RETENTION STUDY OF PACIFIC GAS AND ELECTRIC COMPANY'S 1996 AND 1997 COMMERCIAL ENERGY EFFICIENCY PROGRAMS

1996 -1997 Commercial Lighting & HVAC Fourth Year Retention
Study ID 349R1 & 351R1

March 1, 2001

RETENTION STUDY OF
PACIFIC GAS AND ELECTRIC COMPANY'S
1996 AND 1997
COMMERCIAL ENERGY EFFICIENCY PROGRAMS

1996 -1997 Commercial Lighting & HVAC Fourth Year Retention
Study ID 349R1 & 351R1

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

FOURTH YEAR RETENTION STUDY FOR PG&E'S 1996 & 1997 COMMERCIAL EEI PROGRAM LIGHTING AND HVAC TECHNOLOGIES PG&E STUDY ID #S: 349R1 & 351R1

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 March 1998, 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) for all HVAC and lighting energy efficiency technologies for which rebates were paid in 1996 and 1997 by Pacific Gas & Electric Company's (PG&E's) Commercial Energy Efficiency Incentive (CEEI) Programs. Retrofits were performed under three different PG&E programs, the Retrofit Express (RE), Retrofit Efficiency Options (REO), and Customize Incentives (CI) Programs.

Methodology

The Protocols assert the purpose of a retention study is to collect data on the fraction of installed measures in place and operable in order to produce a revised estimate of its EUL. The ultimate goal is to estimate the EUL (or the median number of years that the measure is still in place and operable), which can be realized by identifying the measure's survival function. For this study, the survival function describes the percentage of measures installed that are still operable and in place at a given time. Survival analysis is the process of analyzing empirical failure/removal data in order to model a measure's survival function. As much as possible, we have attempted to employ classical survival analysis techniques to our study approach.

For this study, the vast majority of measures were in place less than five years (few were installed prior to 1996, and follow-up data collection was conducted no later than the end of 2000). Because the ex ante EUL is 15-16 years for all studied measures, it is very unlikely that our data will be capable of accurately estimating the survival function for the studied measures.

Our overall approach consists of five analysis steps that were used to estimate each of the studied measures' EULs:

1. ***Compile summary statistics*** on the raw retention data. Upon review of the summary statistics, it became clear that such a small percentage of failures and removals had occurred, that it would be difficult to model the equipment's survival function.
2. ***Visually inspect*** the retention data, by simply calculating the cumulative percentage of equipment that had failed in a given month, and plotting the percentage over time. This step clearly illustrated that for each studied measure, there was not enough data over time to support an accurate estimate of the survival function.
3. ***Develop a trend line*** from the survival plots. Using the plots developed in (2) above, a trend line was estimated using standard linear regression techniques. We modeled the trend as a linear and an exponential function. In each case, we used the resulting trend line to estimate the EUL, which was statistically significantly larger than the ex ante estimate.
4. ***Develop a survival function*** using classical survival techniques. We modeled the survival function assuming five of the most common survival distributions: exponential, logistic, lognormal, Weibull and gamma. In each case, we used the resulting survival function to estimate the EUL. In nearly every case, the resulting EUL was either statistically significantly larger than the ex ante EUL, or was not statistically significantly different than the EUL. In only 1 out of 20 cases was the resulting EUL statistically significantly less than the ex ante EUL. In this case the failure events observed during the study period clearly do not provide adequate information for a reliable estimate.
5. ***Develop competing risks models*** that incorporate different distributions for failures, removals, and replacements. Using the LIFEREG procedure in SAS from step 4 above, separate output was generated for failures, removals, and replacements. Then, the best fitting distributions for each event were combined to form one combined survival function. This additional analysis step provided valuable results that have not been previously utilized in retention studies.

Study Results

The exhibit below presents the final EULs for the studied and like measures. Provided are the ex ante and ex post EULs, the 80 percent confidence intervals for the ex post results, the final EUL used for the filing claim, and the realization rate.

PG&E's 1996 & 1997 Commercial Energy Efficiency Incentives Program
Summary of Ex Post Effective Useful Life Estimates
Lighting and HVAC End Uses[1]

Measure Description	Measure Code	EUL Ex Ante	EUL Ex Post	Upper 80% CL Ex Post	Lower 80% CL Ex Post	EUL for Claim -	Realization Rate -
LIGHTING							
T8 Lamps and Electronic Ballasts							
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 4 FT FIXT	**L23**	**16**	**34**	**121**	**-54**	**16**	**1.0**
FIXTURE: T-8, 4-LAMP, 8 FT FIXTURE	L12	16	-	-	-	16	1.0
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 2 FT FIXT	L21	16	-	-	-	16	1.0
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 3 FT FIXT	L22	16	-	-	-	16	1.0
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 8 FT FIXT	L24	16	-	-	-	16	1.0
FIXTURE: 2 FT T-8 W/EL BLST, 1 31-W T-8 U OR 2 17-W T-8	L69	16	-	-	-	16	1.0
FIXTURE: 2 FT T-8 W/EL BLST, 2 31-W T-8 U OR 4 17-W T-8	L70	16	-	-	-	16	1.0
FIXTURE: 2 FT T-8 W/EL BLST, 3 31-W T-8 U OR 6 17-W T-8	L71	16	-	-	-	16	1.0
FIXTURE: 4 FT T-8 W/ELEC BLST, 1 32-WATT T-8 LAMP	L72	16	-	-	-	16	1.0
FIXTURE: 4 FT T-8 W/ELEC BLST, 2 32-WATT T-8 LAMPS	L73	16	-	-	-	16	1.0
FIXTURE: 4 FT T-8 W/ELEC BLST, 3 32-WATT T-8 LAMPS	L74	16	-	-	-	16	1.0
FIXTURE: 8-FT T-8 W/EL BLST, 2 8-FT T-8 OR 4 32-W, 4-FT T-8	L75	16	-	-	-	16	1.0
FIXTURE: 8-FT T-8 W/EL BLST, 1 8-FT T-8 OR 2 32-W, 4-FT T-8	L160	16	-	-	-	16	1.0
FIXTURE: T-8 HIGH-OUTPUT LAMP & ELEC BLST, (FEM or NEW FIXTURE), 8 FT	L184	16	-	-	-	16	1.0
Optical Reflectors w/ Fluorescent Delamp							
REFLECTORS WITH DELAMPING, 4 FT LAMP REMOVED	**L19**	**16**	**5,031**	**28,376**	**-18,313**	**16**	**1.0**
REFLECTORS WITH DELAMPING, 2 FT LAMP REMOVED	L17	16	-	-	-	16	1.0
REFLECTORS WITH DELAMPING, 3 FT LAMP REMOVED	L18	16	-	-	-	16	1.0
REFLECTORS WITH DELAMPING, 8 FT LAMP REMOVED	L20	16	-	-	-	16	1.0
High Intensity Discharge							
HID FIXTURE: INTERIOR, STANDARD, 251-400 WATT LAMP	**L81**	**16**	**29**	**375**	**-318**	**16**	**1.0**
HID FIXTURE: INTERIOR, STANDARD, 101-175 WATT LAMP	L26	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 176-250 WATT LAMP	L27	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 0-100 WATT LAMP	L28	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 101-175 WATT LAMP	L29	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, >= 176 WATT LAMP	L30	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 36-70 WATT LAMP	L79	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 71-100 WATT LAMP	L80	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 36-70 WATTS LAMP, INCANDESCENT	L187	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 36-70 WATTS LAMP, MERCURY VAPOR	L188	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 71-100 WATTS LAMP, INCANDESCENT	L189	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 71-100 WATTS LAMP, MERCURY VAPOR	L190	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 101-175 WATTS LAMP, INCANDESCENT	L191	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 101-175 WATTS LAMP, MERCURY VAPOR	L192	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 176-250 WATTS LAMP, INCANDESCENT	L193	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 176-250 WATTS LAMP, MERCURY VAPOR	L194	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 251-400 WATTS LAMP, INCANDESCENT	L195	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 251-400 WATTS LAMP, MERCURY VAPOR	L196	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 0-100 WATTS LAMP, INCANDESCENT	L197	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 0-100 WATTS LAMP, MERCURY VAPOR	L198	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 101-175 WATTS LAMP, INCANDESCENT	L199	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 101-175 WATTS LAMP, MERCURY VAPOR	L200	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, >= 176 WATTS LAMP, INCANDESCENT	L201	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, >= 176 WATTS LAMP, MERCURY VAPOR	L202	16	-	-	-	16	1.0
HVAC							
A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	**S160**	**15**	**12**	**35**	**-10**	**15**	**1.0**
A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYSTEM (yr<96)	S1	15	-	-	-	15	1.0
A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SINGLE PACKAGE (yr<96)	S2	15	-	-	-	15	1.0
A/C: CENTRAL, >= 135 & < 760 KBTU/HR, AIR-COOLED, SINGLE PKG (yr<96)	S4	15	-	-	-	15	1.0
A/C: CENTRAL, >= 65 & < 135 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	S161	15	-	-	-	15	1.0
A/C: CENTRAL, >= 135 & < 240 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	S162	15	-	-	-	15	1.0
A/C: CENTRAL, >= 240 & < 760 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	S163	15	-	-	-	15	1.0

* Studied Measures are in Bold.

Regulatory Waivers

There were no regulatory waivers filed for this study.

[1] Although negative EUL values are a physical impossibility, the values are presented so that the reader may understand the magnitude of the standard error.

FOURTH YEAR RETENTION STUDY FOR PG&E'S 1996 & 1997 COMMERCIAL EEI PROGRAM LIGHTING AND HVAC TECHNOLOGIES

Study ID#s: 349R1 & 351R1

FINAL REPORT

March 1, 2001

Submitted to

Rafael Friedmann
Market Planning and Research
Pacific Gas & Electric Co.
245 Market Street
San Francisco, CA 94105-1702

Prepared by

QUANTUM CONSULTING INC.
2030 Addison Street
Berkeley, CA 94704

TABLE OF CONTENTS

TABLE OF CONTENTS

(continued)

LIST OF EXHIBITS

LIST OF EXHIBITS

(continued)

LIST OF EXHIBITS

(continued)

LIST OF EXHIBITS

(continued)

LIST OF EXHIBITS

(continued)

LIST OF EXHIBITS

(continued)

LIST OF EXHIBITS

(continued)

ATTACHMENTS TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

This section presents a summary of the retention study results of Pacific Gas & Electric Company's (PG&E's) Commercial Energy Efficiency Incentive (CEEI) Program for lighting and HVAC technologies. The retention study described in this report covers all HVAC and Lighting technologies installed at commercial accounts, as determined by the Marketing Decision Support System (MDSS) sector code, that were included under the RE, REO, and CI programs and for which rebates were *paid* during calendar year 1996 and 1997.

1.1 PROTOCOL REQUIREMENTS

This study was conducted under the rules specified in the "Protocols and Procedures for the Verification of Cost, Benefits, and Shareholder Earnings from Demand Side Management Programs" (the Protocols).[1] This evaluation has endeavored to meet all Protocol requirements.

The retention study results in ex post effective useful lives for each lighting and HVAC measure, and a comparison of realization rates from the ex ante to ex post estimates. The definition of the effective useful life, provided in Appendix A, Measurement Terms and Definitions, of the Protocols is: "an estimate of the median number of years that the measures installed under the program are still in place and operable".

Although there are dozens of measures installed under the Lighting and HVAC programs, the Protocols only require a subset of the measures be studied. The Protocols require the utilities to study either "the top ten measures, excluding measures that have been identified as miscellaneous (per Table C-9), ranked by the net resource value or the number of measures that constitutes the first 50% of the estimated resource value, whichever number of measures is less". For consistency, we will refer to the studied measures as the "Top 50% Measures" throughout this report.

The Protocols state that "measures not included in the … retention studies will be divided into two groups: 'like measures' and 'other measures.' Like measures are defined by the Protocols as measures that are believed to be similar to measures included in the retention studies. We have classified all groups of like measures with similar applications, operating conditions, and operating loads.

Exhibit 1-1 presents the list of studied measures and associated like measures covered under this retention study. In addition, Exhibit 1-1 provides the percent of net resource benefit attributable to each studied measure.

[1] California Public Utilities Commission Decision 93-05-063, Revised March 1998, 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.

Exhibit 1-1
Mapping of Like Measures

Program and Technology Group	Studied Measures	Percent of Total Net Resource Benefit		Measure Grouping
		1996	1997	Like Measures
LIGHTING END USE				
Retrofit Express Program				
T8 Lamps and Electronic Ballasts	L23	31%	34%	L9 - L12, L21, L22, L24, L69 - L75, L117 - L124, L160, L13, L112
Optical Reflectors w/ Fluor. Delamp	L19	16%	18%	L17, L18, L20, L76 - L77
High Intensity Discharge	L81	7%	3%	L25, L78 - L80, L26, L27
HVAC END USE				
Retrofit Express Program				
Central A/C	S160	1%	5%	S1, S2, S4, S160 - S163

1.2 STUDY APPROACH OVERVIEW

As stated above, the Protocols assert the purpose of a retention study is to collect data on the fraction of installed measures in place and operable in order to produce a revised estimate of its EUL. The ultimate goal is to estimate the EUL (or the median number of years that the measure is still in place and operable), which can be realized by identifying the measure's survival function. For this study, the survival function describes the percentage of measures installed that are still operable and in place at a given time. Survival analysis is the process of analyzing empirical failure / removal data in order to model a measure's survival function. As much as possible, we have attempted to employ classical survival analysis techniques to our study approach.

For this study, the vast majority of measures were in place less than five years (few were installed prior to 1996, and follow-up data collection was conducted no later than the end of 2000). Because the ex ante EUL is 15 or 16 years for the studied measures, it is very unlikely that our data will be capable of accurately estimating the survival function for the studied measures.

Our overall approach consists of five analysis steps that were used to estimate each of the studied measures' EULs:

1. ***Compile summary statistics*** on the raw retention data. Upon review of the summary statistics, it became clear that such a small percentage of failures and removals had occurred, that it would be difficult to model the equipment's survival function.

2. ***Visually inspect*** the retention data, by simply calculating the cumulative percentage of equipment that had failed in a given month, and plotting the percentage over time. This step clearly illustrated that for each studied measure, there was not enough data over time to support an accurate estimate of the survival function.

3. ***Develop a trend line*** from the survival plots. Using the plots developed in (2) above, a trend line was estimated using standard linear regression techniques. We modeled the trend as a linear and an exponential function. In each case, we used the resulting trend

line to estimate the EUL, which was statistically significantly larger than the ex ante estimate.

4. ***Develop a survival function*** using classical survival techniques. We modeled the survival function assuming five of the most common survival distributions: exponential, logistic, lognormal, Weibull and gamma. In each case, we used the resulting survival function to estimate the EUL. In nearly every case, the resulting EUL was either statistically significantly larger than the ex ante EUL, or was not statistically significantly different than the EUL. In only 1 out of 20 cases was the resulting EUL statistically significantly less than the ex ante EUL. In this case the failure events observed during the study period clearly do not provide adequate information for a reliable estimate.

5. ***Develop competing risks models*** that incorporate different distributions for failures, removals, and replacements. Using the LIFEREG procedure in SAS from step 4 above, separate output was generated for failures, removals, and replacements. Then, the best fitting distributions for each event were combined to form one combined survival function. This additional analysis step provided valuable results that have not been previously utilized in retention studies.

1.3 STUDY RESULTS

For all studied measures but the L19 Delamping measure, all five approaches discussed above were implemented. The L19 Delamping measure was not put through the competing risks model because there was only one failure type observed during the study period.

The EUL analyses based on current accepted methods are unable to provide results that conclusively indicate that the ex-ante EULs should be modified due to the minute number of "non-operating" equipment. We present here the results of the various currently accepted methods as required by the Protocols, we report a study result, by selecting the best-fit approach as our recommended result.

The recommended results are based on the competing risks model, except for the L19 delamping measure which was based on the best fitting LIFEREG procedure results. Of the three models created, the best fit model is the model of choice. This model is based upon the combination of unique distributions for each event type chosen based upon the maximum of the Log-likelihood estimate generated during the LIFEREG procedure in SAS.

Exhibit 1-2 presents the recommended ex post estimates of the EUL. Because the best fit competing risks model did not provide results that were statistically significantly different from the ex ante results, measured at the 80 percent confidence interval, all of the ex post EULs are based on the ex ante estimates. Also presented are the final study results, and the corresponding upper and lower 80 percent confidence interval. Finally, the program realization rates are provided, which are the ratios of the ex ante and ex post estimates. For all measures, the realization rate is one; i.e., the ex-post EULs fully corroborate using the ex-ante EUL values.

Exhibit 1-2
Final Ex Post EUL Estimates[2]

End Use	Technology	Measure	Ex Ante	Study Results Upper	Median	Lower	Ex Post	Realization Rate
Lighting	T8 Lamps and Electronic Ballasts	L23	16	121	34	-54	16	100%
	Optical Reflectors w/ Fluor. Delamp	L19	16	28,376	5,031	-18,313	16	100%
	High Intensity Discharge	L81	16	375	29	-318	16	100%
HVAC	CAC	S160	15	35	12	-10	15	100%

Exhibit 1-3 presents the final EULs for the studied and like measures. Provided are the ex ante and ex post EULs, the 80 percent confidence intervals for the ex post results, the final EUL used for the filing claim, and the realization rate.

[2] Although negative EUL values are a physical impossibility, the values are presented so that the reader may understand the magnitude of the standard error.

Exhibit 1-3
Final EUL Estimates
For Studied and Like Measures[3]

	Measure	EUL		Upper 80% CL	Lower 80% CL	EUL for Claim	Realization Rate
Measure Description	**Code**	**Ex Ante**	**Ex Post**	**Ex Post**	**Ex Post**	**-**	**-**
LIGHTING							
T8 Lamps and Electronic Ballasts							
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 4 FT FIXT	**L23**	**16**	**34**	**121**	**-54**	**16**	**1.0**
FIXTURE: T-8, 4-LAMP, 8 FT FIXTURE	L12	16	-	-	-	16	1.0
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 2 FT FIXT	L21	16	-	-	-	16	1.0
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 3 FT FIXT	L22	16	-	-	-	16	1.0
FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 8 FT FIXT	L24	16	-	-	-	16	1.0
FIXTURE: 2 FT T-8 W/EL BLST, 1 31-W T-8 U OR 2 17-W T-8	L69	16	-	-	-	16	1.0
FIXTURE: 2 FT T-8 W/EL BLST, 2 31-W T-8 U OR 4 17-W T-8	L70	16	-	-	-	16	1.0
FIXTURE: 2 FT T-8 W/EL BLST, 3 31-W T-8 U OR 6 17-W T-8	L71	16	-	-	-	16	1.0
FIXTURE: 4 FT T-8 W/ELEC BLST, 1 32-WATT T-8 LAMP	L72	16	-	-	-	16	1.0
FIXTURE: 4 FT T-8 W/ELEC BLST, 2 32-WATT T-8 LAMPS	L73	16	-	-	-	16	1.0
FIXTURE: 4 FT T-8 W/ELEC BLST, 3 32-WATT T-8 LAMPS	L74	16	-	-	-	16	1.0
FIXTURE: 8-FT T-8 W/EL BLST, 2 8-FT T-8 OR 4 32-W, 4-FT T-8	L75	16	-	-	-	16	1.0
FIXTURE: 8-FT T-8 W/EL BLST, 1 8-FT T-8 OR 2 32-W, 4-FT T-8	L160	16	-	-	-	16	1.0
FIXTURE: T-8 HIGH-OUTPUT LAMP & ELEC BLST, (FEM or NEW FIXTURE), 8 FT	L184	16	-	-	-	16	1.0
Optical Reflectors w/ Fluorescent Delamp							
REFLECTORS WITH DELAMPING, 4 FT LAMP REMOVED	**L19**	**16**	**5,031**	**28,376**	**-18,313**	**16**	**1.0**
REFLECTORS WITH DELAMPING, 2 FT LAMP REMOVED	L17	16	-	-	-	16	1.0
REFLECTORS WITH DELAMPING, 3 FT LAMP REMOVED	L18	16	-	-	-	16	1.0
REFLECTORS WITH DELAMPING, 8 FT LAMP REMOVED	L20	16	-	-	-	16	1.0
High Intensity Discharge							
HID FIXTURE: INTERIOR, STANDARD, 251-400 WATT LAMP	**L81**	**16**	**29**	**375**	**-318**	**16**	**1.0**
HID FIXTURE: INTERIOR, STANDARD, 101-175 WATT LAMP	L26	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 176-250 WATT LAMP	L27	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 0-100 WATT LAMP	L28	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 101-175 WATT LAMP	L29	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, >= 176 WATT LAMP	L30	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 36-70 WATT LAMP	L79	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 71-100 WATT LAMP	L80	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 36-70 WATTS LAMP, INCANDESCENT	L187	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 36-70 WATTS LAMP, MERCURY VAPOR	L188	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 71-100 WATTS LAMP, INCANDESCENT	L189	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, COMPACT, 71-100 WATTS LAMP, MERCURY VAPOR	L190	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 101-175 WATTS LAMP, INCANDESCENT	L191	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 101-175 WATTS LAMP, MERCURY VAPOR	L192	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 176-250 WATTS LAMP, INCANDESCENT	L193	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 176-250 WATTS LAMP, MERCURY VAPOR	L194	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 251-400 WATTS LAMP, INCANDESCENT	L195	16	-	-	-	16	1.0
HID FIXTURE: INTERIOR, STANDARD, 251-400 WATTS LAMP, MERCURY VAPOR	L196	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 0-100 WATTS LAMP, INCANDESCENT	L197	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 0-100 WATTS LAMP, MERCURY VAPOR	L198	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 101-175 WATTS LAMP, INCANDESCENT	L199	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, 101-175 WATTS LAMP, MERCURY VAPOR	L200	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, >= 176 WATTS LAMP, INCANDESCENT	L201	16	-	-	-	16	1.0
HID FIXTURE: EXTERIOR, >= 176 WATTS LAMP, MERCURY VAPOR	L202	16	-	-	-	16	1.0
HVAC							
A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	**S160**	**15**	**12**	**35**	**-10**	**15**	**1.0**
A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYSTEM (yr<96)	S1	15	-	-	-	15	1.0
A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SINGLE PACKAGE (yr<96)	S2	15	-	-	-	15	1.0
A/C: CENTRAL, >= 135 & < 760 KBTU/HR, AIR-COOLED, SINGLE PKG (yr<96)	S4	15	-	-	-	15	1.0
A/C: CENTRAL, >= 65 & < 135 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	S161	15	-	-	-	15	1.0
A/C: CENTRAL, >= 135 & < 240 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	S162	15	-	-	-	15	1.0
A/C: CENTRAL, >= 240 & < 760 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	S163	15	-	-	-	15	1.0

* Studied Measures are in Bold.

3 Although negative EUL values are a physical impossibility, the values are presented so that the reader may understand the magnitude of the standard error.

The remainder of this report will present our analysis methodology and results. Due to the lack of observed events over the study period, several analysis methods were implemented. The results from each of the methods are presented, which is quite lengthy. This is an attempt to illustrate the effort that was exerted to obtain as much useful information from the empirical data as possible. The conclusion is that no matter what method was selected, the results still would not reject the ex ante EUL this early in the life of the measure.

2. INTRODUCTION

This report summarizes the retention study of Pacific Gas & Electric Company's (PG&E's) Commercial Energy Efficiency Incentive (CEEI) Program for Lighting and HVAC technologies. The evaluation effort includes customers who were paid rebates in 1996 and 1997. Technologies installed under the paid year 1996 and 1997 CEEI Program were covered by three separate program options: the Retrofit Express (RE) Program, the Retrofit Efficiency Options (REO) Program and the Customized Incentives (CI) Program.

