RETENTION STUDY OF PACIFIC GAS AND ELECTRIC COMPANY'S 1994 AND 1995 APPLIANCE ENERGY EFFICIENCY PROGRAMS

1994 -1995 Residential Lighting Sixth-Year Retention
Study ID 384bR2 (1994), 401bR2 (1995)

March 1, 2001

Measurement and Evaluation
Customer Energy Efficiency Policy & Evaluation Section
Pacific Gas and Electric Company
San Francisco, California

All inquiries should be directed to:

Janice Frazier-Hampton
Revenue Requirements
Pacific Gas and Electric Company
P. O. Box 770000, Mail Code B9A
San Francisco, CA 94177

Retention Study of Pacific Gas and Electric Company's 1994 and 1995 Appliance Energy Efficiency Programs

1994–1995 Residential Lighting Sixth-Year Retention Study ID 384bR2 (1994), 401bR2 (1995)

Purpose of Study

This study was conducted in compliance with the requirements specified in "Protocols and Procedures for the Verification of Costs, Benefits, and Shareholders Earnings from Demand-Side Management Programs," as adopted by California Public Utilities Commission Decision 93-05-063, revised June 1999, pursuant to Decisions 94-05-063, 94-10-059, 94-12-021, 95-12-054, 96-12-079, 98-03-063, and 99-06-052.

This study measures the Effective Useful Life (EUL) of lighting measures for which rebates were paid through Pacific Gas and Electric Company's (PG&E) 1994 and 1995 Appliance Energy Efficiency Programs. The EUL is the time at which half the units rebated and installed during the program year are no longer in place and operable.

Methodology

For each measure, this study assumes the time to non-retention follows some parametric distribution. Therefore, the general method of study is to collect retention data from participants and use those data to estimate the parameters of this distribution. The estimated parameters of the distribution for the time to non-retention are then used to estimate the median retention time or EUL.

The data necessary for this study were obtained from the Program tracking data and collected via on-site inspections. The on-site inspection data were collected at two points in time, during 1998 and 2000. A total of 301 sites provide the data for this retention analysis.

The parameters of the distribution of the time to non-retention are estimated by fitting a general linear regression model to the logarithmic transformation of the time to non-retention reported in the data. To estimate the EUL, the estimated parameters are then employed in the survival function. This function is simply one minus the cumulative distribution function for the time to non-retention. The survival function gives the probability of retaining a unit of a measure until at least time t. Therefore, the estimate of the EUL is the time t^* such that the survival probability equals 50 percent.

Study Results

The results of this study are summarized in the table below. In the cases of lighting measures CFL and HID, the adopted *ex post* EUL is larger than the *ex ante* EUL because the *ex ante* EUL is outside the 80 percent confidence interval and smaller than the estimated EUL. The adopted *ex post* EUL is 16 years for these measures, consistent with the use of 16–year EULs for lighting measures in statewide, year 2001 programs. In the cases of both lighting measures CFL and HID, the adopted ex post EUL of 16 years is smaller than the estimated EUL of 78.5 and 36.9 years, respectively, which is likely an overestimate. For lighting measure CFL the realization rate equals 1.60, and for lighting measure HID the realization rate equals 1.07.

1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting
Summary of Effective Useful Life Estimates

Program Year	Measure	End Use	EUL (years) *ex ante*	*ex post* (estimated from study)	Adopted *ex post* (to be used in claim)	*ex post* Standard Error	80% Confidence Interval Lower Bound	80% Confidence Interval Upper Bound	p-value for *ex post* EUL	EUL Realization Rate (adopted *ex post* / *ex ante*)
1994-1995	**CFL**	Lighting	10.0	78.5	16.0	66.7	26.4	233.4	0.02	1.60
1994-1995	**HID**	Lighting	15.0	36.9	16.0	20.7	17.9	76.0	0.11	1.07
1994-1995	**T-8**	Lighting	16.0	35.5	16.0	24.0	14.9	84.6	0.24	1.00

For lighting measure T-8, the *ex ante* EUL of 16 years is inside the 80 percent confidence interval. Therefore, the adopted *ex post* EUL equals the *ex ante* EUL and the realization rate equals 1.00. Again, this EUL is consistent with the 16-year EUL for lighting measures being used for statewide, year 2001 programs.

Regulatory Waivers and Filing Variances

None.

RETENTION STUDY OF PACIFIC GAS AND ELECTRIC COMPANY'S 1994 AND 1995 APPLIANCE ENERGY EFFICIENCY PROGRAMS

FINAL REPORT

1994–1995 RESIDENTIAL LIGHTING
SIXTH-YEAR RETENTION
STUDY ID 384bR2 (1994), 401bR2 (1995)

Prepared for

Pacific Gas and Electric Company
San Francisco, California

Prepared by

XENERGY Consulting Inc.
Oakland, California
Madison, Wisconsin

March 1, 2001

LIST OF TABLES

LIST OF FIGURES

E EXECUTIVE SUMMARY

This report provides the results of the sixth-year retention study of Pacific Gas and Electric Company's (PG&E) 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting, as required by the Measurement and Evaluation (M&E) Protocols of the California DSM Measurement Advisory Committee (CADMAC).

As given in the M&E Protocols, the goal of a measure retention study is to determine "the length of time the measure(s) installed during the program year are maintained in operable condition." This issue is addressed by estimating each measure's Effective Useful Life (EUL). The EUL is defined as the time at which half the units installed during the program year are no longer in place and operable.

Each measure has an *ex ante* estimate of the EUL, which has been used in the earnings claims to date. A measure's *ex post* EUL is the EUL estimated by a retention study. If a measure's *ex ante* EUL is outside the 80 percent confidence interval for the measure's EUL determined by a retention study, the *ex post* EUL will be used for future earnings claims. Otherwise, the *ex post* EUL will not replace the *ex ante* EUL.

E.1 DATA

The measures included in this study are the key lighting measures installed in multi-family common areas. The sites identified to provide the retention data for these measures are the sites included in the third-year retention study of PG&E's 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting. The data necessary for this study were obtained from the Program tracking data and collected via on-site inspections. A total of 301 sites provide the data for this retention analysis.

E.2 STUDY METHODS

E.2.1 Estimating the EUL

For each measure, this study assumes the time to non-retention follows some parametric distribution. Therefore, the general method of study is to collect retention data from participants and use those data to estimate the parameters of this distribution. The estimated parameters of the distribution for the time to non-retention are then used to estimate the median retention time or EUL.

The parameters of the distribution of the time to non-retention are estimated by fitting a general linear regression model to the logarithmic transformation of the time to non-retention reported in the data. To estimate the EUL, the estimated parameters are then employed in the survival function. This function is simply one minus the cumulative distribution function for the time to non-retention. The survival function gives the probability of retaining a unit of a measure until at least time *t*. Therefore, the estimate of the EUL is the time t^* such that the survival probability equals 50 percent.

E.2.2 Confidence Interval for the EUL

The lower and upper bounds of a confidence interval for the EUL are calculated as the exponential of the lower and upper bound values of the confidence interval for the log of the EUL, respectively. In general, the bounds of a confidence interval for a parameter are calculated as the parameter estimate ± the standard error of the parameter estimate times the critical value from the appropriate distribution for the desired level of confidence. The standard error of the log of the EUL estimate employed in the calculation of the confidence interval for the log of the EUL is provided by SAS®, the statistical analysis software used for the analysis. This standard error is a function of the standard errors of the parameter estimates of the general linear regression model. If necessary, the standard error of the log of the EUL estimate provided by SAS is adjusted by the square root of the design effect factor.

Adjustment to the Standard Error

When fitting a general linear regression model to the data for a given measure, an observation is the unit of a measure. The calculation of the standard errors of the parameter estimates assumes each observation is independent. This assumption, however, may be incorrect; because an observation is a unit of a measure being analyzed, which is not the level at which the first and only stage of sampling occurred. The first and only stage of sampling occurred at the site level. Therefore, non-retention of units may be more similar within a site than between sites.

If non-retention of units is more similar within a site than between sites and if the data analyzed for a measure are based on only a sample of sites that obtained a rebate for the measure, it is necessary to adjust the standard error of an estimate for the sample design by the square root of the design effect factor (Kish 1965). If non-retention of units is no more similar within a site than between sites, then the square root of the design effect factor equals one and the unadjusted and adjusted standard error are equal. Usually, however, the design effect factor is greater than one.

E.3 Summary of Results

The results of this study are summarized in Table E-1. In the cases of all measures, the estimated or *ex post* EUL is obtained assuming a Weibull distribution for the time to non-retention of a unit of the measure. Although other distribution assumptions are possible, the Weibull distribution is considered the most appropriate because it is the only distribution for which an estimate of the EUL was obtained known to be consistent with the oldest units of a measure having the highest non-retention rate. Furthermore, there is little to no evidence to justify adopting one of the other distributions over the Weibull distribution.

Table E-1
1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting
Summary of Effective Useful Life Estimates

Program Year	Measure	End Use	EUL (years) *ex ante*	*ex post* (estimated from study)	Adopted *ex post* (to be used in claim)	*ex post* Standard Error	80% Confidence Interval Lower Bound	80% Confidence Interval Upper Bound	p-value for *ex post* EUL	EUL Realization Rate (adopted *ex post* / *ex ante*)
1994-1995	**CFL**	Lighting	10.0	78.5	16.0	66.7	26.4	233.4	0.02	1.60
1994-1995	**HID**	Lighting	15.0	36.9	16.0	20.7	17.9	76.0	0.11	1.07
1994-1995	**T-8**	Lighting	16.0	35.5	16.0	24.0	14.9	84.6	0.24	1.00

In the cases of lighting measures CFL and HID, the adopted *ex post* EUL is larger than the *ex ante* EUL because the *ex ante* EUL is outside the 80 percent confidence interval and smaller than the estimated EUL. The adopted *ex post* EUL is 16 years for these measures, consistent with the use of 16–year EULs for lighting measures in statewide, year 2001 programs[1]. In the cases of both lighting measures CFL and HID, the adopted ex post EUL of 16 years is smaller than the estimated EUL of 78.5 and 36.9 years, respectively, which is likely an overestimate. For lighting measure CFL the realization rate equals 1.60, and for lighting measure HID the realization rate equals 1.07.

