New Evidence of Behavioral Measure Lifetimes– Tracking Retention through Zero

Lisa A. Skumatz, Ph.D., Michael Santulli, and Dana D'Souza, Skumatz Economic Research Associates, Superior, CO, USA

Abstract

Behavioral and social marketing programs are becoming increasingly popular, and there is intense interest in performance statistics for these initiatives. One especially under-studied aspect is the lifetime or length of time the changed behaviors are retained after the program interventions cease. This internally-funded study examines multiple evaluations of Home Energy Report (HER) programs from around the nation, including new data from a Connecticut study that tracked the savings to zero. We examined:

- Savings persistence from multiple locations and program designs.
- The estimated useful lifetimes (EULs) for the programs, using measured savings, formula-based calculations, and model / trend-based projections.
- Whether the programs lead to greater participation in "other" programs and adoption of "deeper" measures (insulation, etc.)
- Whether persistence results can be used to refine delivery to improve savings and cost per kilowatt hour.

We found:

- Program savings are about 1.5%, and are substantially higher for high use customers. Early year retention is about 75%-90%.
- EULs varied from about 2.0 to 5 years, but most hover in the range of 2.5-3.3. EULs based on measured savings (to zero) and model-based projections tended to be a little lower than estimates based on formulae based on average early-year decay and attrition factors.
- The HER program leads to some increase in uptake of a weatherization program, and in one measure (insulation).
- The strong retention opens the door for program "on/off cycling"; computations show substantial gains in simple cost-effectiveness and strong savings.
- Retention studies should be built into budgets for all behavior programs to better estimate cost-effectiveness.

Introduction

Social marketing, feedback, and other behavioral programs are being implemented at more and more utilities, and literature on the impacts of these programs is increasing. Historically, most of the evaluations associated with the programs provided quantitative impacts only at the immediate conclusion of the treatments, which has made it difficult to estimate the effective measure lifetimes associated with these programs, and made it difficult to fully assess the benefits stream for benefit-cost analysis. Without measure life information, program savings are understated, cost-effectiveness is incompletely calculated, and the allocation of funds is not optimal.

This internally-funded research summarizes the results of a number of analyses of behavioral program impacts:

- Persistence of impacts after the treatment year;
- Curve-fitting for decay functions for savings to support estimated useful lifetime (EUL) estimation;
- Computations and comparisons of results for measure lifetimes;
- Use of results for possibly refining program delivery; and
- Analysis of auxiliary effects from program participation.

Even though this measure lifetime work for behavior programs is increasing, more work is needed before lifetimes can be associated with confidence. In addition, retention studies need to be conducted for multiple types or classes of behavioral programs, until a sufficient literature on behavior retention patterns is created. These studies are likely shorter term than measure-based retention studies. Unlike long-lasting measures (HVAC programs, insulation, etc.), these repeated studies don't need to cover decades, but rather, a horizon of perhaps 3-5 years could address most programs.

The Literature on Behavioral Programs

Based on a literature review of more than 150 studies in the wide behavioral sphere, we found that few studies examined savings beyond the "treatment" year, and only a few before 2009 even mentioned retention (Freeman and Skumatz 2012), and even today, work on the statistical analysis of retention is sparse.[1] Recent work (Gillingham, et. al. 2018) identified little behavioral program literature beyond the Home Energy Reports programs discussed in earlier work by this author and others, and retention is not addressed in detail. Omitting retained impacts understates the true cost-effectiveness and makes it hard for potentially important and dynamic education programs to receive high benefit-cost ratios, reducing the likelihood of appropriate funding levels (Skumatz, Khawaja, and Colby 2009).

Description of the HER Program

Home Energy Reports (HER) Programs are the exception; the programs are popular, relatively consistent in design, and well-funded. Impact evaluation studies are quite common (See Figure 1), and more and more retention studies of HER programs are being conducted. The home-energy report-type (HER) is a bill feedback program generally offered to single family households. HER programs are designed to achieve residential electricity savings and customer value to utilities through delivery of a two-page (printed front and back) report. The feedback reports sent to households identify their energy use, and provide comparisons to other households. The group of similar households, referred to as "neighbors", are defined as 100 occupied households similar in size and paying the same rate code as the participating home. HER reports also provide lists of energy-saving tips that differ from month to month and year to year.

These programs are often very large, covering many thousands of households, and use a randomized test/control quasi-experimental design to provide reports to a sample of households, and no reports to another group - a specially-selected "control group".[2] Evaluators compare energy savings of the treatment group to the control group. The program generally uses an "opt-out" design (very few opt out), so the design does not suffer from the self-selection bias that often plagues other energy efficiency program evaluations.[3]

Savings

Energy savings values for the programs are relatively consistent for the high use customers that are the usual targets of these programs. Figure 1 presents a round-up of savings results from programs around the US, showing a relatively common savings figure of about 1.5% electric savings and 0.9% gas savings for the basic program for high use customers. Results from alternative designs are not quite as consistent.

In nearly all cases, the company designs the program to be offered to high use customers, and performance has been relatively consistent at about 1.5%. In a few cases, the company has been asked to deliver the program to more average use customers. Connecticut is one such case, and the program introduced a cohort of about 9,000 "more average" use customers (still substantially higher than average use customers), providing monthly HER reports from July 2012 to December 2016. The monthly average kWh savings per household for "more average" use customers and high use customers are shown in Figure

---

[1] Following up on a similar review conducted in Skumatz and Green 2000 [13]. This lack of retention results was reconfirmed by Mazur-Strammen and Farley, ACEEE, 2013.

[2] Note that the control groups are determined by the vendor and not released, so they have not been able to be independently verified by evaluators. For this reason, some evaluations are adding asterisks to the evaluation results.

[3] Description from NMR Group, Inc. 2016.

2. The savings from the lower use group are much lower. The report (NMR 2017) indicates savings are about 1.2% for the high use customers, and 0.26-0.38% for the more average use customers. We used very simplistic calculations to examine the approximate range of the cost per kWh for the program, simply dividing the utility's direct program fee per household by the savings. Note the results show that the cost per kWh for the average use households is three to five times that of the fee per household. The program is far less cost-effective when delivered outside the high-use cohort of customers. High users show substantially higher savings in kWh and percent, and better "simple" cost-effectiveness than more average use customer groups.

HER program type	Average net electric savings (range)	Average net gas savings (range)
Basic design -opt-out, feedback, social norms (7 studies electric, 6 gas)	1.5% (0.9%-2.1%)	0.9% (0.3%-1.6%)
Adding rewards (2 electric, 1 gas)	0.9% (-2.2%-2.0%)	0.4% (0.3%-0.5%)
Adding on-line portal, enhanced phone (2 electric)	1.6% (0%-2.2%)	
Connecticut program – basic design (NMR 2016) (monthly and quarterly cohorts)	1.88%; 1.7% excluding outlier group (1.57%-3.62%)	

Figure 1: Typical Savings Estimates from HER Programs – Design Variations (Small sample)