2.1 THE RETROFIT EXPRESS PROGRAM

The RE program offered fixed rebates to customers who installed specific electric energy-efficient equipment. The program covered the most common energy saving measures and spans lighting, air conditioning, refrigeration, motors, and food service. Customers were required to submit proof of purchase with these applications in order to receive rebates. The program was marketed to small- and medium-sized commercial, industrial, and agricultural customers. The maximum rebate amount, including all measure types, was $300,000 per account. No minimum amount was required to qualify for a rebate.

Lighting and HVAC end-use rebates were offered in the program for the following technologies:

Lighting Technologies

Halogen lamps

Compact fluorescent lamps

T-12 and T-8 fluorescent lamps

Compact fluorescent lamps and LED's

Electronic ballasts

T-8 and T-10 lamps and electronic ballasts

High-intensity discharge (HID) fixtures

Occupancy sensors, bypass or delay timers, photocells, and time clock controls

Removal of lamps and ballasts

HVAC Technologies

High-efficiency central air-conditioning units in various capacity ranges

Variable speed drive HVAC fans

High-efficiency package terminal air-conditioning units

Programmable thermostats, bypass timers, and electronic timeclocks

Reflective window film

Water chillers of various capacity ranges

Direct evaporative cooler units, evaporative condensers, and evaporative cooler towers

2.2 THE RETROFIT EFFICIENCY OPTIONS PROGRAM

The REO program targeted commercial, industrial, agricultural, and multi-family market segments most likely to benefit from these selected measures. Customers were required to submit calculations for the projected first-year energy savings along with their application prior to installation of the high efficiency equipment. PG&E representatives worked with customers to identify cost-effective improvements, with special emphasis on operational and maintenance measures at the customers' facilities. Marketing efforts were coordinated amongst PG&E's divisions, emphasizing local planning areas with high marginal electric costs to maximize the program's benefits.

The REO program did not include any Lighting measures. Nine HVAC technologies, however, were included, which can be summarized into four general technology groups, described below:

Technology

Variable frequency drive supply fans

Installation of high efficiency water chillers

Variable air volume supply systems, which replace constant air volume supply systems

Evaporative cooling towers

2.3 THE CUSTOMIZED INCENTIVES PROGRAM

The Customized Incentives program offered financial incentives to CIA customers who undertook large or complex projects that save gas or electricity. These customers were required to submit calculations for projected first-year energy impacts with their applications prior to installation of the project. The maximum incentive amount for the Customized Incentives program was $500,000 per account, and the minimum qualifying incentive was $2,500 per project. The total incentive payment for kW, kWh, and therm savings was limited to 50 percent of direct project cost for retrofit of existing systems. Since the program also applied to expansion projects, the new systems incentive was limited to 100 percent of the incremental cost to make new processes or added systems energy efficient. Customers were paid 4¢ per kWh and 20¢ per therm for first-year annual energy impacts. A $200 per peak kW incentive for peak demand impacts required that savings be achieved during the hours PG&E experiences high power demand.

As a result of program design, the measures installed were similar to or the same as those for the RE program, but were installed in larger and more complex projects. Customers were also able to participate under the APOS program. The Lighting measures are the same as those described above for the RE program. For HVAC, the following technologies were rebated in 1996 and 1997:

Technology
HVAC variable speed drive
High efficiency chiller
Energy Management Systems (EMS)

Other miscellaneous Customized Incentives HVAC measures, which included:

- Installation of various energy efficient motors
- Installation of various HVAC controls
- Various technologies (i.e., precoolers and economizers) added to increase overall system efficiency

2.4 STUDY REQUIREMENTS

The retention study described in this report covers all HVAC and Lighting technologies installed at commercial accounts, as determined by the Marketing Decision Support System (MDSS) sector code, that were included under the RE, REO, and CI programs and for which rebates were *paid* during calendar year 1996 and 1997.

This study was conducted under the rules specified in the "Protocols and Procedures for the Verification of Cost, Benefits, and Shareholder Earnings from Demand Side Management Programs" (the Protocols).[1] This evaluation has endeavored to meet all Protocol requirements.

[1] California Public Utilities Commission Decision 93-05-063, Revised March 1998, 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.

The retention study results in ex post effective useful lives for each Lighting and HVAC measure, and a comparison of realization rates from the ex ante to ex post estimates. The definition of the effective useful life, provided in Appendix A, Measurement Terms and Definitions, of the Protocols is:

Effective Useful Life (EUL) – An estimate of the median number of years that the measures installed under the program are still in place and operable.

2.4.1 Studied Measures

Although there are dozens of measures installed under the Lighting and HVAC programs, the Protocols only require a subset of the measures be studied. The Protocols refer to the studied measures as the "Top 10 or Top 50% Measures", which is defined as:

Top 10 or Top 50% Measures – The utility should select the top ten measures, excluding measures that have been identified as miscellaneous (per Table C-9), ranked by the net resource value or the number of measures that constitutes the first 50% of the estimated resource value, whichever number of measures is less.

For the 1996 and 1997 CEEI Program, the number of measures that constitutes the first 50% of the estimated resource value is only three. For consistency, we will refer to these measures throughout the report as the "Top 50% Measures."

For the 1996 and 1997 CEEI Program, HVAC and Lighting comprise the studied end-uses. Among these end-uses, the following four measures shown in Exhibit 2-1 are identified as the "Top 50% Measures", as defined above.

Exhibit 2-1
Top 50% Measures for Paid Year 1996 and 1997

Paid Year	Top 10 Measures	Measure Description	% of Total (Lighting and HVAC) Avoided Cost	Cumulative % of Total
1996	L23	FIXTURE: MODIFICATION / REPLACE LAMPS & BLST, 4 FT FIXTURE	31%	31%
	L19	FIXTURE: MODIFICATION / LAMP REMOVAL, 4 FT LAMP REMOVED	16%	47%
	L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	7%	54%
1997	L23	FIXTURE: MODIFICATION / REPLACE LAMPS & BLST, 4 FT FIXTURE	34%	34%
	L19	FIXTURE: MODIFICATION / LAMP REMOVAL, 4 FT LAMP REMOVED	18%	52%
	L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	3%	55%

In addition to studying the measures identified in Exhibit 2-1, PG&E decided to study one additional HVAC measure, shown in Exhibit 2-2. Adding this measure brings the cumulative net resource benefit up to 55% in 1996 and 60% in 1997.

Exhibit 2-2
Other Studied Measures for Paid Year 1996 and 1997

Paid Year	Other Measures	Measure Description	% of Total (Lighting and HVAC) Avoided Cost	Cumulative % of Total
1996	S160	A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	1%	55%
1997	S160	A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	5%	60%

2.4.2 Like Measures

The Protocols state that "measures not included in the … retention studies will be divided into two groups: 'like measures' and 'other measures.' Like measures are defined by the Protocols as:

> **Like Measures** – measures that are believed to be similar to measures included in the retention studies.

We have classified all groups of like measures with similar applications, operating conditions, and operating loads. Exhibit 2-3 presents the mapping of studied measures to like measures.

Exhibit 2-3
Mapping of Like Measures

Program and Technology Group	Measure Grouping: Studied Measures	Measure Grouping: Like Measures
LIGHTING END USE		
Retrofit Express Program		
T8 Lamps and Electronic Ballasts	L23	L9 - L12, L21, L22, L24, L69 - L75, L117 - L124, L160, L13, L112
Optical Reflectors w/ Fluor. Delamp	L19	L17, L18, L20, L76 - L77
High Intensity Discharge	L81	L25, L78 - L80, L26, L27
HVAC END USE		
Retrofit Express Program		
Central A/C	S160	S1, S2, S4, S160 - S163

The Protocols require that "like measures adopt the same percent adjustment [or realization rate] for the measure effective useful lives of the similar studied measures . . . to adjust their ex ante measure effective useful lives."

Other measures are defined as:

Other Measures – measures that are different from the measures included in the retention study.

Therefore, other measures consist of all HVAC and Lighting measures that are not classified as either studied or like measures. The Protocols require that, for other measures, the ex ante estimate of the effective useful life will be adjusted by the average percentage adjustment [or realization rate] of all the studied measures within that end use."

2.4.3 Combining Program Years

The Protocols also require that two Program Years, 1996 and 1997, be combined and that the studies be conducted on the schedule for Program Year 1996. The Protocols state that combining the two studies "should increase the accuracy of the survival function and decrease the cost of completing the retention studies." Furthermore, "the retention studies shall include data from participant groups from two or more sequential years to increase the robustness of the sample and to allow for the estimation of a survival function for a number of different measures."

Because the Top 50% Measures for the 1997 Program Year are a subset of the 1996 Top 50% Measures, the Protocol's suggestion to combine the two studies will greatly enhance the accuracy of the retention study, without incurring additional cost.

2.4.4 Accepting Ex Post EULs

The Protocols state that "the estimated ex post measure EULs that result from the retention study will be compared to the ex ante EUL estimates. Hypothesis testing procedures will be used to determine if the estimated ex post measure EUL is statistically significantly different from the ex ante measure EUL. If the estimated ex post measure EUL is significantly different than the ex ante measure EUL, the estimated ex post measure EUL will be used. Otherwise, the ex ante estimate will continue to be used. Hypothesis testing will be conducted at the 20% significance level."

2.4.5 Objectives

The research objectives are therefore as follows:

- Collect data on the fraction of the measures that are in place and operable, for all studied measures.

- For each studied measure, calculate the ex post EUL, and the realization rates from ex ante to ex post.

- For each like measure, calculate the ex post EUL, based on a transferred realization rate from the studied measures.

- For each remaining HVAC and Lighting measure, calculate the ex post EUL, based on the average realization rate from all studied and like measures.

- Complete tables 6 and 7 of the Protocols.

2.5 STUDY APPROACH OVERVIEW

As stated above, the Protocols assert the purpose of a retention study is to collect data on the fraction of installed measures in place and operable in order to produce a revised estimate of its EUL. The ultimate goal is to estimate the EUL (or the median number of years that the measure is still in place and operable), which can be realized by identifying the measure's survival function. For this study, the survival function describes the percentage of measures installed that are still operable and in place at a given time. At any given time, the hazard rate is the rate at which measures fail or are removed. Survival analysis is the process of analyzing empirical failure / removal data in order to model a measure's survival function. As much as possible, we have attempted to employ classical survival analysis techniques to our study approach.

Our overall approach was to apply survival analysis to our collected retention data in order to develop a survival function for each of the studied measures. Some of the common survival functions take on the logistic cumulative distribution function. Although there is no documentation to support the ex ante survival function assumptions, discussions with the authors of the Protocols indicated that the ex ante EULs are based on a logistic survival function.

However, the form of the logistic survival function assumed by the Protocol authors is *not* the commonly used form of the logistic model. Generally, in survival analysis, the log-logistic model is used, which is a special form of the logistic distribution. It is this distribution that we used in our analysis. Other commonly used survival functions are based on the exponential, Weibull, lognormal, and gamma distributions. For this retention study, we have examined each of these distributions. We have used the SAS System and the SAS companion guide, "Survival Analysis Using the SAS System[2]," in order to estimate the survival functions based on the retention data for each of our studied measures.

An important issue to keep in mind for this analysis is the definition of survival. Recall that the EUL is defined as the median number of years that the measures installed under the program are still in place and operable. Therefore, to "survive", a measure must not have been removed or have failed. Unfortunately, it is likely that the underlying distribution of measures having failed is very different than the distribution of removals.

There is much literature to suggest, for example, that electronic ballast failures follow an exponential distribution. The exponential survival function has a constant hazard rate. In

[2] Allison, Paul D., "Survival Analysis Using the SAS System, A Practical Guide", SAS Institute, NC, 1997.

other words, the rate at which electronic ballasts fail is constant over time. This belief is founded on the fact that electronic devices are likely to fail at any point in time with equal probability. Because electronic ballasts may have anywhere from 30 to 120 parts, plus more than twice as many solder joints as there are parts, it is likely that the ballast may also fail at any point in time, with equal probability.[3]

However, the removal of an electronic ballast is more dependent on human interaction. For example, consider the act of remodeling, or upgrading the system as new technologies emerge. Both of these actions are likely to occur in the latter stage of the equipment's life. However, if the customer is not satisfied with the technology, the removal may occur early on in the equipment's life. Whatever the case may be, it is likely that the survival function of equipment removal differs from the survival function of the equipment failure.

These reasons have led us to develop a competing risks model that accounts for varying distributions for each event type. The LIFEREG procedure in SAS is used to generate output for each unique event type (failures, removals, and replacements). This output is then used to generate a competing risks model that produces a survival function that is comprised of the best fitting distribution for each event type.

For this study, the vast majority of measures were in place less than five years (few were installed prior to 1996, and follow-up data collection was conducted no later than the end of 1998). Because the ex ante EUL is 15 or 16 years for the studied measures, it was unlikely from the start that our data would be capable of accurately estimating this joint probability density function of failures and removals.

Our overall approach consists of four analysis steps that were used to estimate each of the studied measures' EULs:

1. ***Compile summary statistics*** on the raw retention data. For some measures, it was sufficient to only look at the raw data, because for some measures, all of the sampled equipment was still in place and operable.

2. ***Visually inspect*** the retention data. By calculating the cumulative percentage of equipment that had failed in a given month, and plotting this percentage over time, an empirical survival function emerges.

3. ***Develop a trend line*** from the survival plots. Using the plots developed in (2) above, we estimated a trend line using standard linear regression techniques. We modeled the trend as a linear and an exponential function. In each case, we plotted the resulting trend line and visually compared it to the survival plot developed in (2). Furthermore, we used the resulting trend line to estimate the EUL.

4. ***Develop a survival function*** using classical survival techniques. Using the SAS System and the SAS companion guide, "Survival Analysis Using the SAS System," we modeled the survival function assuming five of the most common survival distributions:

[3] Energy User News, Vol. 23 No. 10, October 1998. Electronics, Energy Products and Life-Cycle Costing, pp. 28.

exponential, logistic, lognormal, Weibull and gamma. In each case, we plotted the resulting distribution and visually compared it to the survival plot developed in (2). Furthermore, we used the resulting survival function to estimate the EUL.

5. ***Develop competing risks models*** that incorporate different distributions for failures, removals, and replacements. Using the LIFEREG procedure in SAS from step 4 above, separate output was generated for failures, removals, and replacements. Then, the best fitting distributions for each event were combined to form one combined survival function. This additional analysis step provided valuable results that have not been previously utilized in retention studies.

The details surrounding each of these steps are provided in Section 3.

2.6 REPORT LAYOUT

This report is divided into four sections, plus attachments. *Sections 1 and 2* are the *Executive Summary* and the *Introduction*. *Section 3* presents the *Methodology* of the evaluation. *Section 4* presents the detailed results and a discussion of important findings. *Attachment 1* provides copies of the Lighting and HVAC retention audit instruments. *Attachment 2* includes retention sample design memos that have been drafted for the CADMAC Subcommittee on Persistence. Finally, *Attachment 3* provides the Protocol Tables 6B and 7B.

3. METHODOLOGY

This section provides the specifics surrounding the methods used to conduct the Retention Study for the 1996 and 1997 Pacific Gas & Electric Company (PG&E) Commercial Energy Efficiency Incentive (CEEI) Programs. It begins with a detailed discussion on the sampling plan for the Retention Study. From there, details regarding the study methodology are presented, along with intermediate results from each of the five approaches implemented.

3.1 SAMPLE DESIGN

3.1.1 Existing Data Sources

PG&E's 1996 and 1997 first year CEEI program impact evaluations established "retention panels" of approximately 350 sites in 1996 and 250 sites in 1997 for the Lighting and HVAC end uses. At each of these sites the rebated equipment was documented by make, model, and location. The total combined data collection effort resulted in a panel of over 300 Lighting, and over 350 HVAC sites.

Exhibit 3-1 provides the available sample frame for each studied measure. The studied measures comprise the four Top 50% Measures which PG&E agreed to study. Nearly every site in the Lighting sample installed at least one of the studied measures. The HVAC sample, however, includes only 118 sites with the studied measure.

3.1.2 Sample Design Overview

As discussed in Section 2, the Protocols require that the Retention Study for the 1996 Paid Year Program combine the retention panel data collected for both the 1996 and 1997 Programs. Although the Protocols provide no requirement on sample size or expected relative accuracy for retention studies, they do require that the ex post estimates of EUL be statistically significantly different than the ex ante estimate, measured at the 80% confidence level, in order to accept the ex post estimate.

Therefore, the sample should be designed in such a manner that if the ex ante and ex post estimates were different, that the ex post estimate would be estimated accurately enough to reject the ex ante estimate at the 80% confidence level. This criteria alone is not sufficient to develop a sample. To do so, one would need to know the underlying distribution of the ex post estimate, and by how much the two means are expected to differ. Furthermore, the sample size that would be calculated would indicate the number of failures or removals needed to be observed, not the number of sites visited. Therefore, another component to this estimate would be the expected rate of failure/removals that would occur per site visited.

To complicate things even more, the unit of analysis for the retention study is not a site, but a unit of measure. For example, for lighting measures, the unit of analysis is generally a ballast. For air conditioners, the unit of analysis is tons. Therefore, a single site may consist of hundreds, or even thousands of units. In this case, each sample unit is not independent of the others. Therefore, the procedures for calculating required sample size is even more complicated.

This has been a major topic of interest for the CADMAC Persistence Subcommittee. Attachment 2 contains a few documents that discuss required sample sizes under certain conditions. The general consensus was that a sample of 30 or so site surveys should be sufficient.

We found that our sample frames were relatively limited for some of our measures, such that obtaining 30 completed follow-up surveys may not even be possible. For example, of our four measures, two had a sample frame greater than 100 sites. In an effort to increase the available sample for this and future retention studies, we chose to augment the sample with sites that had not been visited during the first year impact study. Exhibit 3-1 summarizes the available sample from all sources for each studied measure.

Exhibit 3-1
Available Sample Frame by Studied Measure

Measure	Measure Description	Sites in Retention Panel	Sites Added from MDSS	Total Number of Available Sites
L23	FIXTURE: T-8 LAMP & ELEC BLST, (FEM or NEW FIXTURE), 4 FT FIXT	204	4076	4280
L19	REFLECTORS WITH DELAMPING, 4 FT LAMP REMOVED	65	1066	1131
L81	HID FIXTURE: INTERIOR, STANDARD, 251-400 WATT LAMP	3	211	214
S160	A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	115	1109	1224
Total		387	6462	6849

Our sample design goal was to exceed 30 completed audits for each of the four studied measures. For the measures that had a sufficient sample frame, our goal was to conduct approximately 100 completed audits. For measures with smaller sample frames, our goal was to conduct a census, or up to 50 completed audits if possible.

3.1.3 Final Distribution

Exhibit 3-2 provides the final sample disposition. Shown is the number of sites available in the sample frame, the number of sites surveyed, and the number of surveyed sites that had at least one failure or removal. In addition, we have shown the number of units installed across all sites in both the sample frame and in the completed surveys.

Our sample design goal was met for all four measures. Over 100 audits were completed among the measures with sufficient sample frames (L23 T8, L19 Delamping and S160 Central Air Conditioning) and over 50 audits for the L81 HID 251-400W measure.

Exhibit 3-2
Final Sample Disposition

Top 10 Measures	Measure Description	Retention Unis	Units in Retention Panel	Sites in Retention Panel	Sample Strategy	Units Contacted	Sites Contacted	Sites Contacted with Failures
L23	FIXTURE: MODIFICATION / REPLACE LAMPS & BLST, 4 FT FIXTURE	Ballast	313,749	4,280	122	10,450	134	40
L19	FIXTURE: MODIFICATION / LAMP REMOVAL, 4 FT LAMP REMOVED	Lamp	288,619	1,131	102	6,343	87	8
L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	Fixture	3,511	214	52	1,804	53	10
S160	A / C: CENTRAL, < 65 KBTU / HR, AIR-COOLED, SPLIT-SYS / SNGL PKG	Ton	14,336	1,224	106	2,507	118	2

3.1.4 Data Collection Strategy

The data collection effort surrounding the survival analysis included a combination of telephone and on-site surveys. When possible, these data were gathered using telephone surveys, with alternate data collection using on-site audits where installations were too complex to be supported by self-reported data. Roughly half of the survival analysis surveys were conducted over the telephone, with the other half requiring an on-site visit. In general, on-sites were required for many of the lighting end use installations, while HVAC equipment survival was more readily verified using the telephone interview only. The following outlines the data collection procedures:

A QC auditor contacted each site by telephone to assess whether an on-site audit was necessary, or if a telephone survey would suffice. If the QC auditor determined that the information could be obtained over the telephone, he conducted the telephone survey immediately, or at the customer's earliest convenience. If an on-site audit was deemed necessary, and the participant was willing, the auditor scheduled an appointment and visited the site.

Equipment survival data were collected by the QC auditor, who prompted each site contact to locate the retention technologies using information available from the retention panels. At that time, information was recorded regarding the success or failure in locating the panel-specified equipment.

For each unit of equipment in the retention panel, it was determined whether (1) the equipment was still installed, and (2) if it was operable. If the equipment was not in place or was not operable, it was determined when it was removed or stopped operating according to the owner or operators best recollection. Reasons for removal or failure to operate were also collected. If equipment was replaced, it was determined if the equipment was replaced with a standard, equivalent or higher efficiency technology. Finally, it was determined if replaced equipment was done so under warranty.

3.2 ANALYSIS OVERVIEW

As discussed in Section 2.4, the purpose of a retention study is to collect data on the fraction of measures in place and operable in order to produce a revised estimate of its EUL. The desired result of our approach was to apply survival analysis to our collected retention data in order to develop a survival function for each of the studied measures. However, because our retention data only includes information over the first few years of the measures' lives (which are expected to have median lives of 15-16 years), we were concerned that our data would not support an accurate estimation of a survival function.

Before attempting to estimate a survival function for a given measure, we first evaluated the data collected to see if there was enough data to support an estimate. For this step, for each studied measure, we compiled summary statistics on the raw retention data, and visually inspected the empirical survival function that we observed over the first three to four years.

Next we used the empirical survival function to forecast the survival function using basic linear regression techniques. We analyzed both a linear trend, as well as an exponential trend (which is one of the most common forms of a survival function).

Next, we used classical survival analysis techniques to develop a survival function. This analysis was performed using the SAS System and the SAS companion guide, "Survival Analysis Using the SAS System." As part of this step, we attempted to model the survival function using five of the most commonly used survival distributions: exponential, logistic, lognormal, Weibull and gamma.

Finally, we constructed Competing Risks models that modeled each event with a different distribution. Three different scenarios were developed for each measure: a best-fit model that matched the best fitting distributions (based on the Log-likelihood estimator in SAS), a minimum EUL model, and a maximum EUL model. For example, one distribution may be the most appropriate model for failures, while a second distribution may better represent removals and yet a third distribution may better represent replacements. Statistical methods are employed to determine which distribution best fits the data.

Our overall approach consists of five analysis steps that were used to estimate each of the studied measures' EULs:

1. ***Compile summary statistics*** on the raw retention data.

2. ***Visually inspect*** the retention data.

3. ***Develop a trend line*** from the survival plots.

4. ***Develop a survival function*** using classical survival techniques.

5. ***Develop competing risks models*** that model each event with a different distribution.

The details surrounding each of these methods are provided below.

3.3 SUMMARY STATISTICS

As discussed above, the first step of our analysis was to compile summary statistics on the sample retention data. For each measure in our sample, these statistics include:

- the number of units installed at the site (as documented in the original retention panel);
- the number of units still operable and in place;
- the number of units that had failed, been removed and been replaced;
- the number of failed units that had been replaced under warranty;
- the percentage of units that had failed, been removed or been replaced; and
- the ex ante EUL.

The CADMAC has agreed that failed equipment that is replaced under warranty should be counted as if it is still operable and in place. Exhibit 3-3 summarizes this data at the measure level.

