

Time to Update Those EULs! An Affordable Way Forward

Ann Gibbs and Lisa A. Skumatz, Skumatz Economic Research Associates, Inc.

ABSTRACT

Let's face it – EULs are a critical part of the benefit cost equation, and existing EULs are 20+ years old and very outdated. Given that Benefit Cost calculations include EULs as a key calculation component, a periodic review of these values should be an expected part of evaluation plans. However, the traditional estimation approach needed to update them is cumbersome and expensive. This study suggests that an alternate approach – including multiple years of cohorts of past participants for measure retention and failures – is less expensive, faster, applicable to all types of measures, and performs well.

Three main problems exist with the current approach to EULs: having to wait for years after installations to conduct the first studies, the cost, and the studies have to be repeated several times to nail down the lifetime estimate.

We recently got the chance to pilot-test a new, straightforward approach to the problem, and used it to develop statistically based EUL estimates for thermostats in several states. The technique was then applied in a northeastern state, and they have re-evaluated the measure lifetimes for an array of residential and commercial measures including heat pumps, heating equipment, and a number of other measures.

Since programs have been installing similar equipment for many years now, cohorts of past participation can provide historical failure data. Surveying these groups allow one study to have enough failures and over a long-enough period to provide results now, and inexpensively.

Introduction

Existing energy efficiency (EE) measure lifetimes currently in use are much older than most of the other inputs used in measure screening and benefit-cost test calculations. This paper suggests that a revision to the traditional sampling approach for measure life studies can dramatically cut cost and turnaround time, allow the update of aged estimated useful lifetime (EUL) values, and provide the remaining useful lifetime curves that are now needed for early replacement programs.

The revised approach uses the fact that programs have been in place, delivering similar measures, for a decade and more (depending on measure). Rather than sampling for failures repeatedly from one cohort (or program year of participants), samples can be pulled from the participants that received the measure over the period since the program(s) began installing the substantially-same measures. This avoids the need to repeat the study with time gaps to allow time for failures, and provides estimates quickly and inexpensively. The paper discusses this approach, and provides examples of application of this approach in several studies.

Background on EUL Estimation

A critical part of the benefit cost (B/C) equation used for screening energy efficiency (EE) measures, program, and portfolios, are Measure Lifetimes, or Estimated Useful Lifetimes (EULs).

An EUL is defined as the number of years at which one half of the measures are still in place and operating. To conduct a statistical analysis estimating EULs, there are only two pieces of information needed:

- 1) What share of measures installed under a program have failed vs. are still in place and
- 2) When those failed measures stopped operating

The industry-standard method (for example, Franco et. al 2018, USDOE 2016) for gathering this information is to follow one program year (PY) cohort, with known installation dates for specific measures. Several years after the measures are installed, a first EUL study is conducted (gathering the data listed above). This study is followed by a second EUL study on the same cohort, collecting the same data, conducted several years later to allow time for additional “failures” of the measures.

These studies rely on a site visit, or, for certain measures, surveys can be used to collect the needed data (criteria for measures suitable for surveys are discussed later in the paper). Then statistical modeling is conducted to estimate that median lifetime value, employing hazard functions. Depending on the measure, EE EULs may vary from a few years (behavioral programs) to 25 years or more (insulation or shell measures). The statistical model uses information on: both failure and non-failure (0 and 1), and the elapsed time associated with each condition.

The data is called "right censored" data since the time of failure to the right for some observations are missing (those that have not yet failed), so the lifetimes for all measures in the sample are not known. The statistical approach accounts for the incomplete data set in developing the predicted survival rates for measure based on the censored data. This statistical modeling work has been accomplished for the several decades of EUL studies using well-known statistical packages (e.g., SAS and others). These packages allow multiple distributions to be used for the decay functions (specifically log normal, log logistic, Weibull, exponential, and gamma distributions), and Weibull is often a strong performer on fit. The median lifetime is the estimated EUL.

As mentioned, most EUL studies are conducted a few years after the program installs EE measures and a few years later to allow more time for additional installed measure to have failed. Failures are the observations that are used for the estimation; no failures means the statistical model does not have the observations needed to estimate the relationship between time and failures. Long-lived measures need a longer gap to allow for the failures that help statically nail down the function. Many years may need to elapse to get sufficient failures for some long-lived measures in order to get enough failures to help “identify” and lock-down the statistical curve. Second, to find these failures, especially for long-lived measures, can take a large sample, and the cost is high to conduct large-scale surveys (or sometimes on-sites) needed to find “failures” to support the estimations.

Besides EULs, remaining useful lifetime (RUL) values are also needed. Remaining Useful Lifetime (RUL) is a curve or function that is developed as part of the estimating of an EUL. The curves identify the median life for a measure (the EUL), but also provide information on the years of “life” expected for a measure of any age. For example, the expected remaining useful life of a measure that has already lived past the median EUL is not the EUL; instead it is a greater number, and the estimated curves provide these values across all ages of equipment. This number is needed for programs that allow for “early replacement” of measures; that is, measures

that are incentivized early, or earlier than “replace on failure” (ROF). For early replacement measures, utilities generally claim savings against a two-part baseline: savings relative to the existing equipment that was being replaced for the number of years that equipment would be expected to live (using the RUL curve and the removed equipment’s age), and savings relative to a standard efficiency model of the same measure for the time period of total measure life minus RUL.