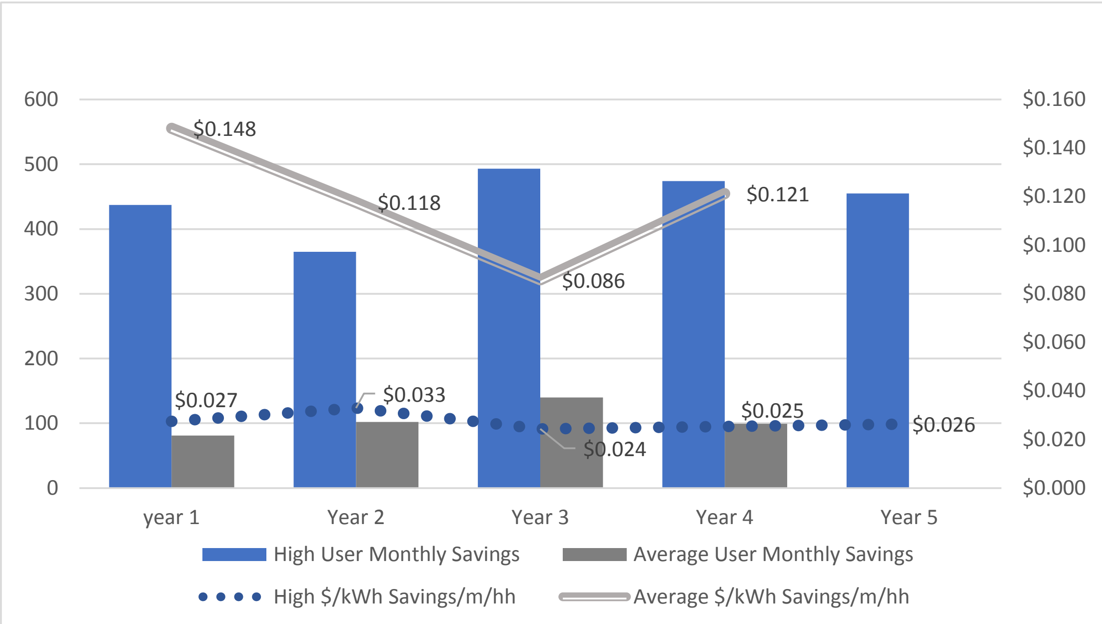

Figure 2: Comparison of Savings and Cost per kWh for High Use vs. "More Average" Use Customers Data from NMR 2017.

In summary, savings for these programs tend to be about 1.5%, with lower savings for some alternate delivery options, and lower for gas programs; however the number of studies is relatively small. Work in Connecticut found savings and a simple cost-effectiveness metric were much better for high-user HER customers compared to "more average" (lower kWh use) HER customers.

Retention and Savings Degradation

We conducted a review of early HER retention studies (Skumatz 2016), to which we have added newer work identified from our updated review of the literature (see Figure 3). This table outlines key aspects of each program (service area, length of treatment / post-treatment). It also shows that the average annual savings decay rates (Column E) vary quite a bit, from a low of 8% to a high of 32%, or about 17% unweighted.

A. Study	B.Service area	C.Treat-ment months	D.Post-treat-ment savings analysis months	E.Average annual savings decay results	F.EUL from Measured Savings to zero savings (in Figure 5)	G.EUL (yrs) (calculated from Formula based on average decay and attrition)	H.EUL from fitted trends (Published / Proj'ns values from Figure 12)
Alcott & Rogers(2014)	Upper Midwest	24-25	26	21%	n.a.	2.9	2.3/2.8
Alcott & Rogers (2014)	West coast	24	29	18%	n.a.	3.3	2.5/3.5
Alcott & Rogers 2014)	West coast	25-28	34	15%	n.a.	4.2	
Integral Analytics (2012)	SMUD	27	12	Savings decay of 32% one-year after treatment stopped	n.a.	2.1	
Statewide Eval Team (GDS+)	PA PUC	48	16	22%	n.a.		2.3/2.8
DOE/LBL	DOE			40%	n.a.		2.3/2.3
DNV-GL(2015)	PSE	24	36	11%	n.a.	5.1	
DNV-GL (2016)	PSE	24	60	8%	n.a.		
DNV-GL (2018)	PSE		Legacy suspended in 2011 declined to about 30% of 2017 group's savings and is at margin of statistical significance.				
ODC (2014)	NGRID-MA	12-24	10	Reduced treatment led to reduced observed savings, with sharper effect for gas cohorts.	n.a.	n.a.	2.6/2.8
NMR (2017; incorporates results from earlier CT studies)	State of CT CT-A[4] CT-B CT-C CT-D[5]	16 mos 5 qtrs 8 mos Varies	44 44 44 44	12% 24% 25% 21% Significant Savings ceased in 4th or 5th year after treatment.	Report-based[6] /Figure 5 2.00 / 2.59 2.50 / 2.53 2.10 /2.14 2.70 / 2.82	2.28 3.80 2.79 3.30 3.3 (range 2.3-3.8)	Average of model, Figure 5 2.63 2.60 2.21 3.17 Range 2.2-3.2
Range and Simple Average		8-48 months	10-60 months	8% - 32%; 17%	2.0-2.8	2.1-5.1	2.2-3.2

(updated from Skumatz, 2016)

Table Note: Nicor Gas study (by Navigant) showed 46% decay (54% retention). We limit the table to electricity programs for better comparison. Published EUL calculates to 1.83.

Table note: All HER reports were delivered monthly and quarterly except the ODC study, which were bi-monthly and quarterly. Attrition for studies not reported in all cases.

Figure 3: Post Treatment Savings from Published HER Program Evaluations[7]

---

[4] These four cohorts are defined below.

[5] The result for Cohort CT-D is a reflection of the weighted average, as it combines the results for the previous 3 cohorts.

[6] The NMR "report"-based results in this column and Column G incorporate consideration of the attrition rate in columns F and G; hence there is not a monotonic relationship between decay rates and lifetimes. The author of the NMR reports that the monthly households used four times as much as the average utility household and the other groups used twice as much, and a sample size issue also enters into the results. The second number in column F omits consideration of attrition.

[7] We exclude attrition (an important factor) because it is not clearly published in the reports generally.

Given a desire to be able to estimate lifetimes for these programs (some of which may not have had their full decay cycle), Khawaja and Stewart (2014) suggested that a measure of effective useful life (EUL) can be estimated using these types of shorter term average decay results. They suggest using EUL = lifetime savings divided by first year savings with lifetime savings calculated using the decay rate and an attrition rate factor.[8] This formula may provide an approximation of lifetimes until sufficient studies are conducted that show when significant savings cease. If the calculation of measure lifetime proposed by Khawaja and Stewart (2014) is applied, the measure lifetimes in Column G of Figure 3 result, with lifetimes ranging from 2.1 years to 5.1 years.

The table includes EULs estimated based on two other methods – both of which are carried out and described later in this paper. Column F develops estimates of lifetimes based on measured savings that are tracked to zero. These data are only available from the Connecticut study by NMR (NMR 2017), and the work is described in the next section. Column H fits statistical trend lines to the first few years of measured or measured / projected data from the non-Connecticut studies until the trend crosses zero; these annual retention figures are then used to compute a lifetime. These modeling efforts are described in a later section of the paper.

Non-Linearity in Decay and Effect on Lifetime Estimates in the Connecticut Study Cohorts

The simplification of the degradation terms in the literature, usually using averages and linear functions, seems inappropriate. Decay functions traditionally are assumed to take on exponential or other forms. To investigate "shapes", we used the data from Connecticut study by NMR 2017 to fit more sophisticated functions than linear to the data. This study is examined in detail because it tracked savings to zero for multiple customer groups. Impact and retention evaluations were conducted (NMR 2017) until the statistical savings disappeared from all groups. Figure 4 shows the pattern of annualized savings for each group. There are four cohorts in the Connecticut Study.