Exhibit 3-3
Summary Statistics on Retention Sample Data

End Use	Technology	Measure	Number of Sites Contacted	Units	Total Number of Units	Number of Units that Failed, were Removed, or Replaced	Number of Units Replaced Under Warranty	Number of Units in Place and Operable	Percent Failed, Removed, Replaced
Lighting	T8 Lamps and Electronic Ballasts	L23	134	Ballasts	10,450	308	40	10,182	2.56%
	Optical Reflectors w/ Fluor. Delamp	L19	87	Lamps	6,343	48	0	6,295	0.76%
	High Intensity Discharge	L81	53	Fixtures	1,804	34	0	1,770	1.88%
HVAC	CAC	S160	118	Tons	2,507	12	0	2,495	0.48%

Exhibit 3-3 clearly demonstrates that for the L19 Delamping and S160 Central Air Conditioning measures, it will be difficult to develop a survival function or an ex post EUL estimate. Both of these measures exhibited a very small number of failures or removals in the sample. We would suggest, for future retention studies, that further analysis be conducted only on those measures that exhibit a significant number of events such that meaningful ex post EUL estimates may be obtained.

Despite the lack of failure/removal data, we had enough data on failures to proceed to the next analysis step. However, due to the minimum amount of failures presented in Exhibit 3-3, we will likely obtain ex post estimates of the EUL that greatly exceed the ex ante.

If we make the assumption that the failure/removal rates provided in Exhibit 3-3 are constant over time, then our survival function would take on the exponential distribution, which is one of the most commonly used distributions in survival analysis. Assuming the failures/removals occurred over a three year period (which is conservative), we can estimate the median EUL. Exhibit 3-4 provides the estimated EULs based on these assumptions.

Exhibit 3-4
Illustrative Ex Post EUL Estimates
Based on Exponential Distribution and Conservative Assumptions

End Use	Technology	Measure	Percent Failed, Removed, Replaced	Annualized Failure, Removal, Replacement Rate^	Mean Life*	Median Life*	Ex Ante EUL
Lighting	T8 Lamps and Electronic Ballasts	L23	2.56%	0.85%	117	81	16
	Optical Reflectors w/ Fluor. Delamp	L19	0.76%	0.25%	396	275	16
	High Intensity Discharge	L81	1.88%	0.63%	159	110	16
HVAC	CAC	S160	0.48%	0.16%	627	434	15

^ Assuming a percentage of failed, removed, replaced occurs over three years.
* Assuming a constant failure rate over time.

Even based on these conservative assumptions, the estimates of median lives greatly exceed the ex ante estimates of EUL.

It is important to note that during some of the follow-up surveys (which were done either on-site or over the phone by an experienced auditor), it was not always possible to identify the exact equipment that was included in the retention panel. In some cases we were unable to identify the exact amount of equipment at the facility, which sometimes lead to larger or smaller estimates of equipment in place and in operation.

Because we obtained counts of the number of units that had failed, been removed or been replaced, we could verify the unit counts in the retention panel. This was done by adding the number of units found to be in place and operable, to the number of units that had failed, been removed or been replaced. In the cases where the number of verified units was smaller than the number of units in the retention panel, we conducted our analysis on only the number that we verified during the survey.

In the cases where the number of units found to be in place and operable was greater than the amount in the retention panel, it was assumed that all of the units in the retention panel were in place and operable.

3.4 VISUAL INSPECTION

For this step, we developed an empirical survival function that was observed from the raw retention data over the first three to four years of the measures' lives. As discussed above, this task was conducted for all measures, regardless of the amount of failures or removals in the sample data.

To develop the empirical function, we calculated for each month the percentage of equipment that was in place and operable. Although this appears to be a straightforward calculation, there were two issues that arose:

- The dates associated with failures and removals were not always well populated.
- Not all customers were surveyed over the same length of time.

Missing Failure Dates

Two common terms used in classical survival analysis are "left-hand censoring" and "right-hand censoring". Left-hand censoring means that it is known that a failure/removal has occurred, but it is unknown when the failure/removal occurred. It is only known that the failure/removal occurred before a certain date.

Right-hand censoring is more common in our data. Right-hand censoring means that at the last time the customer was surveyed, a failure/removal had not occurred, so the time when the equipment will fail or be removed is unknown.

The SAS procedures that are discussed below in Section 3.5 are capable of handling right-hand censored data, and in some cases left-hand censored data. But for this more simplistic task, some assumptions are required.

In order to develop our empirical distribution, we needed to have an estimate of each failure date. We considered four different approaches to estimating the failure dates:

1. Choose the earliest possible date, which would be the date the retention panel was developed. This was usually one year after the installation.
2. Choose the latest possible date, which would be the date the follow-up survey was completed. This could be anywhere from 2 to 5 years after the installation date.
3. Choose the midpoint between the two dates above.
4. Generate a random date between the two dates above, based on a uniform distribution.

It is important to note that approximately 20 percent of the failure dates were missing.

Below in Exhibit 3-7, we present the survival functions based on each of these methods, for the L23 T8 measure. We still needed to resolve the issue of survey length.

Survey Length

The topic of right-hand censoring is directly related to the issue of customer survey length. The issue of having customers surveyed at the same time is not much of a concern. Because our empirical survival function looks only at the percentage of equipment that has failed in each month *since installation*, it is not necessary to have each customer's installation date occur at the same time.

What is more problematic is that some customer follow-up surveys were conducted 36 months after their installation, and others had follow-up surveys conducted 48 months after their installation. Therefore, when we calculate the percentage of equipment in place and operating for, say, month 37 there will be some customers who were last surveyed 36 months (or less) after their installation date. For these customers, if a failure/removal occurred prior to month 37, then we know the unit is not operable and in place during month 37. However, if the equipment did not fail or become removed prior to month 37, we cannot say for certain if the equipment is still in place and operable in month 37. This leaves us with three alternatives for developing our empirical distribution. When we are calculating the percent of equipment operable and in place for month M, but the equipment was last surveyed prior to month M, we can:

1. Not include the equipment at all, regardless if a failure/removal occurred prior to month M.

2. Only include the equipment if a failure/removal occurred prior to month M, because we know that the equipment is still failed or removed in month M.

3. Include the equipment regardless of failure/removal, and assume the equipment is still operable if it has not failed or been removed prior to month M.

Clearly, the third option overstates the percent of equipment that is in place and operable. Also, the second option is likely to understate the percent of equipment that is in place and operable, because you are not counting equipment that was operating up to month M, which is still likely to be operating in month M. Finally, the first option is probably the only unbiased estimate, but has the potential to result in a survival function that violates its non-increasing property. In other words, because the sample size changes for each month, it is possible that in one month the percent operable and in place could exceed the following months percentage (which violates the non-increasing property of a survival function.)

Even with the potential problems suggested with the first option, we feel this is the most accurate method. What we suggest is to only look at the first 30 to 45 months of data, when the majority of the population is still providing usable data, and the survival function is still non-increasing. To be conservative, we also developed empirical functions based on the second option. We did not develop functions based on option three because we felt this to be the most biased of the alternatives, especially in later months. Below, we explore the sensitivity of all of the options discussed above, for both survey length and missing failure dates.

Solutions

Exhibits 3-5 through 3-8 were developed in an attempt to address each of these various issues discussed above. First, Exhibit 3-5 provides the percentage of customers that had a survey length (defined as number of months the follow-up survey was conducted after installation) greater or equal to a given number of months. This illustrates the percentage of the customers that would contribute to the calculated percentage of operating equipment in option one above. Exhibit 3-5 shows that half of the sample (with the exception of the S160 CAC measure) had a survey length of at least 45 months.

Exhibit 3-5
Percentage of Equipment with Survey Length
Greater than or Equal to a Given Month

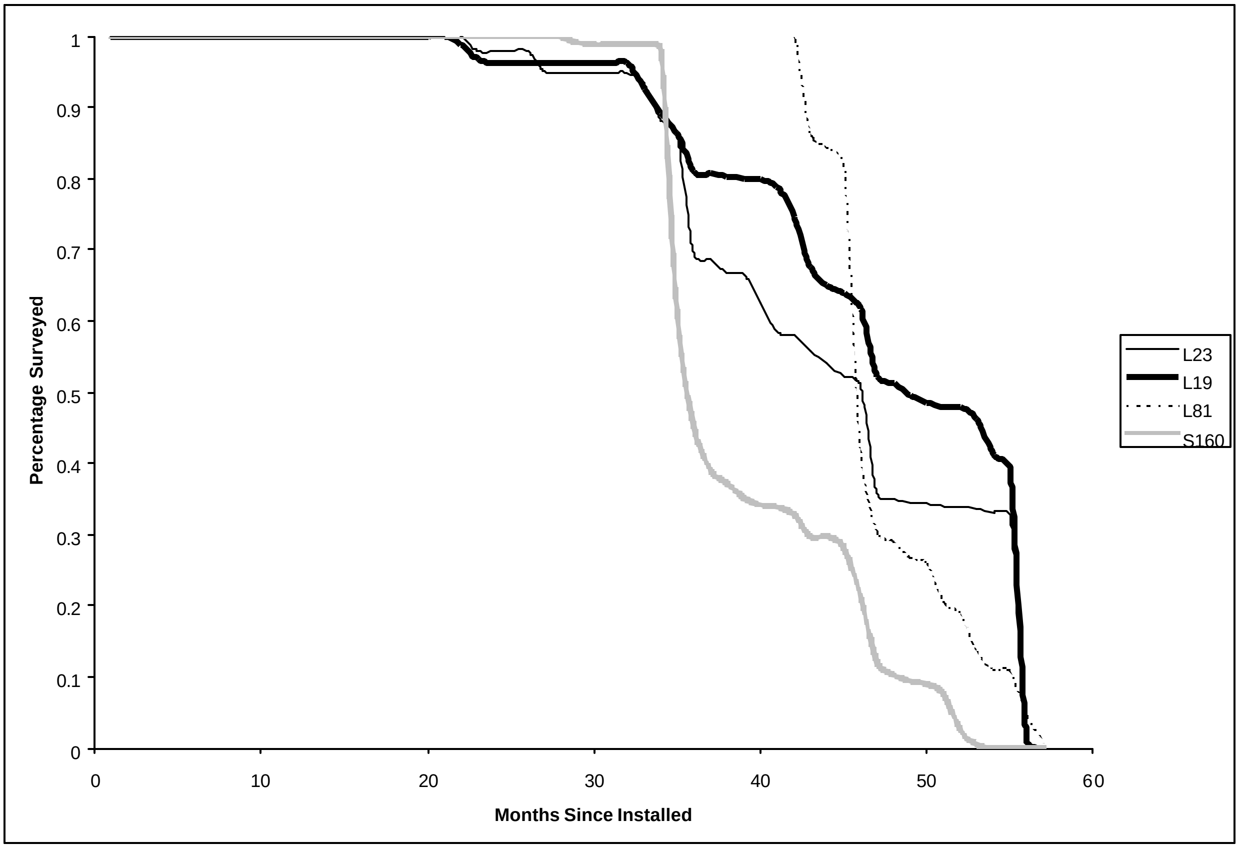


Exhibit 3-6 plots the empirical survival function for all measures under the following assumptions: for missing failure dates, use a random date; also, do not include the equipment if the survey date occurred prior to month M. The purpose of this exhibit is to illustrate how the survival function can become volatile as the sample frame decreases. As stated above, only half of the sample would contribute to the estimate of the survival function in month 45. After this point, we see that the survival function is no longer non-increasing, and has some rather large spikes. For this reason, we have decided to only use the first 45 months to plot the survival function for all measures. Even though the empirical survival function for the S160 CAC measure becomes unstable earlier on, we are not expecting to obtain statistically significant results based on visual inspection of the data.

Exhibit 3-6
Empirical Survival Function for All Measures
All Months

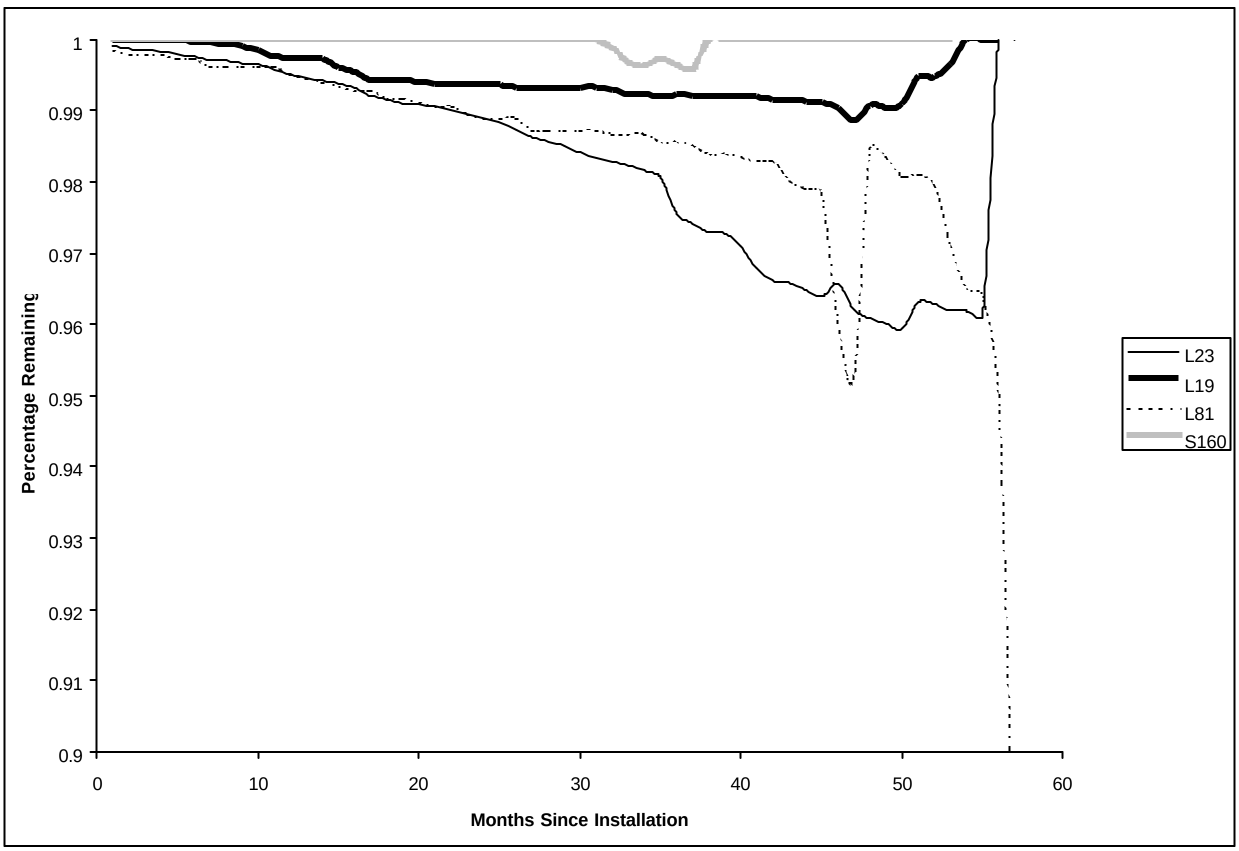


Exhibit 3-7 illustrates the sensitivity of using alternative methods for populating missing failure/removal dates for the L23 T8 measure. Again, we are only plotting the first 45 months for the reasons stated above. In addition, we are not including the equipment in the estimate of the survival function if the survey date occurred prior to month M.

Overall, the survival functions do not vary significantly across the four missing failure date approaches. We have selected the approach of populating missing failure dates with a random date, for conducting our analyses. We have selected this approach for three reasons. First, the random date falls between the earliest and latest dates. Second, the random date is smoother than the others. Third, the random date does not force multiple failure/removals to occur all on the same day, as the other methods would.

Exhibit 3-7
Comparison of Approaches for Populating Missing Failure Dates
L23 T8 Measure

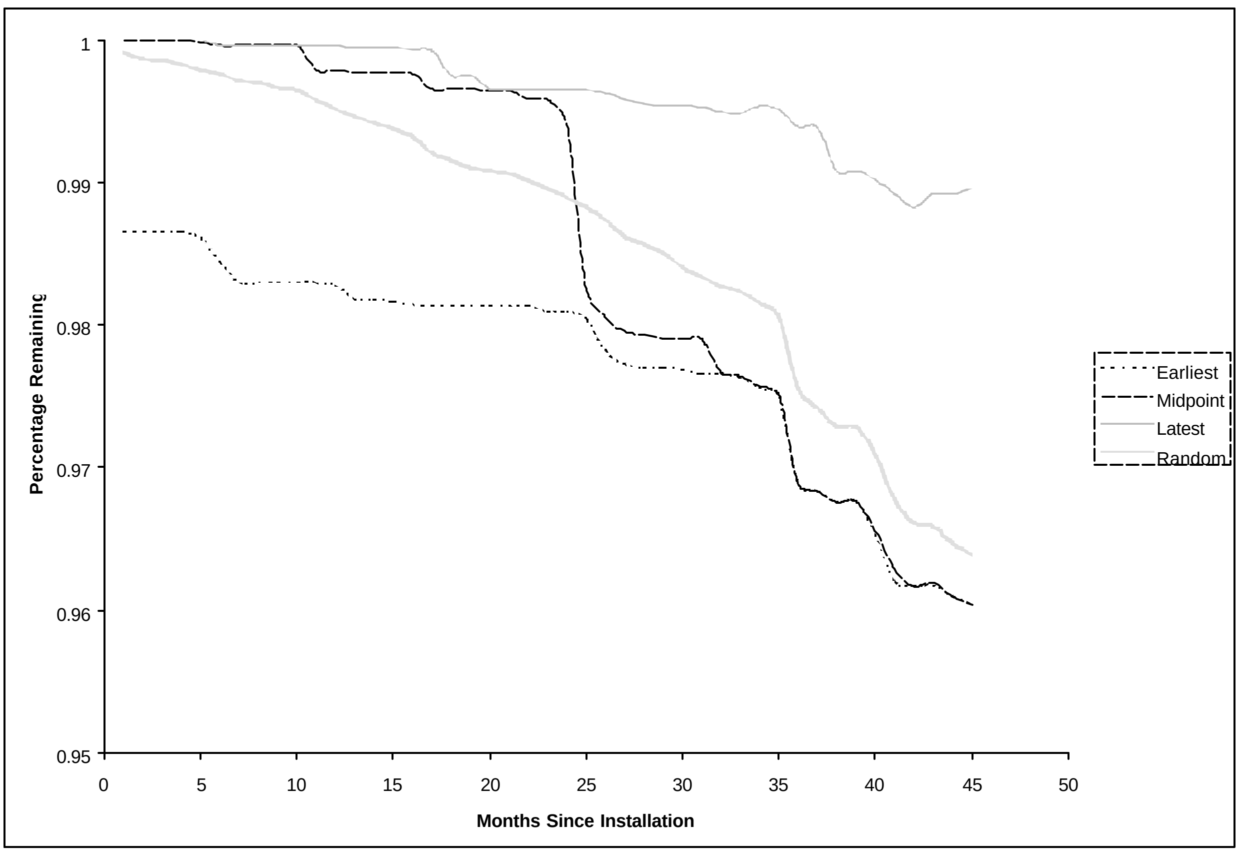


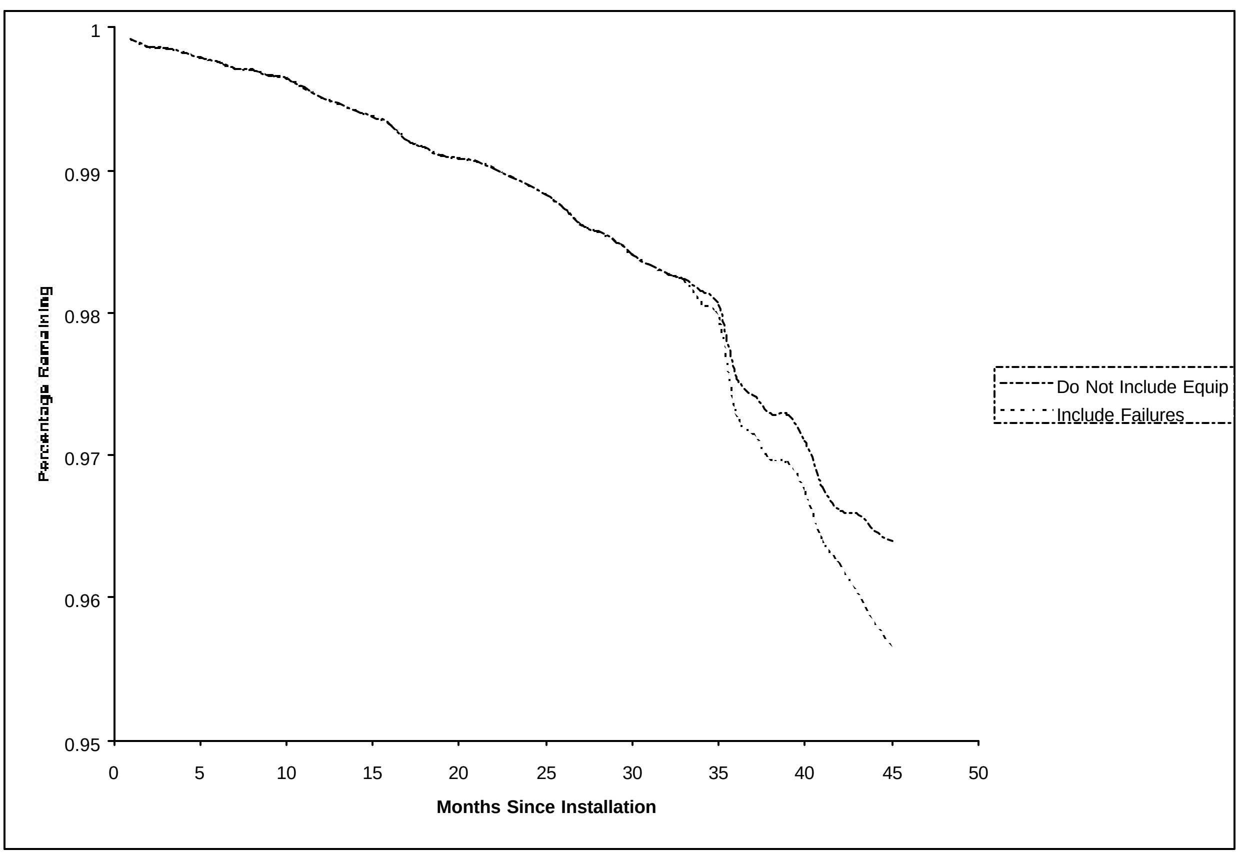
Exhibit 3-8 illustrates the sensitivity of using the two alternative methods for including equipment in the estimate of the empirical survival function if the survey length is less than the month being estimated. Again, these two methods are:

1. Not include the equipment at all, regardless if a failure/removal occurred prior to month M.

2. Only include the equipment if a failure/removal occurred prior to month M, because we know that the equipment is still failed or removed in month M.

Again, we are only plotting the first 45 months, and using a random date to populate missing failure/removal dates. As expected, including equipment if it has failed (option 2), results in a slightly lower survival function. Although this method is clearly biased downward, we see that the survival function is not that sensitive to the method. We have selected the approach of not including the equipment at all, regardless if a failure/removal occurred prior to month M. We feel this is the only unbiased method, and it is not significantly different than the more conservative method of including failed/removed equipment in the calculation.

