REMAINING USEFUL LIFETIMES AND PERSISTENCE–LITERATURE AND METHODS

Lisa A. Skumatz, Ph.D., Skumatz Economic Research Associates, Inc. (SERA), Superior, CO

ABSTRACT

The author examined more than 100 studies on measure lifetimes, and practices in multiple states to identify best practices in estimation methods for measure lifetimes, and gaps in research and values for estimated useful lifetimes (EULs) for residential, commercial, and industrial measures used around the country. We found convergence in lifetimes for many measures, but variations in values for some measures, data gaps for a number of programs and measures, and differences in policies applied to many residential measures.

We examined the issue of 'remaining useful lifetimes" (RULs). Programs that encourage or result in early removal of equipment may warrant assessing savings with two baselines. We review the few studies on this topic, and assess the approaches being employed and considered. This research has implications for both traditional "widget" programs, and for behavioral programs. Information on available primary research, policies, and application around the United States is summarized. Available research on technical degradation was also reviewed.

In addition to the "status quo" review, the study examines the impacts of two complicating factors: 1) transition to more non-measure-based programs (education, advertising), which make it hard to "count" and measure retention, and 2) increased "chatter" in the marketplace, in which consumers may be influenced by any number of local programs as well as outside influences. We examine progress and gaps in research and policies related to these two complicating factors, examine gaps in existing research, and promising techniques for non-measure-based programs, and summarize recommended next steps.

INTRODUCTION

Retention studies, also known as persistence or measure life studies, are a critical and highly useful component of energy efficiency research. There have been established protocols for effective useful life (EUL) studies and, despite some variations in data collection and methodologies employed, the fundamental purpose of measure retention studies is to estimate the amount of time that a measure will be in place, presumably delivering energy efficiency benefits. The measure life provides a limit for the number of years that a program's annual savings will last. Early programs used measure lifetime figures related to laboratory lifetimes. Studies in the early 1990s demonstrated that a combination of factors affect the years over which a measure delivers savings, and it is not well-estimated using laboratory lifetimes. In the commercial sector, business turnover has a strong effect,[1] as well as equipment changeout from periodic remodeling / redecorating, the ability of maintenance staff to keep advanced equipment functioning and other influences. Parallel effects affect *in-situ* retention of equipment in households.

[1] This is a factor that varies dramatically across the non-residential sector. Restaurants may turn over in 6 months or lighting styles; décor changes every couple of years; schools tend to stay schools and keep measures in place until well past their optimal functioning lifetime and dramatically past their optimal economic lifetime!

Current Practices and Uses

Measure lifetimes are one of (at least) three critical components in estimating the savings attributable to energy efficiency programs, along with gross savings, and net savings adjustment factors. The review of the work in persistence finds fairly uniform procedures in place across North America. The statistical and data collection methods and established best practices have largely been reflected in evaluation protocols.[2] The overall approach taken by most measure retention studies in the energy efficiency (EE) field is to estimate the median EUL of the measure in question. The EUL is usually defined as the median number of years[3] that a measure is likely to remain in-place and operable.[4] This amount of time is often calculated by estimating the amount of time until half of the units are no longer in-place and operable. The key data needed to derive these estimates include installation location, measure(s) installed, date installed, and the date that the measure became inoperable or was removed. From these data, a basic measure life study can be conducted.[5]

While this task may seem straightforward at first glance, there are often considerable complications involved with obtaining EUL estimates. Measures often last for a long time, making it impractical to simply wait until half of the units fail in order to determine the median survival time. Measure lives are also frequently interrupted prematurely by the owners or employees of the residence or business in which the measure was installed. Obtaining unbiased EUL estimates, therefore, can require statistical analysis to (1) control for exogenous factors that might affect measure lifetime and (2) predict measure lifetimes based on empirical data. Furthermore, application of these studies requires information on the projected results fairly early into the lifetime of much of the equipment installed as part of various programs, when a set of measures is young and only a relatively small portion of the installations may have failed. For example, protocols that were in place for many years in California required periodic verification of EULs when measures had been installed for fewer than five years. This poses a particular challenge, as EUL estimates are based on failures, and few measures projected to last 20 years or more would be expected to fail so soon. Finally, changes may be needed to the traditional time intervals to address newer behavioral programs, for which it is currently quite unclear what lifetimes may apply. Developing unbiased estimates of EULs under circumstances of limited data early in measure lifetimes is particularly challenging.