- Group CT-A: received monthly HER reports between 1/11-4/12 (1,507 customers)
- Group CT-A*: same as above, but adding savings values for the last two impact years for curve-fitting, even though the small sample size finds savings for these last two years are not statistically significant.
- Group CT-B: received quarterly HER reports 1/11-4/12 (9,374 customers)
- Group CT-C: received monthly HER reports from 1/11-8/11 (3,796 customers)
- Group CT-D: combined discontinued groups from above (14,733)

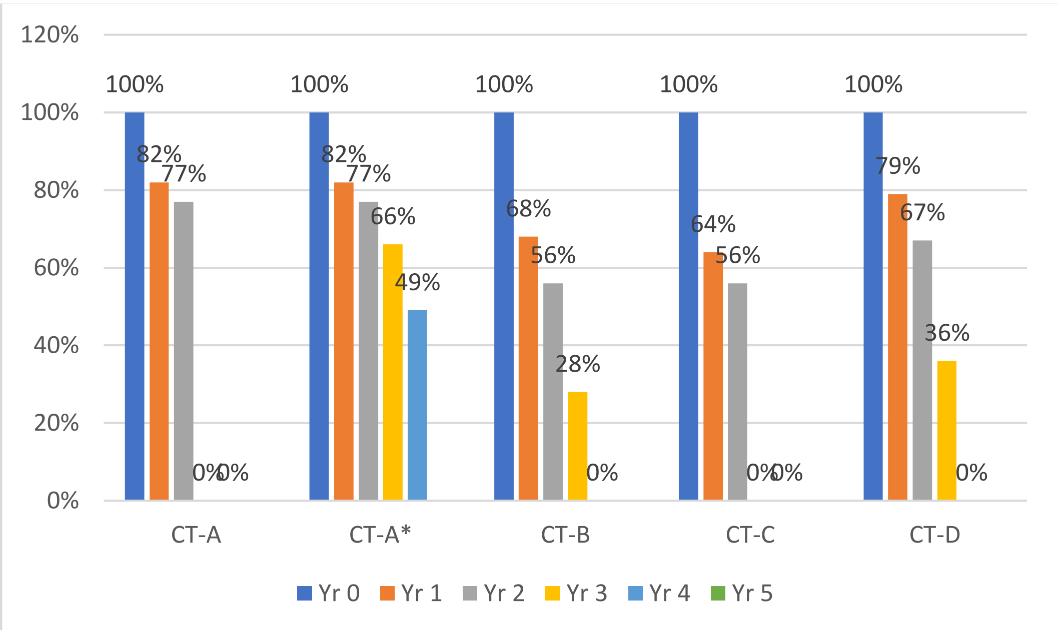

Figure 4: Retention Percent of Annualized kWh/Household Savings for Connecticut Study by Treatment Group

[8] Implied lifetimes calculated per Khawaja, 2014: First year savings / (d+a-d*a) where d=annual decay rate and a=annual attrition rate.

The authors used Excel's regression, log, exponential smoothing, and the other techniques listed in Figure 5 to see which functions would best fit the decay functions evidenced in the Connecticut study (Figure 4). Using these simple functions, the fit statistics (in the last column of Figure 5) and visual inspection of Figure 6 through Figure 10 indicate the exponential smoothing provides the better fit to the data. Figure 5 also identifies the treatment period and the measure lifetimes implied by the fitted function. For illustration purposes, Figures 6 through 10 provide graphical representations of these the measured savings decay and the fitted curves.

	Program Year	1	2	3	4	5	6	7	8	9	10	11	12	EUL	Model "Fit" Statistics
CT - A	**Treatment: 16 months**													**2.63**	<= Avg. of Group's Modeled EULs
	Measured	100%	82%	77%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2.59	Measured EUL
	Linear Trend	**100%**	**80%**	**49%**	**19%**	0%	0%	0%	0%	0%	0%	0%	0%	2.48	R^2=0.7912
	Log	**100%**	**71%**	**47%**	**30%**	**16%**	**6%**	0%	0%	0%	0%	0%	0%	2.69	R^2=0.6506
	Exponential Smoothing	**100%**	**85%**	**79%**	**16%**	0%	0%	0%	0%	0%	0%	0%	0%	2.80	R^2=0.9505
	Regression	**100%**	**80%**	**52%**	**24%**	0%	0%	0%	0%	0%	0%	0%	0%	2.55	R^2=0.8612, F=18.61, t-stat=6.29
	Polynomial	**100%**	**95%**	**64%**	**4%**	0%	0%	0%	0%	0%	0%	0%	0%	2.63	R^2=0.9372
CT - A*	**Treatment: 16 months**													**3.87**	<= Avg. of Group's Modeled EULs
	Measured	**100%**	**82%**	**77%**	***66%***	*49%*	0%	0%	0%	0%	0%	0%	0%	3.74	Measured EUL
	Linear Trend	**100%**	**88%**	**71%**	**54%**	**36%**	19%	1%	0%	0%	0%	0%	0%	3.69	R^2=0.8708
	Log	**100%**	**80%**	**62%**	**49%**	**39%**	**31%**	24%	18%	13%	8%	4%	0%	4.31	R^2=0.7219
	Exponential Smoothing	**100%**	**85%**	**79%**	**69%**	**53%**	0%	0%	0%	0%	0%	0%	0%	3.85	R^2=0.9844
	Regression	**100%**	**88%**	**71%**	**54%**	**36%**	**19%**	0%	0%	0%	0%	0%	0%	3.68	R^2=0.8708, F=26.96, t-stat=9.44
	Polynomial	**100%**	**91%**	**81%**	**63%**	**39%**	7%	0%	0%	0%	0%	0%	0%	3.80	R^2=0.9509
CT - B	**Treatment: 5 quarters**													**2.60**	<= Avg. of Group's Modeled EULs
	Measured	**100%**	**68%**	**56%**	**28%**	0%	0%	0%	0%	0%	0%	0%	0%	2.53	Measured EUL
	Linear Trend	**100%**	**75%**	**51%**	**27%**	**3%**	0%	0%	0%	0%	0%	0%	0%	2.54	R^2=0.9853
	Log	**100%**	**66%**	**42%**	**26%**	**13%**	**2%**	0%	0%	0%	0%	0%	0%	2.49	R^2=0.9299
	Exponential Smoothing	**100%**	**75%**	**60%**	**35%**	**14%**	0%	0%	0%	0%	0%	0%	0%	2.84	R^2=0.9893
	Regression	**100%**	**75%**	**51%**	**27%**	**3%**	**2%**	0%	0%	0%	0%	0%	0%	2.56	R^2=0.9853, F=201.60, t-stat=21.86
	Polynomial	**100%**	**75%**	**52%**	**27%**	**1%**	0%	0%	0%	0%	0%	0%	0%	2.56	R^2=0.9864
CT - C	**Treatment: 8 months**													**2.21**	<= Avg. of Group's Modeled EULs
	Measured	**100%**	**64%**	**50%**	0%	0%	0%	0%	0%	0%	0%	0%	0%	2.14	Measured EUL
	Linear Trend	**100%**	**69%**	**38%**	**6%**	0%	0%	0%	0%	0%	0%	0%	0%	2.14	R^2=0.9564
	Log	**100%**	**60%**	**34%**	**15%**	0%	0%	0%	0%	0%	0%	0%	0%	2.09	R^2=0.896
	Exponential Smoothing	**100%**	**71%**	**54%**	**11%**	0%	0%	0%	0%	0%	0%	0%	0%	2.36	R^2=0.9816
	Regression	**100%**	**69%**	**43%**	**16%**	0%	0%	0%	0%	0%	0%	0%	0%	2.28	R^2=0.9364, F=44.14, t-stat=9.26
	Polynomial	**100%**	**73%**	**41%**	**3%**	0%	0%	0%	0%	0%	0%	0%	0%	2.17	R^2=0.9657
CT - D	Treatment: Varies													**3.17**	<= Avg. of Group's Modeled EULs
	Measured	**100%**	**79%**	**67%**	**36%**	0%	0%	0%	0%	0%	0%	0%	0%	2.82	Measured EUL
	Linear Trend	**100%**	**81%**	**60%**	**40%**	**19%**	0%	0%	0%	0%	0%	0%	0%	3.00	R^2=0.9678
	Log	**100%**	**75%**	**58%**	**45%**	**36%**	**28%**	**22%**	**16%**	**11%**	**7%**	**3%**	0%	4.01	R^2=0.8951
	Exponential Smoothing	**100%**	**84%**	**71%**	**43%**	**13%**	0%	0%	0%	0%	0%	0%	0%	3.10	R^2=0.9903
	Regression	**100%**	**81%**	**56%**	**32%**	**8%**	**7%**	0%	0%	0%	0%	0%	0%	2.84	R^2=0.9647, F=81.94, t-stat=14.51
	Polynomial	**100%**	**84%**	**63%**	**37%**	**5%**	0%	0%	0%	0%	0%	0%	0%	2.89	R^2=0.9819