While the statistical approach to EULs has been known and used regularly since about 1990, the values adopted by utilities are not always statistically-derived. There are a number of other approaches to developing EULs listed below, and unfortunately, many underlie the existing values. Work by Skumatz (2012a) and for the San Diego Gas and Electric and Northwest energy Efficiency Alliance and its Regional Technical Forum (RTF), developed a hierarchy of quality for sources for EULs, shown below.

Figure 1: Hierarchy of Decreasing Quality of EUL sources

Hierarchy of Decreasing Quality of EUL sources • Local statistical studies for the specific measure (with gradations of quality if not local, or not exactly the same measure), • Well-documented values from other well-matching state tables, • ASHRAE service life tables (limited measures), • Stock turnover models (equipment groups only, a la Welch and Rogers 2011), • Manufacturer warranty, Delphi studies, expert discussion / collaborative process - which might be suitable for equipment that is too new (or too few) for a local study.¹
--

Source: Skumatz 2012

The Problems with Existing EULs and Methods

There are several main problems with the existing approach to EULs.

1. The existing values are aged – often more than 20 years old -- yet they are a critical part of the Benefit-Cost and screening methods used for measures, programs, and portfolios for programs across the US. Nearly all the other variables included in the benefit-cost work are updated on a much more regular basis.
2. Many are not statically-derived or based, but draw from previous lists, consensus approaches, or other sources. Sources for many values are not well-documented or source other lists without strong underlying documentation.
3. There are variations in values between utilities that can’t be well-explained with expected causal factors.
4. The age of values may risk assigning lifetimes or technologies that are no longer appropriate.
5. The expense of studies conducted the traditional way is high, leading to infrequent studies, and avoidance of the statistical studies that would provide firmer foundations for EUL values.
6. The gaps between first and second estimation leads to a time lag for the availability of EULs using the traditional estimating methods.

¹ Using these types of substitutes for statistical work in other cases would be inferior, but the priority of an imported value would depend on risk of a poor decision or poor guidance on programs and expenditures, importance of the measure to the portfolio, variability or uncertainty of the lifetime for the measure, and other criteria.

7. The age of studies and incomplete sourcing has left some or many measures without relevant curves to provide values for increasingly-needed RULs.

Age. Work by the author (Skumatz 2007, Skumatz 2012a) has demonstrated that utilities are almost always using information that is more than 20 years old and has been adopted from other utilities for measure lifetimes. Previous studies by the authors (Skumatz 2012a), including work for Southern California and the Northwest’s Regional Technical Forum (RTF), have examined the source of these NEB values in use by states across the US. When traced to their source, some were based on statistical studies conducted between 1990 and 1996 by Bonneville Power Administration or the California utilities², and work by some other states, but a large number were only traceable to ASHRAE³ tables that are even older. Others didn’t use statistical studies, but instead were developed “by consensus”.

Weak sources and methods. Previous research by the authors (Skumatz 2007, 2012a) shows that a majority of the EULs currently used by utilities across the US are often poorly sourced or circular, with even “recent” values in TRMs commonly based on other tables, that then source to other lists and tables. This referencing of references leads to antique sources that are commonly more than 20 or 30 years old. These values may not actually apply to current technological design or obsolescence / failure patterns. Borrowing existing values from previous tables calls into question the statistical validity of the values, which were often based on consensus agreements, not statistics or studies. The statistical EUL studies that do exist are for limited measures, and there are numerous measure gaps.

Variations in values without clear cause. Weak, cascading, or consensus sourcing, has led to other issues. Skumatz (2007) assembled and reviewed values for measure lifetimes used by different states across the nation. It shows that adopted values for same measures in different states can vary by factors of two to four, with no relationship to engineering reasons or logical differences for the variations (climate, hours of use, etc.). The measures that varied are shown in Figure 2 below. These EUL values are used in important equations that influence the expenditures of many millions of dollars. Confidence in these calculations is harmed if the measure lifetimes do not track or compare in sensible ways.

Figure 2: Variations in EUL Values Between States

EULs with variations >50%	Boilers, CFL fixtures, Exhaust fans, Faucet aerators, showerheads, Insulation, Programmable thermostats, Tankless instant water heaters
20%-50%	Central AC, CFLs, Dishwashers, Freezers, Furnaces (Gas & electric), Air Source Heat pumps, Pool pumps, Refrigerators, Room/window AC, Stoves, Tank wrap, Water heaters, Whole house fans, Windows
<10%	Clothes Dryers, Duct sealing, Oil Furnace, Other Heat pumps, Solar screens

Source: Skumatz 2007

² California had an excellent requirement in place in the 1990s, mandating EUL studies for every program, coving the top 4 measure in each program, and assigning two EUL studies, with a gap of 4-5 years between the first and second. This added substantially to the body of work on EULs. The major flaw (which can be easily corrected in future requirements) was that the top four measures were often the same, so the result was many EUL studies of the same measures.