The authors evaluated more than 120 reports and studies addressing EUL methods, research, and primary studies covering a diverse collection of energy efficiency measures.[6] We compared the different data collection, and analysis techniques on the basis of their effectiveness in obtaining meaningful results, their ability to produce reasonable EUL estimates, the degree to which they produced statistical models that fit the data, and the defensibility of the conclusions drawn from them. The review of a large number of studies provided an opportunity to view the range of practices used for small and large, and simple and complex measures over a period of nearly ten years. We found a few problems that arose repeatedly: sampling-based issues (and problems associated with program databases that lacked important data needed for evaluation purposes); data collection issues (the high cost of collecting the needed retention data, particularly for long-lived measures when "failures" are needed to support estimation); and analysis issues, particularly sample size issues, and concerns that studies do not test

---

[2] Although the literature tends to test only one model or distribution (a potential weakness considering the variations in underlying technology and mechanics), and rarely presents comparisons or discussion of *ex ante* and *ex post* values or present comparisons to other studies of the same measures. The methods and EULs have evolved from the initial kernals in work for Bonneville Power Administration, presented in Skumatz and Hickman 1992.

[3] Or other time interval, as appropriate.

[4] "In-place and operable" is at least the most common definition of measure survival. Depending on the specific measure under inquiry, alternative formulations of the definition may be more appropriate.

[5] Enhanced data can improve the estimates; these issues are discussed later in the paper.

[6] Building on the work conducted by Skumatz in work from 2002-2005; see Skumatz 2009 for citations.

different statistical distributions and do not compare results to those from other studies. Best practices suggestions were provided in the study.[7]

Remaining Useful Lifetimes / RULs

Some programs are designed to intervene at the time measures are being replaced, and the years and savings values to be assigned for the lifetime of the savings are fairly unambiguous. However, some programs may be geared toward replacing existing (lower efficiency) equipment with energy-efficient equipment before the old equipment ceases to function or before it would otherwise be replaced – early-replacement programs. These programs can achieve additional savings by accelerating the turnover of long-lived technology stocks.

We could compute two alternative savings increments. The savings could be calculated as the difference between energy use for the old measure replaced and the new EE measure (we'll call this "enhanced delta"), or as the difference between the new standard measures available in the market compared to the extra energy efficiency of the measure induced by the program (we'll call this "standard delta"). Several questions arise regarding early replacement programs: 1) Should programs be able to count higher savings (i.e. enhanced delta) for the early replacement period, in addition to standard delta savings thereafter?[8] 2) If so, how is that transition point estimated?

We conducted interviews with utilities and professionals across the nation on practices regarding RULs. Comments ranged from "we don't use these at all" to "they're used constantly", depending on the region / utility called. Many of the interviewees agreed that RUL is a concept that has potential applicability for early equipment removals of equipment. However, every respondent noted the difficulty of measuring the period in time for the early replacement – and noted it was theoretically do-able, but would require additional information gathering in the program research. None believed it was appropriate to *only* assign to the program the enhanced delta for the period in which the decision was moved forward, a possibility that had come up in some early discussions.[9]

Early Research on RULs

There were only a few primary studies of RULs. One early attempt gathered the age of the air conditioning equipment that was being pulled out (using model numbers) and used the lifetimes associated with that equipment to calculate a mortality table (which properly takes into account the fact that if you've lived to age 90 you stand a better chance of living to 100). These data were used to document the savings stream. A utility in the Northeast is undertaking a survey approach to examine this issue for a few programs.[10] Several other studies made attempts to develope approaches to associate RULs with specific programs and estimated impacts on program-associated energy savings. Preliminary, *ad hoc* rules of thumb assumptions (e.g., one study assumed that 1/3 of lifetimes remain)

---

[7] The core of this study was conducted by Skumatz Economic Research Associates under contract for California Institute for Energy and Environment (CIEE) and the California Public Utilities Commission (CPUC), and contract oversight was by Dr. Edward Vine. Additional research on remaining useful lifetimes (beyond the original study) was conducted for this article.

[8] The total lifetime for the savings computations would still be the EUL for the new equipment. Presumably the RUL period would be assigned based on the difference between the age of the existing equipment removed and the EUL for that equipment, and the number of years over which the "standard delta" savings would be accumulated would be shorter (the EUL of the new equipment less the years attributed to "early removal").

[9] The issue of RULs may also apply to behavioral programs, so if the issue is solved for measure-based programs, the same policy may apply. Consider the following hypothetical. If codes and standards were going to be implemented in a future year that would mandate some behavior (e.g., you may no longer leave outdoor lights on all night – they must be on a timer), and if a program moved that behavior-related impact forward, it is possible that a parallel situation with the measure-based program arises.

[10] Three other interviewees told about related issues. One belatedly found enhanced deltas were recorded for all participants for a program that needed later adjustment; another stated they had issues with first year savings being used throughout the life of the measures (they believed decay functions should be used); and another found out that the auditors were assigning all remaining years of early replacement to the first year – leading to a much over-estimated value for savings. These remain cautionary tales in looking at savings, early replacement, and savings computations and recording.

were also used and/or proposed, but were subject to criticism (e.g., the study did not consider that schools have much longer replacement intervals than other non-residential uses).

