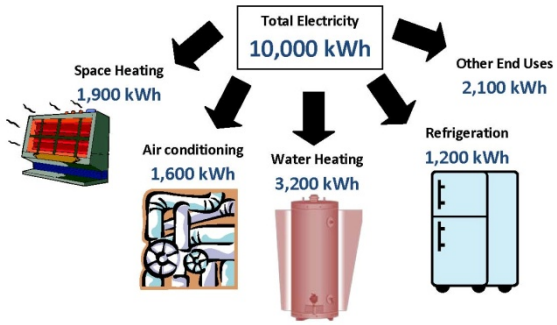
Independent Statistics & Analysis
U.S. Energy Information Administration

February 2013

Residential Energy Consumption Survey (RECS) End-Use Models FAQs

What is an end-use model?

An end-use model is a set of equations designed to disaggregate a RECS sample household's total annual fuel consumption into end uses such as space heating, air conditioning, water heating, refrigeration, and so on.



These disaggregated values are then weighted up to produce population estimates of total and average energy end uses at various levels of geography, by housing unit type, or other tabulations of interest.

Why are end-use models needed?

Information regarding how total energy is distributed across various end uses is critical to meeting future energy demand and improving efficiency and building design. Using submeters to measure exact end-use consumption within each sample household is the ideal way to obtain this information. However, it is prohibitively expensive and time-consuming to implement. Instead, EIA uses a statistical model to decompose total energy consumption into end-use components.

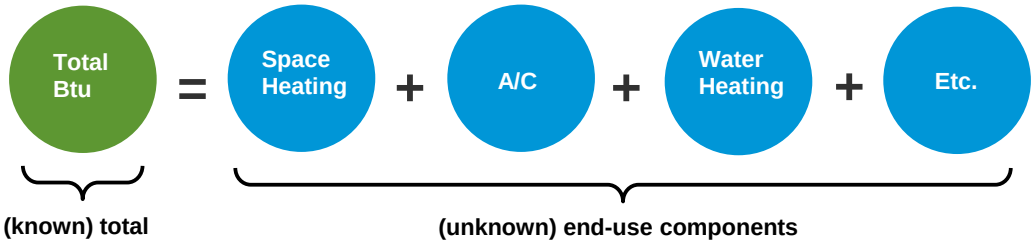
How many end-use models does RECS have?

RECS has five separate, independently-fitted end-use models, one for each of the main fuels: electricity, natural gas, propane, fuel oil, and kerosene.

What is the general structure of the end-use models?

The general model structure specifies total consumption to be a sum of terms, where each term represents an end-use component. Each end-use component, in turn, is composed of model parameters (unknown, to be estimated) and independent (known) variables. By estimating the

model parameters, the end-use components can be individually calculated, yielding the desired disaggregation of total fuel consumption by end use.



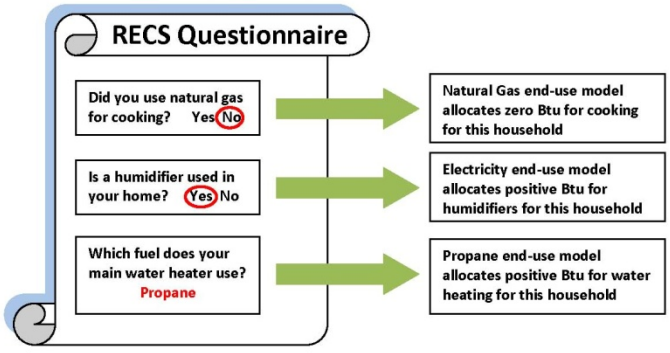
What are the dependent and independent variables in the end-use models?

In the RECS end-use models, annual fuel consumption (in thousand Btu) is the dependent variable. This annual consumption is calculated directly from monthly billing or delivery data obtained from utility companies that supplied fuel to the households. When these data are not available, they are imputed using prior RECS' end-use models. End-use modeling cannot proceed unless every household that used the fuel has a corresponding annual consumption value.

