Retention of Industrial Sector Energy Efficiency Incentive Program Measures

Karen Smith, XENERGY Consulting Inc., Madison, WI
Elsia Galawish, PG&E, San Francisco, CA
Amalia Klinger, PG&E, San Francisco, CA
Fred Coito, XENERGY Consulting Inc., Oakland, CA
Miriam L. Goldberg, XENERGY Consulting Inc., Madison, WI

ABSTRACT

The California DSM Measurement Advisory Committee (CADMAC) Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management Programs (Protocols) require that California investor-owned utilities conduct retention evaluations of targeted measures in commercial, industrial, and agricultural sector energy efficiency incentive (EEI) programs. This paper presents methodologies and results of the retention evaluation of targeted measures in Pacific Gas & Electric Company (PG&E) 1994 and 1995 paid year industrial EEI programs.

The purpose of this retention analysis was to determine the effective useful life (EUL) which has been defined as "the time at which fifty percent of the impacts associated with a group of cohorts for a measure ceased to be realized." The second earnings claim was based on this EUL, along with the first earnings claim values for program costs and recorded participation levels. Earnings claims after the second claim will be based on the first earnings claim values for program costs, participation levels, and effective useful lives; the load impacts from the second earnings claim; and the results of the completed persistence studies.

To estimate the EUL, on-site surveys were conducted at 100 sites with lighting impacts and 34 sites with process impacts from the 1995 program to determine which measures were still in place and operable (sites could have impacts from more than one type of lighting or process equipment). Data from these surveys were combined with previously collected (but not reported) data for 1994 paid year program sites, and used to estimate survival functions. The median EUL was then determined using statistical procedures.

Introduction

This paper describes a retention study of Pacific Gas and Electric Company's (PG&E's) 1994 and 1995 Industrial Energy Efficiency Incentives (IEEI) Programs. The California DSM Measurement Advisory Committee (CADMAC) Protocols specify the impact evaluation procedures for each investor-owned utility DSM program. The impacts from these programs are a key input to the utilities' earnings claims for operating these programs. To determine the lifetime impacts—and implicitly, the associated lifetime benefits--the life of the measure must be known. For this reasons, the Protocols require retention studies at designated intervals, the fourth and ninth years for these industrial programs.

A first earnings claim, prior to the completion of any evaluation studies, is based on *ex ante* estimates of first-year impacts, and *ex ante* estimates of measure life. A second earnings claim, after the first-year impact study is completed, is based on *ex post* estimates of first-year impacts, but still uses the *ex ante* measure life estimate. Third and subsequent earnings claims are based on *ex post* estimates of measure life.

The goals of the measure retention study are to determine (a) the length of time the measure(s) installed during the program year are maintained in operating condition; and (b) the extent to which there has been a significant reduction in the effectiveness of the measures.

The CADMAC Subcommittee on measure retention has agreed that the Protocols require that the first question (a) should be addressed by estimating each measure's Effective Useful Life (EUL). The EUL is defined by CADMAC as the median survival time, that is, the time until half the units are no longer in place and operable. Estimating the EUL is the primary focus of this study. The question of reduced measure effectiveness (b) has been addressed in a separate set of studies.

Each measure has an *ex ante* estimate of the EUL, which has been used in the first and second earnings claims. If the *ex post* EUL determined by the retention study for a particular measure is statistically significantly different from the *ex ante* EUL at the 20 percent significance (80 percent confidence) level, the *ex post* EUL will be used for future earnings claims. If there is not such a statistically significant difference, the *ex ante* EUL will be retained. Whether or not the EUL is revised as a result of this study, the EUL may be revised in the future based on subsequent retention studies required by the Protocols.

In this paper, lighting and process measures rebated to industrial facilities are addressed. The specific measures studied for each end use are indicated in the following table.