Estimating the Years to Assign to Early Removal

The two-part savings calculation is theoretically appropriate. However, the issue is not only one of when the equipment would have died, but also involves a subjective decision by the business or homeowner regarding the estimated time until the owner "would have replaced" the measure. This estimated timing may be even less reliable than the self-report information used in many free ridership and NTG computations. Given that early replacement programs can have an impact on getting inefficient equipment out of the marketplace, reliable techniques are needed to develop estimates of RUL transition points. Direct program- or measure-based research could be used to generate estimates – especially as the research would only require gathering a few more pieces of information at the point of program implementation / installation / replacement.[11] [12] [13] Another approach would use mortality computations / distributions on verified equipment age (or survey information where model information is not available) to develop estimates of early removal periods. This is the approach explored in a recent study (Welch and Rogers, 2011).

The Welch and Rogers paper described a methodology for estimating the RUL of a subset of equipment (residential appliances) using mortality data available from the Survey of Household Energy Use (SHEU) from the Natural Resources Canada (2003) Energy Efficiency Office. The survey collected data on the age of various appliances (including dishwashers, refrigerators, freezers, clothes washers, and clothes dryers) at their point of retirement. The paper hypothesized that the Weibull distribution (commonly used for lifetime analyses) would provide a useful shape to estimate RULs for given EULs and equipment ages. Using the SHEU data, they fitted Weibull shapes for the five equipment types and found the distribution was a good fit to the appliance retirement data, and the outcomes showed very similar shapes when normalized by mean lifetimes. They postulate that the resulting Weibull shape factor (2.34) might be able to be applied to other equipment, particularly residential appliances (and in their study, they use the assumption for the air conditioning equipment of interest for the specific program they were examining). One output of their research is a normalized curve mapping RUL vs. years in service (using the same Weibull shape factor), which can be used to project a remaining life for a piece of equipment (either younger or older than its EUL). They suggest that early retirement programs can use this curve (or this type of curve), plus data on the age of the equipment being replaced (gathered from customer surveys or model data) to estimate the RUL, and from there, the lifetime savings and cost-effectiveness of the early retirement program can be computed. They note two caveats to applying this approach:

- The customer survey data would be less precise than data collected by implementation contractors at the time of equipment replacement;
- Applying this method for *ex ante* planning purposes would be less reliable than for ex post savings calculations because an assumption regarding average age of equipment to be replaced would be needed in advance.

Welch and Rogers also explored an approach they call "system dynamics modeling of technology stocks". This modeling approach used historical data on appliance shipments and total appliance stock

[11] Specifically, equipment model and age, survey questions on when they would have replaced, and age of equipment that can't be traced through models.
[12] One utility interviewed uses the entire savings – old measure to new measure (enhanced delta)– throughout the lifetime of the measures, and they assume that a (majority) share of those installing the new measures (e.g., CFLs) will replace with CFLs again, so their savings go out beyond the initial lifetime. The utility notes that if something like this is not assumed, you should probably be readjusting your demand forecasts.
[13] The case of the very effective "cash for clunkers" early automobile replacement programs may be worth examining. Whether the early replacement period was assigned higher emissions savings than the later periods may suggest a precedent for the policy issue in energy.

to estimate the mean life and mortality shape factor of air conditioners. This approach is advantageous in that it can be used in the absence of EUL and mortality data (and they note it can be used to cross-check EUL estimates from standard databases). The system dynamics approach modeled "…the stock of the technology as a function of the inflows (i.e. new purchases) and outflows (i.e. retirements) of that technology using numerical integration techniques." The fairly complicated estimation work required aging chains of stocks decaying into the downstream stocks, distribution assumptions (they used an Erlang, which is similar in shape to a Weibull), and extensive simulation work. They estimated the aging chain using data from the Air Conditioning, Heating, and Refrigeration Institute (AHRI) and Energy Information Administration (EIA) and historic growth rates. Their fitted results found an EUL of 15.46 years for air conditioners, which compares favorably to the CPUC's DEER estimate of 15 years. As a final step, the paper compared the result of the mortality curve shapes estimated using the two methods (Weibull with shape factor 2.34, and Erlang with shape factor 5) and showed how closely the two curves aligned on a graph. They postulate the shapes may apply to a wide range of residential appliances, and can be applied given estimates of EULs found in the technology databases.

In conclusion, the Welch and Rogers study presents two very promising methods that can be applied in cases in which different types of data are available. It also suggests a default curve that significantly improves on "rules of thumb" for RUL assumptions. That said, additional research is needed on early removal dates for other equipment types, and whether querying current owners on their expected turnover date (without the program) can provide reliable values.