The independent variables vary by fuel and end use. Weather-dependent end uses like space heating or air conditioning use heating or cooling degree-days, as well as heated or cooled square footage. Most of the independent variables are housing unit characteristics such as age of home or type of heating equipment, household demographic variables such as number of household members, or equipment-related characteristics such as size.

How do the models determine what end uses should be included?

The inclusion of end uses is based on the respondent's answers to survey questions. Annual fuel consumption is allocated to the end uses, from traditional ones such as space heating and air conditioning down to individual appliances such as refrigerators.



How are end uses expressed in the models?

Regression equations are written to disaggregate the annual total consumption as a sum of components representing the end uses. These components cannot be negative because they are energy quantities. If the household indicated that the fuel was used for that end use, the component is positive; otherwise the component is zero.

There are many ways to formulate these end-use components mathematically. However, since its inception in the 1980s, RECS has always expressed them in every fuel model as a simple product of a base term and one or more *adjustment terms*.

What is the mathematical form for a base term?

The base term is usually a product of a base parameter and either the number of energy-consuming units or appliances for the end use (if such counts are available), or a binary variable indicating use of such appliances. In this way, the base parameter has a physical interpretation as the average annual fuel consumption per appliance unit.

More complicated end uses such as space heating include terms to account for heating degree-days and heated square footage. The end use component for clothes washers has traditionally included a factor for the number of household members. In these situations, the physical interpretation of the base parameter changes accordingly.

What is the mathematical form for an adjustment term?

Adjustment terms are meant to capture percent increase to or decrease from the base term consumption based on characteristics such as type, age, use patterns, etc, of the equipment or appliance. An adjustment term takes the form 1 plus the product of the associated adjustment parameter and a binary variable. The binary variable has the value 1 if the household satisfies the characteristic for which the adjustment is made, 0 otherwise.

Are the models estimated separately for each end use?

No. Because the dependent variable is total consumption, all base terms and all adjustment terms are included simultaneously for all end uses within a fuel model, by household. Thus, for example, the electricity model does not isolate the air conditioning end use when estimating model parameters. All terms for space heating, water heating, refrigeration, and the various appliances are also there together with the air conditioning term before proceeding with parameter estimation.

Why are the models considered non-linear if the end uses appear additively in the equations?

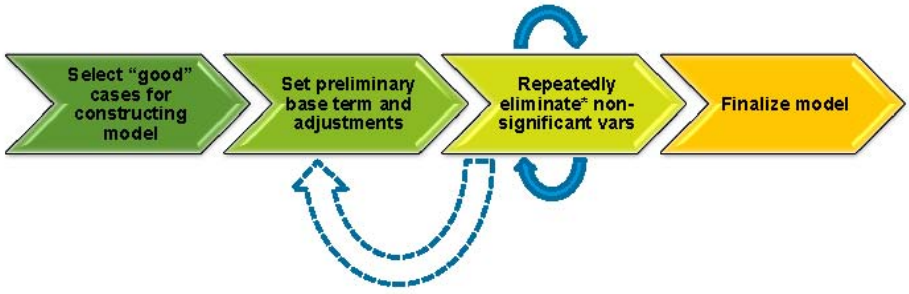
While the end uses do appear additively in the models, one or more end-use components are a product of a base term and at least one adjustment term. This means there is a multiplicative interaction between model parameters, making the models intrinsically non-linear.

What is the form of the error term in the models?

Since the 1990 RECS the error term in the models have been specified as the difference between the fourth roots of the actual and model-estimated total annual consumption. Prior RECS rounds tried other forms of the error term such as the simple difference between actual and modeled consumption or the difference between their logarithms. However, the variance of the error term under these forms tends to increase with the model-estimated consumption. Further, the errors are positively skewed. The present form of the error term comes closest to being normality distributed with constant variance.

How are model parameters estimated?

Because the models are inherently non-linear, RECS uses SAS proc nlin with the Marquardt option for parameter estimation. Proc nlin is an iterative procedure for estimating non-linear regression models. It requires initial values and bounds for the model parameters, and successively replaces them with values that minimize the sum of squares of the error terms. This iterative process stops when the difference between the sum of squared errors in the current step and the previous step is almost zero. When all parameters have valid estimated values at the stopping point, the model is said to have *converged*.