Table 1
Measures Studied

Measure Group	End Use
Fluorescent Delamping, Install Optical Reflector	Lighting
Replace Lamps and Ballasts, 4 ft Fixture	Lighting
Interior HID (High Intensity Discharge), 175 W	Lighting
Interior HID, 251-400W	Lighting
Process Controls	Process
Process Heat Recovery	Process
Process Change/Add Equipment	Process
Process Insulate	Process
Process Other	Process

Study Methods

Survival Analysis

The general survival function. The general method of study for each measure is to collect measure retention data from a sample of participants, and fit a parametric survival function to those data. The survival function is a function $S(t;\theta)$ that gives the probability S of surviving to any positive time t, given the parameters θ. These parameters are estimated from the retention data. Once the survival function parameters are estimated, median lifetime or EUL is determined as the time t^* such that the survival probability $S(t^*;\theta) = 0.5$.

The estimation and application of the survival function requires the specification of the function's parametric form. This form is typically specified in terms of the *hazard function* $h(t;\theta)$. Roughly, the hazard function can be thought of as the instantaneous probability of failing at time t, given that a unit has survived up to that time.

The survival probability $S(t;\theta)$ is one minus the probability $F(t;\theta)$ that a unit will die by time t. Formally, the hazard function is the ratio of the probability density function of the distribution $F(t,\theta)$ to the survival probability $S(t;\theta)$:

$$h(t;\theta) = (dF/dt)/S(t;\theta).$$

Choices of parametric forms for the survival function. Several parametric forms are in common use as hazard functions. Those explored in this study include the gamma, Weibull, exponential, log-normal, and log-logistic.

The gamma function is the most general of these, and includes the Weibull, exponential, and log-normal as special cases. In essence, the gamma function allows certain parameters to be determined by the data that are constrained by each of the other specifications. As a result, the gamma function will be able to follow the empirical data most closely. If one of the other forms is a good description of the data, its results will be similar to those of the less constrained gamma fit. If the other form is not a good match to the data, its results will be at odds with those of the gamma fit. This "goodness-of-fit" can be formally tested by the log-likelihood test.

Similarly, the Weibull also includes the exponential as a special case. The goodness of fit for the exponential form can be tested against the Weibull results, again using the log likelihood test.

The log-normal and log-logistic forms have decreasing hazard functions after an initial peak. That is, failure rates decline over time. This form may be a reasonable fit over a portion of time for certain types of equipment or processes. However, declining failure rates are unlikely to be an accurate representation of the failure pattern several years out.

The exponential form represents a constant hazard function. That is, the chance that a unit will fail in the next time increment, given that it has already survived to the current time, is the same no matter what the current time. This form is often used in survival analysis.

The Weibull form has an increasing hazard function. That is, the failure rate increases as equipment ages. In many respects, this basic assumption is the most reasonable of all the distributions explored.

As noted, the gamma form is the most general. Depending on the empirical data and the resulting parameters estimated, this form may produce an increasing, decreasing, or essentially constant hazard function.

Interpretation of survival model results. At this point in the life of the measures addressed in this study, the observed failure rates are generally low (under 4 percent for all of the lighting measures, and zero for three of the process measure groups). As a result, there is little solid empirical basis for choosing among possible forms of the hazard function. In some cases, it may be possible to match the empirical data reasonably well over the limited domain of the analysis (three to four years since program participation). However, in most cases the resulting estimated median lifetime will be substantially greater than this elapsed lifetime. That is, the EUL estimate entails extrapolating the data far beyond their original range. Such extrapolation is precarious in any modeling exercise. The exception would be if there were a very strong basis for knowing that the model form had been appropriately specified and that its parameters are consistent across the range from the data to the point of extrapolation.

In the present study, there is no such *a priori* basis for specifying the form, and no basis for assuming that the patterns evident so far are retained over extended periods. Consequently, in cases where the estimated EUL's are substantially greater than the four years of observed lifetimes, these estimates should be regarded as indicative, but not definitive. This issue is discussed further in the context of each measure group's analysis.