For lighting measure T-8, the *ex ante* EUL of 16 years is inside the 80 percent confidence interval. Therefore, the adopted *ex post* EUL equals the *ex ante* EUL and the realization rate equals 1.00. Again, this EUL is consistent with the 16-year EUL for lighting measures being used for statewide, year 2001 programs.

[1] *California Measurement Advisory Committee Public Workshops on PY2001 Energy Efficiency Programs*, filed on September 25, 2000 by Sempra Energy, p. 56.

1 INTRODUCTION

1.1 OVERVIEW

This report provides the results of the sixth-year retention study of Pacific Gas and Electric Company's (PG&E) 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting, as required by the Measurement and Evaluation (M&E) Protocols of the California DSM Measurement Advisory Committee (CADMAC)[1]. In this section, the protocol requirements are discussed, followed by a summary of the organization of the report.

1.2 PROTOCOL REQUIREMENTS

As given in the M&E Protocols, the goal of a measure retention study is to determine "the length of time the measure(s) installed during the program year are maintained in operable condition." As agreed within the CADMAC Persistence Subcommittee, this issue is addressed by estimating each measure's Effective Useful Life (EUL). The EUL is defined as the median retention time, that is, the time at which half the units installed during the program year are no longer in place and operable. We refer to "no longer in place and operable" as "non-retention."

The definition of EUL must be considered in the context in which the EUL is used. PG&E uses the EUL in the calculation of the net resource benefit of the Appliance Energy Efficiency Programs, Residential Lighting for a given year. Specifically:

Program net resource benefit =

(Program's First Year Impact) × *(Program Level EUL)* × *(Program Level TDF),*

where

TDF = Technical Degradation Factor.

Together, the program level EUL and TDF address the persistence of energy savings in the calculation of a program's net resource benefit. EUL addresses retention and TDF addresses operational effectiveness.

Each measure has an *ex ante* estimate of the EUL, which has been used in the earnings claims to date. A measure's *ex post* EUL is the EUL estimated by a retention study. If a measure's *ex ante* EUL is outside the 80 percent confidence interval for the measure's EUL determined by a retention study, the *ex post* EUL will be used for future earnings claims. Otherwise, the *ex post* EUL will not replace the *ex ante* EUL. Whether or not a measure's EUL is revised as a result of

[1] California Public Utilities Commission, *Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management Programs*, Decision 93-05-063. Revised June 1999, pursuant to Decisions 94-05-063, 94-10-059, 94-12-021, 95-12-054, 96-12-079, 98-03-063, and 99-06-052.

this study, the EUL may be revised in the future based on subsequent retention studies required by the Protocols.

1.3 REPORT ORGANIZATION

The next section of this report, Section 2, describes the data employed in the study. Section 3 discusses the methods employed to estimate a measure's EUL and the standard error of the estimate. The calculation of both the confidence interval for the EUL and the p-value reported are also discussed in Section 3. Section 4 presents the retention results. Appendix A contains the on-site data collection instrument. Appendices B and C provide Tables 6B and 7B, respectively, required by the M&E Protocols.

2 DATA

This section of the report describes the data used in the retention analysis of PG&E's 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting. A discussion of both the measures and sites included in this study is presented. These discussions are followed by a description of the sources of the data employed in the analysis. The section concludes with the details of preparing the data for analysis.

2.1 MEASURES INCLUDED IN THE STUDY

The measures included in this study are:

- compact fluorescent fixtures (CFL),
- high intensity discharge fixtures (HID), and
- T-8 fixtures and/or ballasts (T-8).

As shown in Table 2-1, the lighting measures CFL, HID, and T-8 account for 89 percent of the total resource value of PG&E's 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting. They are the key lighting measures installed in multi-family common areas.

Table 2-1
Measures Included in the Study

Measure	Percent of Total Resource Value	*Ex Ante* EUL (years)
CFL	61%	10
HID	9%	15
T-8	19%	16
Total	89%	

2.2 SITES INCLUDED IN THE STUDY

A site (identified by PG&E control number) may have received a rebate for more than one measure. Therefore, a site may be included in the analysis of the retention of more than one measure.

The sites included in this study are the same sites visited for the third-year retention study of PG&E's 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting. In the case of the third-year retention study, an on-site inspection was conducted for a sample of 300 sites. In addition, one more site was visited, but an on-site inspection was unable to be conducted.

2.2.1 Sample Design

The sample of sites is stratified by program year, measure, number of units of the measure for which a rebate was obtained, and geographic area. For each program year, the sample of sites is allocated across sample cells to produce the best precision for a given total cost. Table 2-2 gives the population, sample frame, and sample counts by stratum. As compared to the population, the sample frame consists only of those sites for which contact information is available. The sample counts reflect the sites visited for the third-year retention study. For the population, sample frame, and sample, the table provides two counts: the number of sites and the number of projects (a unique site and rebate application combination).

Table 2-2
Population, Sample Frame, and Sample by Stratum

Segment Descriptions			Population		Sample Frame		Sample	
Year	Tech - Size	Area	Sites	Projects	Sites	Projects	Sites	Projects
1994	T8	1	80	84	52	57	26	31
1994	T8	2	64	68	58	64	18	21
1994	T8	3	8	8	8	8	2	2
1994	T8	4	30	32	30	32	4	4
1994	Other - Large	1	124	130	81	85	20	23
1994	Other - Large	2	114	117	108	111	14	15
1994	Other - Large	3	111	112	102	103	11	12
1994	Other - Large	4	60	63	58	61	4	6
1994	Other - Small	1	123	130	65	72	20	24
1994	Other - Small	2	118	124	114	120	15	15
1994	Other - Small	3	116	123	109	115	12	13
1994	Other - Small	4	46	48	46	48	4	4
1994 Subtotal			994	1,039	831	876	150	170
1995	T8	1	48	51	43	46	29	32
1995	T8	2	23	27	24	26	12	12
1995	T8	3	15	15	14	14	6	6
1995	T8	4	19	20	19	20	3	3
1995	Other - Large	1	49	54	41	45	21	21
1995	Other - Large	2	39	41	39	41	12	13
1995	Other - Large	3	59	55	47	48	14	14
1995	Other - Large	4	22	29	20	27	4	4
1995	Other - Small	1	55	59	46	50	18	20
1995	Other - Small	2	56	62	55	61	16	18
1995	Other - Small	3	46	52	40	46	12	13
1995	Other - Small	4	24	26	23	25	4	4
1995 Subtotal			455	491	411	449	151	160
Overall Total			1,449	1,530	1,242	1,325	301	330

The sample of 301 sites is split approximately evenly between the two program years. (As noted earlier, in the case of the third-year retention study, an on-site inspection was conducted for a sample of 300 sites. One more site was visited, but an on-site inspection was unable to be conducted. This site remained in the sample; hence, the sample consists of 301 sites.) The Protocols do not contain a sample size requirement for retention studies. Therefore, as a guideline, the sample is approximately double that required by the Protocols for first-year impact studies (150). Details regarding the sample stratification follow.

Sample Stratification

For sampling purposes, each site was classified as one of three measure and "size" combinations: T8, Other-Large, or Other-Small. This was done to ensure sufficient data were collected to estimate the EUL of each of the key lighting measures. A site was classified as T8 if it received a rebate for at least one T-8. Of the key lighting measures, the fewest rebates were obtained for T-8s. The remaining sites were classified as Other-Large or Other-Small depending on the number of units of CFLs and HIDs for which a rebate was received, where a unit is a lamp. A site that received a rebate for at least 25 CFL or five HID lamps (and zero T-8 lamps) was classified as Other-Large. Otherwise, a site that did not receive a rebate for a T-8 was classified as Other-Small.

To control on-site inspection costs, a site was also classified as one of four geographic areas of varying distances from PG&E's main population center. PG&E's service territory is divided into Divisions. The Divisions were combined to create four sampling geographic areas as follows:

- Area 1: East Bay (EBA), San Francisco (SFO), Mission (MIS), Peninsula (PEN);
- Area 2: Diablo (DIA), De Anza (DEA), San Jose (SJO), North Bay (NBY);
- Area 3: Central Coast (CCO), Fresno (FRE), Kern (KER), Stockton (STO), Sierra (SRA), Sacramento (SAC), North Valley (NVY); and
- Area 4: Los Padres (LOS), North Coast (NCO), Yosemite (YOS).

2.2.2 Sample Disposition

All 301 sites visited for the third-year retention study were visited for this study. An on-site inspection was conducted and all necessary data were collected for 297 of these 301 sites. In the cases of the four remaining sites, the on-site inspection yielded incomplete data for one measure. The data for a measure are incomplete if the inspector could not locate all of the units of the measure installed and indicated some uncertainty about her/his ability to access all areas of the site.

2.3 DATA SOURCES

The data used in this study were obtained from two sources:

1. the PG&E Appliance Energy Efficiency Programs, Residential Lighting tracking data for 1994 and 1995, and
2. on-site inspections conducted for the third-year and for this current retention study.

2.3.1 Program Tracking Data

For each site in the sample, the Program tracking data provides the following information:

- contact information,
- sample stratum; and

for each measure for which a rebate was obtained,

  - the number of units of the measure for which a rebate was obtained (number of expected units for the third-year retention study),
  - the date installed, and
  - the *ex ante* EUL.

In the cases of all measures, a unit is a lamp.

2.3.2 On-Site Inspections

In the case of the third-year retention study, an on-site inspection was conducted for a sample of 300 sites. In the case of the current retention study, an attempt was made to conduct an on-site inspection at all the sites visited for the third-year retention study. For each measure for which a rebate was obtained, the on-site inspection provides the following data:

- of the number of expected units of the measure, the number of units observed to be in place and the percentage of these units that are working;
- if known, in the case of each non-retained unit, the number of months prior to the date of the inspection the unit was not retained; and
- the date of the inspection.

A unit not in place and/or not operable at the time of the inspection is classified as not retained for purposes of this analysis. Therefore, a unit is classified as not retained if it is removed and/or if it fails. When the inspector was able to determine the reason a unit was not retained, this information was recorded as well. A copy of the on-site data collection instrument is provided in Appendix A.