Exhibit 3-8
Comparison of Approaches for Including Equipment
with Survey Lengths Less than Month Estimated
L23 T8 Measure

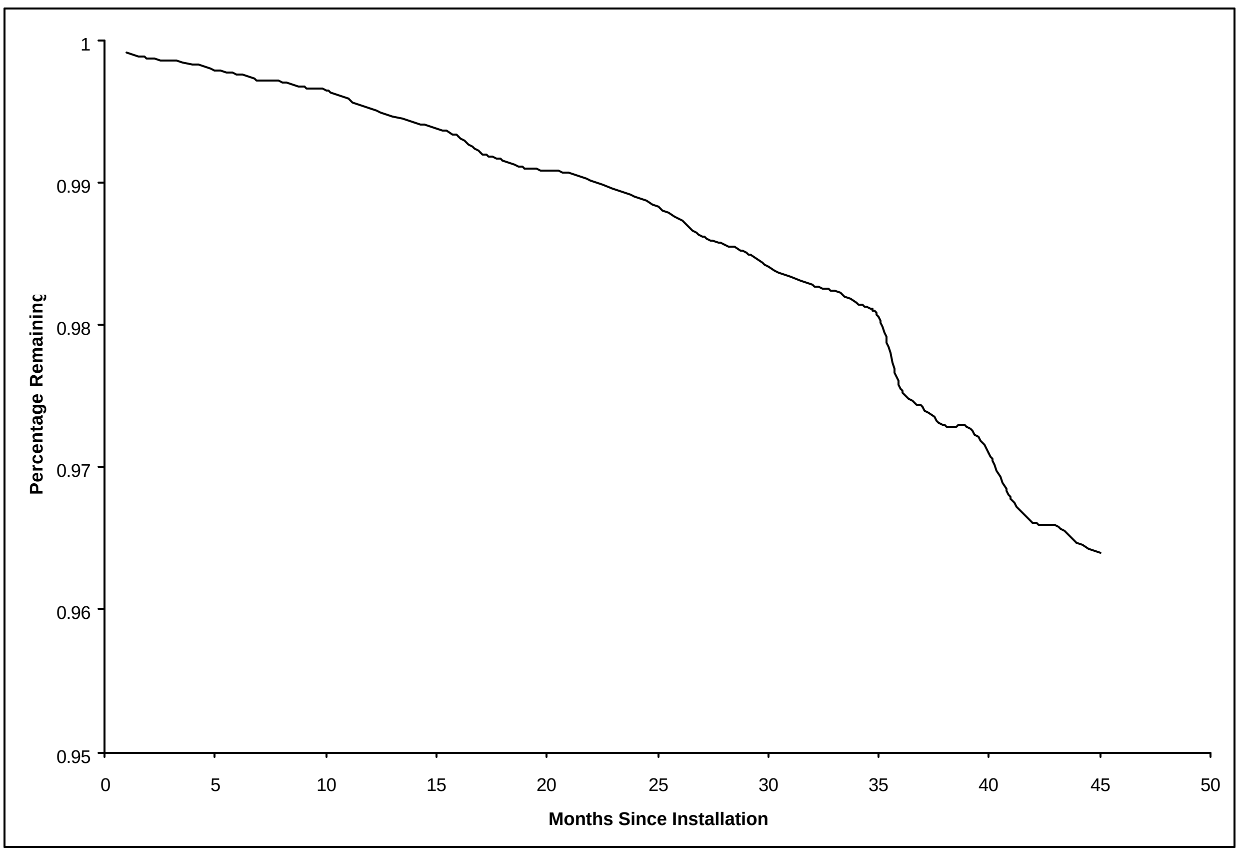


Finally, Exhibit 3-9 presents the final empirical survival function developed for the L23 T8 measure. This survival function is based on the following assumptions:

1. For missing failure/removal dates, generate a random date (based on a uniform distribution) between the date the retention panel was created and date the follow-up survey was conducted.

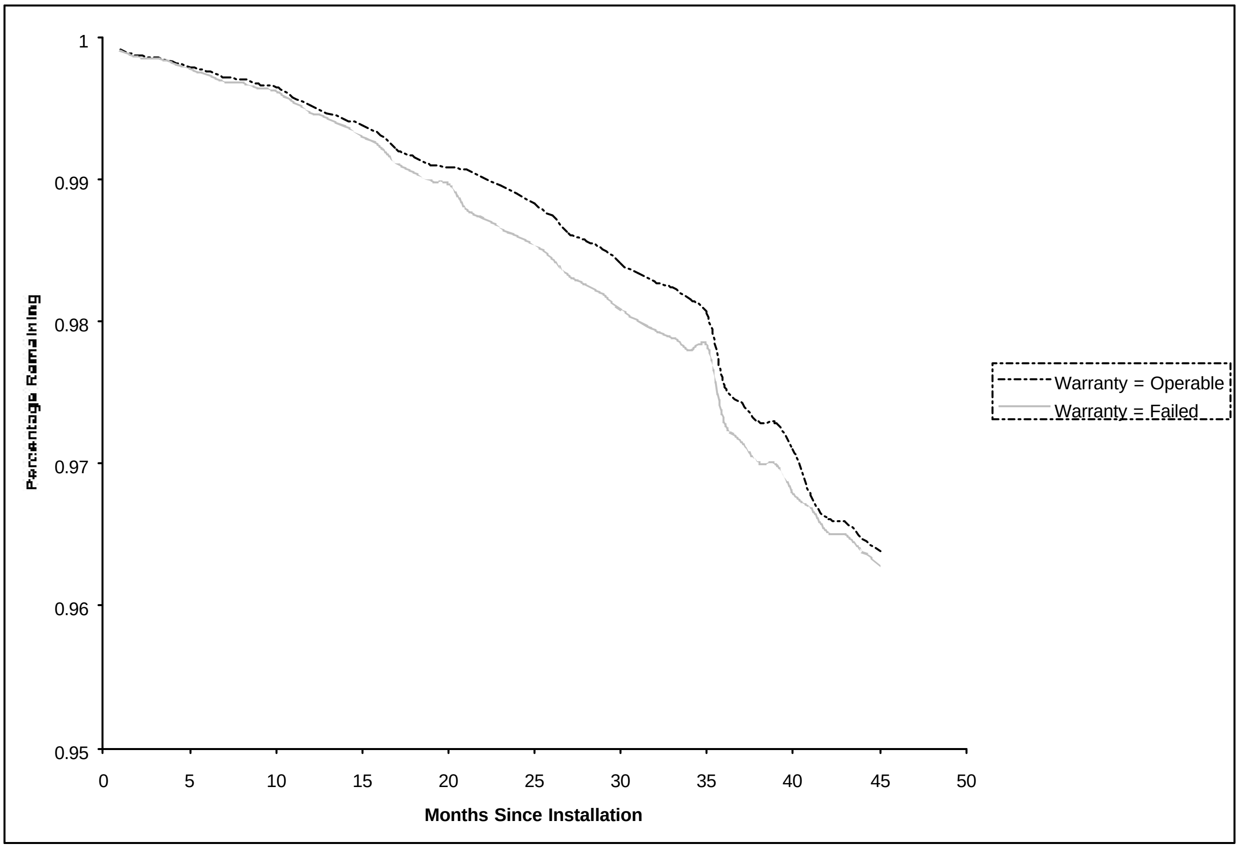
2. To estimate the percentage of equipment operable and in place in month M, do not include the equipment if the survey length is less than month M, regardless if a failure/removal occurred prior to month M.

Exhibit 3-9
Final Empirical Survival Function
L23 T8 Measure



One other interesting issue is that of warrantied equipment. As stated above, failed equipment that is replaced under warranty counts as if it is still operable and in place. For the L23 T8 measure, 13 percent of the failed equipment was replaced under warranty. Exhibit 3-10 compares how the empirical survival function for the L23 T8 measure would change if warrantied equipment did not count as operable and in place.

Exhibit 3-10
Sensitivity to Warranty
L23 T8 Measure



Exhibits 3-11 through 3-13 provide the empirical survival functions for the L19 Delamping, L81 HID 251-400W and S160 CAC measures, based on the same assumptions. The first 45 months of the survival function is plotted for all measures.

Exhibit 3-11
Final Empirical Survival Function
L19 Delamping Measure

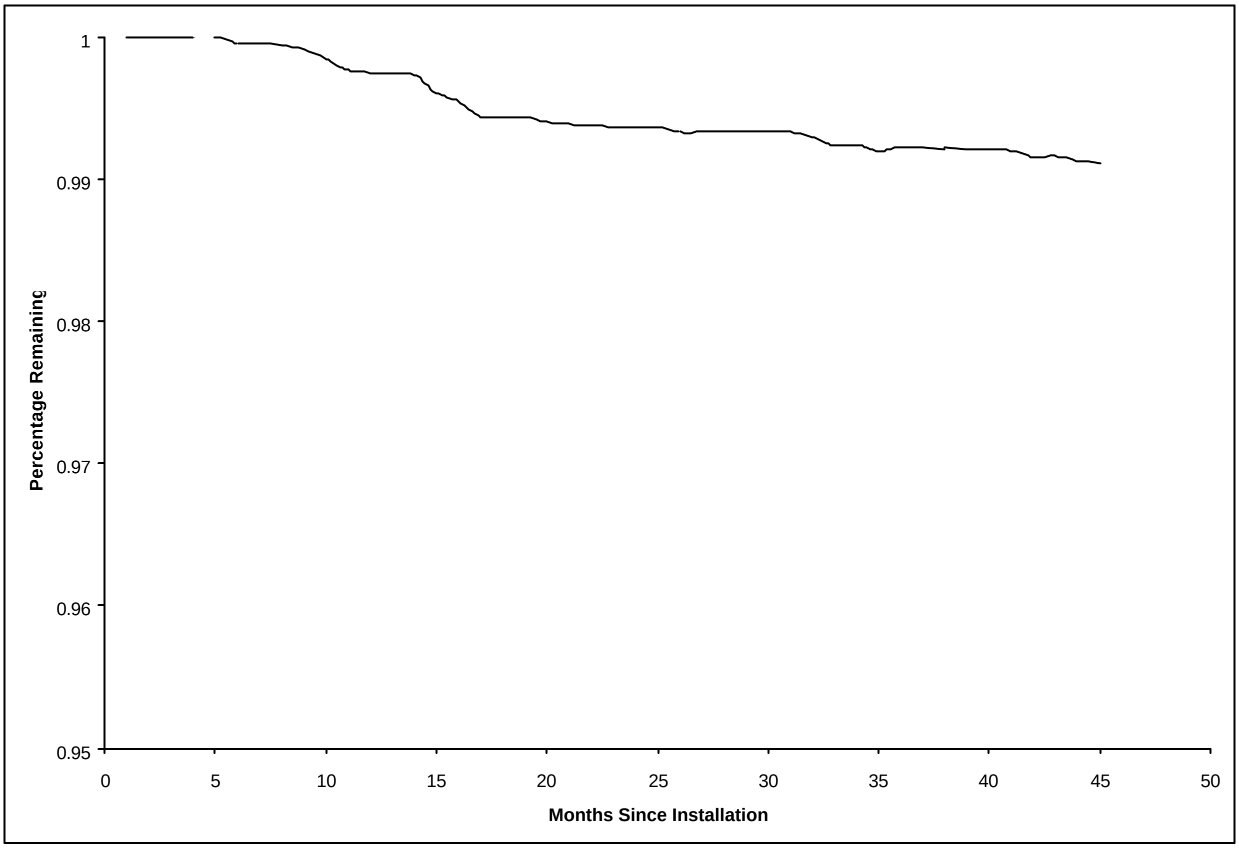


Exhibit 3-12
Final Empirical Survival Function
L81 HID 251-400W Measure

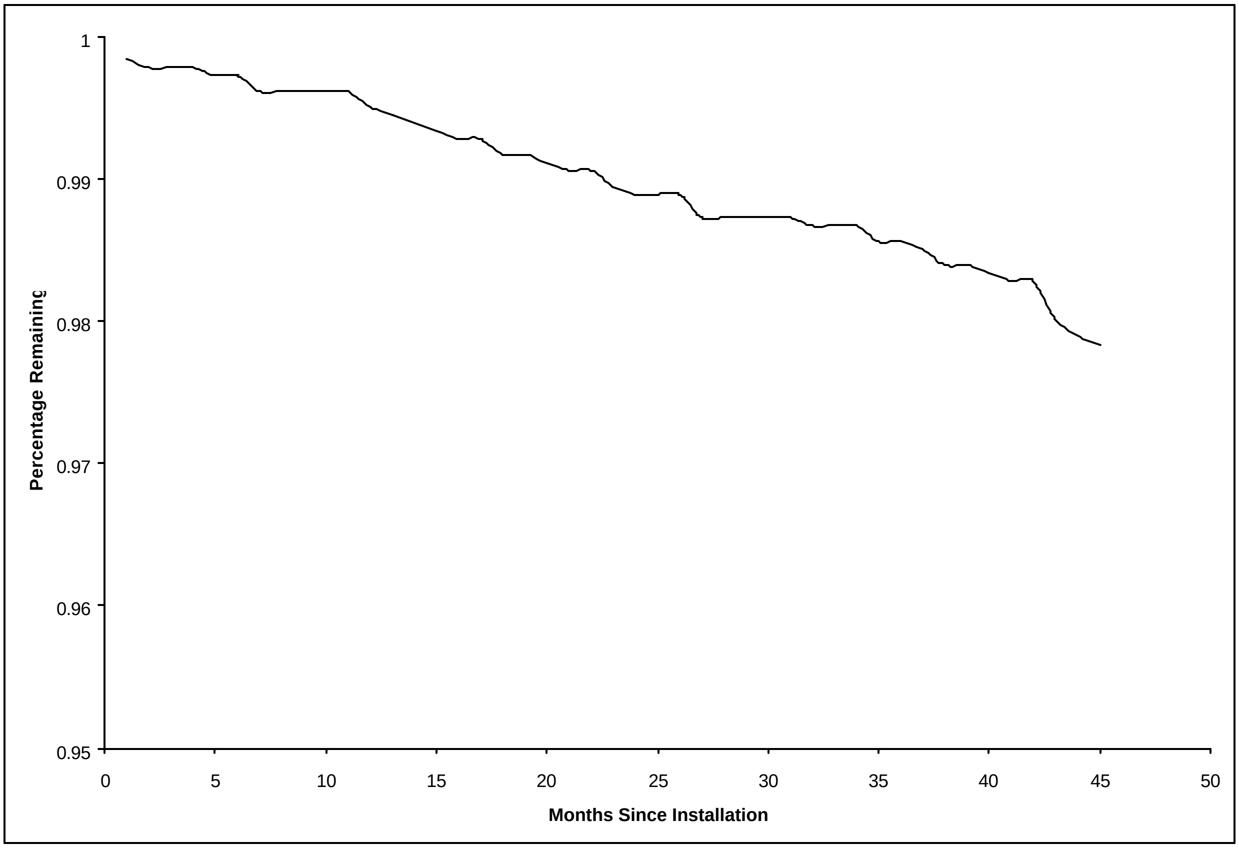
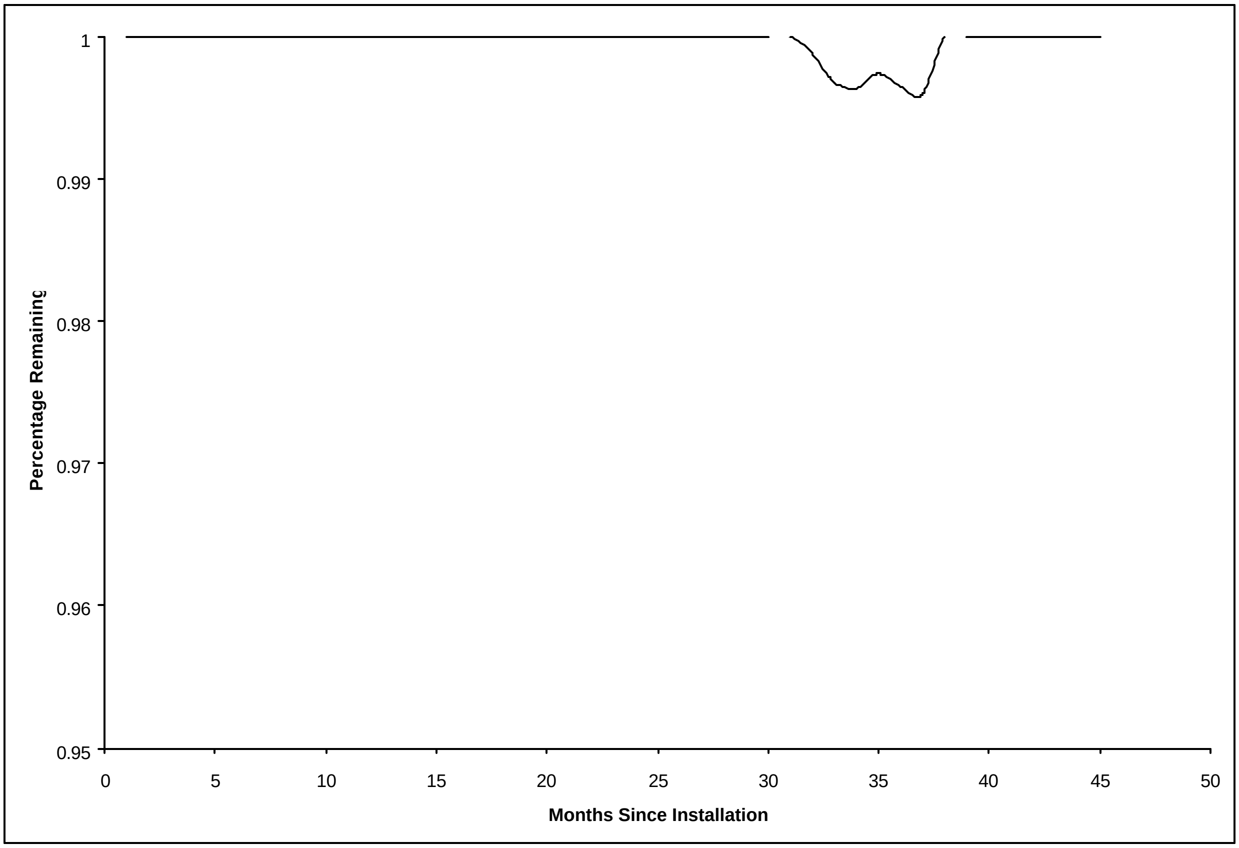


Exhibit 3-13
Final Empirical Survival Function
S160 CAC Measure



The empirical survival function for the S160 CAC measure shown in Exhibit 3-13 above indicates failures, and then goes back to unity. This is due to the issue of survey length discussed in Section 3.4. Referring back to Exhibit 3-6, a survey length of 45 months was selected because it retains the non-increasing property of the survival function for all measures but the S160. Because of the very limited number of events, we did not expect to obtain a statistically significant EUL for this measure. Therefore, rather than conducting an entirely different analysis for this one measure, we accepted the increasing survival function over the last several months. As we will show later in this section, even with this increase late in the study period (that should have the effect of increasing the EUL), the resulting EUL modeled is not reasonable.

3.5 TREND LINES

Based on the empirical survival functions presented above, a trend line was developed to estimate the survival function over the life of the measure, and estimate the measure's EUL. As discussed above, only the first 45 months of the empirical survival functions were used. This was done for the L23 T8, L19 Delamping, L81 HID 251-400W and S160 CAC measures.

Two trend lines were estimated using linear regression:

- The first trend line was assumed to have a linear relationship over time. Therefore, the trend line was developed using a linear regression with the percentage of equipment operable and in place as the dependent variable, and the month as the independent variable.

- The second trend line was assumed to follow the exponential distribution, which is one of the most common distributions used in survival analysis. The trend line was also used with linear regression by making a transformation on the percentage of equipment operable and in place. The natural log of the percentage of equipment operable and in place was used as the dependent variable, and the month as the independent variable. Although the exponential distribution is appropriate for many survival functions, we have doubts about the applicability of the exponential distribution to this data due to the very small hazard rate. Because the exponential distribution asymptotically approaches zero, and the fact that the initially low hazard rates will remain constant, this distribution produces some very large EUL estimates.

The results of these analyses are provided below.

Linear Trends – L23 T8, L19 Delamping, L81 HID 251-400W and S160 CAC

Exhibit 3-14 compares the linear survival function with the empirical function developed above, for the first 45 months of the measure's life for the L23 T8 measure. This exhibit illustrates how well the linear trend compares to the empirical function during the earlier parts of the measure's life. Exhibit 3-15 provides the resulting survival function assuming a linear trend forecasted over 200 months (16.7 years).

Exhibit 3-14
Comparison of Empirical Survival Function and Linear Trendline
L23 T8 Measure

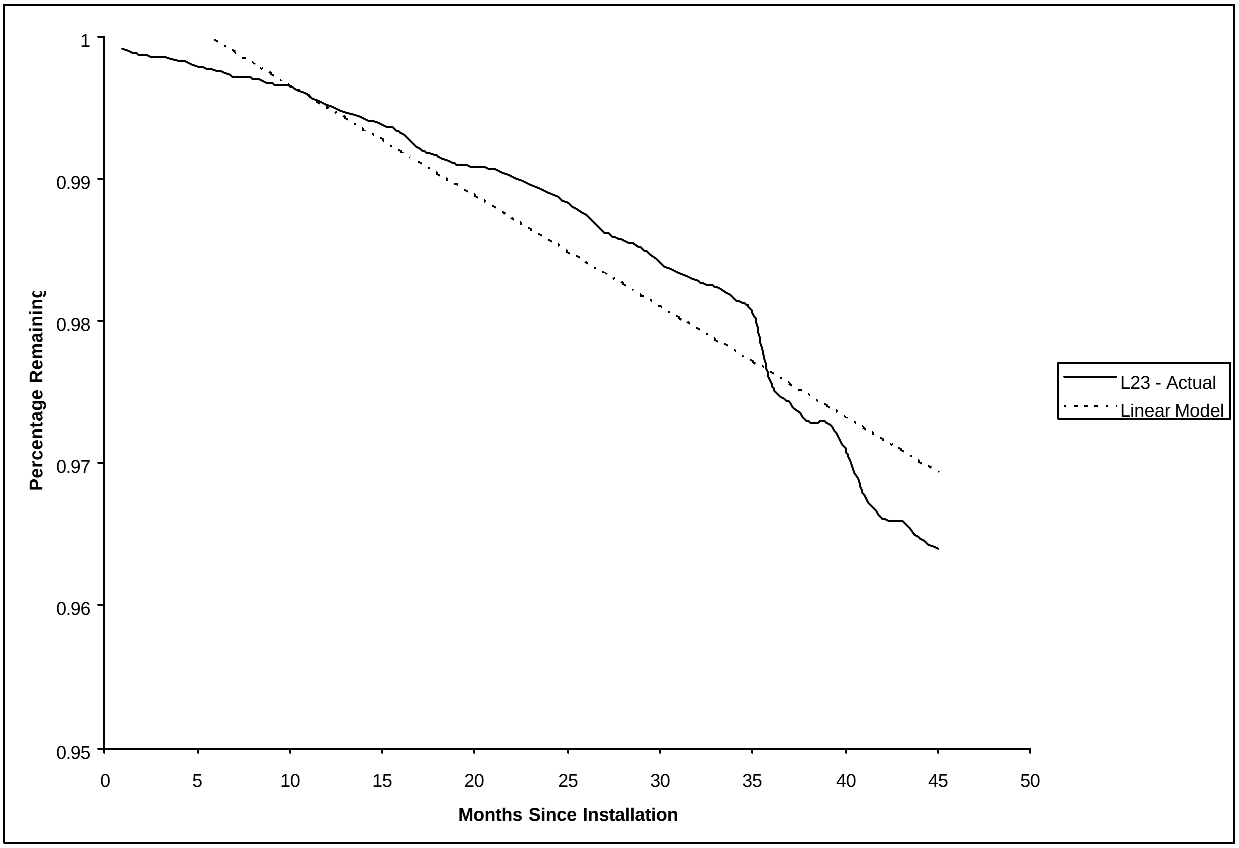
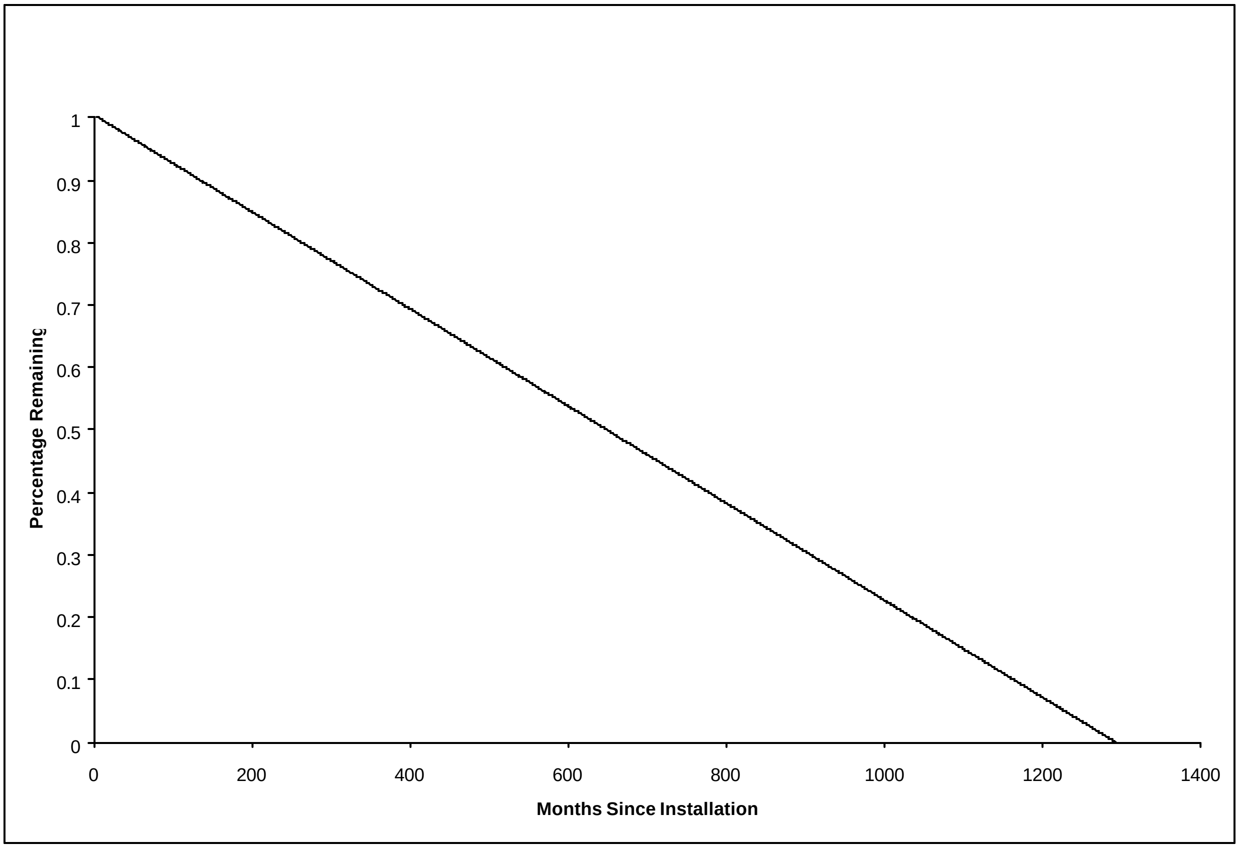


Exhibit 3-15
Survival Function Based on a Linear Trendline
L23 T8 Measure



Similarly, Exhibits 3-16 through 3-21 provide the linear survival functions, and comparisons to the empirical survival functions for the L19 Delamping, L81 HID 251-400W and S160 CAC measures. As discussed earlier, 45 months are shown for all measures.