Weighting. In the survival analysis, each individual unit at each visited site is effectively treated as a separate observation. As a result, without weighting, the apparent sample size for the survival analysis is several thousand. This inflated apparent sample size distorts the calculated standard errors, making the estimates appear to be much more accurate than they are. By assigning the weights as described below, there was no need to adjust the resulting standard errors.

In reality, the analysis has only one observation on each technology for each site. To reflect the actual number of distinct observations for each technology type, the observations were weighted so that the sum of the weights was equal to the total number of sites in the sample with that technology.

For the process measures, individual rebated measures have widely varying sizes in terms of their contributions to the total program savings. To account for varying savings levels across rebated measures, an initial weight was assigned equal to the measure's avoided cost. For lighting, each rebated unit of each technology type was assigned the same initial weight. To reflect the actual number of distinct observations, the initial weights were then rescaled so that the sum of the weights was equal to the total number of sites in the sample for each process technology group. A separate set of final weights was calculated by this means for each technology type studied.

Data Required for the Survival Analysis

The retention data required for the survival analysis are data that indicate for each rebated unit at each sampled premise whether the unit was still in place and operable at the time of the survey. A unit not in place and operable is classified as a "failure" for purposes of this analysis. The unit may not have failed physically, but in terms of the program savings objectives has failed. Wherever possible, the retention data for failed units also include the date when the failure occurred.

In many cases, the failure is reported but the date when the failure occurred is not known. In this case, the observation is said to be left-censored. That is, the unit is known to have failed by a particular date, but the date of its failure is not known. In other cases, indeed the majority in this study, the unit had still not failed at the time the retention data were collected. In this case, the observation is said to be right-censored. The unit will fail at some future, as yet unknown time. The model forms used in this analysis accept both left- and right-censored data.

Data sources. Data sources used in this study were the onsite data collected for this purpose, and program tracking data. Some onsite data were previously collected, but not reported. This study collected data at 100 sites with lighting impacts and 34 sites with process impacts.

A census of 1995 process and lighting participants who were evaluated in the first year impact analysis was attempted. Refusals were dropped from the sample. For out-of-business sites, an attempt was made to find out if a new business had taken their place. These "replacement" businesses were then surveyed. If the site was closed, all measures were considered failures.

For this round of onsite data collection, at each sampled site the inspector determined the number of units currently in place and operable for each of the technology types rebated at that site. Wherever possible, the reason for any shortfall from the rebated number was obtained from a customer respondent. Also obtained if possible was the approximate date any missing equipment was removed or failed, by technology type. Units replaced under warranty were not considered failures.

Program tracking data were used to draw the samples and provide contact information used to recruit sites for the study. For those sites that were visited, the numbers of rebated units of each technology type were provided to the inspectors from the program tracking data. Install dates, when not provided by on-site personnel, were approximated from data in the tracking system.

Results

Summary of Retention Observations

Analysis units. Table 2 and 3 show the number of measures included in the lighting and process analysis respectively, by technology group.

Table 2
Data Included in Lighting Analysis by Technology Group

	Sites		Lamps	
	Number	Percent	Number	Percent
L19 (Delamping/Reflector)	66	28.9%	25,730	25.3%
L23 (Lamps and ballasts-4' Fixture)	86	37.7%	73,028	71.9%
L37 (Interior HID, 175W)	15	6.6%	502	0.5%
L81 (Interior HID, 251-400W)	61	26.8%	2,302	2.3%
Total	**228**	**100.0%**	**101,562**	**100.0%**

Table 3
Data Included in Process Analysis by Technology Group

Measure Description	Sites	Units
Process Controls	12	761
Process Heat Recovery	5	5
Process Change/Add Equipment	16	241
Process Insulate	3	568
Process Other	24	167

Units still in place. Table 4 shows the status at the time of inspection of the rebated lamps used in the analysis. Although lamps are the units of analysis, a measure was considered still installed if the fixture, not the lamp, was still in place. Two lines are shown for HID (High Intensity Discharge) lamps in the 251-400W size range (L81). The second case does not include one site where all the lamps were removed. As shown below, this one site has a strong influence on the estimated EUL.