If an on-site inspection was conducted at a site for the third-year retention study, the number of units of a measure retained based on this inspection is the number of expected units of the measure for the on-site inspection attempted for the current retention study. The on-site inspections conducted for the third-year retention study also provides updated contact information.

2.4 DATA PREPARATION

- In order to combine the Program tracking data and the on-site inspection data, if a site measure combination is not already a unique observation in the Program tracking data, it is made a unique observation. In the case of the on-site inspection data, if the number of expected units of a measure is specified separately for various locations, for example, the data may be entered by location.

- At a given site, more units of a measure may be observed to be in place than are expected. Furthermore, it may be difficult to determine which of the units of the measure observed to be in place correspond to the units of the measure for which a rebate was obtained through PG&E's 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting.

  Therefore, if the number of units of a measure observed to be in place exceeds the number of expected units, the number of observed units is reset to the number of expected units. The number of retained units is then calculated as the revised number of observed units times the original percentage of observed units that are working. And the number of non-retained units is the number of expected less the number of retained units.

- The methods employed in this study (discussed in the next section), namely, the fitting of the general linear regression model, allow inexact measures of the time to non-retention. This is done by specifying both a lower and upper bound for the time to non-retention. The time to non-retention may be inexact for a unit of a measure not retained and it is clearly inexact for a unit of a measure still retained.

  For all units of a measure, the installation date and inspection dates are known. Ideally, in the case of a non-retained unit of a measure, the time to non-retention is calculated exactly using these data and the number of months prior to the relevant inspection the unit was not retained. If it is possible to calculate the exact time to non-retention, the lower and upper bound for the time to non-retention are both equal to this exact time.

  Often, however, the number of months prior to an inspection a unit of a measure was not retained is unknown. If it is only known a unit was not retained some time before an inspection date, the lower bound of the time to non-retention is zero and the upper bound equals the number of years between the installation date and the relevant inspection date. Such observations are said to be left-censored. For SAS®, the statistical analysis software used for the analysis, to recognize an observation is left-censored, the lower bound is set equal to missing.

  If it is also known a unit was retained after some date other than the installation date, the lower bound of the time to non-retention equals the number of years between the installation date and this date and the upper bound equals the number of years between the installation date and the relevant inspection date. Such observations are said to be interval-censored. For example, a unit retained at the time of the third-year inspection but not retained some unknown time before the sixth-year inspection date is interval censored. Left-censoring is a special case of interval-censoring, where the lower bound of the time to non-retention equals zero.

The time to non-retention for a unit of a measure still retained at the time of the latest inspection is inexact. It is somewhere between the number of years between the installation date and the latest inspection date, and infinity. At some time, all units will not be retained. Therefore, in the case of a unit still retained, the lower bound of the time to non-retention equals the number of years between the installation date and the latest inspection date and the upper bound is infinity. Units still retained, then, are said to be right-censored. For SAS to recognize an observation is right-censored, the upper bound is set equal to missing.

3 STUDY METHODS

To analyze retention, this study employs a method commonly referred to as Survival Analysis. The method was first given this name because it was initially used to analyze death rates. The same set of techniques referred to as Survival Analysis is also referred to by several other names depending on the area of application. For example, in Engineering, "Survival Analysis" is Reliability Analysis and in Economics, it is Duration Analysis. In addition, the terminology employed in the analysis may vary depending on the area of application. In this report, we will use the Survival Analysis terminology, but will modify it when appropriate for the application of Survival Analysis to retention.

3.1 SURVIVAL ANALYSIS

3.1.1 The Basics

For each measure, this study assumes the time to non-retention follows some parametric distribution. Therefore, the general method of study is to collect retention data from participants and use those data to estimate the parameters of this distribution. The estimated parameters of the distribution for the time to non-retention are then used to estimate the median retention time or EUL.

The parameters of the distribution of the time to non-retention are estimated by fitting a general linear regression model to the logarithmic transformation of the time to non-retention reported in the data. This model can be written as

$$\ln(T_j) = \mu + \sigma\varepsilon_j,$$

where

T_j = measured time to non-retention,

μ = location parameter or intercept,

σ = scale parameter, and

ε_j = random error term.

The exponential of the error term of this model (e^{ε_j}) is assumed to follow the standardized form of the distribution of the time to non-retention. The general linear regression model is fitted by maximizing the log-likelihood function for the assumed distribution.

To estimate the EUL, the estimated parameters of the distribution of the time to non-retention are then employed in the survival function. This function is simply one minus the cumulative distribution function for the time to non-retention. The survival function $S(t;\theta)$ gives the probability of retaining a unit of a measure until at least time t, given the parameter vector θ.

Therefore, the estimate of the EUL is the time t^* such that the survival probability $S(t^*;\hat{\theta}) = 0.50$, where $\hat{\theta}$ is the vector of parameter estimates.

3.1.2 Distribution Options

This study considers the most common distributional assumptions made when conducting Survival Analysis:

- Gamma,
- Weibull,
- Exponential,
- Log-normal, and
- Log-logistic.

The Gamma distribution is the most general of the distributions listed above. It has three free parameters, location (μ), scale (σ), and shape; whereas the other distributions have only one or two free parameters. In fact, the Gamma distribution includes the Weibull, Exponential, and Log-normal distributions as special cases. The Weibull distribution also includes the Exponential distribution as a special case.

The Weibull, Log-normal, and Log-logistic distributions have two free parameters, location and scale; and the Exponential distribution has one free parameter, location. The Weibull and Log-normal distributions result as special cases of the Gamma distribution when the shape parameter equals one and zero, respectively. The Exponential distribution results as a special case of the Gamma distribution when both the shape and scale parameters equal one or as a special case of the Weibull distribution when the scale parameter equals one.

The Gamma distribution places fewer constraints on the parameters than the Weibull, Exponential, and Log-normal distributions. As a result, the parameter estimates obtained assuming the Gamma distribution will be most based on the data. If one of the other distributions is a good description of the data, its results will be similar to those of the less constrained Gamma distribution.

3.1.3 Distribution Adopted

The selection of the most appropriate distribution is based on several criteria:

- implications for the non-retention rate over time,
- likelihood ratio test,
- analysis of residuals, and
- maximum of the log-likelihood function.

Non-Retention Rate Over Time

The distributional assumption has implications for the non-retention rate over time. These implications are seen via the hazard function $h(t;\theta)$. Roughly, the hazard function can be thought of as the instantaneous probability of not retaining a unit at time *t*, given that the unit has been retained up to that time. Formally, it is the negative ratio of the survival probability density function *dS/dt* to the survival function,

$$h(t;\theta) = -\frac{dS/dt}{S(t;\theta)}.$$

An increasing hazard function means the non-retention rate increases as a unit of a measure ages, whereas a decreasing hazard function means the non-retention rate decreases as a unit of a measure ages. If the hazard function is constant, the non-retention rate remains constant as a unit of a measure ages. The non-retention rate most likely increases as a unit of a measure ages. Also, the non-retention rate most likely increases at an increasing rate as a unit of a measure ages. Therefore, the distributional assumption should probably be consistent with a hazard function that is always increasing, preferably at an increasing rate.

The hazard function of the Gamma distribution may have a variety of shapes, including always increasing at an increasing rate. Unfortunately, however, it is often difficult to determine which possible shape the hazard function of the Gamma distribution actually takes on.

The Weibull distribution produces a hazard function that is either always decreasing or always increasing. If the scale parameter is greater than one then the hazard function is decreasing, whereas if the scale parameter is less than one then the hazard function is increasing. Recall, a Weibull distribution with scale parameter equal to one corresponds to the Exponential distribution. The Exponential distributional results in a constant hazard function.

If the hazard function of the Weibull distribution is increasing (the scale parameter is less than one), the rate of increase depends on the value of the scale parameter. If the scale parameter is between 0.5 and 1, the hazard function is increasing at a decreasing rate; if the scale parameter equals 0.5, the hazard function is increasing at a constant rate; and if the scale parameter is between 0 and 0.5, the hazard function is increasing at an increasing rate.

The Log-normal distribution produces a hazard function that increases to a peak then decreases. The larger the scale parameter, the sooner the hazard function reaches its peak and begins to decrease. A hazard function that is increasing then decreasing means that for some period of time after a unit of a measure is installed, the non-retention rate increases as the unit of the measure ages then, after some point, the non-retention rate decreases as the unit of the measure ages. This pattern may be reasonable up to a point if there is initially more non-retention because of immediate dissatisfaction and removal of units of a measure. The clear problem with assuming a Log-normal distribution is that once the non-retention rate decreases as a unit of a measure ages it does so thereafter.

The hazard function of the Log-logistic distribution may increase to a peak then decrease or it may be always decreasing. If the scale parameter is less than one then the hazard function is increasing then decreasing, whereas if the scale parameter is greater than or equal to one then the hazard function is always decreasing.

Likelihood Ratio Test

If a distribution is a special case of another distribution, the appropriateness of the former versus the latter can be formally tested using the likelihood ratio test. Therefore, it is possible to compare the appropriateness of the Weibull, Exponential, and Log-normal distributions versus the Gamma distribution. It is also possible to compare the appropriateness of the Exponential distribution versus the Weibull distribution.

Analysis of Residuals

According to Allison (1995), Cox-Snell residuals are commonly used in Survival Analysis and are defined as:

$$e_j = -\log(S(t_j;\hat{\theta})),$$

where

e_j = the residual at the observed time to non-retention t_j and

$S(t_j;\hat{\theta})$ = the estimated survival function at time t_j.

A residual will be right-censored, interval-censored, left-censored, or uncensored, if the observed time to non-retention it is associated with is right-censored, interval-censored, left-censored, or uncensored, respectively.

If the fitted general linear regression model is appropriate, the residuals have an approximate exponential distribution with scale parameter one. To determine whether or not this is the case, a general linear regression model is fitted to the logarithm of the residuals assuming the exponential of the error term follows the standardized form of the exponential distribution. An estimated scale parameter not statistically different from one at a 10 percent level of significance or better, suggests the general linear regression model fitted to the logarithmic transformation of the time to non-retention may be appropriate.