³ American Society of Heating, Refrigerating, and Air Conditioning Engineers. The tables are based on (self) reports of equipment age, not statistical studies, are very old, and are calculated as averages, while measure lifetimes for benefit-cost work are defined as median ages.

Technology changes. Sourcing can cause difficulties in justifying some of the EUL values, but even if the sources are statistical studies, the results may be too old and actually apply to very different technologies than may currently be used for energy efficient measures. Given that Benefit Cost calculations include EULs as a key calculation component, a periodic review of these values should be an expected part of evaluation plans, like updating impact, or NTG or other inputs.

Cost and Time Lag. Conducting two or more (lagged) EUL studies to get enough failures to nail down a survival curve and the associated EUL value is generally going to be about twice as expensive as getting the work accomplished in one study (because the desired sample sizes will not vary). Conducting studies using surveys rather than on-sites will further amplify the savings. EUL studies conducted using the new approach have cost on the order of \$50,000 (e.g., email-based surveys with valuations and some use of photos, for both the main case studies listed below). The savings for similar measures would be at least another \$50,000, and studies for commercial measures will likely save even more. The savings for conducting one, rather than two, on-site surveys are multiple times these values, relating to the much higher cost-per-complete for on-site work than web (or phone) surveys. Furthermore, studies can be designed to cover more than one measure, enhancing the savings per EUL. Finally, the results from one study are available in a much more timely way than one that has two components, offset by 4-5 years.

RULs. Remaining Useful Lifetime (RULs) are needed for early retirement programs and measures, and these values are lacking for many measures, because the underlying EUL sources were not statistical, or are based on statistical studies that did not provide these tables or curves. For EULs derived from weak, non-static, old, or consensus sources, these RUL curves are non-existent or difficult to find. As early replacement programs increase the missing RUL curves are needed. The updated approach we are suggesting provides these values.

The Key Elements and Benefits of the Recommended Approach

The proposed approach resolves all these issues. This study proposes that the studies can change the sampling. The population could be defined to increase the time frame for the sample to include multiple program years or cohorts that have installed the measures over time. This can include sample from population of cohorts of multiple program years as far back as available and for which the (reasonably) same measure was installed. This longer-term population is used as the sample from which to conduct a measure life study. This immediately provides the timeframe for multiple years of failures. Furthermore, if the equipment is just the right type, it might not need site visits, and inexpensive web surveys could suffice. This isn't rocket science, but for some reason, it just hasn't been done routinely. This alternative approach provides a number of benefits.

1. Has lower cost per completed EUL study: It is less expensive than the traditional two-study lagged approach (conducting multiple studies, years apart). Coupled with survey-based data collection, it can be especially cost-effective, combining savings from a single study (vs. lagged multiple studies) and survey-based (vs. on-site). Using either survey or on-site methods, it is suitable for the wide range of residential and commercial measures.

2. Produces more immediate results: Studies can be conducted immediately and to provide statistically-based and updated replacement EULs for PSDs / TRMs without waiting multiple years for failures.
3. Allows more frequent and more defensible studies that provide statistically-based values, including RULs. The cost and timing (and elimination of the gap between studies) makes it feasible for more frequent statistical studies of EULs, which are important components of measure, program, and portfolio screening. It also provides much-needed cumulative lifetime curves needed to specify “remaining useful lifetimes” for measures – a value that is increasingly important for application to “early removal” rather than “replace on failure” programs or “events”.

These basic characteristics can resolve all the issues mentioned above as weaknesses of current EUL methods and adopted values for a wide range of measures. It also allows more frequent updates, appropriate to updated technology, and it also allows more utilities to conduct their own studies, rather than borrowing values from utilities that may not have the same climate or other factors that affect operating hours or operating conditions. Conducting studies in this manner will allow utilities to replace old, somewhat dodgy-in-origin, and potentially inappropriate, EULs.

The remainder of this paper presents best practice guidelines for studies using this method, including considerations in design and suitability, and case studies where the method has been applied.

The Conditions for Success: Sample, Data Needs, Measure Types, and Failures

Major users of his method will be utilities, in association with energy efficient measures, working to be provided more verified numbers for the measure lifetimes. However, the method can also be applied to broad samples of businesses or households, and used to identify lifetimes of a variety of measures. Measure life estimation projects not tied to utility direct install or other programs can also be successful (as noted in the thermostat example below). They can be used to identify lifetimes of non-EE measures (see uses of these result below) or EE measures.

The major change for this approach is the ample; specifically using multiple years. However, given that approach can be coupled with surveys to deliver particular savings, we also address needs associated with the use of surveys. The conditions for success for the utility and broader case are described below.

Sample Source and data needs:

- Utility Programs: Studies conducted for utility programs need contact information for current and previous cohorts of program participants receiving the measures, and date of installations from program records. The contact information will need to be sufficient to reach a respondent that can verify the presence and/or failure date reasonably reliably. Tenure (for all) and job title (for commercial measures) would also potentially be expected affect the ability to respond reliably.
- Sample and extra data needs for measures not tied to energy efficiency programs: A broad sample of the population can be achieved using purchased contact information for use with post card or email to web survey, or purchased panel sample approaches.