Figure 5: Measured and Simple Fitted Values for Connecticut HER Cohorts
The fitted functions are included at the end of the report.

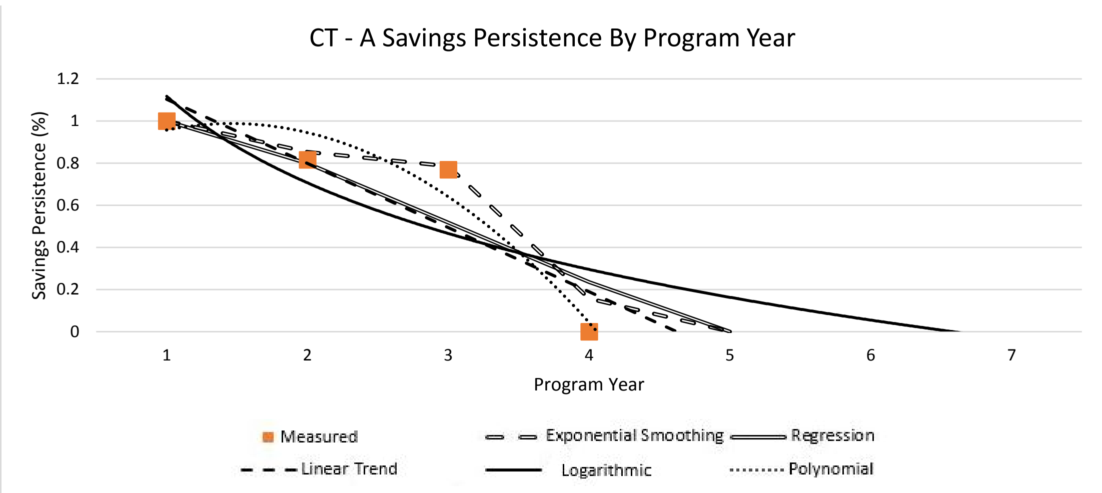

Figure 6: Cohort CT-A Persistence of Savings over Time – Measured and Modeled

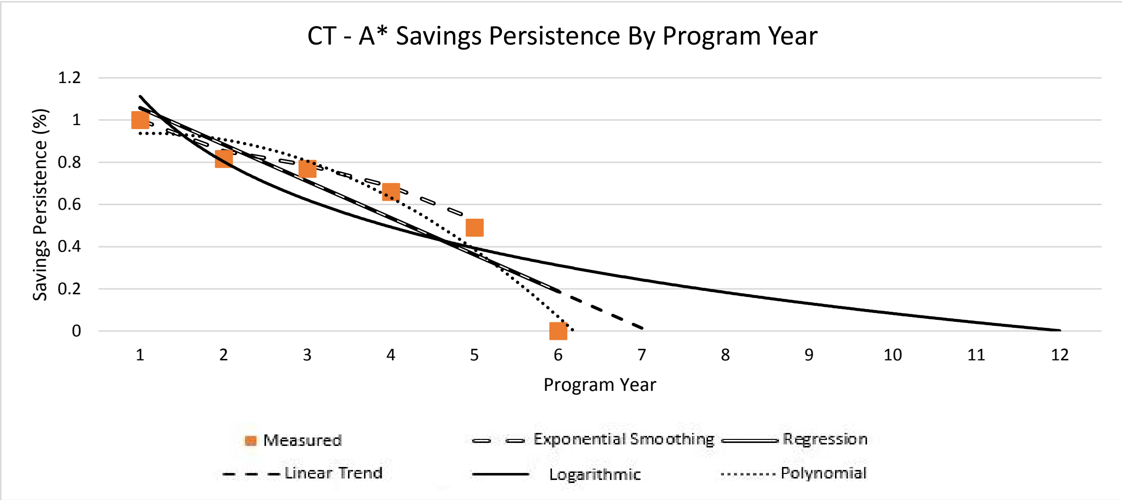

Figure 7: Cohort CT-A* Persistence of Savings over Time – Measured and Modeled

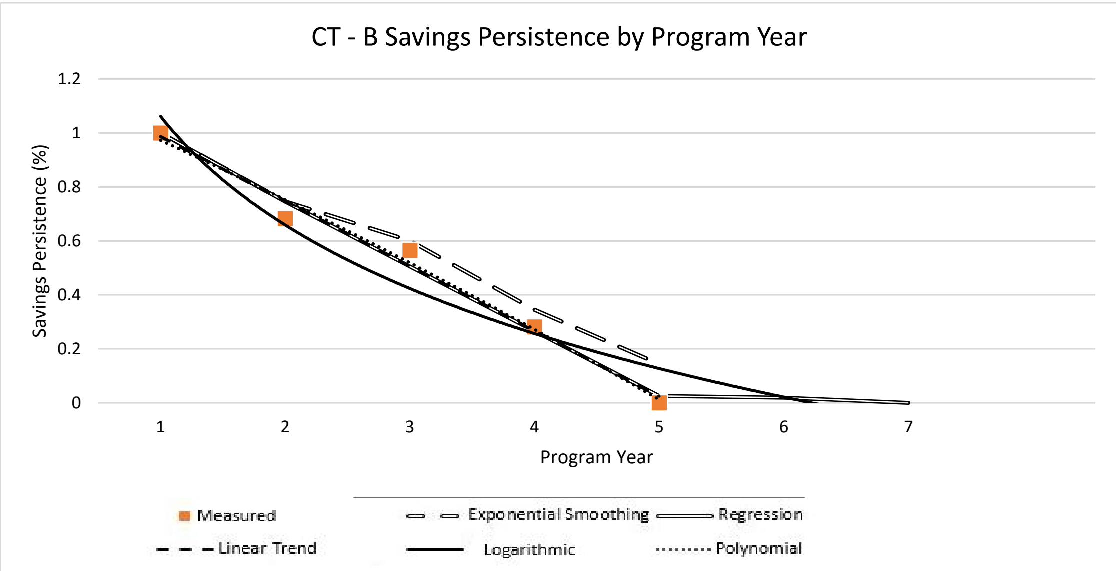

Figure 8: Cohort CT-B Persistence of Savings over Time – Measured and Modeled

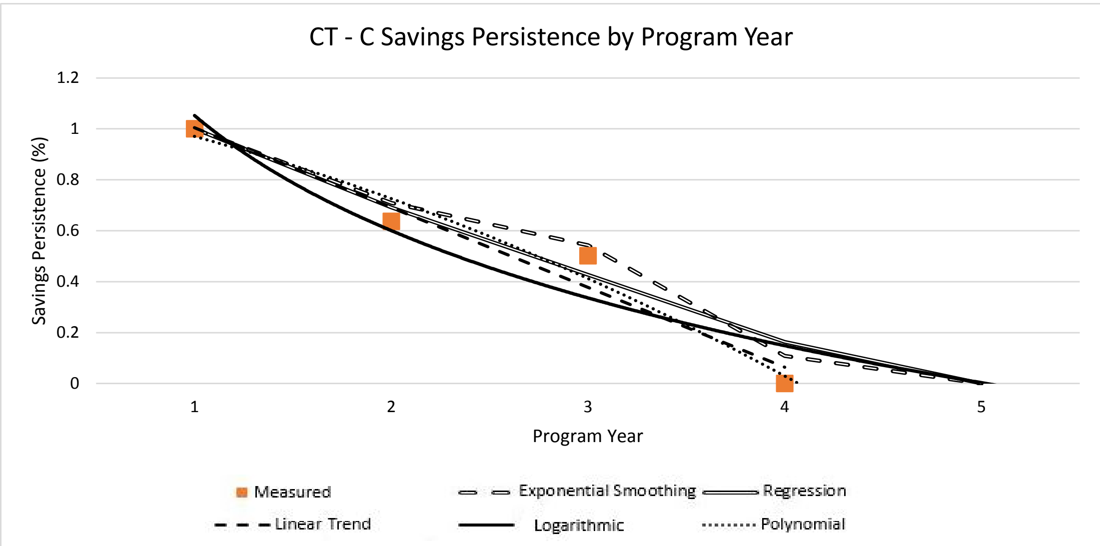

Figure 9: Cohort CT-C Persistence of Savings over Time – Measured and Modeled

Figure 5, and the illustrations of its models, can be used to develop estimates of the effective lifetime for the behavioral savings, here essentially a multiplier that can be applied to the first year savings to estimate the lifetime savings from the behavior change. This is similar to, but not the same as a widget- or mechanical-equipment-type measure life, but has a similar function in the benefit-cost equation and assessment of the measure or program's cost-effectiveness. However, here even if we have an EUL of, say, 3 years, we could expect to find actual savings occurring in year 5; this is not the same expectation as equipment EULs.

We would hope to find the models and shapes can derive lifetimes that are similar to the measured savings, and if so (or even better if evidence of similar "shapes" can be found over time), it may be that models can be fit to early year behavioral savings as proxies for lifetimes in early stages of programs.[9] In Figure 5 we show the EUL based on "Measured" data from the impact evaluation, as well as EULs resulting from our exploratory work using a variety of fitted curves and trends (linear trend, log, exponential smoothing, regression, and polynomial)[10]. For each group (CT-A through CT-D), the average of the "modeled" EULs is relatively close to the measured EUL. The best R-squared value (and the best visual fit in Figures 6-10) is associated with the exponential smoothing model. However, in most cases, the EUL that is closest to the "measured" EUL is the regression model. The least squares fit also seems to deliver the best estimate of EUL.