Table 4
Status of Rebated Lamps

Measure	Lamps in Place	Lamps Removed	Total Lamps	Percent Removed
L19 (Delamping/Reflector)	24,783	947	25,730	3.7%
L23 (Lamps and ballasts-4' Fixture)	71,453	1,575	73,028	2.2%
L37 (Interior HID, 175W)	491	11	502	2.2%
L81 (Interior HID, 251-400W)	2,255	47	2,302	2.0%
L81 (Interior HID, 251-400W)[1]	2,221	13	2,234	0.6%
Total	**98,982**	**2,580**	**101,562**	**2.5%**

1. L81 with the extreme site removed. This line is not included in the total.

Table 5 shows the status at the time of inspection of the rebated process equipment used in the analysis.

Table 5
Status of Rebated Process Equipment

Measure Description	Units in Place	Units Removed	Total Units	Percent Removed
Process Controls	533	228	761	30.0%
Process Heat Recovery	5	0	5	0.0%
Process Change/Add Equipment	224	17	241	7.1%
Process Insulate	568	0	568	0.0%
Process Other	167	0	167	0.0%
Total	**1497**	**245**	**1742**	**14.1%**

Survival Analysis Results

Lighting measures. Table 6 presents the estimated median lifetime or EUL in years for the studied lighting measures, and the corresponding standard error, for various distributional assumptions. Missing values indicate that the model did not converge (there were not enough removal data to estimate an effective useful life under this model assumption). Results are presented for large HID's (L81) with and without the influential site.

Table 6
Estimated EUL's and Standard Errors for Various Hazard Functions
(years)

	L19 (Delamping/Reflector)		L23 (Lamps and ballasts-4' Fixture)		L37 (Interior HID, 175W)		L81 (Interior HID, 251-400W)		L81 (Interior HID, 251-400W) - Outlier Removed	
ex ante EUL	16		16		20		16		16	
	EUL (Years)	Standard Error	EUL (Years)	Standard Error	EUL (Years)	Standard Error	EUL (Years)	Standard Error	EUL (Years)	Standard Error
Weibull	12.1	14.9	27.1	52.5			7.3	2.5	13.5	36.3
Gamma			22.7	416.7					10.9	233.9
Exponential	168.3	166.4	166.1	145.0	116.3	202.9	119.5	107.7	425.2	725.0
Log-normal	18.8	31.8	195.9	668.3			7.5	3.2	19.5	71.3
Log-logistc	13.3	17.8	32.5	68.6			7.4	2.8	14.6	42.1

Table 7 shows the corresponding 80 percent confidence intervals. Also indicated in the table are the estimates that are statistically significantly different from the *ex ante* EUL at this confidence level. Formally, the CADMAC Protocols indicate that the *ex ante* EUL's should be replaced by the *ex post* results in these cases. However, the range of results across the different hazard function forms, and the conceptual appropriateness of these forms, suggest that such replacement would be premature. This issue is discussed further below.

Table 7
Estimated EUL's and Confidence Intervals for Various Hazard Functions
(years)

	L19 (Delamping/Reflector)		L23 (Lamps and ballasts-4' Fixture)		L37 (Interior HID, 175W)		L81 (Interior HID, 251-400W)		L81 (Interior HID, 251-400W)-Extreme Outlier Removed	
	EUL	80% Confidence Interval	EUL	80% Confidence Interval	EUL	80% Confidence Interval	EUL	80% Confidence Interval	EUL	80% Confidence Interval
Weibull	12.1	(0.0 , 31.3)	27.1	(0.0 , 94.4)			7.3	(4.1 , 10.5) *	13.5	(0 , 60.0)
Gamma			22.7	(0.0 , 556.8)					10.9	(0 , 310.8)
Exponential	168.3	(0.0 , 381.6)	166.1	(0.0 , 352.1)	116.3	(0.0 , 376.4)	119.5	(0.0 , 257.6)	425.2	(0 , 1,354.7)
Log-normal	18.8	(0.0 , 59.5)	195.9	(0.0 , 1052.7)			7.5	(3.4 , 11.5) *	19.5	(0 , 110.9)
Log-logistic	13.3	(0.0 , 36.1)	32.5	(0.0 , 120.4)			7.4	(3.8 , 11.0) *	14.6	(0 , 68.6)