However, first, to be feasible, the measure must be in place in a large share of households or businesses, or the qualifier questions will reject a large sample, driving up costs. In addition, additional data will be needed from this sample. Respondents will need to be asked for an estimate of installation date for the equipment, not just removal date or continued retention of the measure.

Suitable measure by data collection method. On-site work verifying presence and operability of the measure can be accomplished with almost any measure, for utility or non-utility situations. If surveys are going to be used, only certain measure types can be routinely successful.

- Suitable measures for utility programs - surveys: The list can be fairly broad. The questionnaire can refer to an installation date, looking for failures (or not) after that date. Photos can be submitted (for a valuation sample or full sample) to verify efficiency levels, and the survey can drill down on failures and replacements. The best measures are some subset of unitary, identifiable, important enough to notice, and visible for inspection. Unique features can also help.
- Suitable measures for non-utility installations – surveys: For *broad* samples of respondents to be able to answer intelligently about an *efficient* measure's dates of installation and failure, that measure and its installation will need to be memorable (or uniquely describable), visible, and identifiable, or measures that may have pride associated (solar, etc.). Measures with regular interaction (smart thermostats) or measures associated with memorable retrofits (insulation, heating equipment) can also suit. However, unless samples are drawn from purchasers, the frequency of the measure in the population must be considered, or low incidence will drive up costs dramatically.

Ability to define and respond on “failure” clearly. For an EUL study to be conducted successfully, the respondent needs to be able to identify the condition “in place and operable” vs. not (“failure”). This determination does not change based on new method proposed for the sampling, but it deserves mention. For both utility and non-utility samples, “in place and operable” can be defined in a manner or with parameters that are reasonably clear for the knowledgeable field staff likely to be conducting on-site inspections. However, this may be less clear or more difficult for survey respondents. The condition needs to be able to be observable or determined with a reasonable degree of clarity for the typical survey respondent. Some measures may fail clearly; for instance, it will no longer function dramatically, like water heaters, or lamps that no longer turn on. Measures that no longer operate “efficiently” may be much harder for survey respondents to identify and these potential nuances in definition affect all EUL studies. If this is part of the definition set for “failure”, these conditions will need to be observable for a measure to be suitable for a survey-based EUL study. EUL studies have historically taken the definition as in place and operating, without nuance as to how well or how efficiently the equipment is still operating. A series of follow up questions may help illuminate the nuance (repair frequency, etc.) but the observability of the condition will need to be considered for a measure to be suitable for a survey-based EUL.

A non-exhaustive list indicating examples of types of measures that are suitable for each situation and data collection type are summarized in Figure 3 below. The study project manager will need to consider these types of factors to determine which data collection approaches are feasible.

Figure 3: Examples of Suitable Measures by Data Collection Method and Utility / Non-Utility Case

Case	Suitable for Surveys	Not suited to survey
Utility Program Measures	Full banks of LEDs; thermostats, identifiable motors and pumps; HVAC equipment including the variety of heat pumps, furnaces, boilers; kitchen equipment; water heaters; process measures, insulation ⁴ , and many others	Screw-in LEDs, ⁵ unless replaced widely; Faucet aerators and showerheads, air compressor O&M measures,
Broad Population Measures	Advanced, smart thermostats; programmable thermostats; general HVAC equipment types, insulation, possibly water heaters ⁶ and others. Suitability affected by low incidence: Solar equipment; Certain commercial equipment (dimming refrigerator cases, etc.); others	Screw-in LEDs, efficient faucet aerators and showerheads

Prioritizing suitable measures. Criteria to keep in mind when selecting measures appropriate for web survey data collection will include which measures had sufficient participants currently and back a number of years that could make this multi-cohort process really powerful.

Concerns and Potential Issues

A study's use of a long-term ample and survey-based methodology can result in much lower fielding costs than two studies conducted using in-person site visits to verify the operation and age of equipment of the same number of respondents. Certainly, there are potential issues and complications affecting the suitability of including longer-term samples, as well as the risk that relying on self-reported information exposes the study possible respondent error.

- Contact information and tenure influences: The most relevant contact information for the *current* occupant of the home (or business). This contact information will rarely be available on a large scale, so this modification will bias the respondents toward those with long, stable tenure in the location.
- Time patterns in response rates: Response rates from older cohorts may be different and there may be concern this might affect the estimates. Changes or failures that happened long in the past may not be well-recalled, and the date of failure may be harder to pin down. The failure dates are important for the estimation work. Some errors, if random, may not be problematic, but if there is a pattern, (e.g., tending to state the failures as more recent, or earlier than reality), this will affect the estimates.
- Equipment that changes over time: Also, the equipment installed over time may, in fact, have changed in ways that affect the measure lifetime. It may have switched underlying technology, or used cheaper components in later years, or other changes. Including years of cohorts with measures not comparable to current measures will produce biased EULs.

⁴ Insulation can work if the occupants can be asked about remodeling to the wall / ceiling and absence of water damage.