The measured and modeled EULs for these Connecticut HERs cohorts vary from about 2.2-3.2, with a longer EUL for the group CT-A*. Recall this is the group that included savings values for the two tail years that were not statistically significant, but helped outline the shape of the curve for research interest. This group's estimated measure life was 3.9 years. We find these values are generally lower than the EULs estimated shown in Column G of Figure 3, which were estimated using the literature-proposed formula based on decay and attrition.

From this small sample, the longer the period of treatment for these cohorts – whether quarters or months – the stronger the retention and the longer lifetime is associated. These shapes may also be useful in helping extrapolate the pattern of savings for other programs that have only one or two years of retention data. The exception is CT-A*, the group that received 16 months of HER reports – but including the estimated – but insignificant - retention values for years 3 and 4. The lifetimes for this group were longer; however, the relatively small sample made it more difficult to produce significant results for savings.

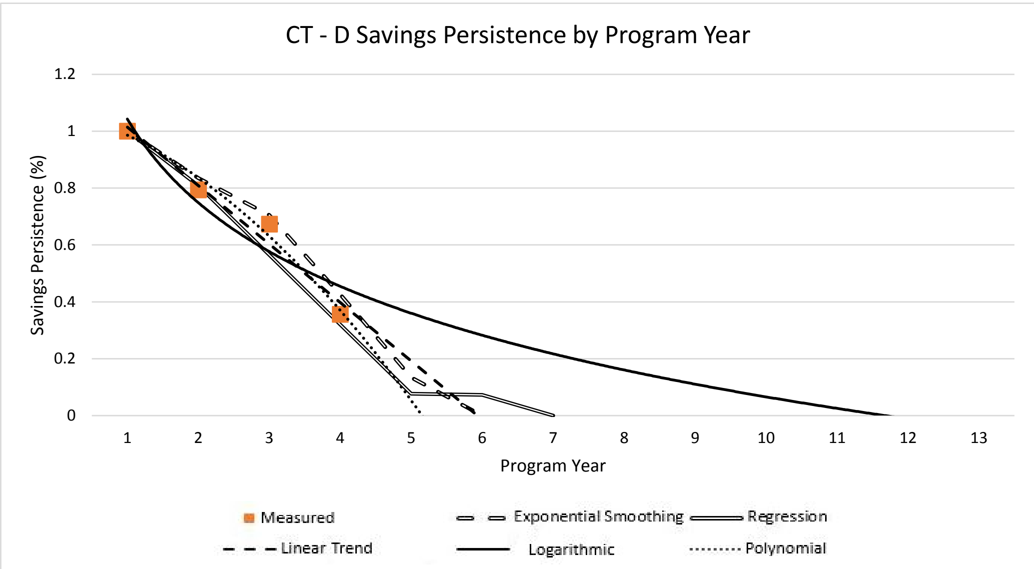

Figure 10: Cohort CT-D Persistence of Savings over Time – Measured and Modeled

In summary, the Connecticut report series estimated savings from the HER program annually until the savings became insignificantly different from zero. The first year's measured retention ranged from 64%-82%, the second year retention ranged between from 50%-77%, and third year retention ranged from

[9] There are several simplifying assumptions being made here. Measure-based EULs are medians; these are average figures. Attrition is being ignored for simplicity and because the figures were not available for many studies. .

[10] In more extensive work, additional functions would certainly be tested.

0%-28% (Figures 4 and 5). Average percentage points of fall-off (CT-D) were 21%, 12%, 31%, and 36% across the years. The associated lifetimes based on the measured savings from the repot (tracked to "zero") and the fitted values averaged 2.7 years (range 2-2.5) and tended to be lower than values computed using the Khawaja and Stewart average-decay-and-attrition-based formula for estimating lifetimes. Some of the estimated models tracked well against the measured EULs, including the exponential smoothing model, but the EULs computed based on the regressions most closely matched the measured EL values.