Table 7 shows that none of the *ex ante* EUL's would be rejected for any of the lighting technology groups with any of the hazard functions explored. The exception is the large HID lamps, whose results are driven down by a single site, as noted. Thus, there is insufficient information at this point in the life of the measures to re-estimate the median lifetime, at least with the sample sizes used for this study. This finding is not surprising, given that the overall failure rates were under four percent for all the technology types as of three to four years since installation.

For delampling (L19), the exponential distribution is rejected by the log likelihood test. The remaining distributions, Weibull, log-normal, and log-logistic, give median lifetimes ranging from 12 to 19 years, consistent with the *ex ante* value of 16 years.

For four-foot lamp-ballast-fixture combinations (L23), the estimated EUL's vary widely across the distributions explored. None of these distributions can be rejected. The conceptually most appropriate distribution, the Weibull, gives an EUL estimate of 22 years with a very large standard error.

For small HID's, only the exponential form converged. This form is conceptually inappropriate, and also gave EUL estimates substantially larger than the more appropriate Weibull distribution for the other lamp types studied.

For large HID's with all observations included, the exponential form is rejected against the Weibull. This distribution as well as the log-normal and log-logistic distributions give EUL's around 7 years. This result is suspect despite its consistency across these distributions. Only 2 percent of the units studied failed within three to four years from installation. Calculating that half will be gone in another three to four years means that a steeply sloping hazard function is being extrapolated far beyond the range of the data.

Inclusion of the outlier case for large HID's is arguably appropriate: dissatisfaction leading to comprehensive measure removal is one of the factors that the measure retention study must reflect. However, the dramatic difference in the estimated EUL when this one site out of 61 is excluded indicates that a broader base of observation is required before the EUL can be estimated accurately for this measure.

Based on these findings, none of the estimated EUL's from this study is recommended as a basis for revising the *ex ante* values. For purposes of the next filed incentive claim, the *ex ante* values will be retained for all lighting measures.

Process. Table 8 presents the estimated median lifetime or EUL in years, and the corresponding standard error for various distributional assumptions for process controls and process change/add equipment. Survival analysis could not be performed at this time on the other three measures, because no units had failed within the study period. Missing values in Table 8 indicate that the model did not converge.

Table 8
Estimated EUL's and Standard Errors for Various Hazard Functions
(years)

	550-Process Controls		569-Process Change/Add Equipment	
Distribution	**EUL**	**SE**	**EUL**	**SE**
Weibull	6.8	8.4	67.2	497.8
Gamma				
Exponential	37.0	43.5	13.7	8.6
Log-normal	9.0	13.9	124.6	1115.8
Log-logistic	7.4	10.1	91.9	752.0

Table 9 shows the corresponding 80 percent confidence intervals. For all the distributions explored for both measures, the 80 percent confidence intervals include the *ex ante* EUL. That is, the estimated EUL is not significantly different from the *ex ante* value at the 80 percent confidence (20 percent significance) level under any distributional assumption. Thus, all *ex ante* EUL's are retained.