⁵ Not suited to lightbulbs unless the bulbs are marked at installation

⁶ Hot water heaters are identifiable and singular in households, but *efficient* measures are not always easily identified, unless the installed measures were not wrapped and efficiency stickers were left in place. In this case photos of the model numbers would need to accompany responses, and a larger number of responses would be needed to achieve sufficient efficient models and their failures to conduct the EUL study.

- Survey data collection: The appropriate use of surveys as part of the long-term will be affected by whether self-reports are faulty, or too much less accurate than the on-site visits often used for EUL studies.

To allay the reader's potentially growing concerns, validation is a common method of measuring and correcting sources of bias, and several of these considerations have already been researched, at least to some degree.

- Survey data collection can be accurate: Work as early as Skumatz and Bordnere (1996) used phone surveys rather than on-site as the basis of measure life studies for residential hot water heaters installed under a PG&E program. That study revisited a sample of homes in-person to confirm presence / absence of the program installed measure and drill-down on failure dates. The study found very few variations and the survey method was strongly recommended for similar types of measures (one unit, easily-identified, clear failures) for other studies.
- Follow-ups and photos can be used to uncover or correct bias: The thermostat case studies discussed below used a sample of on-sites and re-surveys with photos to confirm number, age, type, and presence / absence / failure dates of the measure. Correction factors (which were small) were reported and used to adjust various means and statistics in the studies.
- Validation work indicates absence of time patterns in responses: The Connecticut study mentioned below conducted validation work on some of these possible sources of bias. They used photograph attachment / uploads and follow-up verification interviews to confirm presence / absence of failures, and found little error in the self-reported information. They also used the interviews with early cohorts to examine directions and size of possible failure dates, and did not find evidence of directional or size patterns. This was very reassuring information. The Connecticut study is also planning to conduct estimates over subsets of cohorts to test whether measure lifetimes changed over time, potentially due to physical changes in the equipment over time.
- In addition, the thermostat case study demonstrates that estimates can be derived in the non-utility case, which poses challenging sample conditions and the requirements for customers to remember additional information.

These indications were determined in case studies that were carefully designed and considered – with thought to the appropriateness of both longer sample and the selected measure(s). Going forward, other studies should also conduct validation work to explore evidence of possible influencing factors; this is best practices for surveys and studies generally.

Examples of the Proposed Approach in Action - Case Studies

Case Study 1: Series of State Thermostat Studies

From 2009-2015, the authors (Skumatz Economic Research Associates / SERA) were tasked with identifying the amount of mercury that would be coming off walls in the form of old mercury thermostats annually, so state or industry recycling programs could be sized and

recovery goals could be set. One step of the work⁷ was to estimate measure lifetimes of traditional thermostats, which was coupled with the age and count of existing mercury thermostats to identify the annual natural market outflow and resulting tons of mercury. The work started in California (12/2009), and later was replicated in Illinois, Rhode Island, Maine, and New York in 2014 and 2015. Key elements of this study follow:

- The studies were not associated with a utility program, so no contact information from participants was available. The studies sampled from the general public, and from general business populations, using purchased samples of addresses for homes and businesses, and emails where available for businesses. The sample stratification included single-family, multifamily, small and medium commercial, and large commercial (stratified by employees).
- Unlike a utility study, the study sponsor had no direct relationship with the respondents. This harmed response rates compared to a utility study (1.9% - 6.5%, depending on group and state), but the response rates were higher than open opinion surveys (like Pew Research, etc.). Multiple mail-outs, and phone follow-ups, and drawings for gift cards were used to generate sufficient responses.
- The studies used a survey-based approach. The authors were confident this could be successful without the need for on-site visits because of a study conducted by the authors in 1995 (Skumatz and Bordner 1996). This study used a phone survey approach to conduct a residential water heater lifetime study in California. Occupants were asked about whether the measures were still in place and operable; staff were then sent to a sample of respondents to verify the measure on-site. In the case of such a visible measure, the survey approach performed well, and the verification found very little respondent error.
- The studies did not know which types of thermostats were in place in any particular building, or the install date of the equipment. The thermostats included all manner and make of thermostat, and photos were needed for households and businesses to identify thermostat types needed by the study. Up-front field work identified which thermostat types included mercury, and these were included in the photos. Respondents were not asked whether their thermostats included mercury, partly to avoid alarming them, and partly because the study did not want to incur safety liability from respondents opening the thermostats. These questions were answered successfully, based on on-site verification of a sample of residential and commercial respondents.

Essential to the success of the study was the fact that virtually every building has at least one thermostat. The series of studies demonstrated that data on dates of installation, measure type, removal dates, and other needed data could be collected from at-large samples, with the right measure. Households and businesses responded and provided data that were by and large reasonable, even when we were asking difficult questions about the past. Sufficient data could be collected, and the estimation methods worked. The data, with some corroborating questions,

⁷ Steps estimated average number of thermostats per building in each strata, percent containing mercury, age of the existing thermostats, a measure life curve, natural annual market outflow, share that would be mercury, and tons of mercury flowing out per year. The sequence and research steps were outlined in the papers cited and a 2014 Western Economics Association conference paper and presentation

were viable and sufficient to develop EULs that were reasonable, compared to estimated 25-year EUL for efficient programable thermostats in DEER and CPUC sources.⁸

Figure 4: Results from Case Study Thermostat Studies

Samples and response rates	Resulting EUL estimates for traditional thermostat
Single family: 385-584, response rates of 4.3% - 5.5% • Multifamily: 57-469, response rates 2.6%-6.5% • Small commercial: 60-168, response rates 1.7% -2.4% • Medium commercial: 49-181, response rates 1.3%-2.4% • Large commercial: 13-46, response rates 4.1-4.5% • All residential: 492-1053, 4.1%-5.8% • All commercial: 230-477, 1.9%-2.4%	California: 27.5 years Illinois: 27.5 years Rhode Island: 36 years ⁹ Maine: 28.6 years New York: 32 years.