A Note on Estimating Savings after the RUL Transition Point – "Standard Delta"

The other part of the equation is the savings to be assumed for the period *after* the equipment would otherwise have been removed. The new baseline value may be:

- "standard" efficiency at the point of early replacement
- "standard" efficiency at the future date when removal would have occurred
- codes and standards level now or at that future date
- standard practice now or at the future date, or
- some other baseline

The first option is used now, for programs with or without the issue of early removal. For short-lived measures, there is no important difference. However, if measures are longer-lived, this RUL issue is more important to consider. Identifying the standard efficiency at that future date is far from straightforward. The recommendation would be "standard practice", but practical methods to estimate a useful proxy for that metric would be needed to estimate the most accurate savings. The most practical alternative in the meantime (current "standard") would deliver an overestimate of savings. Using a more conservative alternative would help minimize criticism and skepticism when EE savings are compared to generation alternatives. Research into other alternatives (adoption curves, incorporation of known standards upgrades, etc.) would be beneficial to see if any are applicable to this question.

RULs and Educational / Behavioral Programs

This concept of RUL carries over directly to education and behavioral programs. Bringing forward in time a behavior that would tend to be generally expected (or mandated) in the future has near-term value. With more and more "green" education coming through a variety of mass media channels, greener behaviors are likely to become (more) standard, including energy saving behaviors.

The measurement issues are even more complicated, but just as necessary to examine if these programs are to become an increasing share of portfolios. Current behaviors will not remain the baseline forever. This will tend to decrease savings associated with programs; however, the net result will depend on whether program-induced behavioral changes are longer than baseline new behavior adoption.

Needless to say, no work has yet been conducted on this topic. Again, some kind of adoption curves may serve as a proxy; but research is needed. Many questions arise, such as: what would be assumed for timing? What would be assumed for the ultimate efficiency of the behaviors? How many different behaviors? Policy-wise, early adoption of new behaviors is an appropriate concept. Periodic billing analyses may have potential as an option; measurement is a significant issue needing further study.

Technical Degradation / TDFs

We also explored the topic of Technical Degradation Factors (TDF), which are addressed in the California EM&V protocols. Another factor affecting how much savings are being delivered is whether program-installed measures perform at the new efficiencies consistently over time, or whether their efficiency performance degrades over time (or potentially in given installations). Unexpected decay in performance could be an important issue, particularly for measures for which savings are assumed to accrue for upwards of 15 years. Unfortunately, in reviewing more than 100 EUL and TDF studies, we found very few TDF papers within the last decade that were based on primary data. A paper by Jump et. al. (2008) applied 1998 laboratory measurements of lamp median life (from Rennselaer Polytechnic Institute Lighting Lab) to residential logger data collected by KEMA in 2003-2004, and derived an average CFL normalized lamp life. Using the rated life from the lamp packaging allows computation of an observed life.[14] An engineering study (Blasnik 1997) suggests that only a few measures may be affected in a positive or negative way relative to the decay in performance of standard measures.[15] Primary research could well be justified, particularly for measures with technical or engineering changes that may affect the degradation of specific equipment types relative to the degradation that would be expected with older technology – or measures accounting for large shares of portfolio savings.

Of course, it is worth noting that performance degradation is the combination of two effects – technical degradation, as well as a behavioral / operational component, including the quality of use and quality of upkeep of the equipment [see California EM&V Protocols on EUL]. Studies that look at degradation *in situ* need to account for the influence of both these factors. Engineering studies that only examine potential technical / mechanical degradation causes may miss equally important behavioral changes. Therefore, setting priorities for future TDF studies will need to examine both technical and behavioral elements.

We do not separately address TDF related to behavioral programs, as we consider the concept in tandem with EUL. The behavior (as a measure or a performance) decays and ceases. Studies of both elements are needed, but the topic was addressed under EUL.[16]

Retention Results for Measure-Based Programs

The authors conducted a review of half a dozen recent studies of EUL summaries, as well as examining more than 100 EUL studies conducted on a host of programs in California, and we examined

[14] This is a result distinct from an EUL because EULs include early burnout or removal, which is not captured by this method.
[15] See Skumatz (2009) for more detailed information on this topic.
[16] Technically, each behavior for each person educated by the program has a presence, in place and operating, parallel to an EUL. There is also a TDF associated with the behavior – for example, when that person does the behavior only a share of the time or begins to forget the learned behavior. However, for ease, we treat it all under EUL. Given partial adoption, both issues will need to be considered as part of any credible EUL or TDF study.

measure lifetime assignments for hundreds of measures. We found that deemed measure lifetimes or EUL values used in different areas of the country often have similar values, as illustrated in Table 1.