Non-Connecticut Patterns of Decay – Projections from Studies that did not Track to Zero

Figure 3 included results from multiple HER savings and retention studies conducted around the country. Most of these studies have not (yet) followed the cohorts until the savings cease, and some of the studies with shorter follow-on months have larger degradation factors. In addition, many of the reported numbers are "average of decay values", which may be too simplistic a function to accurately portray the degradation of savings. The first set of columns in Figure 11 tally the actual measured decay factors published for a number of these studies – excluding the Connecticut study already addressed. Note most only measure one year after the year of discontinuation of the reports (except Pennsylvania), and retention varies from about 40% to 90%. The second group of columns include the near-term projections some of these reports made regarding decay rates (grey header in Figure 11). Some projected retention out one additional year, and two projected their expected retained savings through a third and fourth year. These values are illustrated graphically in Figure 12. Note one used a simplistic assumption that the retention would stay steady, and the graph is horizontal for that study.

In the last set of columns in Figure 11, we again fit functions to the study data. In most cases we used linear models and in a few the fit appeared better using the polynomial smoothing model. Comparing EULs, we find that of course, the one-year measurement understates the likely EUL, and the projections improve that estimate. Our modeling work estimates that the measure lifetimes for these programs would be expected to range from 2.3-3.5 – not dissimilar to the Connecticut results of 2.5-3.3. The outlier is the value of 1.8 years for the sole gas program in the table.

	Measured Savings Persistence				Published Measure Persistence by Study						Measured Persistence with SERA Trending to Zero							
Study	Year 0	Year 1	Year 2	EUL	Year 0	Year 1	Year 2	Year 3	Year 4	EUL	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	EUL
Alcott & Rogers	100%	75%		1.75	100%	75%	56%			2.31	100%	75%	56%	33%	11%	0%	0%	2.
Alcott & Rogers	100%	82%		1.82	100%	82%	67%			2.49	100%	82%	67%	50%	34%	18%	1%	3.
Pennsylania PUC	100%	78%	56%	2.34	100%	78%	56%			2.34	100%	78%	56%	34%	12%	0%	0%	2.
DOE	100%	60%		1.60	100%	60%	36%	22%	13%	2.31	100%	60%	36%	22%	13%	0%	0%	2.
Opinion Dynamic	100%	40%		1.40	100%	40%	40%	40%	40%	2.60	100%	40%	40%	40%	40%	16%	4%	2.
Nicor Gas	100%	54%		1.54	100%	54%	29%	0%		1.83	100%	54%	29%	0%	0%	0%	0%	1.

Figure 11: Post Treatment Decay in Savings from Non-CT Studies: Numeric Data and Linear / Polynomial Projections to Zero

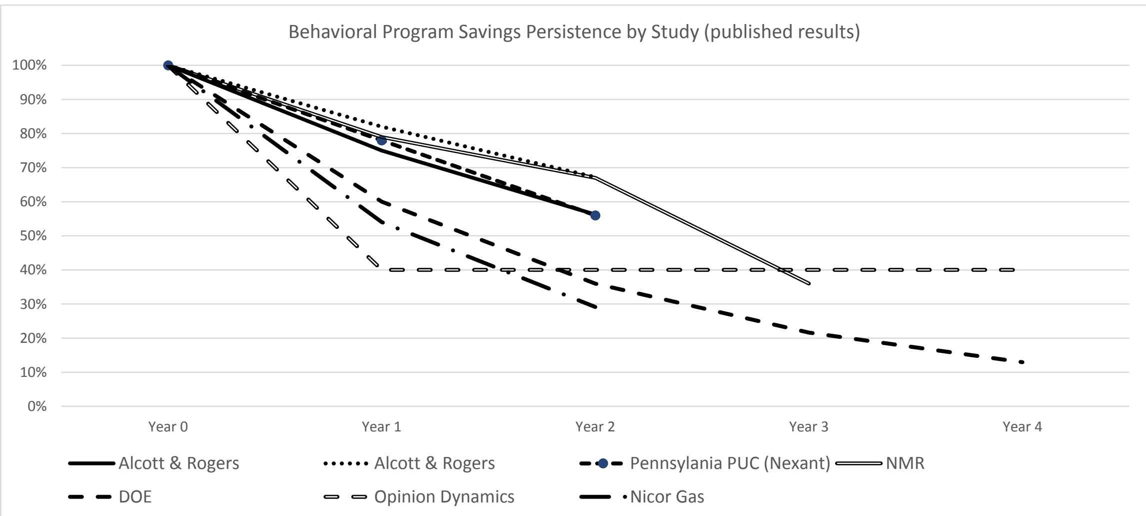

Figure 12: Post-Treatment Measured and Published / Projection Results from Non-Connecticut Reports

In summary, the retention of savings one year after treatment varied a great deal between studies. Values ranged from about 40% to 90%. The resulting lifetime estimate, using a simplified formulae, vary from 2.1 years to 5.1 years for all programs the table, or 2.9-3.3 for those included in this estimation work (from Figure 3 column G), and our modeling efforts suggest these EULs may range from 2.3-3.5 years. Again, some of the modeling efforts may provide opportunities to estimate lifetime savings after perhaps two years of post-measurement.

Indicative Patterns of Lifetimes by Length and Frequency of Treatment

A simple review of the results for lifetimes shows some expected patterns, and some counterintuitive results. This remains a small sample of evaluations, so these are indicative results only, and an increasing body of literature can only help improve the reliability of the findings.

Duration of Treatment Period:

The average EUL increases with the duration of the treatment period (i.e. total length of time the HERs are sent out). We present splits for short, medium, and longer periods, as well as an option that collapses short and medium to increase sample sizes. The average annual decay does not show a consistent pattern with the duration of treatment period (value reflects removal of one outlier of 10%) but the average first year decay decreases as the duration of treatment period increases.

	Average Lifetime Years	Average Annual Decay	Average First Year Decay
Short (<1 year) (2 obs)	2.29	33%	40%
Medium (1 year to <2 years) (3 obs)	2.34	30%	36%
Long (2 years plus) (4 obs)	2.75	36%	31%
Combined Short & Medium (5 obs)	2.32	31%	38%

Figure 13: Patterns by Duration of Treatment Period – Indicative / Small Samples

Treatment Frequency:

The average EUL is equal for monthly and quarterly treatment frequencies. The average annual decay is lower for those receiving quarterly HERs (unexpected result), and the average first year decay is lower for the customers receiving quarterly HERs (unexpected direction). There are few results published for quarterly frequency programs, so these results are small sample comparisons. More research on programs with quarterly reports would be useful to explore how much difference frequency makes in behavior change and retention.

Treatment Frequency	Average Lifetime Years	Average Annual Decay	Average First Year Decay
Monthly (7 obs)	2.48	35%	37%
Quarterly (2 obs)	2.49	20% (only 2 observations)	27%

Figure 14: Patterns by Treatment Frequency– Indicative / Small Samples

In summary, these small-sample results indicate lifetimes end to be longer for groups with longer treatment periods, and first year retention is also higher for these groups. The patterns are less clear annual average decay. The small-sample results do not show consistent differences or consistent patterns for households receiving HERs reports on a monthly versus quarterly basis.

Other Impacts: Do HER reports Increase Adoption of Deeper Measures?

The HER program encourages behavior change in use of equipment. However, it may also encourage customers to invest in additional energy efficient equipment. These behaviors may be considered attributable to the behavior change program, or may be counted under another measure-related energy efficiency program; however if that equipment is purchased under another utility program, the risk arises that the savings may be double-counted – and inappropriately attributed to both the HER and the other program.