What initial values and bounds are used for parameter estimation?

The initial values and bounds come from the final parameter values from the prior RECS models, if they appeared in the final form. Otherwise, reasonable initial values suggested by energy experts at EIA are used.

Because of the positivity requirement for end uses, all base parameters are positive. Adjustment parameters associated with characteristics expected to increase or decrease the base consumption are given positive or negative bounds, respectively. For example, Energy Star

appliances could be reasonably expected to consume less energy than non-Energy Star ones, so adjustment parameters for end uses involving Energy Star appliances use negative bounds. For the same reason that end uses must be positive, negative adjustment parameters cannot have values less than or equal to -1.

How do the models know which base or adjustment terms to include for an end use?

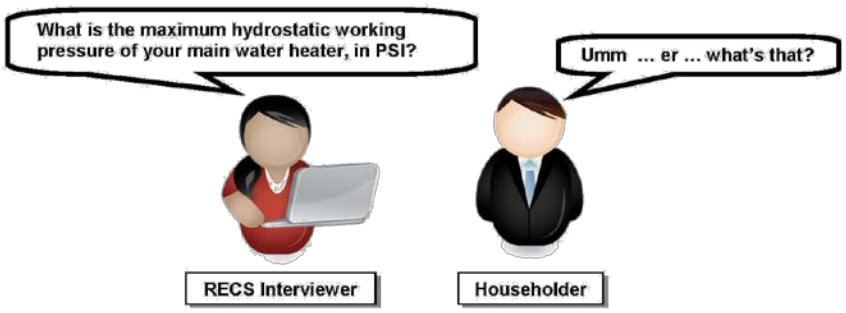
The models start from the final form of the end-use models from the prior RECS round, with updates to account for new survey variables or variables that no longer exist. The over-arching principle is to include, for every end-use component, the largest possible set of relevant information from RECS.

Once the initial equations are specified, each run of proc nlin in the parameter estimation phase may retain or remove base or adjustment parameters and their associated terms. The retention or removal of model parameters is strictly based on statistical criteria embedded in proc nlin. As such, the survey data ultimately determine which base or adjustment terms make it to the final form of the end-use models.

If a base or adjustment term is dropped by proc nlin, it is not forced back into the model. A modified version of the term, however, may meet the criteria for inclusion. For example, the variable indicating presence of an unusually high ceiling in the housing unit by itself rarely makes it to the final form of any RECS end-use model. When combined with another building envelope variable, however, it could find its way into the final form.

Do any of the base or adjustment terms include engineering quantities?

No. RECS does not have a mechanism for capturing important engineering quantities because householders cannot answer technical questions like "What is the R-value of your wall insulation?" or "What's the SEER rating of your air conditioning unit?" Indirectly, however, the impact of these engineering quantities on energy consumption is reflected in highly-correlated variables captured by RECS and retained in the end-use models.



How are the final estimates of end-use consumption calculated?

After the final models are specified, the model-estimated annual consumption, calculated by adding the estimated end uses, will most likely be unequal to the actual annual consumption for a household. For this reason, the estimated end uses cannot be used directly. Instead,

normalization is done by multiplying each estimated end use with the ratio of the actual total consumption to the estimated total. The results of this calculation become the final end-use consumption estimates for the household.

For example, consider a household with an actual annual consumption of 40,000 million Btu of electricity. Assume that the household uses electricity only for air conditioning, refrigeration, and other appliances. After end-use modeling, the estimated end uses may be, 16,800 for air conditioning, 10,500 for refrigeration, and 14,700 for other appliances. The sum of these estimated end uses is 42,000 which is not equal to the actual total of 40,000. The final end-use estimates for this household are then calculated as:

16,800 * (40,000 / 42,000) = 16,000 for air conditioning,
10,500 * (40,000 / 42,000) = 10,000 for refrigeration, and
14,700 * (40,000 / 42,000) = 14,000 for other appliances, all in thousand Btu.