Table 1: Range of EUL Values Used in the US

Residential Measures	Commercial Measures
• Lighting – CFL Bulbs: 6-8 years, with some recent work starting to incorporate variations based on assumptions about hours per day that the bulb operates • Hardwired fixtures – 15-20 years for interior and exterior fixtures • Lamps (table or touchier) – 5-10 years for most studies[17], depending on type • Occupancy sensors – 10-15 years • HVAC replacement – 15-25 years • HVAC and water heating in Energy Star – 15-25 years • Room A/C – 11-15 years • Programmable thermostat – 10-12 years • Whole house fans – 25 years • Attic ventilation fans with thermostat controls – 19 years • Duct sealing and air sealing – each 15-20 years • Insulation – 20-25 years • Duct insulation – 20 years • Windows – 20-35 years • Pipe wrap – 10-20 years • Tank temperature turn down – 4-7 years (needs study) • Weatherization (combination measures) – 20-25 years (13 years in CA)	• Lighting – CFL Bulbs – 3.4-6 years, with some recent work starting to incorporate variations based on assumption on hours per day bulb that operates in business locations • Fluorescent fixture – 11-16 years • Hardwired CFL – 10-15 years • HID (interior and exterior) 13-15 years • Occupancy sensors – 8-15 years • Daylighting dimming – 9-10 years (16 years in CA) • Packaged AC/Heat Pump – 12-15 years • Chillers 19-23 years • Economizers – 7-15 years • Programmable thermostat – 5-10 years • Energy Management Systems (EMS) – 10-15 years • Motors – 13-20 years.

Our review of EULs identified several issues:[18]

- Process equipment (e.g. plastic extrusion equipment, and the array of other equipment associated with specific industries) lacks EUL studies in many cases, largely because each specific measure has a small sample size. Some utilities or agencies "assign" a 10 year lifetime, assuming that progress in the industry leads to reconfiguring of equipment on that kind of schedule. This issue may bear additional research, especially since lifetimes are likely dependent on the pace of innovations in the particular industry.[19]
- Some equipment may require evaluations of operating assumptions: for example, CFLs and other lighting equipment in commercial establishments, variable speed drives (VSD)s when applied to agricultural milking that endure harsh outside conditions, etc.[20] Lighting logger studies are particularly important given that huge shares of utility programs and savings are based on lighting measures.
- Reliable EUL estimates are missing in many key end uses: e.g., cooking, air compressor equipment, chillers, adjustable speed drives (ASDs)/VSDs, refrigeration equipment and freezers in some sectors. In addition, there is only limited information available on the increasingly important – and targeted – plug loads sector (e.g., copiers and office equipment) and unless very short lifetimes are assigned, these measures may need to have EUL studies conducted to provide justifiable savings estimates.

---

[17] But longer for California (9-16 years). All California numbers from the Database on Energy Efficiency Resources (DEER).
[18] Based on our review of national and California EUL studies (Skumatz and Gardner 2005) and others.
[19] Think of the difference between high-tech computer chip manufacture vs. traditional steel or paper manufacture, as hypothetical extremes.
[20] In addition, some lifetimes may specifically need to be adjusted based on the influence of behavioral programs. For instance, if a program suggests relying on daylighting and leaving lights off until really needed, the operating hours for CFLs may need to be adjusted in accordance with the success of such a hypothetical program.

- There are few retention studies on building shell measures. Building shell measures are not generally assumed to be subject to widespread failure / removal, but this assumption should be verified, potentially in different parts of the country.[21]
- There has been a trend in the field to move toward simplified EUL tables, but this is a problem. Even some of the earliest research (Skumatz and Hickman 1992) found significant variations in business turnover by business type, and this turnover has a direct effect on retention of measures (particularly lighting). These variations are important factors in program savings computations and program design.[22]

The need for additional EUL research in specific measures should be weighted by the expected future savings to be derived from the measures. For those that are rare and offer low savings, the priority is low. Similarly, for measures unaffected by operating hours and climate (e.g., exit signs), priority for investment of additional research budget should probably also be low. Measures subject to climate and operating hours assumptions may be higher priority (e.g., HVAC).[23] Again, waiting for natural failures impairs the timeliness of EUL studies, making the results less applicable to current and next generation measures that are being installed.

Retention for Non-Widget-Based Programs - Education / Training / Behavioral

Probably the single biggest gap in lifetime studies is the virtual non-existence of studies examining the retention of education, training, and behavior-focused "measures". On the behavioral side, programs tend not to get energy savings credit, so EULs / retention / persistence has not been much studied, even though the programs and their outcomes presumably do have lifetimes. Reviewing more than 100 studies in education / training (Skumatz and Green 2000; Freeman and Skumatz 2009), we found only a couple that even mentioned the topic of the retention of savings. Almost all studies examined savings for the first year of the program, which makes it hard for potentially important and dynamic education programs to receive high benefit/cost ratios, reducing likelihood of funding.

