189	244	WA	266	399
ME	203	229	WI	194	302
MI MSA	113	191	WV	187	219
MI non-MSA	96	197	WY	246	256
MN	185	286			
			Total	**14,379**	**19,998**

Determining Starting Sample and Sample Selection

To achieve the number of completed cases at the strata levels indicated in Table 2, frame addresses must be selected to account for losses due to ineligibility and nonresponse. Using yield rates from the 2020 RECS, a starting sample of 56,776 HUs will be expected to meet the targeted 19,998 responding HUs.

Within each stratum, the frame will be sorted (i.e., implicit stratification) before performing systematic sample selection. The implicit stratification helps to ensure against selecting a sample that is nonrepresentative or disproportionate to the frame distribution of the variables used for sorting, and it is expected to improve the precision of the estimates that are correlated with the sorting variables. Below is the list of sort variables, in order, for systematic sample selection:

- Multifamily dwelling unit indicator (address level, from CDS file)
- Urban-Rural indicator (county level, from U.S. Census)
- Median income (census block group level, grouped into quintiles within each stratum)
- Heating degree days, base temperature 65F (census tract level, grouped into quintiles within each stratum)
- ZIP Code
- Carrier route (for mail delivery)
- Walk sequence (mail delivery sort order within carrier route)
- ZIP+4 (for addresses that do not have a walk sequence; otherwise, this does nothing)

As discussed in Supporting Statement Part A, there will be two phases of data collection for the 2024 RECS *Household Survey*. Phase 1 will consist of 20% of the selected addresses and Phase 2 will consist of the remaining 80% of the addresses. The Phase 1 and Phase 2 addresses will be selected simultaneously prior to the start of Phase 1 data collection using Chromy's minimum replacement technique.[6] This is a systematic selection technique that selects sample units from successively ordered zones created on a sorted sampling frame.

A reserve sample of approximately 8,000 additional addresses will be selected for possible use in an additional data collection phase. A portion or all reserve sample cases may be released in select domains if sufficient surveys are not completed, or EIA determines that precision requirements will not be met. The reserve sample will be selected at the same time as the samples for Phase 1 and Phase 2, but they will only be released as needed during an additional data collection phase as a risk mitigation strategy.

Weighting and Estimation Procedures

The estimation of household energy-use metrics from RECS is based on the responding sample. Each completed case of the responding sample will have a final weight associated with it, which reflects the number of occupied, primary housing units it represents in the target population. After sample selection, the initial base weight will be calculated based on the addresses selected for Phase 1 and Phase 2, which is the reciprocal of the probability of selection within each stratum. These base weights

[6] Chromy, J.R. 1979. "Sequential sample selection methods." In Proceedings of the American Statistical Association, Section on Survey Research Methods, pp. 401-406.

will then be adjusted for ineligibility, nonresponse, and benchmarking to result in a final weight. The sum of the final weights for all completed cases will equal the number of occupied, primary housing units at the U.S., state, and selected domains (e.g., by housing unit type) from the 2024 American Community Survey (ACS). The final weights will be used to produce a wide range of population estimates, such as average household energy consumption, average energy expenditure, percent of households using dishwashers, and so forth. In addition, a set of replicate weights will be created for variance estimation.

B.3. Maximizing Response Rates

The 2024 RECS will utilize robust contact, data collection, statistical analysis, and risk mitigation strategies to maximize the response rate and produce representative samples of key household energy subpopulations.

Maximizing Unit Response and Coverage in the Household Survey

Contact Strategy and Protocol

The 2024 RECS Household Survey data collection strategy will be similar to the 2020 RECS strategy and use self-administered Web and paper modes. The contact protocol calls for up to six mailings to sampled households over approximately seven weeks, including postcards and invitation letters.

Selected households will first be invited to respond via the Web instrument only (i.e., a "push-to-Web" strategy). All sampled households will be provided $5 cash as a token of appreciation at the first invitation, and responding households will be provided a $10 cash promised incentive.

For the second and third invitations, households will be offered the choice of responding via Web or a paper form. (The paper form will be mailed with the second and third invitations.) Web respondents to the second and third invitations will be provided a bonus $10 incentive for a total of $20 promised incentive. Paper respondents will receive the $10 incentive. This "Choice+" protocol was tested during pilot studies and used exclusively for the 2020 RECS Household Survey. As demonstrated for the RECS pilot tests, the Choice+ protocol results in higher response rates than Web-only or traditional Choice protocols where the same incentive amount is provided for all modes. The Choice+ protocol also incentivizes Web response and results in lower per-complete cost compared to the traditional Choice protocol.

All contact materials will include the same text in both English and Spanish. The Web questionnaire will be available in both English and Spanish, with the URLs for both versions included in all mailings. The paper questionnaire will be available in both English and Spanish and a RECS helpline (also in both English and Spanish) will be available throughout the data collection period for respondents to call with any questions or concerns.

Risk Mitigation Strategy using Responsive Design

The 2024 RECS Household Survey will use a responsive design with two phases of data collection. Phase 1, which will consist of 20% of the starting sample, will include two experiments to test whether alternative strategies can increase response rates and/or improve coverage:

- Experiment 1: Including QR codes on contact materials that link to the respondent Website.
- Experiment 2: Eliminating the pre-notice postcard and adding an additional reminder.

After Phase 1 is completed, EIA will analyze the results of the experiments and apply the optimal strategy to all of the Phase 2 sample. Phase 1 will also allow for a full end-to-end test of the data collection, monitoring, and data management processes to ensure they are performing as expected prior to releasing the majority of the sample in Phase 2.