Maximum of the Log-Likelihood Function

Recall, under each assumed distribution, the general linear regression model is fitted by maximizing the log-likelihood function. A larger maximum value of the log-likelihood function suggests a better model fit.

3.2 STANDARD ERROR OF THE EUL ESTIMATE

3.2.1 Calculation of the Standard Error

Because the general linear regression model fitted is the log of the time to non-retention, the parameters thus estimated and employed in the survival function will directly produce the log of the EUL estimate such that the survival probability is 0.50. The estimate of the EUL is then obtained by calculating the exponential of this log value ($e^{ln(EUL\ estimate)}$). Calculating the standard error of the EUL estimate, however, is not as simple because the logarithmic transformation is non-linear.

If the distribution of the log of the EUL estimate is known, it may be possible to calculate the exact standard error of the EUL estimate. However, this distribution is unknown in this study, as it is in most studies. Therefore, the approximate standard error is calculated by SAS® using a first order Taylor expansion around the EUL estimate of the log of the time to non-retention. This approximation is a function of the log of the EUL estimate and the standard errors of the parameter estimates of the general linear regression model.

3.2.2 Adjustment to the Standard Error

When fitting a general linear regression model to the data for a given measure, an observation is the time to non-retention of a unit of the measure. This unit in the cases of all of the lighting measures is a lamp. The calculation of the standard errors of the parameter estimates assumes each observation is independent. This assumption, however, may be incorrect, because the level of an observation is a unit of a measure being analyzed and not the level at which the first and only stage of sampling occurred. The first and only stage of sampling occurred at the site level. Recall, sites were selected for data collection. Therefore, the time to non-retention of a unit of a measure may be more similar within a site than between sites.

Several factors may cause the time to non-retention of a unit of a measure to be more similar within a site than between sites. Dissatisfaction or renovations may lead to the simultaneous removal of a large number of units, perhaps even all the units, of a measure at a site. Site-specific measure installation practices and operating conditions may affect the time to non-retention. The units of a measure installed at the same time at a site are likely to be of a similar quality and, therefore, have a similar time to non-retention. In addition, the time to non-retention of a unit of a lighting measure may be more similar within a site than between sites because the unit is a lamp and one fixture may hold more than one lamp. Consequently, non-retention of one fixture may account for non-retention of more than one lamp.

While the time to non-retention of a unit of a measure may be more similar within a site than between sites, it is not expected to be identical within a site. Dissatisfaction or renovations do not necessarily lead to the simultaneous removal of all the units of a measure at a site. Site-specific measure installation practices and operating conditions and measure quality may result in similar but not necessarily identical times to non-retention.

If the time to non-retention of a unit of a measure is more similar within a site than between sites and if the data analyzed for the measure are based on only a sample of sites that obtained a rebate for the measure, it is necessary to adjust the standard error of an estimate for the sample design by the square root of the design effect factor (Kish 1965). If the time to non-retention of a unit of a measure is no more similar within a site than between sites, then the square root of the design effect factor equals one and the unadjusted and adjusted standard error are equal. Usually, however, the design effect factor is greater than one.

If it is possible to obtain data from all the sites that obtained a rebate for a measure, it is not necessary to adjust the standard error of the EUL estimate. If all the units of a measure are included in the analysis, that the data collection occurred at the site level has no consequences and it is not necessary to adjust the standard error of the EUL estimate by the square root of the design effect factor.

The Design Effect Factor and Rho

In sampling terminology, a site is a cluster. In the case of a one-stage sample and assuming each cluster has the same number of units, the design effect factor can be expressed as

$$deff = 1 + rho(\overline{n} - 1)$$

where

rho = an estimate of the intra-cluster correlation *Rho* and

$\overline{n}$ = the average number of rebated units of a measure per sample site (the total number of rebated units of the measure across all sample sites included in the analysis of the measure divided by the total number of sample sites included in the analysis of the measure).

The equation for the population intra-cluster correlation (also known as the rate of homogeneity) *Rho* is

$$Rho = \frac{\sigma_b^2 - \dfrac{\sigma_w^2}{\overline{N} - 1}}{\sigma_o^2},$$

where

σ_b^2 = between-cluster population variance = $\dfrac{\sum_{i=1}^{C} N_i(\overline{T}_i - \overline{T})^2}{\sum_{i=1}^{C} N_i}$,

σ_w^2 = within-cluster population variance = $\dfrac{\sum_{i=1}^{C}\sum_{j=1}^{N_i}(T_{ij} - \overline{T}_i)^2}{\sum_{i=1}^{C} N_i}$,

σ_o^2 = overall population variance = $\dfrac{\sum_{i=1}^{C}\sum_{j=1}^{N_i}(T_{ij}-\overline{T})^2}{\sum_{i=1}^{C} N_i}$,

$\overline{N}$ = average number of units per cluster,

C = number of clusters,

N_i = number of units in cluster *i*,

$\overline{T}_i$ = average time to non-retention of a unit in cluster *i*,

$\overline{T}$ = average time to non-retention of a unit over all clusters, and

T_{ij} = time to non-retention of unit *j* in cluster *i*.

Noting that the overall population variance σ_o^2 equals the sum of the between- and within-cluster population variances σ_b^2 and σ_w^2, respectively, limit values of *Rho* can be determined and interpreted as follows:

- *Complete homogeneity within clusters* implies $\sigma_w^2 = 0$ and therefore $\sigma_b^2 = \sigma_o^2$ which leads to *Rho* = 1. *Rho* = 1 results in the largest design effect factor possible and, therefore, the largest adjustment to the standard error.
- *Extreme heterogeneity within clusters* implies σ_w^2 takes the largest possible value, σ_o^2, and, therefore, $\sigma_b^2 = 0$, which leads to *Rho* = $-1/(\overline{N}-1)$.
- *Units within a cluster no more closely related than units between clusters* implies $\sigma_b^2 = \sigma_w^2/(\overline{N}-1)$, which leads to *Rho* = 0. If this is the case, the design effect factor is one and the standard error obtained directly from the fit of the general linear regression model is correct.

In practice, *Rho* takes a value somewhere between zero and one. Negative values rarely happen. Thus, the design effect factor is usually larger than one.

Estimating Rho by Measure

In this study *Rho* is estimated separately for each measure as

$$rho = \frac{\hat{\sigma}_b^2 - \dfrac{\hat{\sigma}_w^2}{\overline{n}-1}}{\hat{\sigma}_o^2}$$

where

$\overline{n}$ = as defined earlier (see equation for *deff*),

$\hat{\sigma}_b^2$ = estimate of the between-cluster population variance = $\dfrac{\sum_{i=1}^{c} n_i (p_i - p_o)^2}{\sum_{i=1}^{c} n_i}$,

$\hat{\sigma}_w^2$ = estimate of the within-cluster population variance = $\dfrac{\sum_{i=1}^{c} n_i p_i (1 - p_i)}{\sum_{i=1}^{c} n_i}$,

$\hat{\sigma}_o^2$ = estimate of the overall population variance = $\hat{\sigma}_b^2 + \hat{\sigma}_w^2 = p_o(1 - p_o)$,

c = number of sites included in the analysis for which a rebate was obtained for the measure,

n_i = number of rebated units of the measure for site i,

p_i = proportion of rebated units of the measure not retained as of the latest on-site inspection date for site i,

p_o = proportion of rebated units of the measure not retained as of the latest on-site inspection date over all sites i = 1, 2, … , c.

Note, $\hat{\sigma}_b^2$ employs $\sum_{i=1}^{c} n_i$ as a divisor rather than the more typical $\sum_{i=1}^{c} n_i - 1$, because σ_b^2 employs $\sum_{i=1}^{C} N_i$ as a divisor.

Ideally $\hat{\sigma}_b^2$, $\hat{\sigma}_w^2$, and $\hat{\sigma}_o^2$ would be based on time to non-retention of a unit of the measure data. However, because the exact time to non-retention of a unit of a measure are difficult data to collect, typically, these data are the same for all non-retained units of a measure at a site. Therefore, $\hat{\sigma}_b^2$, $\hat{\sigma}_w^2$, and $\hat{\sigma}_o^2$ are instead based on the event of not retaining a unit of the measure data. That is, the proportion of rebated units of the measure not retained. The design effect factor is expected to be similar for the event of non-retention as the time to non-retention.

Value of rho Employed in the Analysis

The value of *rho* is calculated by measure, as just discussed, for all measures for which at least one unit of the measure was not retained. The average *rho* over all lighting measures is then calculated, and this average is used in the calculation of the design effect factor. Therefore, the design effect factor and, hence, the adjustment to the standard error of the EUL estimate (square root of the design effect factor), is the same for all lighting measures.

The design effect factor is calculated by measure type, lighting, employing the average *rho* over all lighting measures because it is likely the *rho*s for individual lighting measures contain information for each other. This is likely because data on all measures are limited by the difficulty of collecting data on the exact time to non-retention. Calculating a single design effect for lighting measures minimizes this limitation by allowing all available data on the time to non-retention for lighting measures to inform the value of the design effect factor. Also, the *rho*s for individual lighting measures may contain information for each other because the same site may be included in the analysis of more than one lighting measure.