Exhibit 3-16
Comparison of Empirical Survival Function and Linear Trendline
L19 Delamping Measure

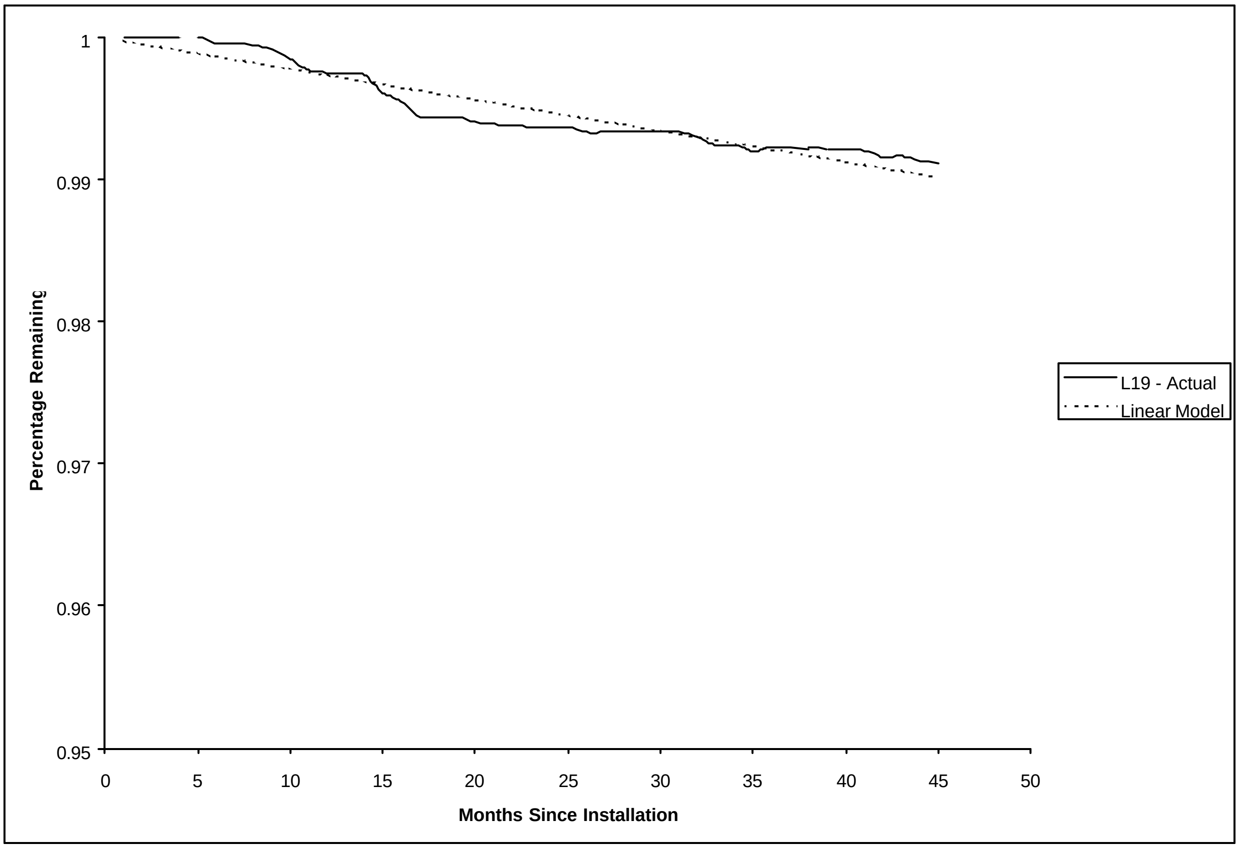


Exhibit 3-17
Survival Function Based on a Linear Trendline
L19 Delamping Measure

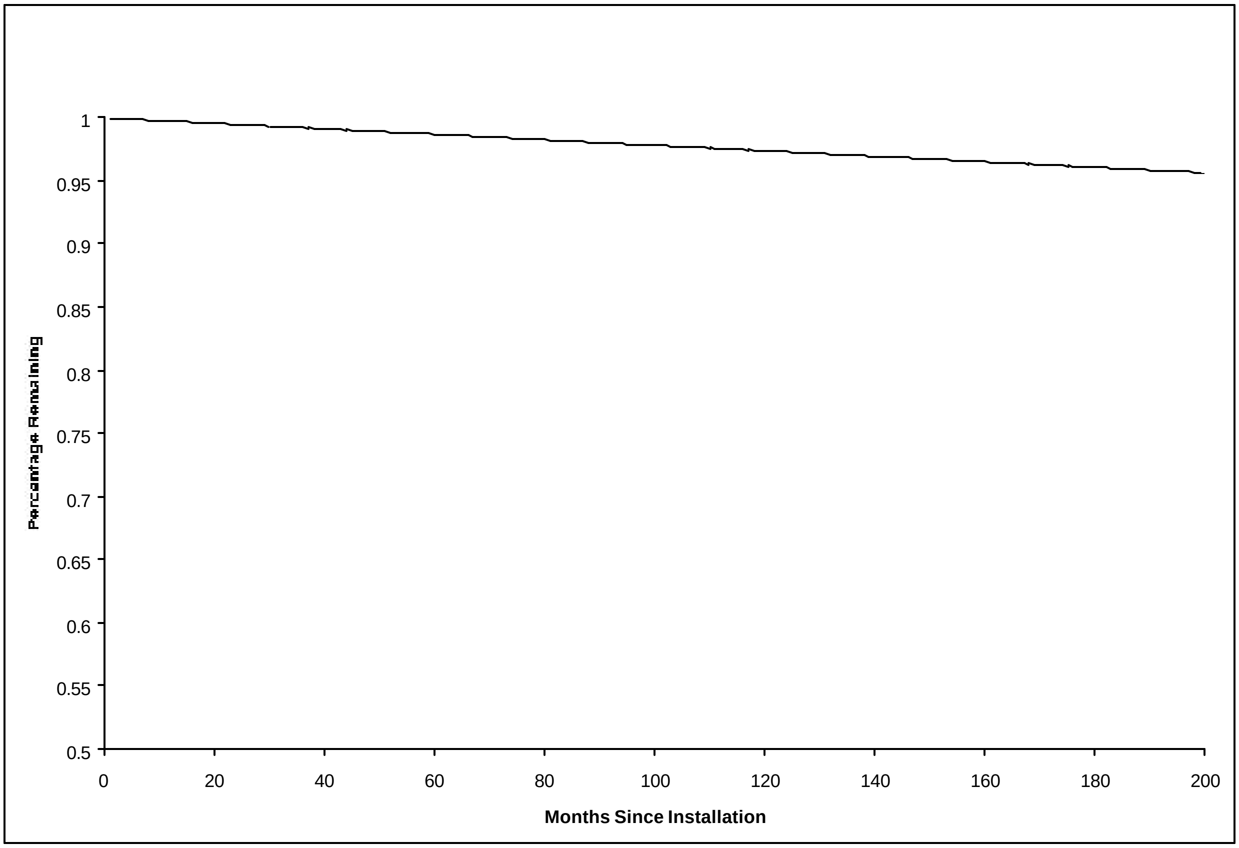


Exhibit 3-18
Comparison of Empirical Survival Function and Linear Trendline
L81 HID 251-400W Measure

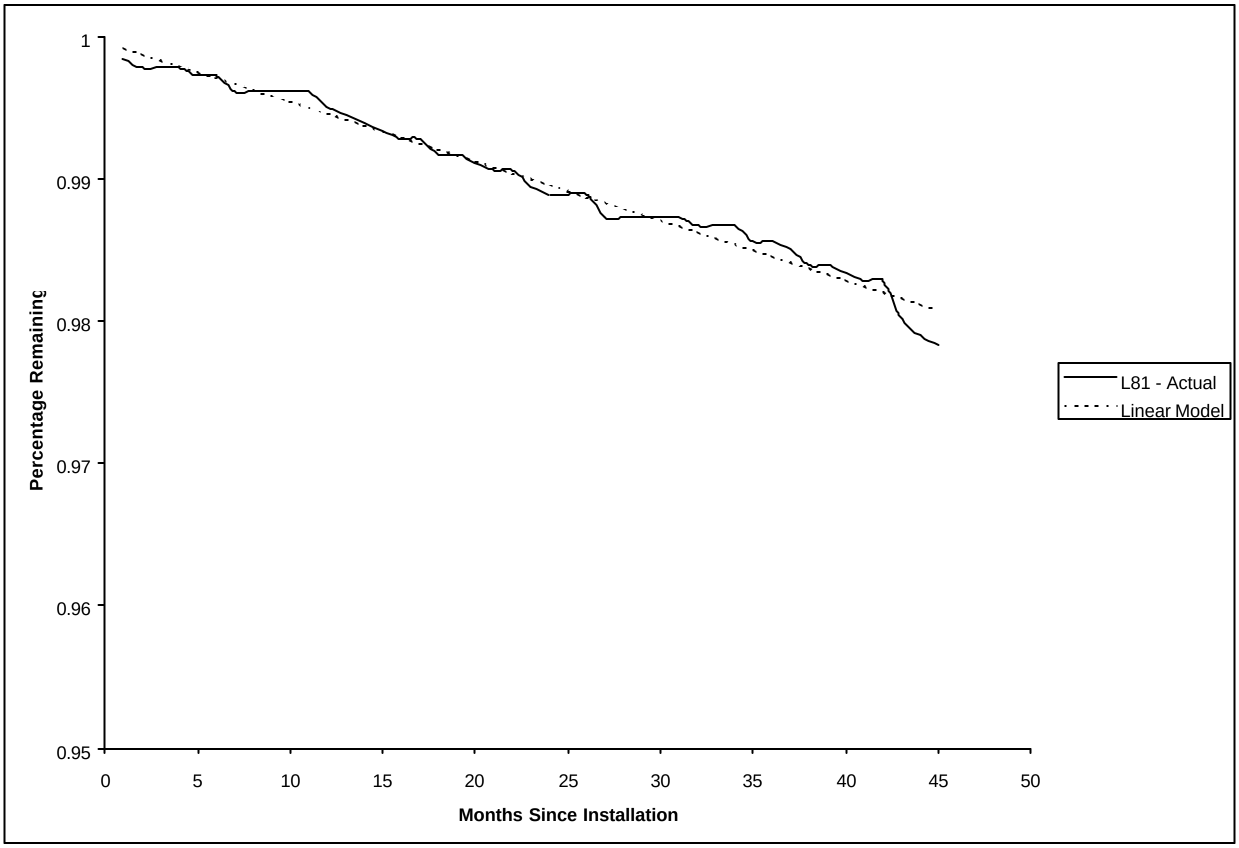


Exhibit 3-19
Survival Function Based on a Linear Trendline
L81 HID 251-400W Measure

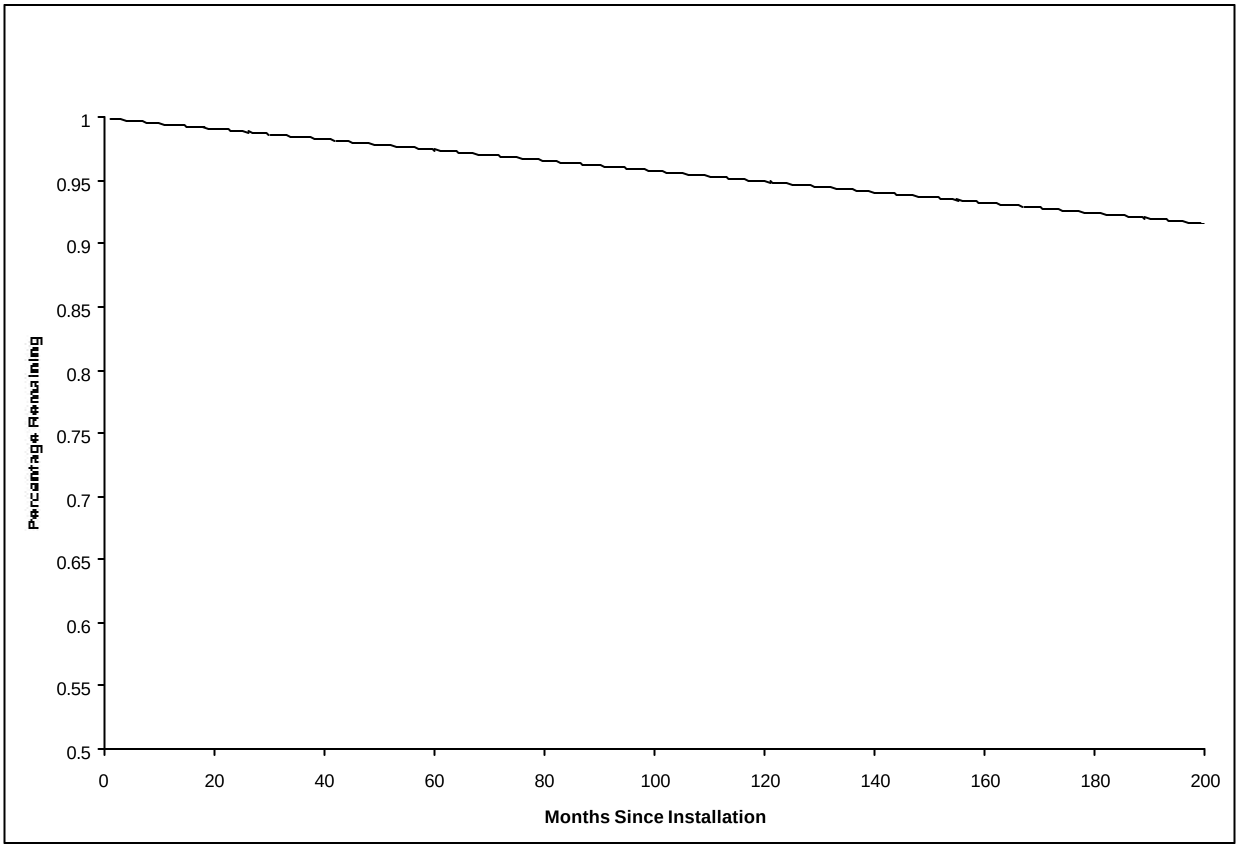


Exhibit 3-20
Comparison of Empirical Survival Function and Linear Trendline
S160 CAC

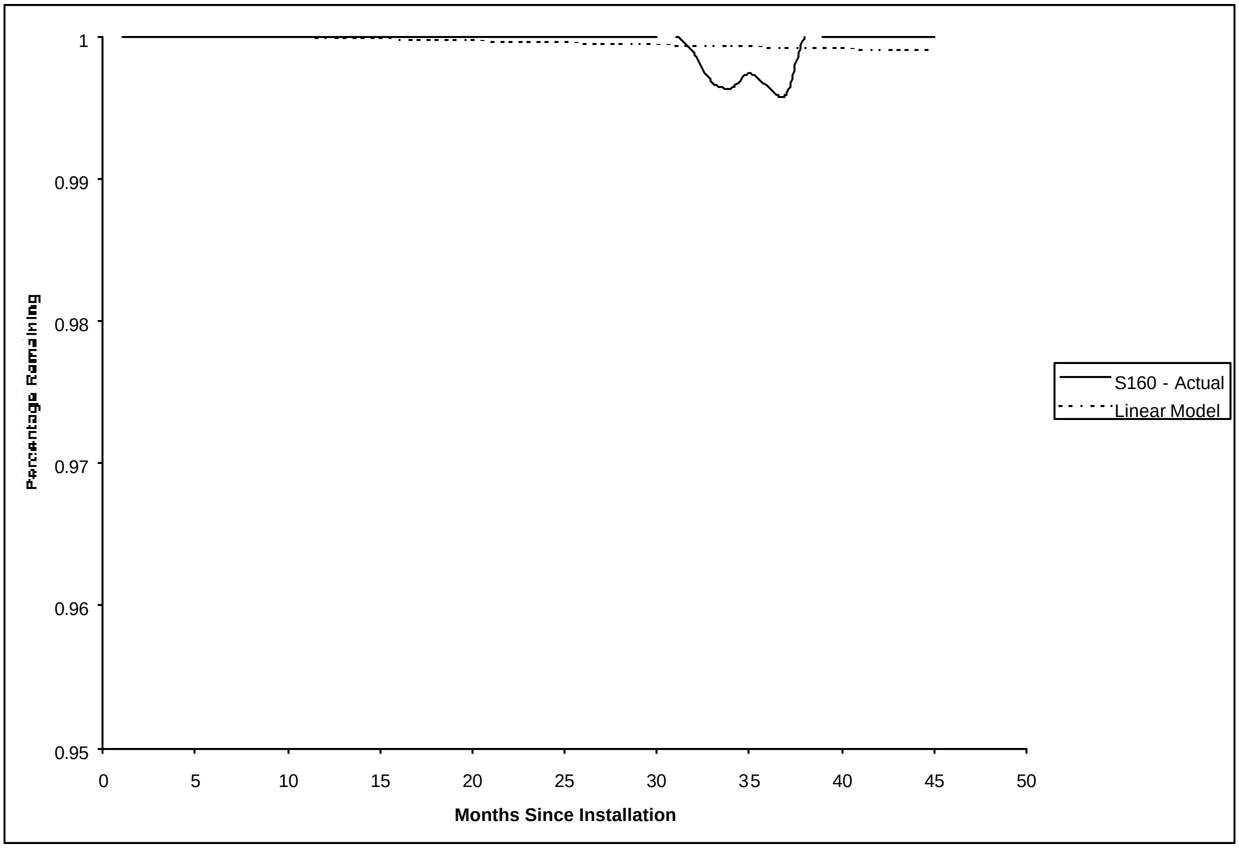
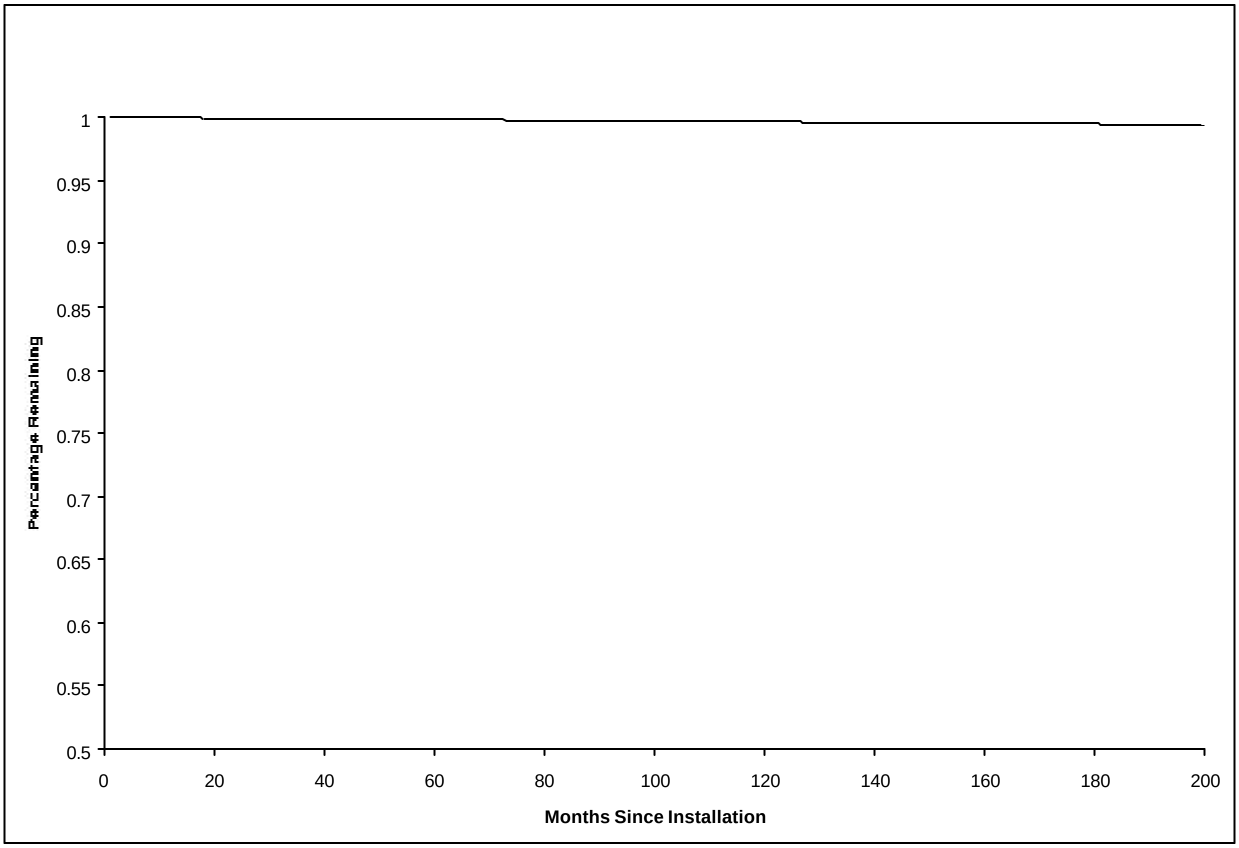


Exhibit 3-21
Survival Function Based on a Linear Trendline
S160 CAC Measure



The results of the linear regressions are provided in Exhibit 3-22 for each of the four measures. Also provided in Exhibit 3-22 is the estimated EUL for each measure. For a linear survival function, the EUL (median life) is calculated as:

EUL = (0.5 – intercept) / slope

Exhibit 3-22
Regression Results of Linear Trendline
and Resulting Ex Post EUL Estimates

Measure	Measure Description	Intercept	t-Statistic	Slope	t-Statistic	EUL
L23	FIXTURE: MODIFICATION / REPLACE LAMPS & BLST, 4 FT FIXTURE	1.0043	1,103	-0.0008	-22.53	54
L19	REFLECTORS WITH DELAMPING, 4 FT LAMP REMOVED	0.9999	3,554	-0.0002	-20.43	191
L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	0.9996	4,141	-0.0004	-45.71	100
S160	A / C: CENTRAL, < 65 KBTU / HR, AIR-COOLED, SPLIT-SYS / SNGL PKG	1.0003	3,135	0.0000	-2.38	1,450

Clearly, the results of the linear trendline estimate indicate that the ex post EUL estimate is significantly larger than the ex ante estimates (which are all 16 years for lighting measures and 15 years for the S160 CAC measure). Each of these results would easily reject the ex ante estimate at the 80 percent confidence level.