All or portions of the reserve sample may be deployed as a risk mitigation strategy in case the overall target of 20,000 completed interviews or the precision targets are not met after Phase 1 and Phase 2 data collection. This phase can address any shortfalls in the areas that have lower-than-expected yield rates.

During Household Survey data collection, a robust dashboard of metrics will be used to monitor the data collection outcomes and data quality in real time. A variety of key metrics will be tracked, including comparisons with the frame and/or other external benchmarks, such as those from the ACS, to ensure data adequacy and sample representativeness to the population.

Item Nonresponse

Item non-response occurs when respondents do not know the answer or refuse to answer a question, therefore, no data is provided to a request. Item nonresponse has been generally low for prior RECS Household Survey data collections. For the 2024 RECS Household Survey, item non-response will be corrected using the hot-deck imputation method, which will preserve the distribution of the outcome variables and variance structure in the data. Hot-deck imputation uses the non-missing values of a variable as donors to impute a missing item. This method requires first identifying a set of "core" variables (e.g., housing unit type) that are potentially related to all items with missingness. Additional "class" variables are then identified for each individual item. These core and class variables are put into decision trees to determine imputation pools for each item. A recipient case borrows a value from a random donor within each pool. The procedure will be done using the Cyclical Tree-Based Hot Deck (CTBHD) imputation system.[7]

The CTBHD implements classification or regression tree analysis to select variables for construction of imputation classes and uses weighted sequential hot deck (Cox, 1980) to select donors within imputation class. Under CTBHD, variables with missing values are imputed sequentially, which accounts for questionnaire skip patterns caused by the relationship between a gate question and subsequent follow-up questions. In addition, the CTBHD implements a cycling process which can help stabilize imputed values.

[7] Martin, P., Wang, J., Frechtel, P., Sukasih, A., Lewis, K., Deng, G., & Kinyon, D. 2017. "Three-based hot deck imputation cycling —does cycling help?" Proceedings of the 2017 Joint Statistical Meeting, Baltimore, Maryland.

Nonresponse Bias Analysis

At the end of data collection, EIA will assess whether characteristics of survey non-respondents differ from respondents to the 2024 RECS. EIA will compare respondent data with external benchmark data of other federal surveys and data from prior RECS studies.

EIA will compare the distributions of key 2024 RECS estimates to those of the ACS, the American Housing Survey (AHS), and the 2020 RECS. Before unit non-response adjustments and final post-stratification adjustments are done, the estimates will be compared for a net estimate of bias at the lowest level of aggregation possible. Minimally, we will compare the type of occupied housing units, main space heating fuel, age of home, and household income data from 2024 RECS respondents with ACS, AHS, and a more robust set of variables from the 2020 RECS study.

In addition, during data collection, a real-time dashboard accessible by all project staff will be used to continuously monitor nonresponse bias for key variables.

Energy Supplier Survey

Energy billing data for household respondents will be requested from the utilities and other energy suppliers identified on Household Survey questionnaires. EIA uses its mandatory data collection authority to collect information from the energy suppliers on Forms EIA-457D, E, F, and G. The 2024 RECS Energy Supplier Surveys will occur after all phases of the household data collection are complete and employ similar methods for contacting energy suppliers as were used in the 2020 RECS ESS. These strategies include utilizing existing EIA contact information (e.g., 2020 RECS ESS, EIA electricity and natural gas supply surveys, etc.), sending advanced notification to suppliers to alert them of the data collection, and identifying key respondents within their organizations. All suppliers will then receive an official data request. This request will include instructions on accessing the survey Website and submitting data, and inform suppliers of the mandatory requirement for this phase of the RECS. Nonresponse follow-up procedures will include reminder phone calls and emails, as well as late notices from contractor staff and EIA. The estimated number of respondents for the ESS is shown in Supporting Statement Part A.

B.4. Test Procedures and Form Consultations

As part of determining the 2024 RECS questionnaire content, EIA consulted with stakeholders, reviewed lessons learned from prior rounds, and conducted pretesting activities. In general, changes to the questionnaire content from the 2024 RECS focused on:

- adding questions to reflect current household energy-related technologies, behaviors, and emerging topics,
- revising questions to improve response quality
- removing questions that were outdated or had not performed well in prior rounds

Under *EIA-882T: Generic Clearance of Questionnaire Testing, Evaluating, and Research, OMB 1905-0186 (Expiration 9/30/2025)*, EIA conducted online cognitive interviews for the 2024 RECS Household Survey.

EIA conducted 39 unmoderated cognitive interviews in December 2023 and January 2024 on select RECS topics. These topics included new or revised questions on space heating, appliance usage, electric vehicles, solar panels, and swimming pools. After the pretesting was completed, EIA drafted a comprehensive report of the results, including recommendations about specific questions. EIA finalized the 2024 RECS household questionnaire content based on these recommendations. The final questionnaire content reflects some updates from the changes proposed in the 60-day Federal Register Notice.

Also, under EIA-882T, EIA and its contractor are conducting usability testing on the general navigation, formats, and functionality of edits in the Web instrument.

B.5. Statistical Consultations

The principal EIA official directing the 2024 RECS sample design is Grace Deng, who can be reached at (202) 586-5329 or by email at shaofen.deng@eia.gov. The principal EIA official directing the 2024 RECS data collection, processing, and dissemination is James (Chip) Berry, who can be reached at (202) 586-5543 or by e-mail at james.berry@eia.gov.