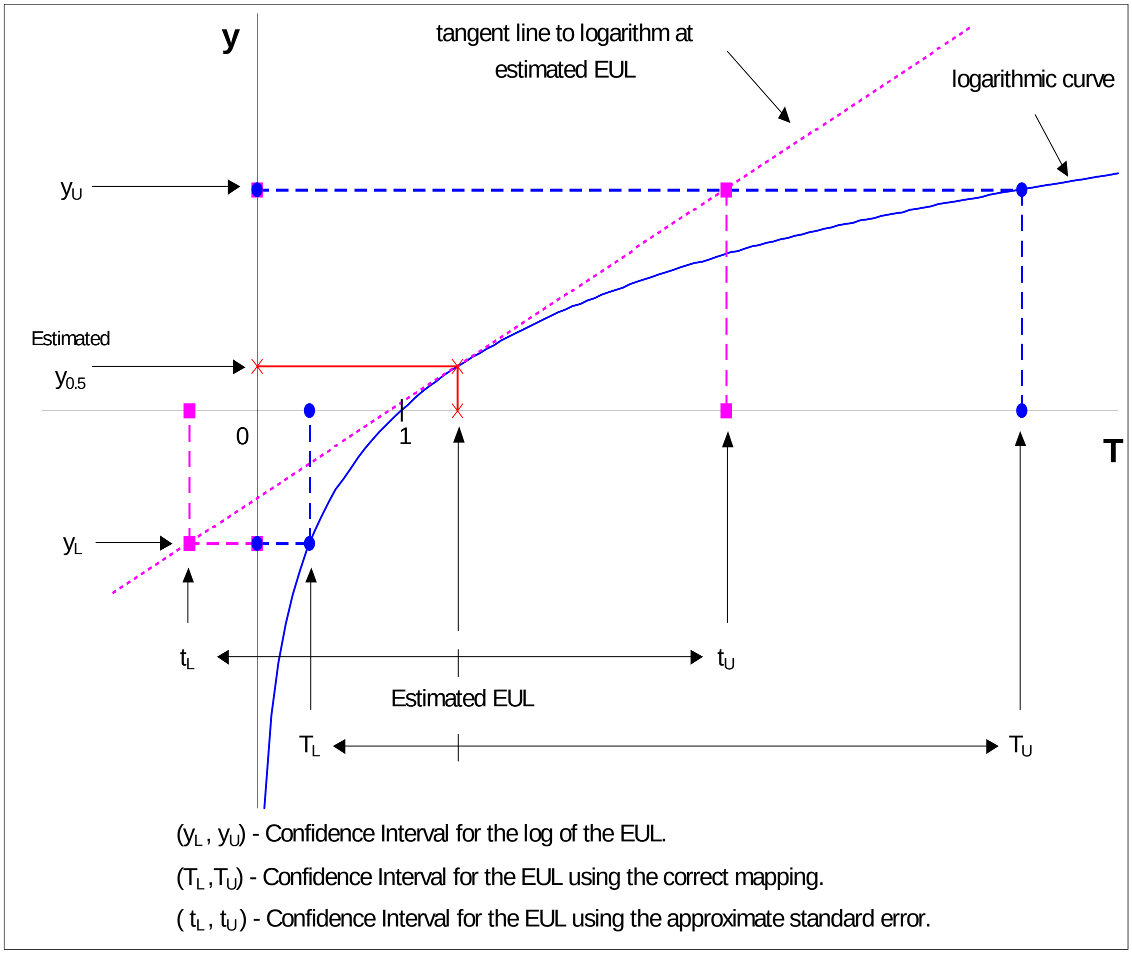
3.3 CONFIDENCE INTERVAL FOR THE EUL

Recall, it is only possible to calculate an approximate standard error of the EUL estimate. This is because it is the log of the EUL estimate that is directly obtained and the distribution of the log of the EUL estimate is unknown. A confidence interval for the EUL can be calculated using the adjusted, if necessary, approximate standard error of the EUL estimate. However, a more accurate confidence interval for the EUL can be obtained from the confidence interval for the log of the EUL. The lower and upper bounds of a confidence interval for the EUL equal the exponential of the lower and upper bound values of the confidence interval for the log of the EUL, respectively. This study calculates and reports this more accurate confidence interval for the EUL.

The lower and upper bounds of a confidence interval for the EUL based on the approximate standard error of the EUL estimate are the same distance from the EUL estimate. The confidence interval for the log of the EUL is similarly symmetric about the log of the EUL estimate. However, the confidence interval for the EUL based on the confidence interval for the log of the EUL is not symmetric about the EUL estimate. This result occurs because the logarithmic transformation is non-linear, explaining why the confidence interval for the EUL based on the approximate standard error of the EUL estimate is less accurate than the confidence interval for the EUL based on the confidence interval for the log of the EUL. The larger the approximate standard error of the EUL estimate, the greater the consequences of the non-linearity of the logarithmic transformation and the less accurate the confidence interval for the EUL based on the approximate standard error of the EUL estimate.

The non-linearity of the logarithmic transformation also explains why the confidence interval for the EUL based on the approximate standard error of the EUL estimate may contain negative values, which are clearly impossible. The confidence interval for the EUL based on the confidence interval for the log of the EUL will never contain negative values. The two methods of calculating a confidence interval of the EUL are illustrated in Figure 3-1.

Figure 3-1
Two Methods of Calculating a Confidence Interval for the EUL



3.3.1 Confidence Interval for the Log of the EUL

In general, the bounds of a confidence interval for a parameter are calculated as the parameter estimate ± the standard error of the parameter estimate times the critical value from the appropriate distribution for the desired level of confidence. The standard error of the log of the EUL estimate employed in the calculation of the confidence interval for the log of the EUL is provided by SAS. This standard error is a function of the standard errors of the parameter estimates of the general linear regression model. If necessary, the standard error of the log of the EUL estimate provided by SAS is adjusted by the square root of the design effect factor.

The log of an estimate of the EUL is assumed to be approximately normally distributed. Therefore, the critical value employed in the calculation of a confidence interval for the log of the EUL is approximated using the value from the Student distribution for the appropriate degrees of freedom and desired level of confidence. The degrees of freedom equals the effective sample size n_{eff} minus one, where n_{eff} is the number of units of the measure employed in the analysis divided by the design effect factor. The value of n_{eff} may be a non-integer.

3.4 The P-Value

The p-value reported is for the null hypothesis: the *ex ante* and *ex post* EULs are equal, and the alternative hypothesis: the two EULs are not equal. In this study, a p-value of less than or equal to 0.20 would cause the null hypothesis to be rejected.

The p-value is calculated based on the value of the following test statistic:

$$\frac{\left|\ln(ex\ post\ EUL) - \ln(ex\ ante\ EUL)\right|}{adjusted,\ if\ necessary,\ standard\ error\ of\ the\ \ln(ex\ post\ EUL)}.$$

The log of the *ex post* EUL is assumed to have an approximate normal distribution with mean ln(EUL) and unknown variance. Therefore, this test statistic has an approximate Student distribution with degrees of freedom equal to the effective sample size n_{eff} minus one, per the earlier distributional assumption regarding the log of an estimate of the EUL.

3.5 Weights

The relative importance of a site in the retention analysis of a measure depends on two factors: the energy costs the site avoids by installing the measure and the site's sample weight. If one or both of these factors per unit of a measure varies across sites, it is necessary to employ weights that reflect the different levels of one or both of these factors when estimating the general linear regression model. In the cases of each lighting measure, although the energy costs avoided per unit of the measure are the same across sites, the sample weights may vary across sites.

This analysis employs standard sample weights. The sample weight of site i in stratum k equals

$$\frac{B_{k(i)}}{b_{k(i)}},$$

where

$B_{k(i)}$ = population of sites in stratum k and

$b_{k(i)}$ = number of sample sites in stratum k.

The population counts used to calculate the sample weights are given in Table 2-1 in section 2. The sample counts are derived directly from the on-site inspection data.

In the retention analysis of a lighting measure, the weight w_i applied to each rebated unit j of the measure for site i in sample stratum k is calculated as

$$\left(\frac{B_{k(i)} / b_{k(i)}}{n_i} \right) \times \left(\frac{\sum_{i=1}^{c} n_i}{\sum_{i=1}^{c} B_{k(i)} / b_{k(i)}} \right),$$

where

$B_{k(i)} / b_{k(i)}$ = sample weight of site i in stratum k,

n_i = as defined earlier (number of rebated units of the measure for site i), and

c = as defined earlier (number of sites included in the analysis that received a rebate for the measure).

To obtain the correct unadjusted standard error of the EUL estimate, the sum of the weights must equal the number of observations included in the analysis. This is achieved by multiplying the component of the weight that reflects the different sample weights per unit of a measure $\left(\frac{B_{k(i)} / b_{k(i)}}{n_i} \right)$ by $\left(\frac{\sum_{i=1}^{c} n_i}{\sum_{i=1}^{c} B_{k(i)} / b_{k(i)}} \right)$.

3.6 References

Allison, Paul D. 1995. *Survival analysis using the SAS® system: A practical guide*. Cary, NC: SAS Institute Inc.

Kish, Leslie. 1965. *Survey sampling*. New York: John Wiley & Sons, Inc.

XENERGY Consulting Inc. 2000. *1996 Residential Lighting Third-Year Retention. PG&E Study ID 372R1*. Madison, Wisconsin.

XENERGY Consulting Inc. 1999. *1994–1995 Industrial Energy Efficiency Incentive Programs Third-Year Retention Study. PG&E Study IDs 311R1, 328R1, 314R1, 325R1*. Madison, Wisconsin.

4 LIGHTING RETENTION RESULTS

This section of the report presents the retention analysis results for PG&E's 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting. Recall, for each measure, the ultimate objective of this study is to estimate the median retention time or EUL. To begin, data descriptive of the measure data employed in the analysis are provided. Next, the estimate of *Rho* used in the adjustment of the standard error of an EUL estimate obtained from survival analysis is reported. Lastly, the results of the survival analysis are discussed.

4.1 DESCRIPTIVE STATISTICS

Table 4-1 reports various statistics regarding retention in the analysis data by measure. This table includes only those units of a measure that were inspected in the sixth year. Of particular interest in these tables is the percentage of units not retained. When third-year and sixth-year data are taken into account, less than five percent of each of the lighting measures CFL and T-8 has not been retained, and approximately twelve percent of lighting measure HID has not been retained.

Table 4-1
Survey Data by Measure

Measure	Initially Installed	At 3rd-Year Study		At 6th-Year Study	
		Units Retained*	% Installed Not Retained	Units Retained	% 3rd-Year Not Retained
CFL	13,320	12,880	3.3%	12,806	0.6%
HID	1,873	1,710	8.7%	1,651	3.5%
T8	6,101	5,978	2.0%	5,923	0.9%

* These unit counts are greater that those shown in the 3rd-year study report because a number of observations were removed from the 3rd-year study because of inconclusive survey results. During the 6th-year study, surveyors were able to confirm that many of the measures in questions were either in place or removed. These observations were, therefore, not removed from the 6th-year study.

Also interesting to note in Table 4-1, a larger percentage of units were not retained during the period between the initial purchase and the third-year inspection than during the period between the third-year and sixth-year inspections. Early fixture removals are often due to initial failure of defective equipment and early customer dissatisfaction with lighting performance. The low non-retention rates shown in Table 4-1 contribute to the relatively large EULs estimated in this study and presented in Section 4.3.