The Connecticut work (NMR 2016) examined the issues of the degree of potential double-counting. The project showed that participation in the utility's basic Home Energy Saver (HES) program was higher for HER participants than non-participants (4.7% vs. 3.9%). Participation in other programs was not significantly different. Inspection for differences in investment in deeper measures incentivized by the program found a higher rate of installation of insulation for high users (8.9% vs. 7.1%), but no significant difference in uptake of other measures (furnaces and boilers, other HVAC, refrigerators and freezers, heat pump water heaters, windows). The study concluded that, at this time, the double-counting issue was not a substantial problem.

In summary, the HER program led to a 0.8 percentage point increase in uptake in HES program participation (20.5% increase), and led to higher rate of installation for insulation (1.8 percentage points, 25%).

Implications and Cost-Effectiveness

The short term retention rates for the programs in Figure 3 are quite high – ranging from about 60% to more than 90% after program investment stops. This implies dollar savings might be achieved by "cycling" customers on and off the program, saving all program costs during the "off" years. The result of four simple options are presented in Figure 7. Those options are: continuous treatment of the homes with HER reports; 1 year on and 1 year off; 1 year on and 2 years off; 1 year on and 3 years off. Figure 16 shows the detailed computations for low, medium, and high decay programs. Figure 15 provides a graphical illustration of the medium retention model. Figure 15 presents the total savings for all years involved in each scenario (dark bar), and the average savings per year (lighter bar). It also shows the average cost per kWh, attributing only the utility's direct (contracted) cost per household.

Figure 15 shows that average savings per year do not decline nearly as much as does the cost per kWh, and "on/off cycling" of customer groups shows promise for maximizing cost-effectiveness. The program may be able to start with somewhat higher savings if the program has been administered for possibly two years before the cycling begins, according to some of the results from the literature; a brief ramping-up period in the program is claimed. Khawaja and Stewart (2014b) note over four years of delivery, the programs deliver 73%, 93%, 96%, and 100% of the maximum savings, and for a three year program the ramp-up was shown as 57%, 87%, and 96%. In addition, we would assume the program would include only high users to provide greater value.

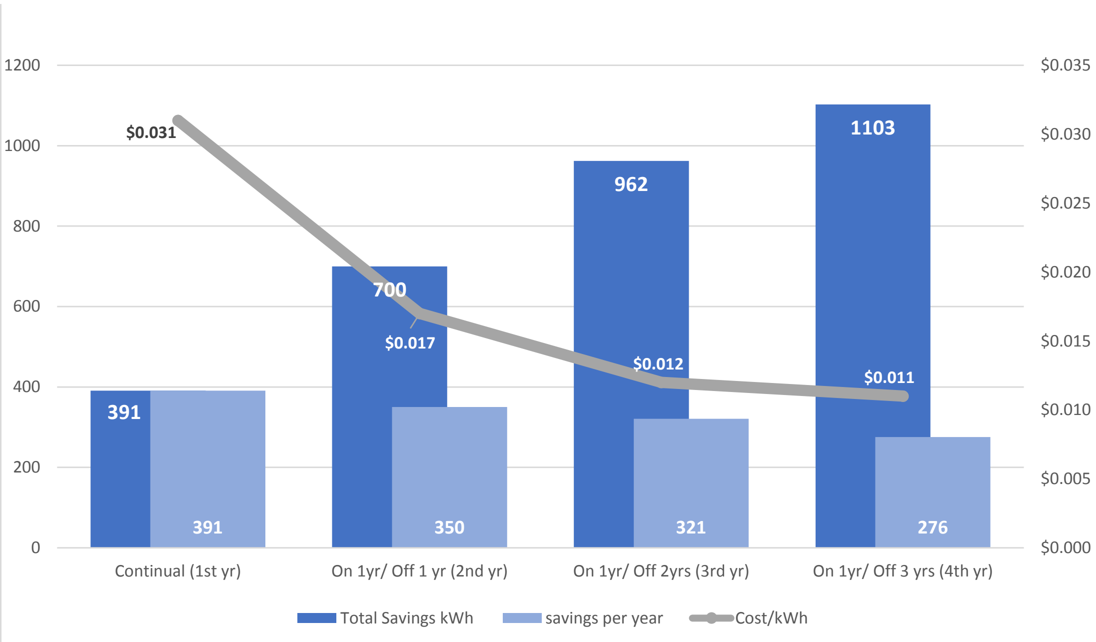

Figure 15: Effect of Behavioral Program "Cycling" on Savings and Costs per kWh

Again, Figure 16 includes detailed computations of three "decay-level" scenarios. Simplified sample computations for the range and average savings retention values from Figure 3 are used to derive the three “decay rate” columns presented in Figure 16. Note we demonstrate the computations for high, “typical / CT” results, and low decay rate cases. The computations show:

- Cycling among customers may allow savings to be achieved at a considerably lower cost per kWh.
- Reviewing the overall column, two or even three years off can achieve a high percentage of savings at a fairly substantial savings (on the order of 75% of the savings at 38% of the cost of continual treatment.
- If the full amount of total annual savings must be achieved, second cohorts, signed on during “off years”, will make up the difference in kWh. This is still less expensive, and “touches” more customers.
- The preference for cycling at a particular utility may depend on the costs relative to other programs, and the degree to which savings goals are stated as maximum savings, or most cost-effective savings, or other statements of goals.
- These are simplistic calculations. Other cycling designs may be created (2 years on, 1 year off, 2 on / 2 off, etc.). Note that these figures are approximate, based on annualized computations and rounded data to show that benefits accrue, even for retention for the high decay case.

	Low Decay Range (8-11%)	Average Literature (17%) (uses CT 2017)	High Decay Rate (24-32%)
Year 1 savings kWh per household (HH)	391	391	391
Year 2 savings (% of Yr1)	91%	79%	68%
Year 3 savings (% of Yr1)	81%	67%	36%
Year 4 savings (% of Yr1)	72%	36%	0%
Cost/hh for treatment	$12.00	$12.00	$12.00
Continual treatment			
Total savings 1 year	391	391	391
Cost/kWh	$0.031	$0.031	$0.031
% of $/kWh compared to continual treatment	100%	100%	100%
1 year on, 1 year off			
Savings	745	700	657
Cost/kWh	$0.016	$0.017	$0.018
% of $/kWh compared to continual treatment	52%	56%	60%
Percent of kWh compared to similar years continual treatment	95%	90%	84%
1 year on, 2 years off			
Savings	1062	962	798
Cost/kWh	$0.011	$0.012	$0.015
% of $/kWh compared to continual treatment	37%	41%	49%
% of kWh compared to similar years continual treatment	91%	82%	68%
1 year on, 3 years off			
Savings	1341	1103	798
Cost/kWh	$0.009	$0.011	$0.015
% of $/kWh compared to continual treatment	29%	35%	49%
% of kWh compared to similar years continual treatment	86%	71%	51%

Figure 16: Calculations of Cost and Savings Incorporating Customer Cycling

In summary, the strong retention of behavior change from the HER programs for several years after the reports imply the programs could be cycled on and off and achieve similar savings at lower cost, and could touch more customers on a fixed budget. There are issues related to ramp-up, and other considerations, but the indications are fairly clear that program delivery optimization can be explored with good financial effects.