Even in somewhat challenging conditions, as these studies experienced, this one-time survey approach can work, either to estimate EULs, or to verify ranges for EULs currently in use, or to derive remaining useful lifetime curves (RUL) that may be missing from current EULs in state Technical Reference manuals (TRMs).

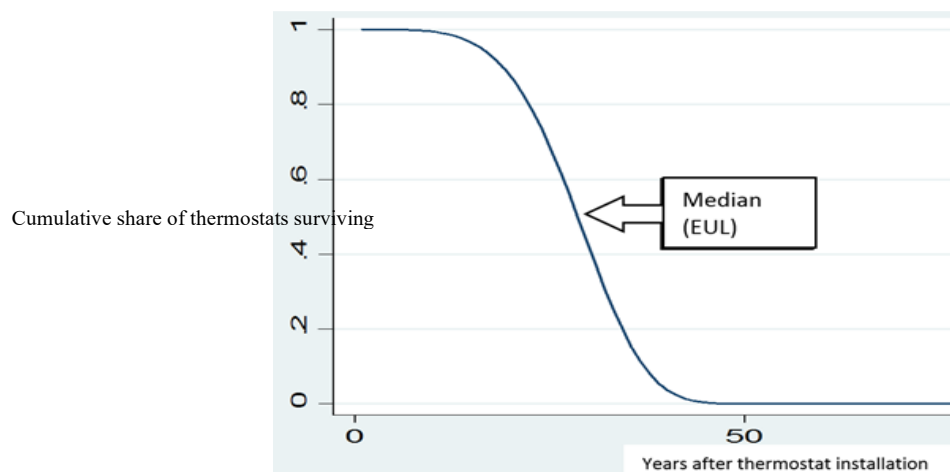


Figure 5: Determining Median EUL.: Percent of Thermostats Surviving by Age of Thermostat (Vertical axis is share of thermostats surviving, horizontal axis provides number of years after thermostat installation)

Case Study 2 – Residential Measures in State of Connecticut

A few utilities within the last several years are revisiting measure lifetime estimates, including Connecticut, but this long-overdue exercise is not yet common. The author is lead evaluation oversight contractor, responsible for developing the State's Evaluation Plan for the State's Energy Efficiency Board. We worked with stakeholders to commission a study specifically using this new multi-cohort method for a multi-measures residential and commercial EUL / RUL study. The authors designed and oversaw the study; the detailed design work and

⁸ Note California has newer housing stock than at least New York and Rhode Island.

⁹ Rhode Island was an outlier for multiple variables in the study, including a bulge in thermostat age that may link to past building permit bulges found.

implementation work was conducted by Michaels Engineering.¹⁰ Key elements of the study follow.

- The study was designed to target the most appropriate measures, (see Figure 6). These were prioritized, mindful of which measures were appropriate or not for web survey data collection, which were responsible for large or increasing savings, and which measures had sufficient participants currently and back a number of years (to 2011) for similar technology and with contact information – conditions that make this multi-cohort process powerful.

Figure 6: EULs and RUL curves currently being estimated in CT

Residential measures	Commercial measures
• Air Source Heat Pumps • Ductless Heat Pumps • Heat Pump Water Heaters • Natural Gas Furnaces • Insulation • Central ACs	• Split system and unitary ACs and HPs, including rooftop units. (This covers most typical commercial heating and cooling equipment that is not a chiller)

- The overall response rate for residential measures was 10%, ranging from 8% for insulation to 18% for HPWHs. Initial verification work indicates no significant bias in responses based on age of response (old cohorts of installed measures vs. newer) or type of equipment.
- To make the estimates as cost-effective as possible, the following procedures were used. Homes with multiple measures installed were asked about each of measures installed in the same year. If a single participant installed measures in multiple years, they were asked about the equipment installed in the earliest year to maximize the response’s value on the analysis (longest period to identify retention or failure).
- Data were collected that supported the basic definition of “operable”, but additional information was also collected that would allow nuanced analysis as well.
- Two types of verification work was conducted to assess potential respondent error in self-reported information: photographs of installed equipment and follow-up verification interviews. For both verification activities, we found very little error in the self-reported information, indicating that the information provided by respondents is accurate. The Evaluation Team primarily chose these verification approaches because they were cost-effective and replicable. However, the study’s timing also coincided with the COVID-19 pandemic which limited the ability for in-person site visits.
- The study’s methodology led to cost savings by analyzing multiple years of installations (2011-2019) instead of multiple studies of single years of installations, and by using surveys rather than on-site fielding. When conducted in this way, the cost of a single residential EUL measure is well under \$50K, and the incremental cost decreases with additional measures. This benefits from the fact that multiple measures are often installed in a single household as part of program efforts. Printing / mailing costs and the survey completion incentives that are commonly needed these days to achieve reasonable response rates accounted for more than half the budget. Studies that used single year

¹⁰ Many thanks to Jake Millette of Michaels Engineering, for information on this study.

cohorts, repeated multiple times (twice, with a lag) would cost approximately twice this amount.