Two studies have addressed retention of educational messages and installation of low-cost energy-efficiency measures delivered through energy education programs. The Energy Smart Program conducted in Oregon with low-income households found strong to mild retention (about 40% after 3 years) of behavioral changes. Especially successful have been those energy education efforts that provide quality education over a longer period of time. Three energy education programs delivered in schools (the Kentucky NEED Program, the Iowa LivingWise Program and the Washington Energy Education in Schools Program) show the importance of quality education and reinforcement of behavioral change messages over time. Of these three programs, the greatest behavioral changes were observed in Washington where teachers conducted at least three different classroom sessions and one assembly with kids over the course of an entire school year. A study by Peters and McCrae (2001) followed participants for one to four years after participation in a Builder Owner Certification in the Northwest, which is considerably longer than the usual six-month follow-up. They did not specifically estimate persistence, but reported fairly consistent findings with each successive follow-up. These efforts, along with an early study by Harrigan and Gregory (1994), which found 85%-90% of the savings from the education portion of a weatherization program was retained after three years, seem to

[21] Some interviewees expressed concern that within the TRC environment, new construction, design and shell measures are generally assigned lifetimes no longer than about 20 years, even though many of these measures last perhaps 100 years or longer. The 20-year horizon was considered a problematic artifact and that the future is too discounted to reflect the actual climate imperative.
[22] Without consideration of variations by business type, programs could keep replacing measures continually in the same business types, and fixed or deemed EULs that don't vary by business type would miss this effect and keep counting streams of savings that never materialize over time.
[23] The question arises whether lifetimes for HVAC equipment <u>should</u> be similar between two very different areas of the country; say, the Northwest vs. Florida. Behavioral considerations should be expected to matter.

be the only studies that have conducted primary data analysis of the topic. Even well-funded multi-year statewide outreach programs have not examined the persistence of behavior change.
For this reason, many utilities assign retention values no higher than three years in most cases.[24] Also of concern is that the savings and potentially the persistence may be highly variable depending on the specific program, specific media, quality of the campaign, and many other factors. It may be that every program will require its own persistence study for at least a while, until there is time to develop reliable best practices and "template" programs. The behavioral persistence topic is gaining interest,[25] and should be among the highest priorities for new research.[26]

Attributing behavioral changes or energy savings effects to particular campaigns or programs is becoming more complex as more agencies work toward similar energy efficiency behavior changes. Generally, this factor has minimal effect on the measurement of energy savings lifetimes; however, it does tend to affect in a significant way the estimate of (the share of) behavior-induced energy savings that can be clearly attributed to a specific program or intervention.[27] Research is exploring several options for behavioral programs, as noted below.

One avenue is identifying cases for which it is suitable to apply the measure-based "Best Practices" methods to the development of measure life estimates for behavioral programs. This may work, in general, with revisions to questions to ask about the persistence / presence of behaviors.[28] There are nuances related to "partial retention" (e.g., some household members), but, conceptually, this approach can apply to some programs.[29] However, there may be problems in using this approach. For example, there may be more problems with bias[30] since behaviors may not be as observable as measures. Also, the costs for conducting this type of research may be even higher than traditional EUL work, because behavioral programs may not be easily associated with specific businesses or homes. Large-scale survey approaches may be one of the few data collection options available, and these are costly.[31] Finally, the traditional EUL approach is most suited to longer-lived measures (assuming behavioral measures are less long-lived than measure-based programs).[32] Given that the lifetimes may be short, data collection might also have to be more frequent. In conclusion, simplified approaches – perhaps as straightforward as the kind of retention study conducted by Harrigan and Gregory (1994) – may be more appropriate for lower-budget programs. And random assignment, follow-up of test and control groups, and similar methods to estimate retained shares of savings and behaviors are the types of analytical approaches that are needed. Large scale surveys of households or business populations may be needed until reasonable estimates can be derived and some kind of convergence in results by type of program emerges.

[24] Although it is unclear if a median EUL of 3 years can be justified given that there is minimal research for this estimate.
[25] It is gaining mention in more and more policy documents, and, for example, the Canadian Association of Evaluators has established a working committee on the topic.
[26] The associated issue of technical degradation (TDF) is probably best represented in behavioral programs as "retention", and there are certainly no extant studies of this topic separately.
[27] This topic is the subject of Skumatz (2009).
[28] For behavioral or market-based outreach / education programs that influence a home or business to purchase a measure (e.g., Energy Star programs advocating purchase of CFLs or Energy Star refrigerators, etc.), the traditional approach is appropriate.
[29] However, from a data analysis point of view, it may provide more failure data, which can assist model-fitting!
[30] Bias may depend on who in the house / business is being interviewed, and how the information is obtained (e.g., surveys where the respondent may be trying to please the interviewer).
[31] Residential appliance saturation surveys could be expanded to include behaviors, or large-scale surveys incorporating interviews on behaviors from several programs could be conducted. Surveys of rolling segments of the population may also be appropriate.
[32] This seems sensible because people move from the home (and potentially the service territory) about every 5 years, taking their program-influenced behaviors with them. In contrast, most measure-based programs are permanent to the home (except refrigerators and CFLs) and remain after the occupant moves.