Exponential Trends – L23 T8, L19 Delamping, L81 HID 251-400W and S160 CAC

Exhibit 3-23 compares the exponential survival function with the empirical function developed above, for the first 45 months of the measure's life for the L23 T8 measure. This exhibit illustrates how well the exponential trend compares to the empirical function during the earlier parts of the measure's life. Exhibit 3-24 provides the resulting survival function assuming an exponential trend over the first 200 months (16.7 years).

Exhibit 3-23
Comparison of Empirical Survival Function and Exponential Trendline
L23 T8 Measure

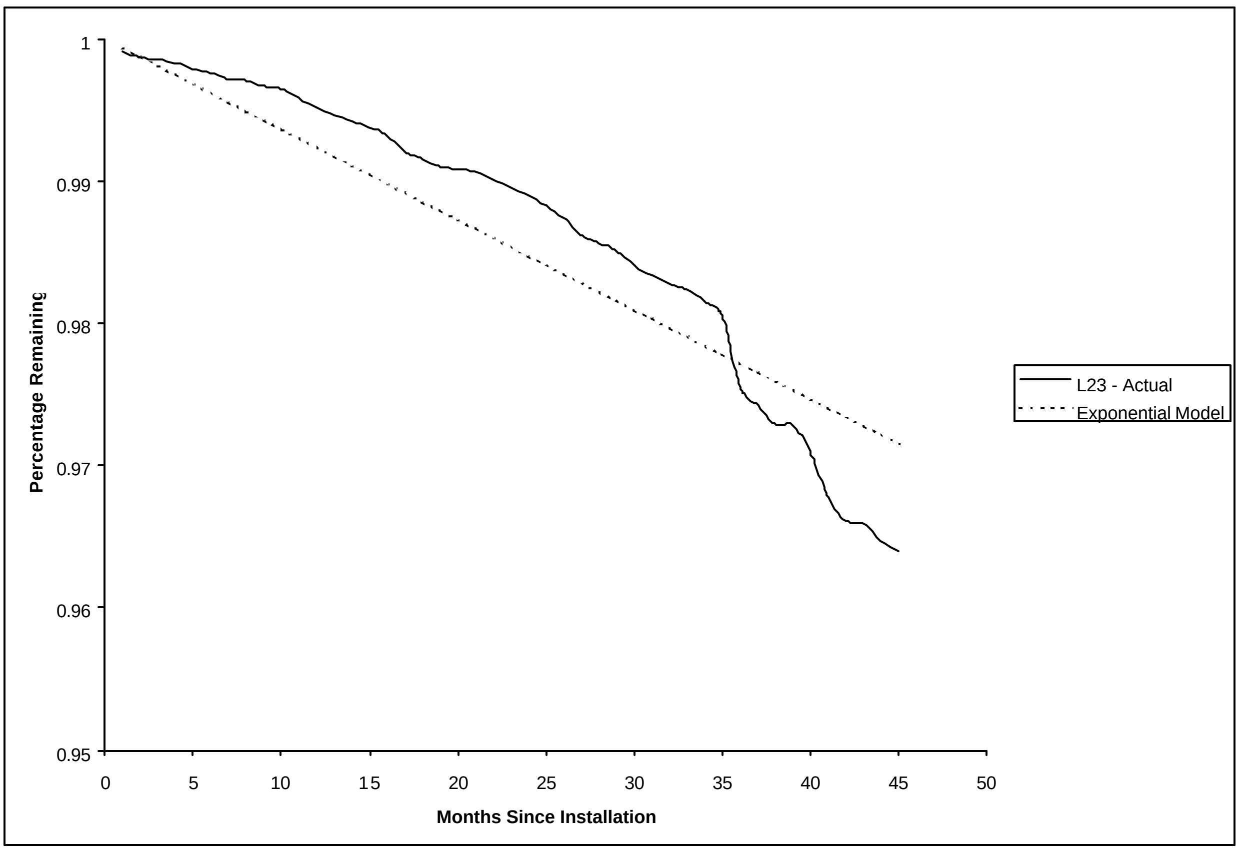
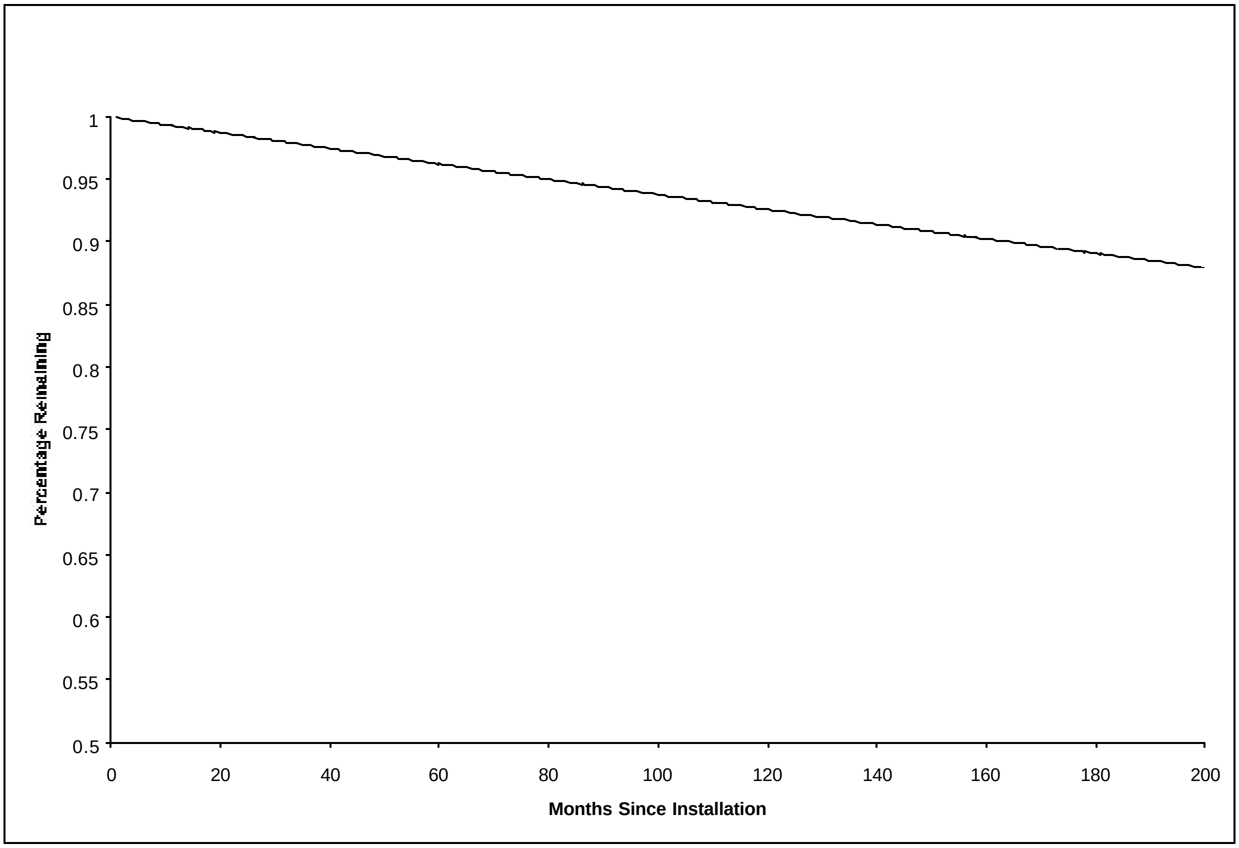


Exhibit 3-24
Survival Function Based on an Exponential Trendline
L23 T8 Measure



Similarly, Exhibits 3-25 through 3-30 provide the exponential survival functions, and comparisons to the empirical survival functions for the L19 Delamping, L81 HID 251-400W and S160 CAC measures.

Exhibit 3-25
Comparison of Empirical Survival Function and Exponential Trendline
L19 Delamping Measure

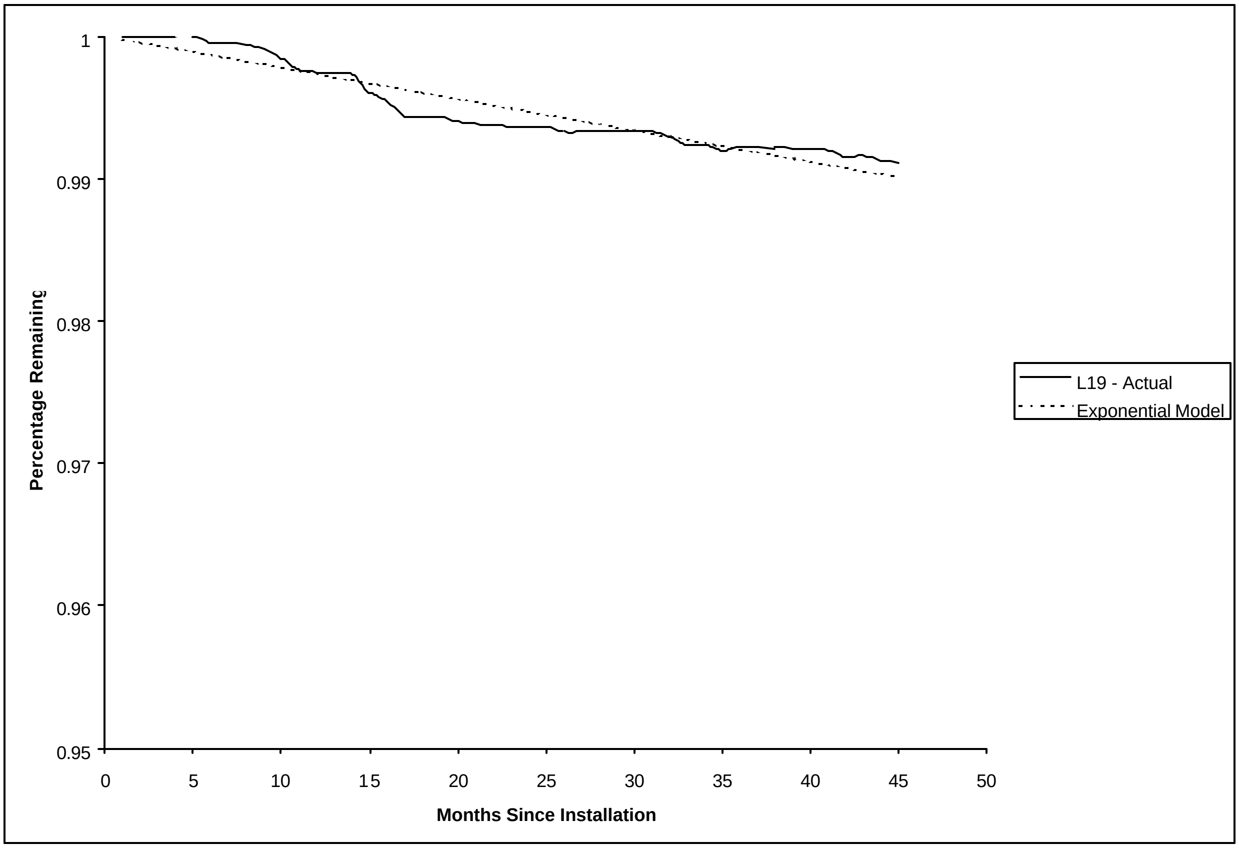


Exhibit 3-26
Survival Function Based on an Exponential Trendline
L19 Delamping Measure

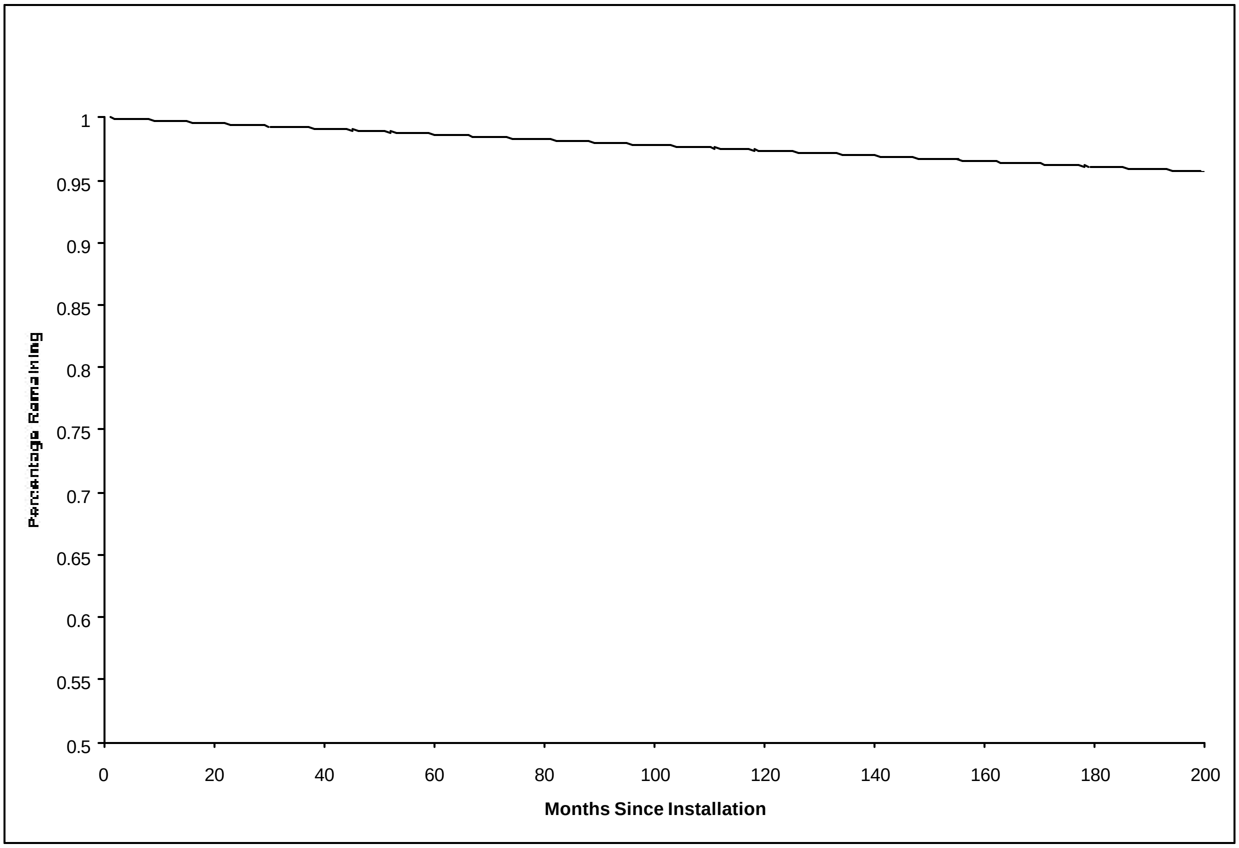


Exhibit 3-27
Comparison of Empirical Survival Function and Exponential Trendline
L81 HID 251-400W Measure

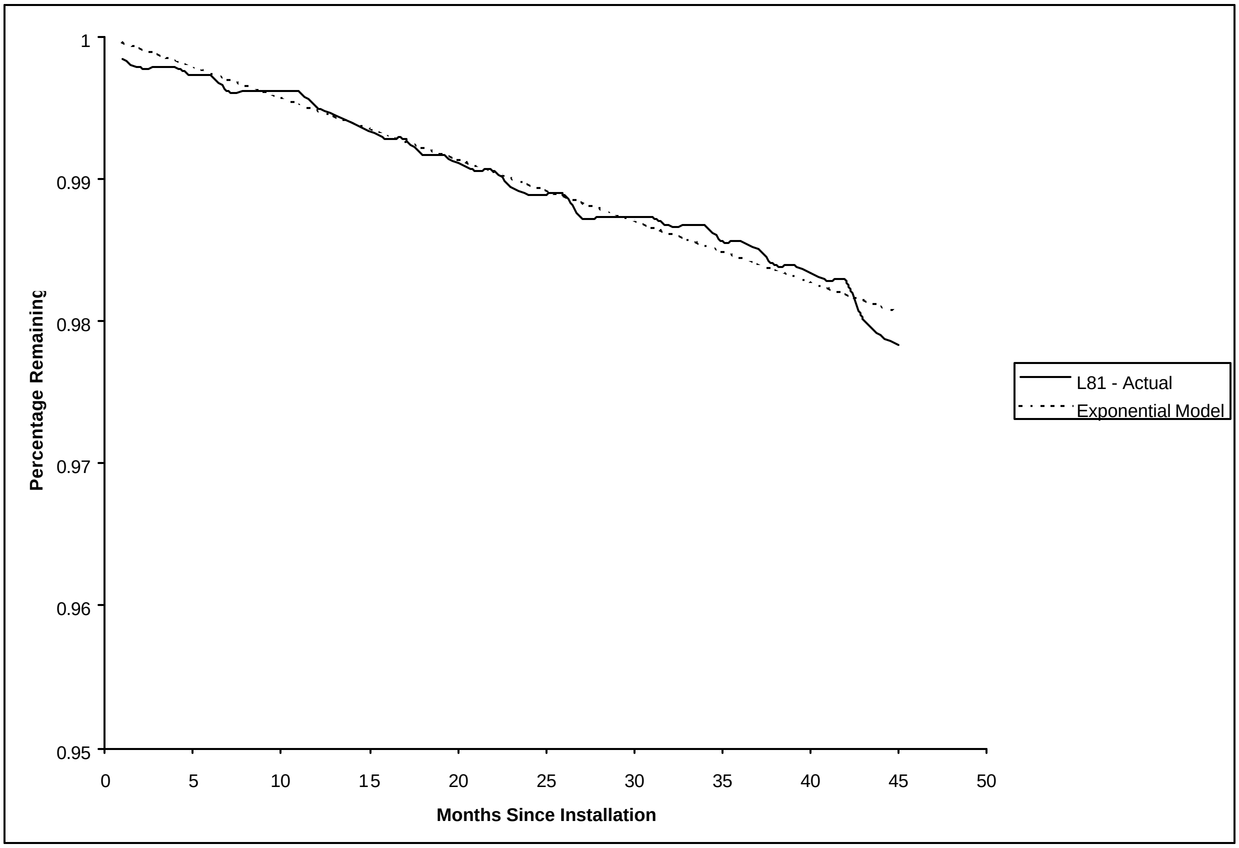


Exhibit 3-28
Survival Function Based on an Exponential Trendline
L81 HID 251-400W Measure

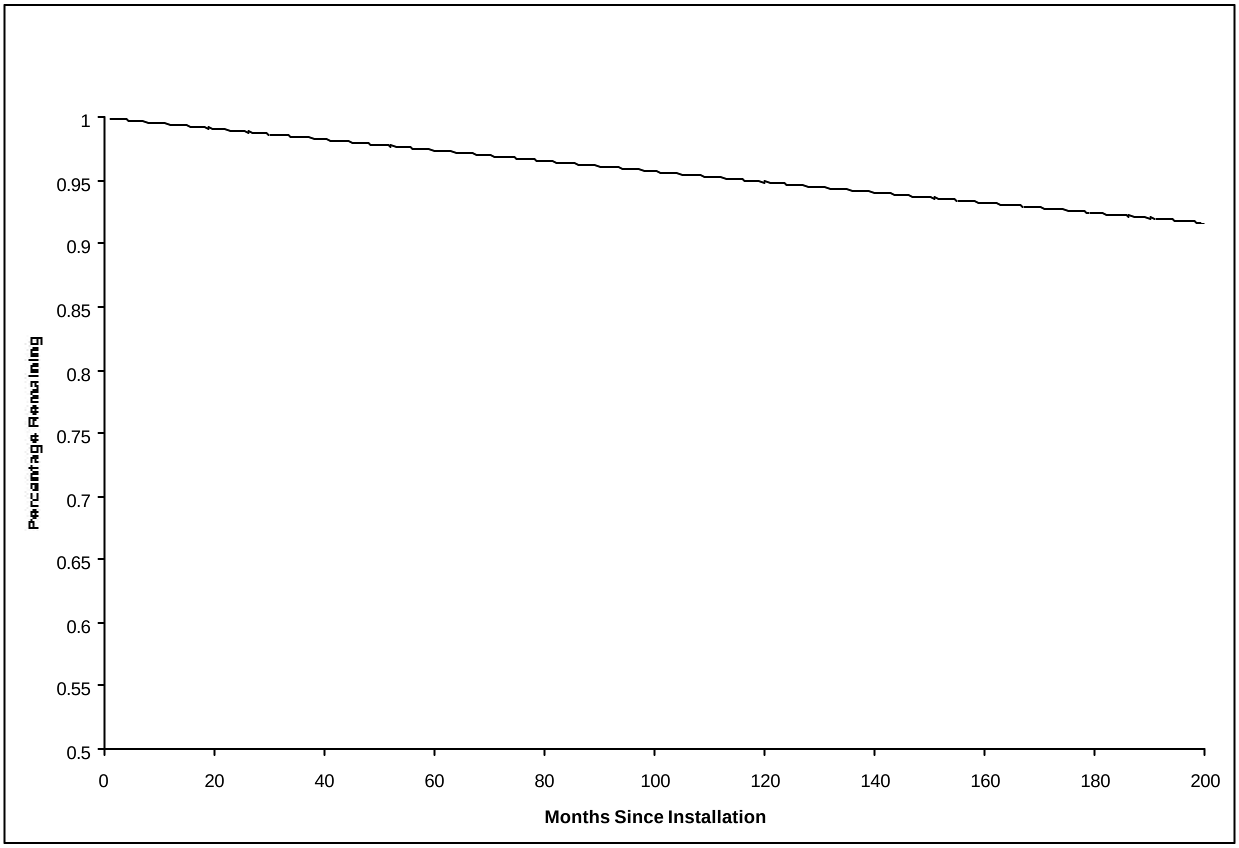


Exhibit 3-29
Comparison of Empirical Survival Function and Exponential Trendline
S160 CAC Measure