The residential EULs are complete (Michaels 2022) and are shown in Figure 7. The highlighted numbers were recommended for the State's latest PSD update. Gas furnaces and insulation had large confidence intervals, so updates to the PSD were not recommended. RULs were also provided by the study. A commercial study is underway.

Figure 7: Estimated EULs for CT Residential Measures

Measure Life (yrs)	New Estimate	Old PSD
Air Source Heat Pump	20	18
Ductless Heat Pump	17	18
Heat Pump Water Heater	15	13
Central Air Conditioner	25	18
Gas Furnace	28	20
Insulation	38	25

Source: Michaels 2022

The results from the residential study shows that this approach can provide statistically-based updated results for reasonable budgets, and provide remaining useful lifetime (RUL) curves for use in early retirement programs. The study shows updated values were merited for some measures, but the EULs are also not dissimilar to current adopted values.

Questions this method helps investigate. Traditional EUL studies are expansive and examine result from individual cohorts. Thee reduced cost and testing changes for the recommended approach opens the door to investigating other researchable questions. Several examples follow, although researchers may identify many others.

- Lifetimes for standard efficiency equipment could be conducted using households (or businesses) that did not participate in the program. Those lifetimes can be compared with the results for EE measures using participants to determine if there are differences. It is generally assumed that the lifetimes are similar, but a study of this type could be used to make sure the market doesn't know something EE planners don't know, and haven't planned for. If they are not, that may affect baselines, incentives, program outreach materials, etc.
- Lifetimes for early cohorts can be compared to those for later cohorts to examine whether technologies (and lifetimes) have changed over the period efficiency measures have been included in programs. This can be used to explore additional effects of technology changes, and can identify whether maintaining constant lifetimes over decades has led to misestimates of savings. The Connecticut study plans to examine this question. It may also provide guidance on minimums for how often EULs may need to be updated, or the conditions that should trigger these studies.

- A review of results from a number of early EUL studies using this approach may help indicate the importance of replacing (or not) existing measure lifetimes at other utilities.
- Separate EUL studies can be conducted for some measures for different climate zones, to test if these conditions may affect lifetimes, and measure screening may merit more refined lifetimes.

Other researchable questions might also be of interest, but have not been explored because of concern about cost and complexity of EUL studies.

Conclusions and Recommendations

Given that Benefit Cost calculations include EULs as a key calculation component, on an ongoing basis, a periodic review of these values should be an expected part of evaluation plans, routinely updated, like savings, NTG or other inputs. In the immediate term, re-estimating should be instituted for those values currently in place that are from weak sources, (e.g., not statistically-derived, borrowed, consensus, etc.) or are elderly. This sampling / estimation approach provides an affordable, statistical approach to accomplishing these goals for many existing program measures.

This paper proposes a non-standard¹¹ but straightforward method for conducting statistical EE measure lifetime studies. It demonstrates successful use of multiple (more than a decade of) cohorts of program-installed measures to estimate EULs in an efficient and cost-effective way, coupled with survey-based approaches – for certain types of measures. The tests showed strong results from half a dozen studies across the US.

Although this approach may not be suitable to all conditions or measures, the advantages of this approach are several.

- **Fast turnaround:** The approach can be conducted immediately using current and past participants and can provide statistically-derived, updated replacement EULs for PSDs / TRMs without waiting multiple years for failures since they've already occurred.
- **Cost savings:** The delay and repetition of studies, and the traditional reliance on onsite visits are key reasons traditional EUL studies can be expensive. This new approach provides a longer timeframes for potential failures, eliminates the need for repeated surveys with time gaps, and halves the costs of a traditional study that would have occurred twice. The approach can be applied to multiple types of residential and commercial measures, so budget can be saved for more complicated measures that need on-site or more detailed EUL studies.
- **Removes barriers to more frequent EUL updates by utilities:** More importantly, the survey approach provides current, updated, and statistically-derived estimates of EULs which is a key benefit-cost and screening parameter that is often outdated or poorly sourced.
- **Provides RULs:** Updated RUL curves that are needed for utility early retirement programs, not just EULs for “replace on failure” programs are provided as well. It

¹¹ It should not be implied that no previous utility EUL studies used cohorts from multiple years. However, relying on this approach as the primary method for estimating EULs is a change in practice.

provides updated numbers which utilities can then adapt and adjust with market information to develop the market-adjusted- EULs that are needed for properly integrating baseline into EULs.