Upstream

The previous discussion largely considered "direct" behavioral / educational programs – those related to occupants of the home or business. However, there is also the issue of retention of "upstream" behavioral / educational programs. Degradation of upstream technical training programs offered to agents that do not actually operate the measures (e.g., vendors, manufacturers, commissioning agents, builders, and architects and engineers) is another matter. To the extent that these programs work to influence second and third round and future savings, then EUL and TDF is an important consideration and very hard to measure. The TDF may decay, but presumably, it may increase as the builder / agent is inspired to take on more and more (self-) education and measures as a result. This delves into the realm of spillover, but it can also be viewed under the subject of EUL / TDF. This is a topic that has not been studied and represents a particular challenge for developing credible methods.

Summary

Table 2 summarizes key patterns in EUL results. The research findings are summarized below.

Table 2: Variations in EULs by Program Type and Region

	EULs
General results	After early work in the Northwest, results have gravitated toward values fairly similar to those in California's protocols, with some variations elsewhere. The State of California required *ex post* statistical verification, leading to minor refinements. There are a number of measures for which there are missing or inadequate data; the most glaring example is the nearly complete omission of retention information or estimates for behavioral programs.
Variations by Program type	Almost all EUL results are by measure, not by program design or incentive provided. Therefore, although measures have EULs, there are no variations for measures installed from programs designed as rebate vs. codes / standards, etc. Any program delivering a measure receives basically the same retention value for that measure.
Variations for behavioral vs. measure-based programs	There is almost no information for retention of behavioral programs including education / training, commissioning training, and similar programs. Widget-based programs have fairly thorough EUL information, with omissions for some measures (cooking, some shell, others).

- **Best practice suggestions for effective useful life (EUL) studies:** Assure that implementation databases are structured to support evaluation research; use appropriate sampling approaches when bundled programs are implemented; use phone data collection only when measures are unique or memorable; use panel surveys if possible; more enhanced modeling that supports the incorporation of tests of multiple model specifications; and, most importantly, benchmark the results against the findings for earlier years of the program and for similar programs around the nation.
- **Results and gaps in EULs**: A review of results from measure-based EUL studies around North America showed that measure lifetimes are fairly consistent for many commercial, residential, and industrial programs. Relatively similar EUL values are being assigned by utilities across the country – perhaps with not enough recognition of the variation in operational hours by climate zone. The review also shows a lack of depth in studies in process equipment; some shell measures; and specific end-uses like cooking, refrigeration, and air compressors. The study also identifies measures for which there are insufficient studies to confirm or develop reliable estimates – particularly there are important gaps in the areas of: cooking, air compressors, ASD/VSD, refrigeration / freezers in some sectors, plug loads, building shells measures (which at least need verification), and a few others. These lack reliable study; EULs for some other measures have been estimated repeatedly. There has also been a trend toward simplified tables (e.g. one value across all business types), but this omits

important turnover differences between business types, and investment differences between customer groups (e.g. the dramatically different replacement schedules for lighting between schools vs. restaurants).

- **RUL issues**: Regarding the topic of remaining useful lifetimes (RULs), some utilities argue RULs are critical to certain programs; others don't feel the estimation complexity is a worthwhile expenditure. The jury is still out on the policies to be applied broadly, but if a program is designed as early replacement, a credible case could be made that its savings pattern is significantly altered from end-of-lifetime programs. A recent study provides promising research on two estimation methods with practical options for estimating RULs; additional research is needed to explore the potential of these (or other) methods for expansion to other measures, and to behavioral programs as well.
- **Technical degradation**: The issue of technical degradation was discussed, and there is a shortage of primary research on this topic. Certainly, engineering-type studies can help to identify research priorities to some extent, noting which technologies have undergone engineering, mechanical, or process changes that will more likely significantly change their performance relative to standard equipment. However, equipment with significant changes in behavioral (operational or upkeep) elements may also see changes in performance. Priority-setting for new research on this topic should take both factors into account (mechanical and behavioral), and resulting figures should be verified periodically.
- **Retention of behavioral changes results and needs**: Of particular note is the virtual absence of studies addressing retention or persistence of education / outreach / behavioral programs. This is an important gap, as behavioral and market-based programs have become a larger and larger share of utility / agency portfolios. Further research in best practices for the array of behavioral programs or "types" would be a useful addition to the literature, and agencies should consider requiring new behavioral programs to conduct retention assessments every year or two for a period reaching on the order of three or more years out. This may be the only way to gain enough information to develop credible estimates of the persistence of savings from behavioral programs and to allow more serious consideration of them as reliable resource substitutes. The issue of retention of behaviors and savings for "upstream" education and training programs is particularly troublesome, and, to the degree that these programs are part of portfolios, retention work is needed where there currently is none. Finally, EUL measurement approaches will need to be tested and applied to a variety of behavioral programs. Some may parallel traditional EUL estimation best practices, but the application of statistical approaches to some programs may be challenging. EULs for behavioral programs will have to consider issues related to how to treat partial retention,[33] examine alternate measurement methods considering the potential short lifetimes of some programs, examine issues of frequency of data collection, retention of "upstream" behaviors, large surveys and random assignment, among other items not examined much to date in association with measure-based programs. This research should be a priority for the near term.