Conclusions

Measure lifetimes are important and often neglected features of behavioral programs. The lack of these studies makes it difficult to properly allocate expenditures among energy efficiency program alternatives, and hampers the development of optimal designs for delivery of these programs. The programs are either being implemented every year, or in some cases, ad hoc lifetimes, unsupported by analyses, are assigned. An increasing body of literature addressing the Home Energy Reports (HER) program offered by multiple utilities, provided information on decay (and retention) for multiple years. Relatively few of the studies have tracked savings to zero; more of these studies will be useful in developing clearer estimates of the level of retention, and the pattern of retained savings until program impacts cease to be significant. Our review of multiple retention studies of HER programs around the county found:

- Program savings are about 1.5%, and are substantially higher for high use customers. Early year retention is about 75%-90%. Persistence and lifetimes may be higher when the program has been delivered longer, but there was little evidence for differences for use of quarterly vs. monthly HER Reports.
- EULs varied from about 2.0 to 5 years, but most hover in the range of 2.5-3.3. Savings last into the fifth year in some cases, but the total savings impact is about 2.5-3.3 times first year savings. EULs based on measured savings (to zero) and model-based projections tended to be a little lower than estimates based on formulae based on average early-year decay and attrition factors.
- The HER program leads to some increase in uptake of a weatherization program, and in one measure (insulation).
- The strong retention opens the door for program "on/off cycling"; computations show substantial gains in simple cost-effectiveness and strong savings. In some cases, 90% of the savings can be achieved for 56% of the cost per kWh, and even better results obtain if retention is higher than the middle ranges.
- More research is needed, and retention studies should be built into budgets for all behavior programs to better estimate cost-effectiveness.

Defensible estimates of retained savings from behavioral programs is needed to properly attribute program savings, refine program design, and seriously consider the programs in cost-effectiveness calculations with lifetime numbers that are stronger than guess-work or *ad hoc* deemed values.

This paper argues that behavioral program lifetimes matter and are useful in refining program design. We present some quasi-real-world calculations of the implications of retention data, indicating that program cycling (stopping and starting the program or delivering the program to different subgroups in turn) may be a cost-effective way to deliver this, and possibly other, behavioral programs, depending on the utility's goals and energy efficiency mandates. More work on quarterly programs would be useful; few studies on this option for less-frequent reports are published. The most important conclusion of the study is that interesting results from some published evaluations are now available, but retention studies should be planned into every behavioral program at the start, or at least those with demonstrated significant first year savings. Once the project is complete, the measurement opportunity is lost, and the cost-effectiveness is highly dependent on retention of behaviors. Without that, the programs are short-changing their cost-effectiveness results.[11] We're may be underinvesting in behavioral programs – but we wouldn't know!

Appendix

CT - A	**Formula**	**CT - B**		**CT - D**	
Linear	y = -0.1739x + 1.2314	Linear	y = -0.2401x + 1.2262	Linear	y = -0.2054x + 1.2191
Log	y = -0.593ln(x) + 1.1178	Log	y = -0.58ln(x) + 1.0615	Log	y = -0.424ln(x) + 1.0426
Polynomial	y = -0.1463x2 + 0.427x + 0.6764	Polynomial	y = -0.0067x2 - 0.2x + 1.1793	Polynomial	y = -0.0277x2 - 0.0669x + 1.0805
CT - A*		**CT - C**			
Linear	y = -0.1739x + 1.2314	Linear	y = -0.3136x + 1.3184	Exponential Smoothing	
Log	y = -0.447ln(x) + 1.1129	Log	y = -0.652ln(x) + 1.0524	Projections were estimated using Excel's Data	
Polynomial	y = -0.0361x2 + 0.0789x + 0.8942	Polynomial	y = -0.0346x2 - 0.1408x + 1.1457	Analysis toolpak which does not provide formulae	

Figure 17: Fitted Functional Forms underlying Figure 5 Estimates for the Connecticut Models

[11] Many thanks to Lisa Wilson-Wright and Chris Russell of NMR Group for the work on the useful document (NMR 2016) and its predecessors, and for helpful comments on this article. In addition, sincere thanks to the Connecticut Energy Efficiency Board and its Evaluation Subcommittee who funded the important series of studies on the State's HER program.

References

Arnold, H., Massachusetts Cross-Cutting Evaluation: Home Energy Report Savings Decay Analysis. Opinion Dynamics. September 2014.

Department of Energy 2014 (M. Li and A. Todd), “Evaluation, Measurement, and Verification of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendation, April 3.

DNV-GL, 2015, “Residential Energy Efficiency Special Projects 2014 Impact Evaluation of Home Energy Reports Program for Puget Sound Energy”, Bellevue WA, July 13.

DNV-GL, 2016, “PSE Home Energy Reports Program 2015 Impact Evaluation – Final Report for Puget Sound Energy”, Bellevue WA, October 26.

DNV-GL, 2018, “PSE Home Energy Reports Program 2017 Impact Evaluation – Final Report for Puget Sound Energy”, Bellevue WA, April 27.

Freeman, J., and L. A. Skumatz, 2012. “Widgets versus Actions: Measuring the Role of Behavior Change in DSM Programs”, Proceedings of the IEPEC Conference, Rome, June.

Gillingham, K., A. Keyes, and K. Palmer, 2018. “Advances in Evaluating Energy Efficiency Policies and Programs”, Working Paper Resource for the Future, Revised February.

Illume Advising, Indicia Consulting, and Dr. E. Vine, “Energy Efficiency Behavioral Programs: Literature Review, Benchmarking Analysis, and Evaluation Guidelines”, conducted for Minnesota Department of Commerce Division of energy Resources, May 4, 2015.

Integral Analytics with Building Metrics Incorporated and Sageview, 2012. ”Impact and Persistence Evaluation Report: Sacramento Municipal Utility District Home Energy Report Program”, Sacramento CA, Prepared by May Wu, November.

Khawaja, S. and J. Stewart, 2014a, “Long Run Savings and Cost-Effectiveness of Home Energy Report Programs”, The Cadmus Group, Winter 2014/2015.

Khawaja S. and J. Stewart, 2014b, “Long Run Savings and Cost-Effectiveness of Home Energy Report Programs”, ACEEE-SEE Webinar, December 3.

Mazur-Strammen, S. and K. Farley, 2013, “ACEEE Field Guide to Utility-Run Behavior Programs”, ACEEE Research Report B132, American Council for an Energy Efficient Economy, Washington DC, December 19.

NMR Group, Inc. 2016, “Evaluation of Persistence in the Eversource Customer Behavior Program (R32), Connecticut Energy Efficiency Board, March.

NMR Group, Inc. 2017, “1606 Eversource Behavior Program Persistence Evaluation”, Connecticut Energy Efficiency Board, October.

Skumatz, Lisa A., Ph.D. 2014. “What’s the Evidence on Behavioral Impacts, Effectiveness, and Retention”, Behavioral Energy and Climate Change (BECC) Conference, Washington DC, December.

Skumatz, Lisa A., Ph.D., 2016, “Persistence of behavioral programs: New information and implications for program optimization”, *The Electricity Journal,* 29 (2016) 27-32.

Skumatz, L. A. and J. L. Green. 2000, “Evaluating the Impacts of Education / Outreach Programs – Lessons on Impacts, Methods, and Optimal Education”, Proceedings of the ACEEE Summer Study on Buildings, Asilomar, CA, August.

Skumatz, L.A., Ph.D., M. S. Khawaja, and J. Colby, 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”, CIEE, Sacramento CA, November.

Smith, J., 2013, “Puget Sound Energy HER Program Persistence”, Puget Sound Energy, in Goodwin, McCrae, Babbitts, and Smith, “Behavioral-Based Energy Efficiency Programs”, presentation to Emerging Technologies Showcase, July 17.

Statewide Evaluation Team (2015) (GDS, Nexant, Research Into Action, Apex Analytics), “Residential Behavioral Program Persistence Study”, Prepared for Pennsylvania Public Utility Commission, Harrisburg PA, December 15.