Table 9
Estimated EUL's and Confidence Intervals for Various Hazard Functions
(years)

	550-Process Controls		569-Process Change/Add Equipment	
Ex Ante EUL	**12.1**		**18.9**	
Distribution	**EUL**	**80% Confidence Interval**	**EUL**	**80% Confidence Interval**
Weibull	6.8	(0.0 , 18.4)	67.2	(0.0 , 750.1)
Gamma				
Exponential	37.0	(0.0 , 96.3)	13.7	(1.9 , 25.4)
Log Normal	9.0	(0.0 , 28.1)	124.6	(0.0 , 1655.4)
Log Logistic	7.4	(0.0 , 21.3)	91.9	(0.0 , 1123.6)

Summary of Results

The results of this study are summarized in Table 10. The table shows the estimates for the most appropriate distribution for which results were obtained. Conceptually, as discussed above, the Weibull or gamma distributions are most appropriate. However, the gamma distribution failed to converge for three of the four lighting measures and the Weibull distribution failed to converge for interior HID, 175W. That is, the available data were insufficient to allow an estimate to be developed with these forms for these measures. Therefore, for 175W interior HIDs, the results for the exponential distribution are shown.

For three of the process measures, there were no observed failures. As a result, no model could be estimated and no *ex post* EUL is available.

For the other two process measures, and for three of the lighting measures, the *ex ante* EUL was not significantly different from the *ex post* EUL at the 80 percent significance level for any of the hazard distributions. Therefore the *ex ante* EUL is retained for these measures.

Table 10
Summary of EUL Findings
(years)

						80% Confidence Interval		
End Use	**Measure Group**	**Measure Codes**	***ex ante* *EUL***	**Distribution for *ex post* EUL**	***ex post* EUL**	**Lower**	**Upper**	**EUL for claim**
Lighting	Delamping/Reflector	L19	16.0	Weibull	12.1	0.0	31.3	16.0
	Lamps and ballasts-4' Fixture	L23	16.0	Weibull	27.1	0.0	94.4	16.0
	Interior HID, 175W	L37	20.0	Exponential	116.3	0.0	376.4	20.0
	Interior HID, 251-400W	L81	16.0	Weibull	7.3	4.1	10.5	16.0
Process	Process Controls	550	12.1	Weibull	6.8	0.0	18.4	12.1
	Process Heat Recovery	560	28.9	n/a	n/a	n/a	n/a	28.9
	Process Change/Add Equipment	569	18.9	Weibull	67.2	0.0	750.1	18.9
	Process Insulate	590	10.1	n/a	n/a	n/a	n/a	10.1
	Process Other	599	17.0	n/a	n/a	n/a	n/a	17.0

The *ex ante* useful life of 16 years (20 years for L37) does not fall within the 80 percent confidence interval. Formally, the *ex ante* EUL would be rejected in favor of the new estimate.

An *ex post* EUL significantly different from the *ex ante* value was found only for large HID lamps (L81) with the extreme site included. However, the projection of the fitted model far beyond the available data implied by this result, together with the very different results obtained when the single extreme site is removed make this result questionable. Thus, the *ex ante* value is retained for this measure as well.

Measure life is a key component to the determination of life-cycle benefits of efficiency measures, and their associated economic value. Determining measure life by direct observation can require a long wait for long-lived measures. In the meantime, decisions regarding new measure installations must be made, and payments for measures installed must be awarded.

Survival analysis is the appropriate tool to allow estimation of failure rates over time from early failure data. However, application of this tool requires distributional assumptions. In this study, the elapsed time was too short relative to the lives of the measures under study to provide a solid basis for choosing among alternative assumptions. In many cases, alternative assumptions for the distributional form result is very different estimates of the median time to failure. A broader base of experience—over time and over greater numbers of units—may help, without the need to wait 20 years to confirm or refute the *ex ante* estimates.

The next retention study for this program is required by the Protocols in another five years. With additional data covering a longer span of time, there will be a better possibility of identifying appropriate distribution forms and parameters, and obtaining more reliable *ex post* estimates of median measure life. However, for measures with very few participants, such as the process measures, the sample sizes may remain too small for accurate estimation even with a longer observation period for the survival analysis.