Measure lifetimes are a key element in the computation of program savings. However, if measure lifetimes, technical degradation factors, and other factors are known for some programs and unknown up front for others, there will be a bias away from developing new (more uncertain) programs. Risk is an issue affecting investment and development.

Risk needs to be considered from two perspectives – providing up-front information on computational elements encourages program development. "True-up" is needed for credibility and reliability of savings estimates for EE relative to generation capacity. One suggestion may be that new

[33] A few members of a household keep the behavior but others don't- perhaps parallel to Technical Degradation associated with a measure.

programs are assigned a deemed lifetime by general "type" up front, and then after 1-2 years, a true-up is prepared that does not readjust program incentives retroactively, but does refine the estimate of future savings from a resource perspective.

Identifying the lifetimes or EULs of behavioral or information programs is complicated as more media messages on behaviors and education bleed across territories. This affects retention of the messages and behaviors because behaviors originally attributable to the program may be "refreshed" from other sources. It may not be possible to separate these out cleanly; research is required to determine the extent of this problem. The priority depends on the ranking of estimated savings and costs from these programs. In addition, results on measure lifetimes, and any remaining useful lifetime (RUL) and technical degradation factor (TDF) research should be accumulated in a database and updated continuously so comparisons and tracking are facilitated.

REFERENCES[34]

Blasnik, Michael, 1997, "Persistence of Savings: An Assessment of Technical Degradation", Proceedings of the IEPEC Conference, Chicago, November.

Harrigan, Merrilee, and Judith M. Gregory, 1994, "Documenting Energy Savings Enhancements from Energy Education Components of a Low Income Weatherization Program", Proceedings of the 1994 ACEEE Summer Study, Asilomar, CA, August.

Jump, Corina, James Hirsch, Jane Peters, and Dulane Moran. 2008. "Welcome to the Dark Side: The Effect of Switching on CFL Measure Life", Proceedings of the ACEEE Summer Study on Building Conference, Asilomar, CA, August.

Peters, Jane S., Ph.D., Marjorie R. McRae, 2001, "Regional Building Operator Certification - Venture: Final Market Progress Evaluation Report", Prepared for Northwest Energy Efficiency Alliance, September.

Skumatz, Lisa A., 2009. "Lessons Learned and Next Steps in Energy Efficiency Measurement and Attribution: Energy Savings, Net to Gross, Non-Energy Benefits, and Persistence of Energy Efficiency Behavior", prepared for CIEE, December. Available at: http://uc-ciee.org/energyeff/energyeff.html.

Skumatz, Lisa A., and Juri Freeman, 2011, "Measuring the Impacts of Social Marketing – What is the "Bang for the Buck"? Is it Worth It?", Proceedings of the EEDAL Conference, Copenhagen, May.

Skumatz, Ph.D., Lisa, and John Gardner, 2005. "Best Practices in Measure Retention and Lifetime Studies: Standards for Reliable Measure Retention Methodology Derived from Extensive Review", Proceedings of the IEPEC Conference, New York, August.

Skumatz, Lisa and John Green. 2000. "Evaluating the Impacts of Education/Outreach Programs - Lessons on Impacts, Methods, and Optimal Education". Proceedings from ACEEE Conference, Asilomar, CA, August.

Skumatz, Lisa A., and Curtis Hickman, 1992, "Measure Life Study: The Effect of Commercial Building Changes on Energy Using Equipment", Proceedings of the ACEEE Building Summer Study, August.

Welch, Cory and Brad Rogers, 2011, "Estimating the Remaining Useful Life of Residential Appliances", Proceedings of the AESP Conference, Orlando, January.

.

[34] The long list of references underlying this project is cited in Skumatz 2009, the full report from which this article is adapted.