4.2 ADJUSTMENT TO THE STANDARD ERROR OF THE EUL ESTIMATE

The standard error of the EUL estimate is a function of the log of the EUL estimate and the standard errors of the parameter estimates of the general linear regression model. The calculation of the standard errors of the parameter estimates assumes each observation is independent. This assumption, however, may be incorrect because when fitting a general linear regression model to the data for a given measure, the level of an observation is of a unit of the measure, whereas the first and only stage of sampling occurred at the site level. Therefore, if the data analyzed for a measure are based on only a sample of sites that obtained a rebate for the measure, it is necessary to adjust the standard error of the EUL estimate. This is the case for all the lighting measures.

It is necessary to correct the standard error of an EUL estimate to the extent the time to non-retention of a unit of a measure is more similar within a site than between sites. The extent to which the time to non-retention of a unit of a measure is more similar within a site than between sites and, therefore, the extent of the adjustment to the standard error, is reflected by the value of *rho*. Typically, *rho* ranges between zero and one. The closer *rho* is to one, the more similar the time to non-retention of a unit of a measure is within a site than between sites and the larger the adjustment to the standard error.

The value of *rho* is smallest for lighting measure T-8, 0.36 and largest for lighting measure HID, 0.73. The average *rho* for lighting measures, which is used in the adjustment of the standard error of the EUL estimate for all lighting measures, is 0.54. These data are reported in as well as data used in the calculation of *rho* by lighting measure. The parameter *rho* and its components are dimensionless. The components are dimensionless because they are estimated for the non-retention event, rather than for the time to non-retention.

Table 4-2
***rho* by Lighting Measure and Overall**

Measure		CFL			HID			T8		
Overall proportion of non retention (p_0)		3.98%			11.85%			2.90%		
Variance Components	Within-site variance (σ_w^2)	0.0178	Min.	0.0000	0.0260	Min.	0.0000	0.0177	Min.	0.0000
			Max.	0.2500		Max.	0.2484		Max.	0.2500
	Between-site variance (σ_b^2)	0.0204			0.0785			0.0104		
	Overall variance (σ_o^2)	0.0382			0.1045			0.0282		
rho		0.5256			0.7325			0.3608		
Average rho (overall rho)		0.54								

4.3 SURVIVAL ANALYSIS RESULTS

The results of the survival analysis for each measure are presented in Table 4-3. Results are presented for each distribution for which it was possible to fit a general linear regression model. The standard errors reported in Table 4-3 are the corrected standard errors. For each measure, if the *ex ante* EUL is outside the 80 percent confidence interval, it is smaller than the estimated or *ex post* EUL.

Table 4-3
Survival Analysis Results by Measure

Measure	Distribution	Maximum of Log Likelihood	Selected Parameter Estimates: Scale	Selected Parameter Estimates: Shape	*ex ante* EUL (years)	*ex post* EUL (years)	80% Confidence Interval (years)	Standard Error (years)
CFL	*Exponential*	-2448.5	1.00 [a]	-	10.0	107.6	(78.8 , 146.9)	26.1
rho = 0.54	*Log-logistic*	-2445.9	0.88	-		102.2	(31.0 , 337.2)	95.1
n_{eff} = 459.7	*Log-normal*	-2439.0	2.05	-		226.5	(51.6 , 995.2)	261.2
	Weibull	-2446.6	0.89	-		78.5	(26.4 , 233.4)	66.7
HID	*Exponential*	-643.0	1.00 [a]	-	15.0	45.5	(34.1 , 60.8)	10.3
rho = 0.54	*Gamma*	-642.4	0.37	2.48		30.1	(15.7 , 57.6)	15.2
n_{eff} = 227.4	*Log-logistic*	-642.5	0.87	-		46.4	(20.8 , 103.5)	29.0
	Log-normal	-642.2	1.89	-		78.2	(30.0 , 204.2)	58.4
	Weibull	-642.4	0.90	-		36.9	(17.9 , 76.0)	20.7
T-8	*Exponential*	-2306.6	1.00 [a]	-	16.0	38.6	(28.3 , 52.6)	9.3
rho = 0.54	*Log-logistic*	-2305.3	0.92	-		43.9	(16.8 , 114.5)	32.7
n_{eff} = 179.2	*Log-normal*	-2303.5	1.89	-		66.2	(21.5 , 203.4)	57.8
	Weibull	-2306.3	0.96	-		35.5	(14.9 , 84.6)	24.0

[a] In the case of the Exponential distribution, the scale parameter is taken to equal one, it is not estimated.

4.3.1 Distribution Adopted

For each measure, this study must make a recommendation regarding the most appropriate distributional assumption for the survival analysis. For all measures, we recommend focusing on the survival analysis results when a Weibull distribution is assumed. Primarily because:

1. Of the distributions for which an estimate of the EUL was obtained, only the Weibull distribution is known to be consistent with the oldest units of a measure having the highest non-retention rate.
2. There is little to no evidence to justify adopting one of the other distributions over the Weibull distribution.

In the case of the Weibull distribution, the estimated scale parameter is less than one. This means the non-retention rate increases as a unit of a measure ages. Therefore, of the distributions for which an estimate of the EUL was obtained, only the Weibull distribution is known to be consistent with the oldest units of a measure having the highest non-retention rate.

The likelihood ratio test and analysis of residuals provide no evidence to justify adopting one of the other distributions over the Weibull distribution. On the basis of the likelihood ratio test, no distribution is determined to be more appropriate than the Weibull distribution. Furthermore, in the case of lighting measure CFL, the Weibull distribution is more appropriate than the Exponential distribution at a 5.2 percent significance level.

The results of the residual analysis do not suggest any distribution is more appropriate than the Weibull distribution. In fact, this analysis suggests none of the distributions may be appropriate.

Of the four criteria used to evaluate the appropriate distribution for the survival analysis, the maximum of the log-likelihood function is the least informative. The maximum of the log-likelihood function is the only criterion considered by which another distribution may be determined to be more appropriate than the Weibull. The Weibull distribution produces only the third largest maximum of the log-likelihood function.

The Log-normal distribution produces the largest maximum of the log-likelihood function for each measure. This distribution produces an increasing then decreasing hazard function, which is consistent with the data shown in Table 4-1, higher removal rates in the early period, followed by lower removal rates in the later period. The problem with the Log-normal distribution is that after a given point in time, the non-retention rate of a measure continues to decrease as the unit of a measure ages. This non-retention pattern over time contributes to the very high EUL estimates provided by the Log-normal distribution. A more reasonable assumption is that the non-retention rate of a unit of a measure increases as unit of a measure ages.

Thus given the performance of the Weibull distribution in three of the four evaluation criteria used in the survival analysis, it was determined that the results regarding the maximum of the log-likelihood function provide insufficient evidence to justify adopting one of the other distributions over the Weibull distribution.

4.3.2 Ex Post EUL Adopted

Lighting Measures CFL and HID

Both the retention data and survival analysis results suggest the EUL for lighting measure CFL is larger than its *ex ante* EUL of 10 years. After six years, only 4.0 percent of lighting measure CFL has not been retained. Consequently, it is very unlikely after only four more years, 50 percent of lighting measure CFL will not retained.

In addition, both the retention data and survival analysis results suggest the EUL for lighting measure HID is larger than its *ex ante* EUL of 15 years. After six years, only 11.9 percent of lighting measure HID has not been retained. Consequently, it is unlikely after nine more years, 50 percent of lighting measure HID will not be retained.

In the cases of lighting measures CFL and HID, the *ex ante* EUL is always outside the 80 percent confidence interval and smaller than the *ex post* EUL. Also in the cases of both these measures, at this time, the best estimate of the EUL is obtained when a Weibull distribution is assumed. For lighting measure CFL this EUL is 78.5 years and for lighting measure HID this EUL is 36.9 years. Although we recommend adopting an *ex post* EUL larger than the *ex ante* EUL in the cases of both lighting measures CFL and HID, we do not recommend adopting an *ex post* EUL as large as 78.5 and 36.9 years, respectively. Our reasons are as follows:

- In the case of the Weibull distribution, the estimated scale parameter is less than one, but greater than 0.5. This means the non-retention rate increases as a unit of a measure ages but at a decreasing rate. In contrast, the non-retention rate most likely increases at an increasing rate as a unit of a measure ages. Therefore, the estimated or *ex post* EUL of 78.5 years for lighting measure CFL and 36.9 years for lighting measure HID are likely to be overestimates.
- It is difficult to estimate when non-retention will be 50 percent (i.e., the EUL), when the percentage of units of a measure not retained to date is so small. This is evidenced by the relatively large bounds of the 80 percent confidence interval when a Weibull distribution is assumed. For lighting measure CFL these bounds are 26.4 and 233.4 years, and for lighting measure HID these bounds are 17.9 and 76.0 years. Recall, to date, only 4.0 percent of lighting measure CFL has not been retained and only 11.9 percent of lighting measure HID has not been retained.

Therefore, at this time, we recommend adopting an *ex post* EUL of 16 years in the cases of both lighting measures CFL and HID. Statewide, year 2001 programs are using a 16-year EUL for lighting measures.[1] In the cases of both lighting measures CFL and HID, a 16-year EUL is larger than the *ex ante* EUL (10 and 15 years, respectively) and smaller than the best estimate of the EUL (78.5 and 36.9 years, respectively), which is likely an overestimate.

Lighting Measure T-8

For lighting measure T-8, the *ex ante* EUL is outside the 80 percent confidence interval and smaller than the estimated or *ex post* EUL when an Exponential, Log-logistic, or Log-normal distribution is assumed. However, in the cases of the Log-logistic and Log-normal distributions, the *ex ante* EUL is not far outside the 80 percent confidence interval. Furthermore, the *ex ante* EUL is inside the 80 percent confidence interval when a Weibull distribution is assumed. Therefore, at this time, we recommend adopting an *ex post* EUL equal to the *ex ante* EUL of 16 years for lighting measure T-8. Again, year 2001 programs statewide are using a 16-year EUL for lighting measures.