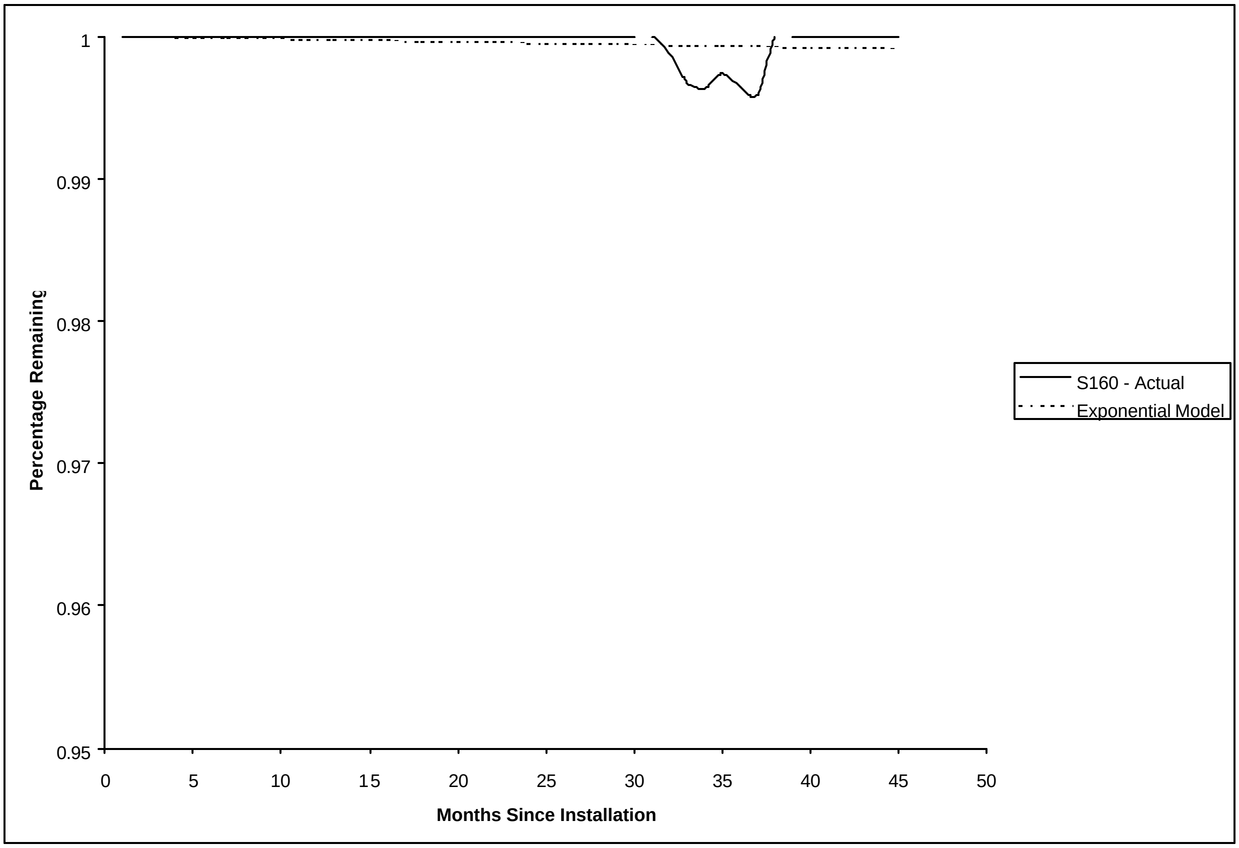
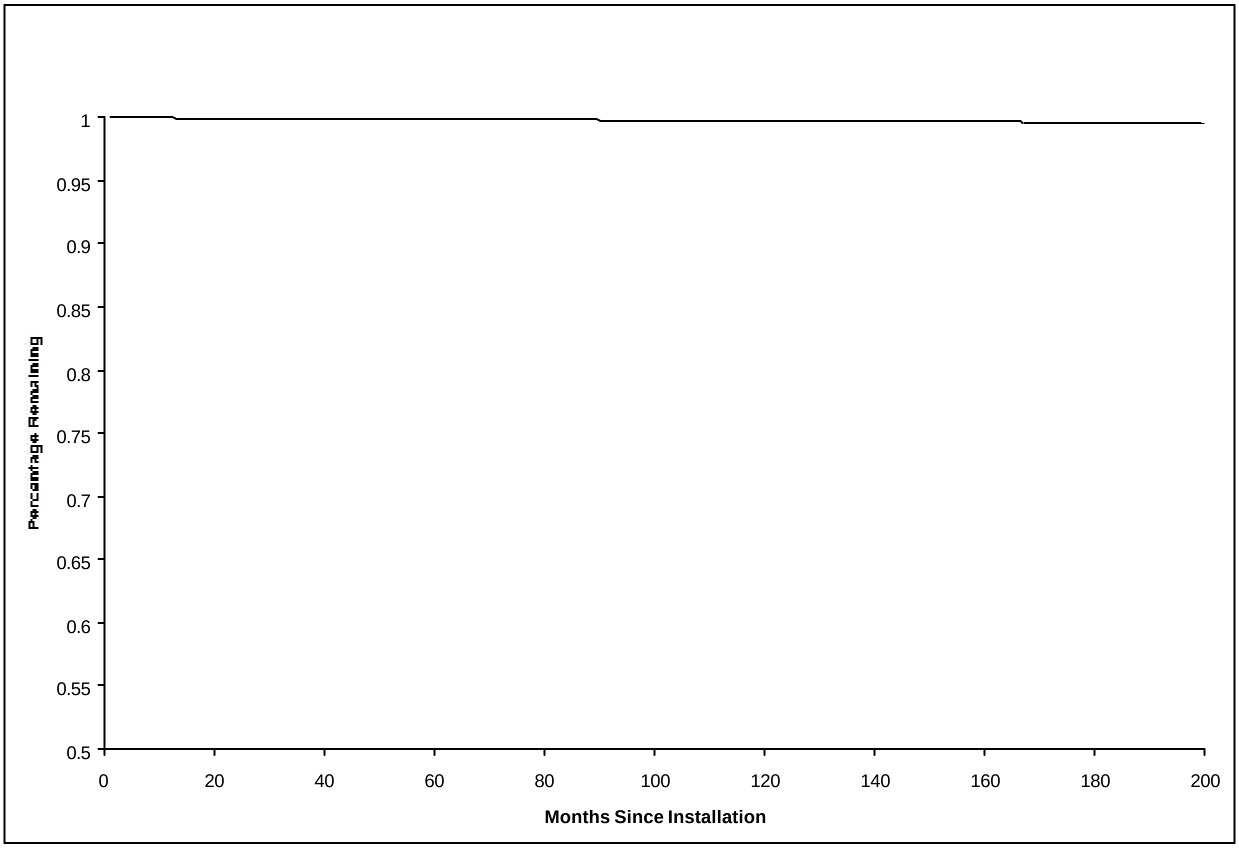


Exhibit 3-30
Survival Function Based on an Exponential Trendline
S160 CAC Measure



The results of the exponential regressions are provided in Exhibit 3-31 for each of the four measures. Also provided in Exhibit 3-31 is the estimated EUL for each measure. For an exponential survival function, the EUL (median life) is calculated as:

EUL = ln(2) / slope

Exhibit 3-31
Regression Results of Exponential Trendline and Resulting Ex Post EUL Estimates

Measure	Measure Description	Slope	t-Statistic	EUL
L23	FIXTURE: MODIFICATION / REPLACE LAMPS & BLST, 4 FT FIXTURE	0.0006	30.00	90
L19	REFLECTORS WITH DELAMPING, 4 FT LAMP REMOVED	0.0002	42.68	261
L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	0.0004	93.04	133
S160	A / C: CENTRAL, < 65 KBTU / HR, AIR-COOLED, SPLIT-SYS / SNGL PKG	0.0000	3.45	2,827

The results of the exponential trendline estimates are even more dramatic than for the linear trendline estimates. Again, these results clearly indicate that the ex post EUL estimate is significantly larger than the ex ante estimates. Each of these results would easily reject the ex ante estimate at the 80 percent confidence level.

The exponential distribution has some important assumptions that should be addressed. Most importantly, the exponential distribution assumes a constant hazard rate. Although this distribution works well to explain certain data, this assumption is not believed to be valid for many technologies. If this were the case, then study results indicate that the L23 T8 fixtures purchased through the program (the measure with the smallest EUL estimate) would have an EUL of 90 years.

As we will discuss in more detail in Section 4, this approach is not recommended for the final study results. In addition to the concern of the exponential distribution having properties that are not in line with our expectations, developing a trend line on empirical data in this manner is not optimal. The empirical data is interval and right hand censored, meaning that for some failures / removals, the time of the event is unknown; and it is also unknown when currently operating equipment may fail. This trendline approach does not statistically correct for censored data in the way that classical survival analysis approaches do, as discussed in the following section.

3.6 CLASSICAL SURVIVAL ANALYSIS

This step in our approach is founded on applying classical survival analysis techniques to the retention data in order to develop a survival function. Using the SAS System and the SAS companion guide, "Survival Analysis Using the SAS System," we have modeled the survival function assuming five of the most common survival distributions: exponential, logistic, lognormal, Weibull and gamma. In each case, we have plotted the resulting distribution and visually compared it to the empirical functions developed above. Furthermore, we have used the resulting survival function to estimate the EUL.

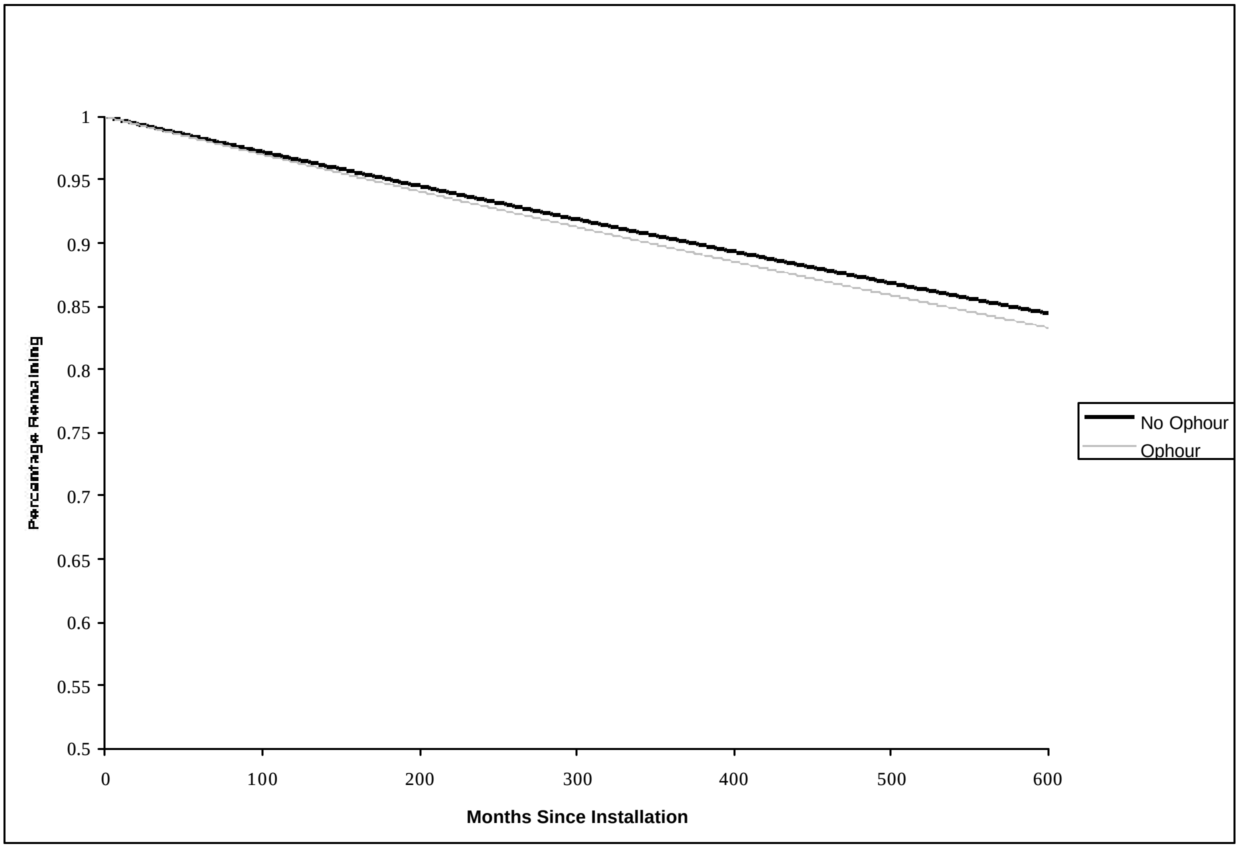
Some of the same issues we faced when developing the empirical survival function need to be addressed here as well. The problem of right-hand censoring is not an issue for SAS. The LIFEREG procedure, which we used for all of our modeling in this step, is capable of handling right-hand censored data.

SAS is also capable of handling left-hand censored data. In fact, our retention data is actually not left-hand censored, but interval censored. The true definition of left-hand censoring is that we know that an event occurred earlier than some time t, but we don't know exactly when. Interval censoring occurs when the time of failure occurrence is known to be somewhere between two times, but we don't know exactly when. Left censoring can be seen as a special case of interval censoring.

Although the LIFEREG procedure is capable of handling both left and interval censoring, interval censored data is more predictive than left hand censoring. Another commonly used survival analysis procedure in SAS in PHREG. Unfortunately, this procedure cannot handle either left or interval censored data. Therefore, we only conducted our analysis using the LIFEREG procedure.

Another important feature of the LIFEREG procedure is the use of covariates. This feature enabled us to use other predictive variables to help estimate the survival functions. For example, it would be expected that the EUL for a T8 is dependent on the number of hours that it is used during a year. So, an obvious covariate would be the inclusion of operating hours for each of our customers in the retention sample. Exhibit 3-32 compares the estimated survival function for the L23 T8 measure using the LIFEREG procedure without covariates, and with operating hours as a covariate. Here, we are using modeling the survival function with an exponential distribution.

Exhibit 3-32
Comparison of Survival Functions
Modeled without Covariates and with Operating Hours as a Covariate
L23 T8 Measure



The two survival functions are relatively similar, with the model using operating hours as a covariate resulting in an EUL. The parameter estimate on operating hours is negative, as expected, indicating that the more hours in operation, the more likely a failure will occur. We decided to use the model that incorporates operating hours as a covariate, because it adds more information to the model. Only the S160 CAC measure did not have a sufficient sample size to utilize operating hours.

As discussed above, the LIFEREG procedure was used to model the survival function for the L23 T8, L19 Delamping, L81 HID 251-400W and S160 CAC measures. Exhibit 3-33 compares the empirical survival function, with both the LIFEREG estimate of the exponential survival function and the exponential trendline, over the first 45 months of the measure's life. Although the exhibit cannot show a high enough level of detail, the two curves are virtually on top of each other.

Exhibit 3-33
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline versus Empirical Function
L23 T8 Measure

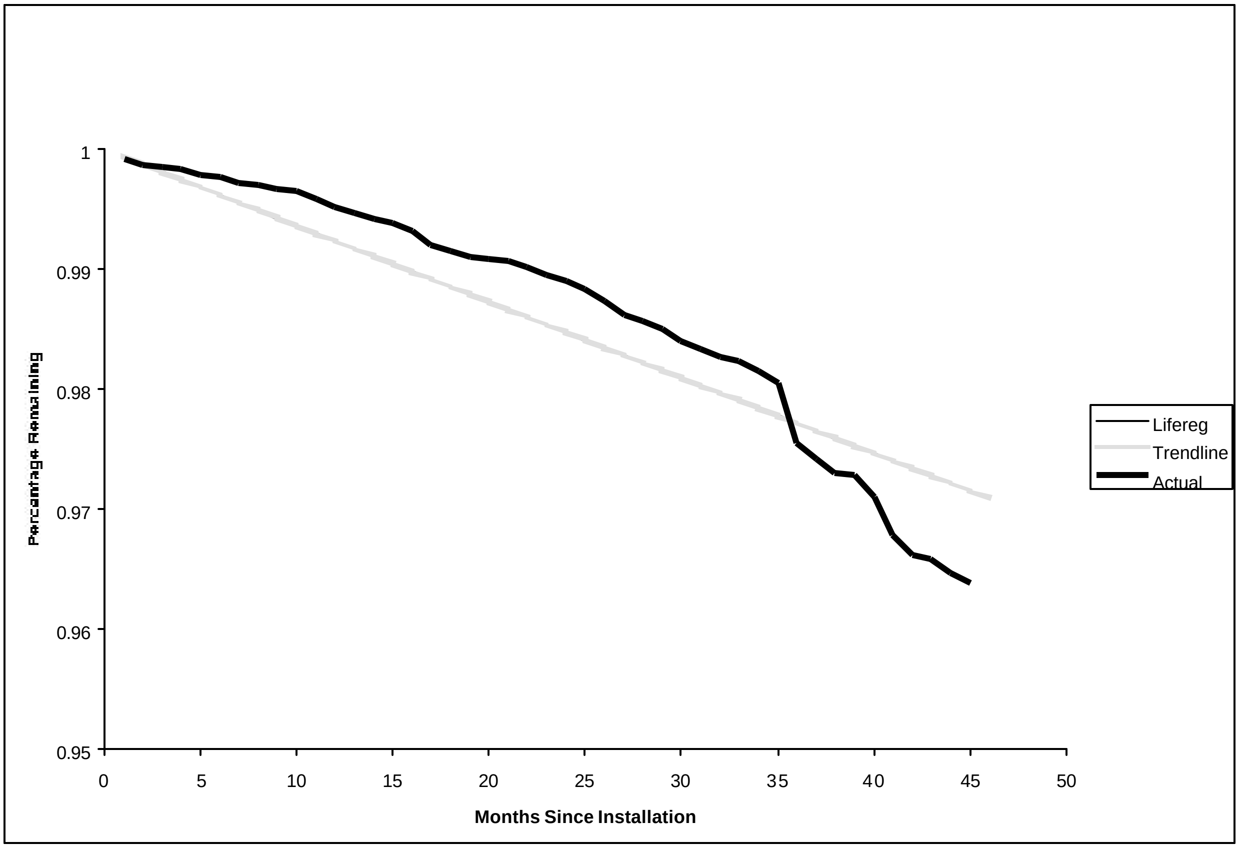


Exhibit 3-34 compares the exponential survival function versus the exponential trendline that was estimated based on the empirical survival function discussed above. Again, the LIFEREG and trendline curves are virtually identical.

Exhibit 3-34
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline
L23 T8 Measure

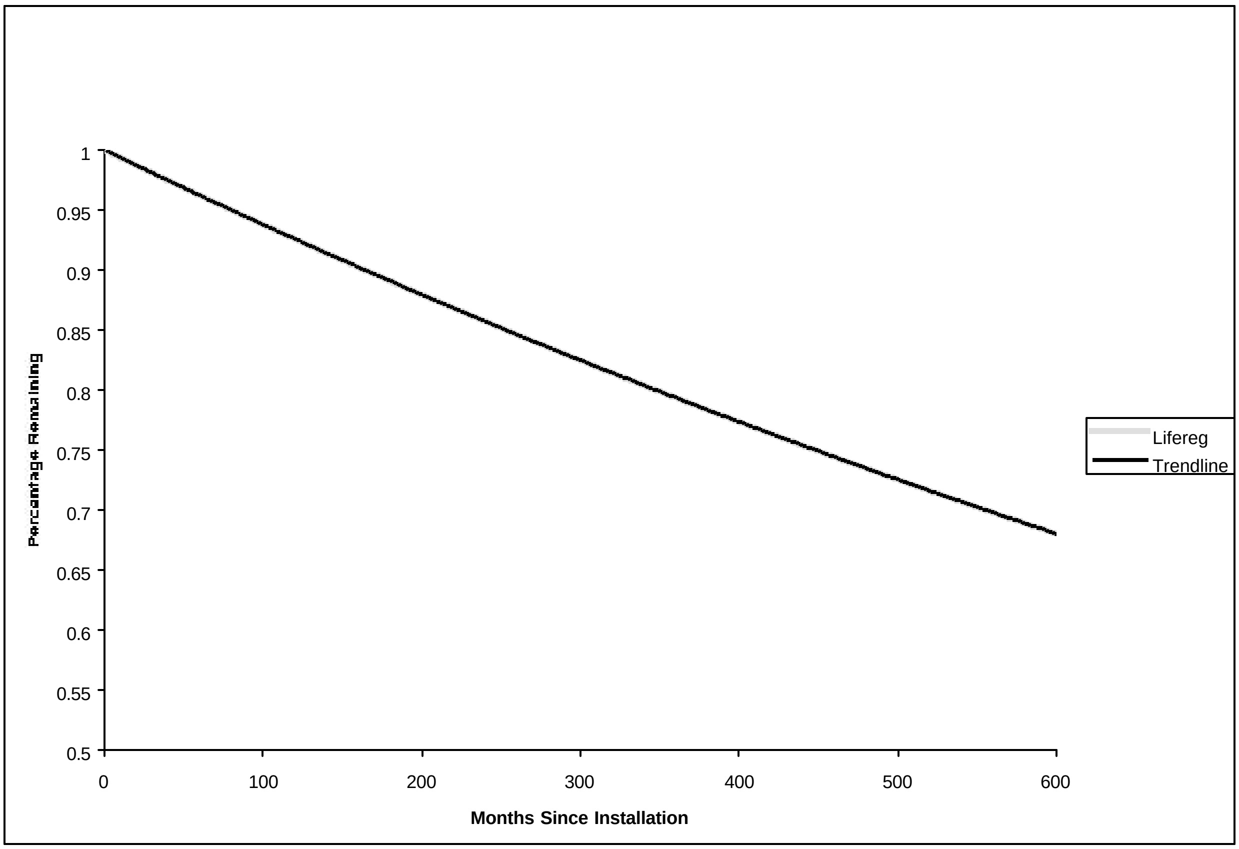


Exhibit 3-35 compares the survival functions based on the exponential, logistic, lognormal, Weibull and gamma distributions, estimated for the L23 T8 measure using the LIFEREG procedure with the empirical survival function, over the first 45 months of the measure's life. Exhibit 3-36 provides forecasts generated by these five survival functions over the first 600 months.