Finally, as illustrate by the thermostat examples, this approach works with populations that are not participating in utility programs, and it is reasonable cost. This opens the door to investigating other researchable questions. For instance, lifetimes for standard efficiency equipment could be conducted using households (or businesses) that did not participate in the program. Those lifetimes can be compared with the results for EE measures using participants. It is inherently assumed that the lifetimes are similar. If they are not, that may affect incentives needed, program outreach materials, etc. Other researchable questions might also be of interest, but have not been explored because of concern about cost and complexity of EUL studies.

References

- Franco, F., Y. Bennani, J. Ke, E. Cubero, and A. Lekov. 2018. *Estimating Residential Appliance Lifetime for Energy Efficient Policy Analysis*. Energy Analysis and Environmental Impacts Division, Lawrence Berkeley National Laboratory
- Millette, J. 2022. *X2001A: Connecticut Measure Life/EUL Update Study, for Connecticut Energy Efficiency Board*. Michaels Energinnering, June 3, 2022.
- Navigant. 2018. *ComEd Effective Useful Life Research Report*.
- Prinja, S., N. Gupta, and R. Verma. 2010. *Censoring in clinical trials: review of survival analysis techniques*. Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine, 35(2), 217–221.
<https://doi.org/10.4103/0970-0218.66859>
- Skumatz, L. 1997. *Recognizing All Program Benefits: Estimating the Non-Energy Benefits of PG&E's Venture Partner's Pilot Program (VPP)*. Proceedings of the 1997 Energy Evaluation Conference, Chicago, Illinois, 1997.
- . 2007. *Updated Measure Lifetime Estimates: EULs Based on 10 Years of Studies*. Proceedings of the 2007 IEPEC Evaluation Conference, Chicago, IL.
- . 2009. *How Long Do Program Savings Last? Issues in Measure Lifetimes and Retention of Savings and Behaviors*. EEDAL, Berlin, June 2009
- . 2011. *The Tri-Fecta: Comprehensive Review and Best Practices in Impact, Attribution, and Retention of Behavioral Programs and Beyond*. Proceedings of the EEDAL Conference, 2011
- . 2011a. *Remaining Useful Lifetimes and Persistence – Literature and Methods*. Proceedings of the IEPEC Conference, 2011
- . 2012. *What Makes a Good EUL? New Protocols for Estimating Useful Lifetimes and An Analysis of Existing Estimates*. IEPEC International Conference, June 2012, Rome Italy.

- . 2012a, Guidelines for the Development of Electric Energy Efficiency Measure Lifetimes”, Prepared for the Regional Technical Forum, Portland, OR., March.
 - . 2014. *Estimated Annual Outflow of Mercury-Containing Thermostats in the State of Rhode Island*. Available from:
https://www.nrdc.org/sites/default/files/hea_14021901a.pdf
 - . 2014a. *Mercury-Containing Thermostats – Population, Lifetimes, Decay, and Policy Implications: A Study in Several States*. Western Economics Association Conference, Denver, CO, July 2014.
 - . 2016. *Persistence of Behavioral Programs: New Information and Implications for Program Optimization*. The Electricity Journal, June 2016.
 - . 2016a. *Estimated Annual Outflow of Mercury-Containing Thermostats in the State of New York*. Prepared for Clean Water Fund and New York Public Interest Research Group Fund (NYPIRG Fund), by Skumatz Economic Research Associates, Inc., Superior, CO, August.
 - . 2018. *New Evidence of Behavioral Measure Lifetimes—Tracking Retention through Zero*. (with co-authors), Proceedings of the IEPPEC Conference, Vienna, Austria, 2018.
- Skumatz, L. and R. Bordner. 1997. *Evaluation of PG&E’s Venture Partner’s Pilot Program (VPP)*. SERA, Prepared for Pacific Gas and Electric Company, San Francisco.
- Skumatz, L., K. Lorberau, and C. Breckinridge. 1991. *Bonneville Power Administration Measure Life Study I*. Prepared for BPA, Portland, OR.
- Skumatz, L., D. Whitson, D. Thomas, K. Geraghty, B. Dunford, K. Lorberau, and C. Breckinridge. 1994. *Bonneville Power Administration Measure Life Study II*. Prepared for BPA, Portland, OR.
- Skumatz, L., R. Woods, and S. Dimetrosky. 2004. *Review of Retention and Persistence Studies for the California Public Utilities Commission, (CPUC)*. Final Report, October 20, 2004, California Public Utilities Commission, San Francisco, CA.
- U.S. DOE. 2016. *Technical Support Document: Energy Efficiency Program for Consumer Products and Commercial and Industrial Equipment: Residential Furnaces*. And *Technical Support Document: Energy Efficiency Program for Consumer Products and Commercial and Industrial Equipment: Commercial Warm Air Furnaces*.
<https://www.regulations.gov/document?D=EERE-2014-BT-STD-0031-0217>
- Welch, C. and B. Rogers. 2010. *Estimating the Remaining Useful Life of Residential Appliances*. Proceedings of the ACEEE Conference, Asilomar, CA
- Woods, R. and L. Skumatz. 2007. *Incorporating Real-World Data Into Measure Lifetime Estimates: How Long Does Energy Efficient Equipment Really Last On-Site?*. Proceedings for the European Council for an Energy Efficient Economy, France, June 2007.