[1] *California Measurement Advisory Committee Public Workshops on PY2001 Energy Efficiency Programs*, filed on September 25, 2000 by Sempra Energy, p. 56.

4.4 Alternative Adjustments to the Standard Error of the EUL Estimate

For the survival analysis results obtained when a Weibull distribution is assumed, we tested the sensitivity of the results to the value of *rho* employed in the adjustment of the standard error of the EUL estimate. We consider the two extreme values of *rho*, zero and one, and a value in the middle, 0.5. The closer *rho* is to one the more similar the times to non-retention of units of a measure are within a site than between sites and the larger the adjustment to the standard error. The results of the sensitivity test are given in Table 4-4. For purposes of comparison, this table also includes the results for the value of *rho* estimated from the data and used in the analysis, 0.54. The results of the sensitivity test support our earlier conclusions for all measures.

Table 4-4
Sensitivity Test Results by Measure

Meas.	CFL			HID			T-8		
ex ante EUL	10.0			15.0			16.0		
Dist.	Weibull			Weibull			Weibull		
rho	*ex post* EUL (years)	80% Confidence Interval (years)	Standard Error (years)	*ex post* EUL (years)	80% Confidence Interval (years)	Standard Error (years)	*ex post* EUL (years)	80% Confidence Interval (years)	Standard Error (years)
0.00	78.5	(64.2 , 95.8)	12.3	36.9	(28.7 , 47.5)	7.2	35.5	(30.6 , 41.1)	4.1
0.50	78.5	(27.4 , 224.4)	64.2	36.9	(18.4 , 74.2)	20.1	35.5	(15.4 , 81.9)	23.1
0.54	**78.5**	**(26.4 , 233.4)**	**66.7**	**36.9**	**(17.9 , 76.0)**	**20.7**	**35.5**	**(14.9 , 84.6)**	**24.0**
1.00	78.5	(18.0 , 342.7)	90.0	36.9	(14.2 , 96.1)	27.4	35.5	(10.9 , 115.3)	32.4

For lighting measure CFL, we recommend adopting an *ex post* EUL larger than the *ex ante* EUL because the *ex ante* EUL is outside the 80 percent confidence interval and smaller than the *ex post* EUL when a Weibull distribution is assumed. In the cases of all values of *rho* tested, the *ex ante* EUL of 10 years remains outside the 80 percent confidence interval.

For lighting measure HID, we also recommend adopting an *ex post* EUL larger than the *ex ante* EUL because the *ex ante* EUL is outside the 80 percent confidence interval and smaller than the *ex post* EUL when a Weibull distribution is assumed. In the cases of all but one value of *rho* tested, the *ex ante* EUL of 15 years remains outside the 80 percent confidence interval. The one value of *rho* that produces a result different from the value of *rho* used in the analysis is an extreme value of rho, one. A value of *rho*=1 means all units of a measure at a site have the same time to non- retention, which is unlikely. It seems reasonable to expect the times to non-retention of units of a measure to be more similar within a site than between sites. This expectation is supported by the value of *rho* estimated from the data, 0.54.

For lighting measure T-8, we recommend adopting an *ex post* EUL equal to the *ex ante* EUL because the *ex ante* EUL is inside the 80 percent confidence interval when a Weibull distribution is assumed. In the cases of all but one value of *rho* tested, the ex ante EUL of 16 years remains inside the 80 percent confidence interval. The one value of *rho* that produces a result different from the value of *rho* used in the analysis is an extreme value of rho, zero. A value of *rho*=0 means the times to non-retention of units of a measure are no more similar within a site than between sites, which is unlikely. Again, it seems reasonable to expect the times to non-retention of units of a measure to be more similar within a site than between sites, and this expectation is supported by the value of *rho* estimated from the data, 0.54.

A ON-SITE SURVEY INSTRUMENT

[This page intentionally left blank.]

[This page intentionally left blank.]

PG&E Residential Appliance Efficiency Incentive Lighting Program
1994-1995 Sixth-Year Multifamily Measure Retention Study

PG&E Account Number	Name of Owner / Checkname	Tracking #
XXX9999999	**JOHN DOE PROPERTIES**	**9999999**
Name of Contact Person	*Contact Phone*	*Segment*
JOHN DOE	**(999) 999-9999**	**1-96-CEN**

Name of Complex/Customer:	**JOHN DOE PROPERTIES**	*PG&E Division:*	**XXX**
Address:	**9999 ANYSTREET**	*PG&E Local Office:*	**Anycity**
City/State/Zip:	**ANYCITY, CA 99999**	*Billing System Phone:*	**(999) 999-9999**

Third-Year Retention Survey Date: **1/15/1998**

	Area Code	Application Code	Check Date	Measure Code	Measure Description	Number Purchased	Number Expected	Number Observed	Control Code	Discrep Code	Removal Code	Yrs Since Removal
1	7	XXX9999	9/9/99	L86	COMPACT FLUORESCENT: HARDWIRE FIXTURE, 5-13 WATTS (RES. LIGHTING)	13	**10**					
2	5	XXX9999	9/9/99	L101	FIXTURE: T-8 FIXTURE & BALLAST, 4 FT, 2-LAMP	15	**15**					
3	10	XXX9999	9/9/99	L89	HID FIXTURE: 0-70 WATTS	11	**11**					
4												
5												
6												
7												
8												

Notes:

Table 1-Area Codes

Code		Description
A	1	Hallway
	2	Storage/utility
	3	Office
	4	Recreation area
	5	Parking lot
	6	Laundry room
	7	Exterior walkway
	8	Exit
	9	Kitchen
	10	Other

Table 2-Control Codes

Code		Description
C	1	Manual switch
	2	Photosensor
	3	Occupancy sensor
	4	Timer

Table 3-Observed/Expected Discrepancy Codes

Code		Description
D	1	Removed, not replaced
	2	Removed, replaced with different (describe)
	3	Never installed, stockpiled
	4	Temporarily taken out of operation
	5	Could not locate
	6	Other (describe)

Table 4-Removal Codes

Code		Description
R	1	Equip failed, not replaced
	2	Remodeled/Equipment replaced
	3	Unable to locate equivalent replacement
	4	Change of use
	5	Other (describe)

B

TABLE 6B

[This page intentionally left blank.]

[This page intentionally left blank.]

Protocol Table 6B
Results of Sixth-Year Retention Study
Pacific Gas and Electric Company's 1994 and 1995 Appliance Energy Efficiency Programs, Residential Lighting
Study ID 384bR2 (1994), 401bR2 (1995)

Item 1			Item 2		Item 3	Item 4	Item 5	Item 6		Item 7	Item 8	Item 9
			EUL (years)									
								80% Confidence Interval				
Measure	End Use	Measure Description	*ex ante*	Source of *ex ante*	*ex post* (estimated from study)	Adopted *ex post* (to be used in claim)	*ex post* Standard Error	Lower Bound	Upper Bound	p-value for *ex post* EUL	EUL Realization Rate (adopted ex post / ex ante)	"Like" Measures Associated with Studied Measures
CFL	Lighting	COMPACT FLUORESCENT: HARDWIRE (RES. LIGHTING) COMPACT FLUORESCENT: HARDWIRE FIXTURE, 5-13 WATTS (RES. LIGHTING) COMPACT FLUORESCENT: HARDWIRE FIXTURE 14-26 WATTS (RES. LIGHTING) COMPACT FLUORESCENT: HARDWIRE FIXTURE, 27-50 WATTS (RES. LIGHTING)	10.0	a	78.5	16.0	66.7	26.4	233.4	0.02	1.60	None
HID	Lighting	HID FIXTURE: 35-70 WATTS HID FIXTURE: 0-70 WATTS HID FIXTURE: >= 71 WATTS HID FIXTURE: 35-100 WATTS (RES. LIGHTING) HID FIXTURE: >= 150 WATTS (RES. LIGHTING)	15.0	a	36.9	16.0	20.7	17.9	76.0	0.11	1.07	None
T-8	Lighting	FIXTURE: T-8 FIXTURE & BALLAST, 4 FT, 1-LAMP FIXTURE: T-8 FIXTURE & BALLAST, 4 FT, 2-LAMP FIXTURE: T-8 FIXTURE & BALLAST, 4 FT, 3-LAMP FIXTURE: T-8 FIXTURE & BALLAST, 8 FT, 2-LAMP FIXTURE: T-8 FIXTURE, 1-LAMP (RES. LIGHTING) FIXTURE: T-8 FIXTURE, 2-LAMP (RES. LIGHTING) FIXTURE: INCAND TO FLUOR CONVERSION W/ES BLST (RES. LIGHTING) FIXTURE: INCAND TO FLUOR CONVERSION W/ELEC BLST (RES. LIGHTING) FIXTURE: REPLACE LAMP & BLST, 2 FT, T-8 & ELEC BLST FIXTURE: REPLACE LAMP & BLST, 3 FT, T-8 & ELEC BLST FIXTURE: REPLACE LAMP & BLST, 4 FT, T-8 & ELEC BLST FIXTURE: REPLACE LAMP & BLST, 8 FT, T-8 & ELEC BLST FIXTURE: T-8 FIXTURE & BALLAST, 2 FT, 2-LAMP FIXTURE: T-8 FIXTURE & BALLAST, 2 FT, 4-LAMP BALLAST: ELECTRONIC (RES. LIGHTING)	16.0	a	35.5	16.0	24.0	14.9	84.6	0.24	1.00	None

[a]PG&E Advice Letter 1867-G/1481-E. 1995 DSM Program Activity and Expected Earnings. As approved by the California Public Utilities Commission May 8, 1995.

C TABLE 7B

C.1 OVERVIEW INFORMATION

a. Study Title and Study ID Number

Study Title: Retention Study of Pacific Gas and Electric Company's 1994 and 1995 Appliance Energy Efficiency Programs, 1994-1995 Residential Lighting Sixth-Year Retention.

Study ID Number: 1994, 384bR2 and 1995, 401bR2.

b. Program, Program Years, and Program Description

Program: Appliance Energy Efficiency, Residential Lighting.

Program years: 1994 and 1995.

Program description: The Multifamily Property Rebate Program provides financial incentives to owners and managers of multifamily dwellings for the installation of selected energy efficiency measures in common areas of their complexes.

c. End Uses and Measures Covered

This study covers lighting end uses. Table C-1 lists the lighting measures covered.

Table C-1
Measures Included in the Study

Measure	Measure Description
CFL	COMPACT FLUORESCENT: HARDWIRE (RES. LIGHTING)
	COMPACT FLUORESCENT: HARDWIRE FIXTURE, 5-13 WATTS (RES. LIGHTING)
	COMPACT FLUORESCENT: HARDWIRE FIXTURE 14-26 WATTS (RES. LIGHTING)
	COMPACT FLUORESCENT: HARDWIRE FIXTURE, 27-50 WATTS (RES. LIGHTING)
HID	HID FIXTURE: 35-70 WATTS
	HID FIXTURE: 0-70 WATTS
	HID FIXTURE: >= 71 WATTS
	HID FIXTURE: 35-100 WATTS (RES. LIGHTING)
	HID FIXTURE: >= 150 WATTS (RES. LIGHTING)
T-8	FIXTURE: T-8 FIXTURE & BALLAST, 4 FT, 1-LAMP
	FIXTURE: T-8 FIXTURE & BALLAST, 4 FT, 2-LAMP
	FIXTURE: T-8 FIXTURE & BALLAST, 4 FT, 3-LAMP
	FIXTURE: T-8 FIXTURE & BALLAST, 8 FT, 2-LAMP
	FIXTURE: T-8 FIXTURE, 1-LAMP (RES. LIGHTING)
	FIXTURE: T-8 FIXTURE, 2-LAMP (RES. LIGHTING)
	FIXTURE: INCAND TO FLUOR CONVERSION W/ES BLST (RES. LIGHTING)
	FIXTURE: INCAND TO FLUOR CONVERSION W/ELEC BLST (RES. LIGHTING)
	FIXTURE: REPLACE LAMP & BLST, 2 FT, T-8 & ELEC BLST
	FIXTURE: REPLACE LAMP & BLST, 3 FT, T-8 & ELEC BLST
	FIXTURE: REPLACE LAMP & BLST, 4 FT, T-8 & ELEC BLST
	FIXTURE: REPLACE LAMP & BLST, 8 FT, T-8 & ELEC BLST
	FIXTURE: T-8 FIXTURE & BALLAST, 2 FT, 2-LAMP
	FIXTURE: T-8 FIXTURE & BALLAST, 2 FT, 4-LAMP
	BALLAST: ELECTRONIC (RES. LIGHTING)

d. Method and Models Used

In the cases of all measures, the final model specification used for the study assumes a Weibull distribution. See the Study Methods section (3) for a complete discussion of the methods employed in this study. Also see the Lighting Retention Results section (4) for the results of the final model specification as well as the other model specifications considered. Section 4.3.1 discusses the reasons the Weibull distribution is adopted for all measures.

e. Analysis Sample Size

Table C-2 shows the analysis sample sizes by measure. This table shows both the number of sites and the number of units of a measure included in a measure's analysis data set. Sites were selected for data collection and a unit of a measure is the level at which the data are analyzed. Third-year on-site inspections were conducted January through December 1998, and sixth-year on-site inspections were conducted August through October 2000. In the cases of all measures, a unit is a lamp.

Table C-2
Analysis Sample Sizes by Measure

Measure	Sites	Rebated Lamps
CFL	252	13,780
HID	130	1,873
T-8	98	6,226
Total	480 (301 unique)	21,879

C.2 DATABASE MANAGEMENT

a. Data Sources and Elements

Program tracking data:
TRAK9495.SD2 SAS dataset.

Third-year on-site inspection data:
RETLIT.SD2 SAS dataset.

Sixth-year on-site inspection data:
LITSURV3.SD2 SAS dataset.

See section 2.3 for a list of the data elements obtained from each of these sources.

b. Data Attrition

In the case of each measure, an attempt was made to conduct an on-site inspection of only a sample of sites that obtained a rebate for the measure. All 301 sites visited for the third-year retention study were visited for this study. An on-site inspection was conducted and all necessary data were collected for 297 of these 301 sites. In the cases of the four remaining sites, the on-site inspection yielded incomplete data for one measure. The data for a measure are incomplete if the inspector could not locate all of the units of the measure installed and indicated some uncertainty about her/his ability to access all areas of the site.

For each measure, Table C-3 shows the number of sites in the population and analysis data set. This table also shows for the population and analysis data set the number of units of the measure rebated. All sample sites and rebated lamps are included in the analysis data set because between the two on-site inspections all necessary data were collected for each sample site at least one time. Therefore, if a sample site obtained a rebate for a given measure, the site is included in the measure's analysis data set.

Table C-3
Analysis Data by Measure

Measure	Population		Analysis Data	
	Sites	Rebated Lamps	Sites	Rebated Lamps
CFL	1,232	66,291	252	13,780
HID	628	9,029	130	1,873
T-8	287	21,082	98	6,226
Total	2,147 (1449 unique)	96,402	480 (301 unique)	21,879

c. Data Used to Merge Data Sets

The Program tracking data and on-site inspection data were merged by project, measure, and specific measure component. A project is a unique site—identified by PG&E control number--and rebate application combination. Both the Program tracking data and on-site inspection data employed common codes for the measures and their components. The components of each measure are listed in Table C-1's measure description column.

d. Data Collected Specifically for the Analysis but not Used

All data collected specifically for the analysis were used.

C.3 SAMPLING

a. Sampling Procedures and Protocols

The sample sites included in this study are the same sample sites visited for the third-year retention study of PG&E's 1994–1995 Residential Lighting Efficiency Incentives Programs. See section 2.2.1 for a complete discussion of the sampling procedures and protocols.

b. Survey Information

The on-site data collection instrument is provided in Appendix A. The sample disposition is discussed earlier in section C.2.b. All 301 sites visited for the third-year retention study were visited for this study and an on-site inspection was at least partially completed at 300 sites. Therefore, no effort was made to test or correct for non-response bias.

c. Statistical Descriptions

Table C-4
Survey Data by Measure

Measure	Initially Installed	At 3rd-Year Study		At 6th-Year Study	
		Units Retained*	% Installed Not Retained	Units Retained	% 3rd-Year Not Retained
CFL	13,320	12,880	3.3%	12,806	0.6%
HID	1,873	1,710	8.7%	1,651	3.5%
T8	6,101	5,978	2.0%	5,923	0.9%

* This table includes only those units of a measure inspected in the sixth year. In contrast, Tables C-2 and C-3 include units of a measure inspected in the sixth year as well as units inspected in only the third year.

C.4 DATA SCREENING AND ANALYSIS

a. Treatment of Outliers and Missing Data Points

The residuals of each general linear regression model fit were examined for the presence of any outliers or influential data points. In the cases of all the fitted models, no data points appeared to have an inordinate influence on the model fit.

b. Background Variables

Background variables such as economic and political activity may affect the time to non-retention. However, the collection and analysis of data clearly relevant to non-retention are sufficiently challenging to justify the omission of background variables.

c. Data Screens

If a sample site obtained a rebate for a given measure and either the third or sixth-year on-site inspection was at least partially completed for the measure, the site is included in the measure's analysis data set.

d. Model Statistics

The standard model statistics for all final general linear regression models are provided in Table C-5. The table provides the corrected standard errors and the approximate p-value associated with the corrected standard errors. The p-value for the intercept corresponds to a test of the hypothesis that the intercept equals zero. A p-value is not provided for the scale parameter because the distribution of the scale parameter is presumably unknown. Each general linear regression model was fitted using the SAS LIFEREG procedure.

Table C-5
Final General Linear Regression Model Statistics

Measure	Distribution	Intercept(μ) Estimate (ln (years))	Intercept(μ) Standard Error (ln (years))	Intercept(μ) p-value	Scale(σ) Estimate (adimensional)	Scale(σ) Standard Error (adimensional)
CFL	Weibull	4.69	0.95	<0.01	0.89	0.28
HID	Weibull	3.94	0.65	<0.01	0.90	0.26
T-8	Weibull	3.92	0.79	<0.01	0.96	0.33

Table C-6 presents EUL estimates developed using the final models.

Table C-6
Summary of EUL Estimates

Measure	Distribution	EUL (years) *ex post* (estimated from study)	EUL (years) *ex post* Standard Error	EUL (years) 80% Confidence Interval Lower Bound	EUL (years) 80% Confidence Interval Upper Bound	p-value for *ex post* EUL
CFL	Weibull	78.5	66.7	26.4	233.4	0.02
HID	Weibull	36.9	20.7	17.9	76.0	0.11
T-8	Weibull	35.5	24.0	14.9	84.6	0.24

e. Specification

See the Study Methods section (3) for a complete discussion of the methods employed in this study. Also see the Lighting Retention Results section (4) for the results of the final model specification as well as the other model specifications considered.

1. Heterogeneity

The number of units rebated and installed of a measure may vary across sites. The heterogeneity of sites is recognized and addressed in the model specification and estimation procedures by employing a unit of a measure as an observation in the analysis. Therefore, the number of observations on a site included in the analysis is equal to the number of units rebated and installed at the site.

2. Omitted Factors

Again, the collection and analysis of data central to non-retention are sufficiently challenging to exhaust the limited resources available to conduct this study. However, future studies of retention may want to consider collecting and including in the analysis data on some broad indicators of retention. Such broad indicators may include:

- whether or not the same firm occupies the space and
- whether the space is being used for the same or a different purpose.

f. Error in Measuring Variables

There are no particular concerns regarding error in measuring variables. All 301 sites visited for the third-year retention study were visited for this study and an on-site inspection was at least partially completed at 300 sites. Therefore, no effort was made to test or correct for non-response bias. In addition, the methods employed are well suited to handle imprecise measures of the time to non-retention.

g. Influential Data Points

See C.4 Data Screening and Analysis, a. Treatment of Outliers and Missing Data Points.

h. Missing Data

There are effectively no missing data. If a sample site obtained a rebate for a given measure and either the third or sixth-year on-site inspection was at least partially completed for the measure, the site is included in the measure's analysis data set. Also, as just stated in C.4.f, the methods employed are well suited to handle imprecise measures of the time to non-retention.

i. Precision

See sections 3.2 and 3.3.