Exhibit 3-35
Comparison of Survival Functions
Exponential, Logistic, Lognormal, Weibull and Gamma versus Empirical Function
L23 T8 Measure

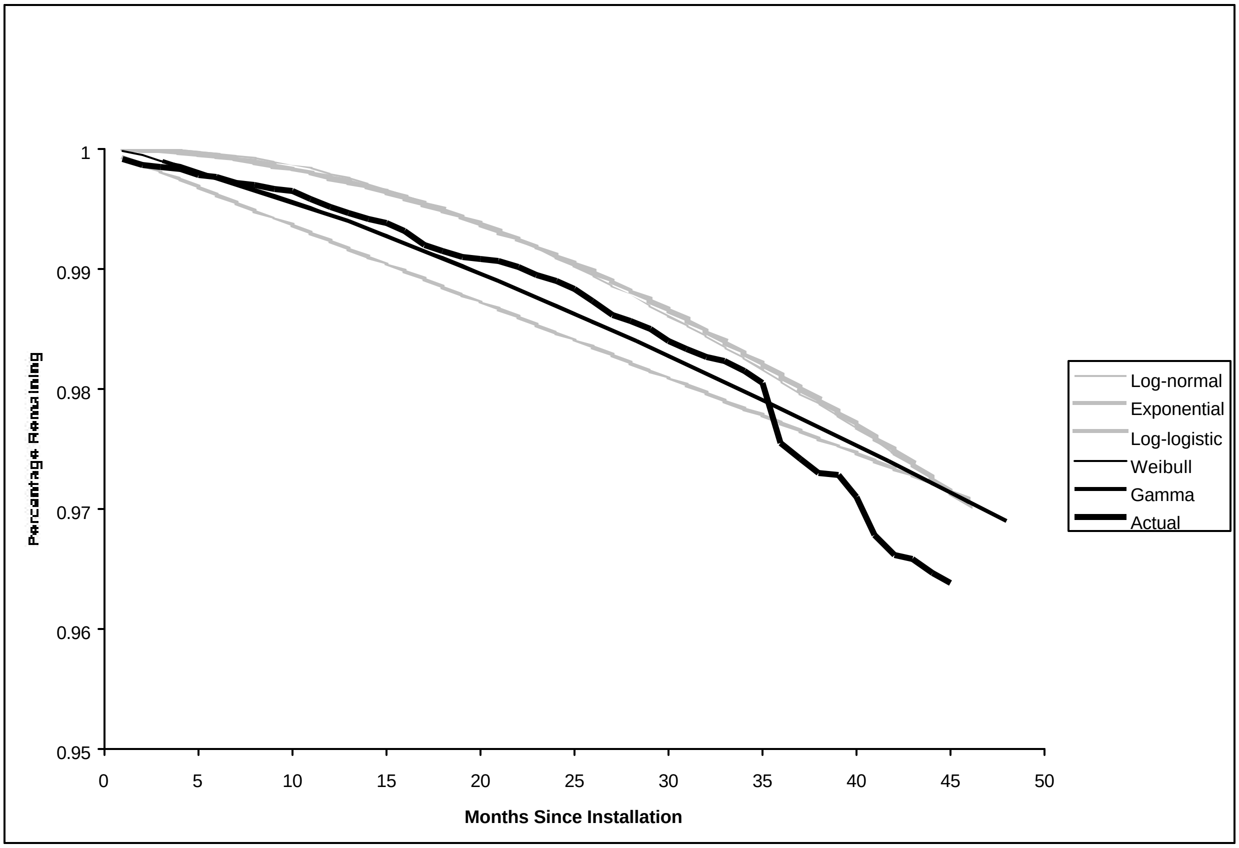


Exhibit 3-36
Exponential, Logistic, Lognormal, Weibull and Gamma Survival Functions
Based on LIFEREG Procedure
L23 T8 Measure

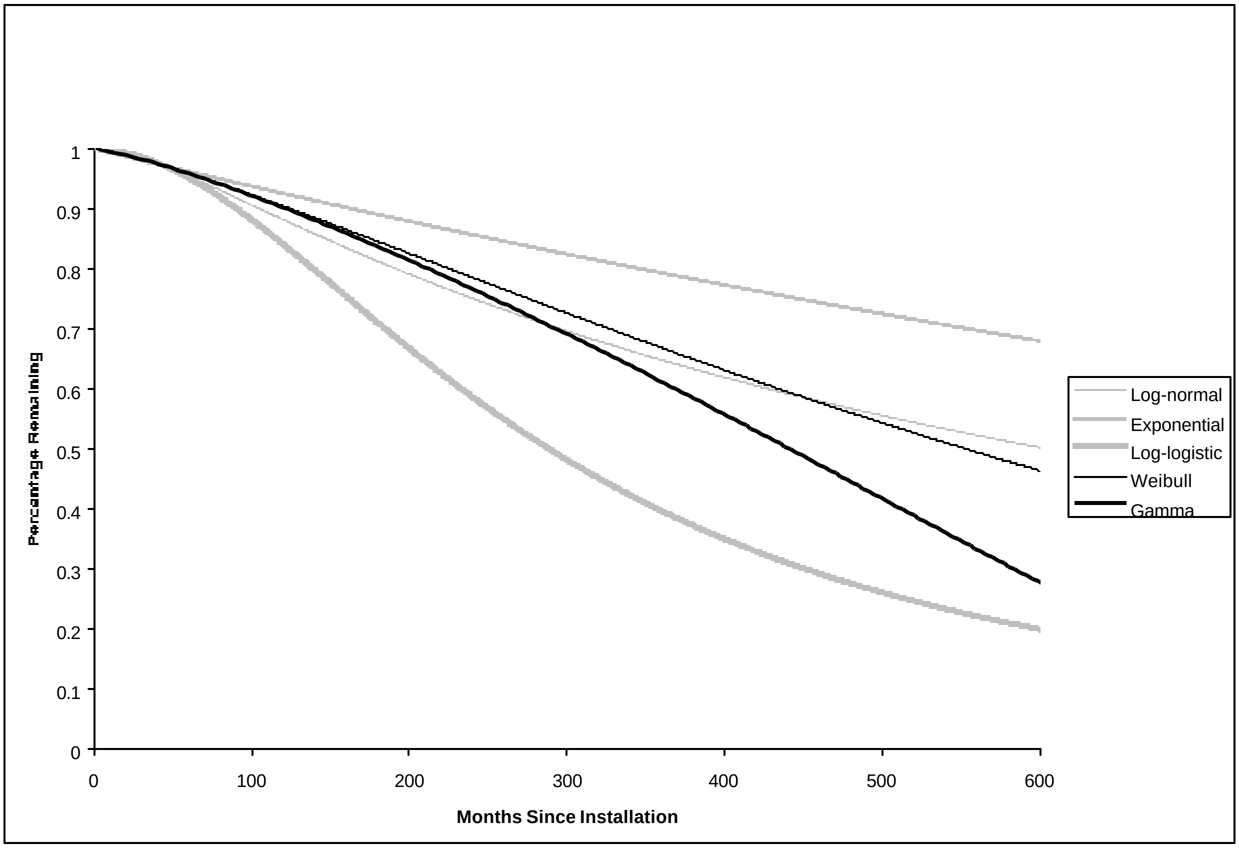


Exhibit 3-37 compares the empirical survival function, with both the LIFEREG estimate of the exponential survival function and the exponential trendline, over the first 45 months of the measure's life. Exhibit 3-38provides the estimated exponential survival function for the L19 Delamping measure, and compares it with exponential trendline that was estimated based on the empirical survival function discussed above over the first 600 months.

Exhibit 3-37
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline versus Empirical Function
L19 Delamping Measure

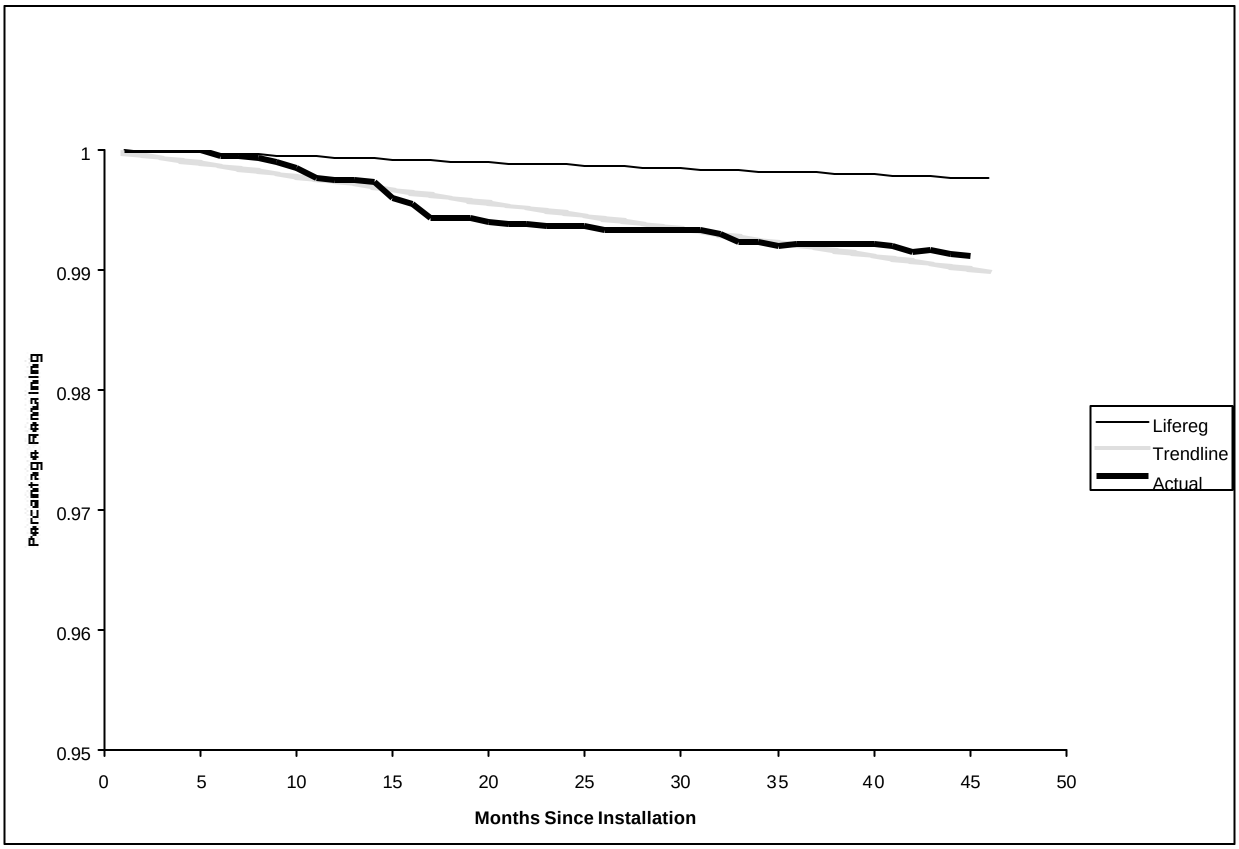


Exhibit 3-38
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline
L19 Delamping Measure

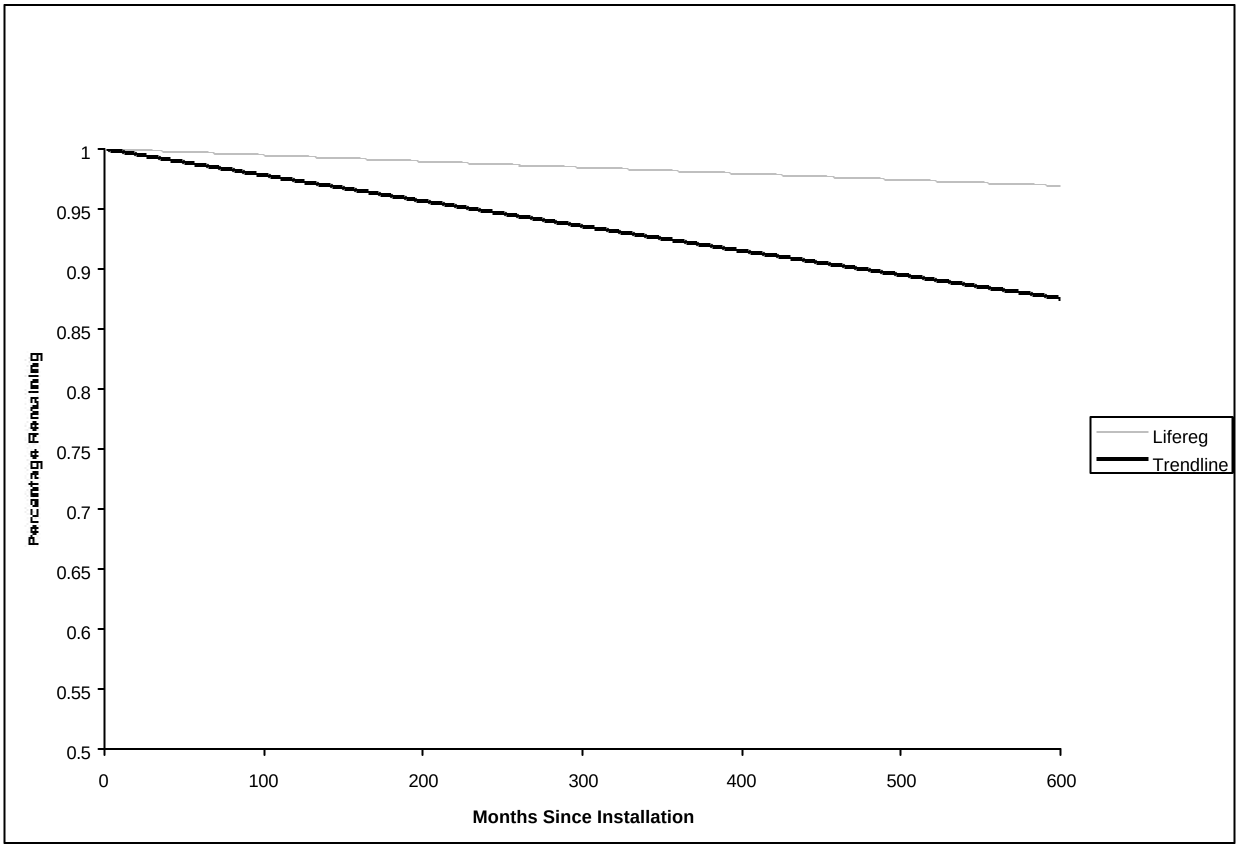


Exhibit 3-39 compares the survival functions based on the exponential, logistic, lognormal, Weibull and gamma distributions, estimated for the L19 Delamping measure using the LIFEREG procedure with the empirical survival function, over the first 45 months of the measure's life. Exhibit 3-40 provides forecasts generated by these five survival functions over the first 600 months.

Exhibit 3-39
Comparison of Survival Functions
Exponential, Logistic, Lognormal, Weibull and Gamma versus Empirical Function
L19 Delamping Measure

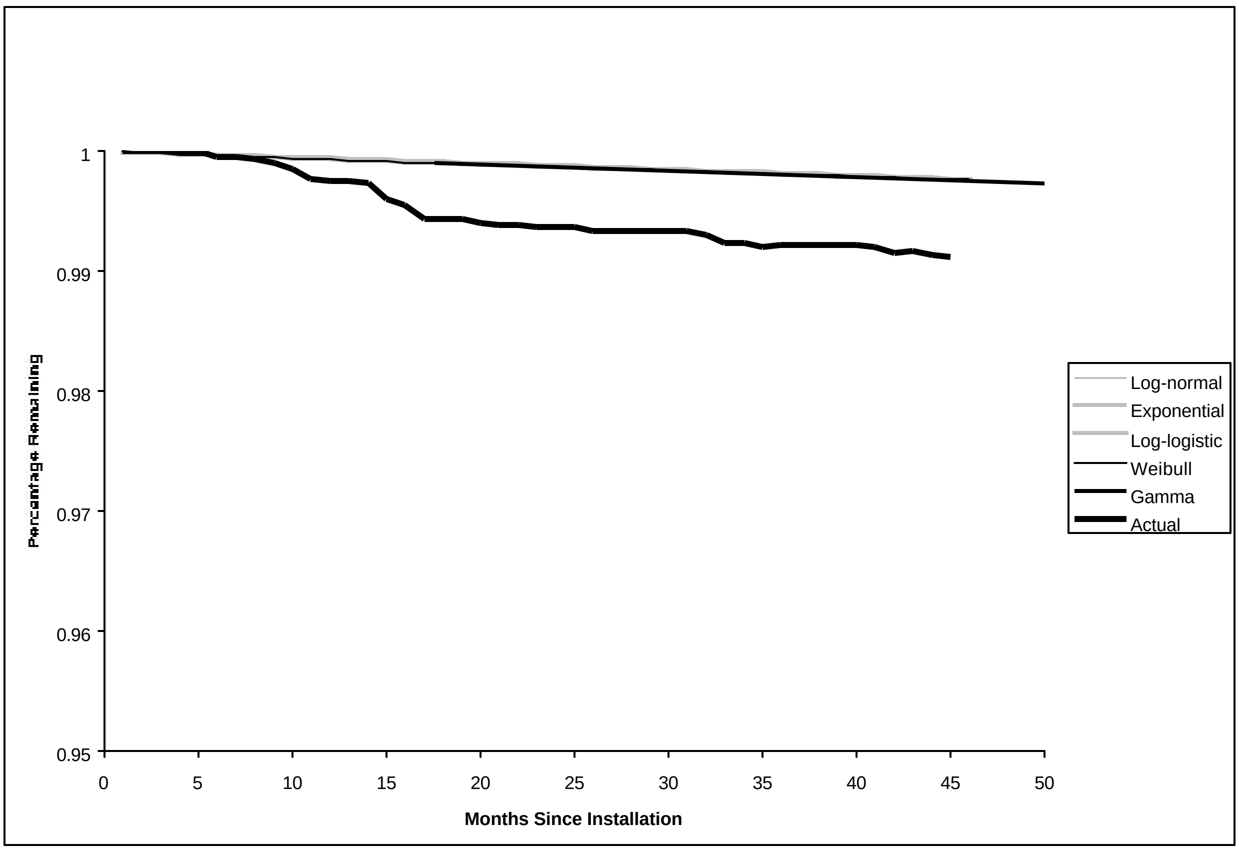
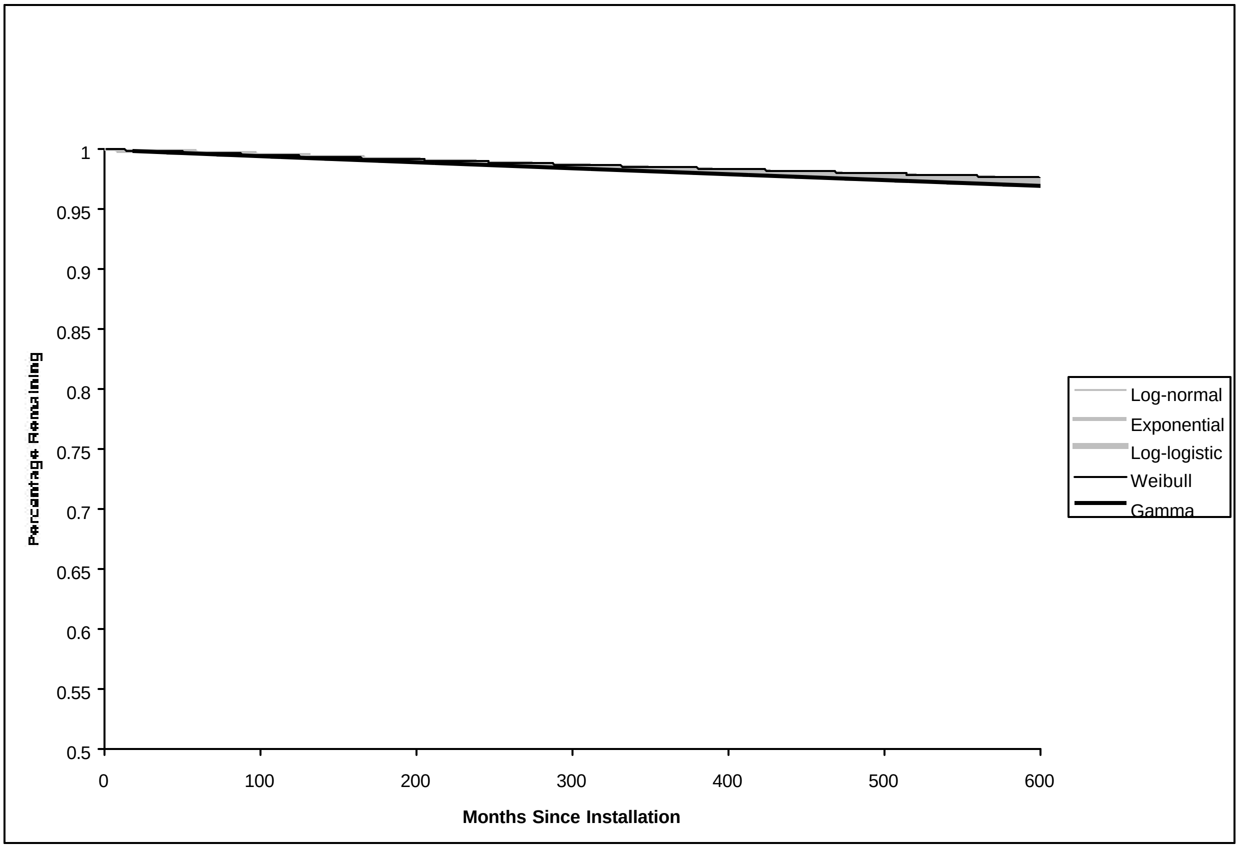


Exhibit 3-40
Exponential, Logistic, Lognormal, Weibull and Gamma Survival Functions
Based on LIFEREG Procedure
L19 Delamping



Below, we provide the results of the L81 HID 251-400W measure. Exhibit 3-41 compares the empirical survival function, for the L81 HID 251-400W measure, with both the LIFEREG estimate of the exponential survival function and the exponential trendline, over the first 45 months of the measure's life. Exhibit 3-42 provides the estimated exponential survival function, and compares it with the exponential trendline that was estimated based on the empirical survival function discussed above.

Exhibit 3-41
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline versus Empirical Function
L81 HID 251-400W Measure

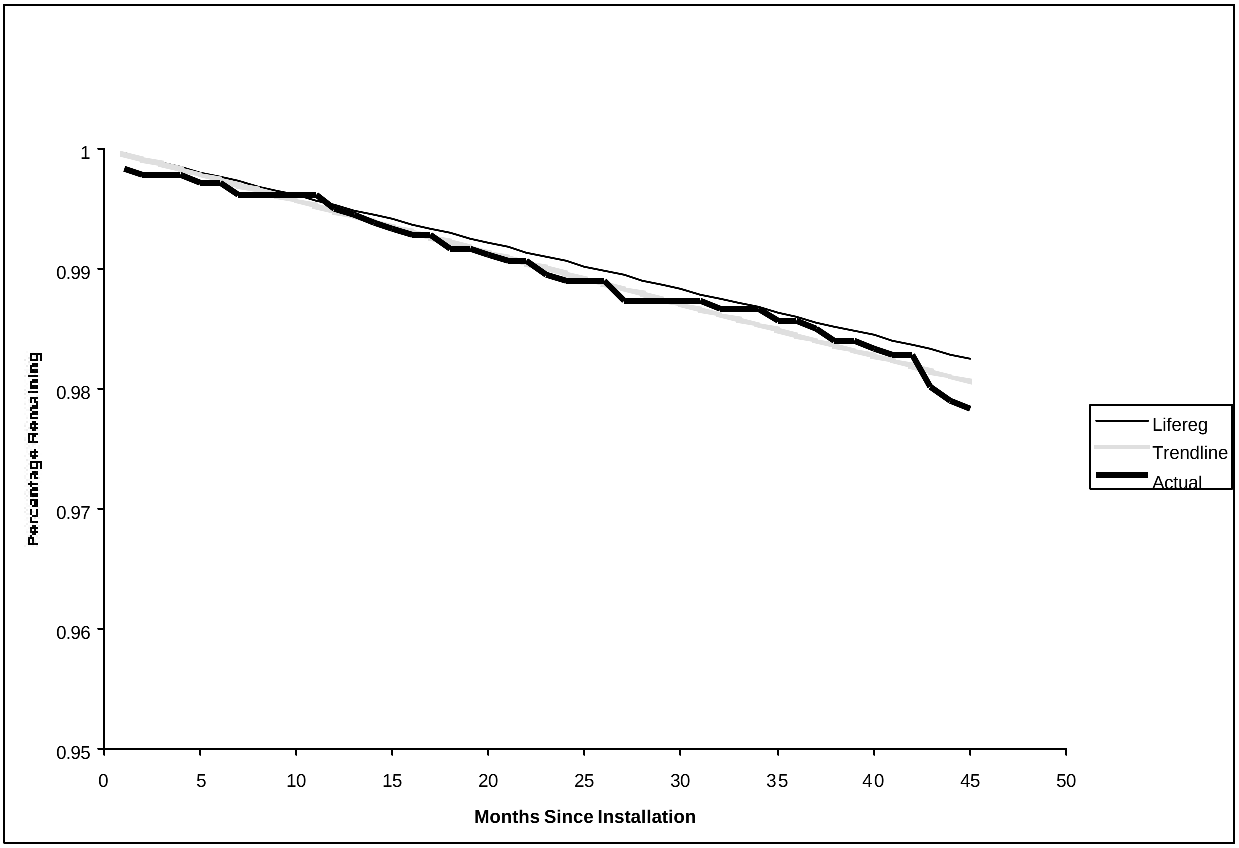


Exhibit 3-42
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline
L81 HID 251-400W Measure

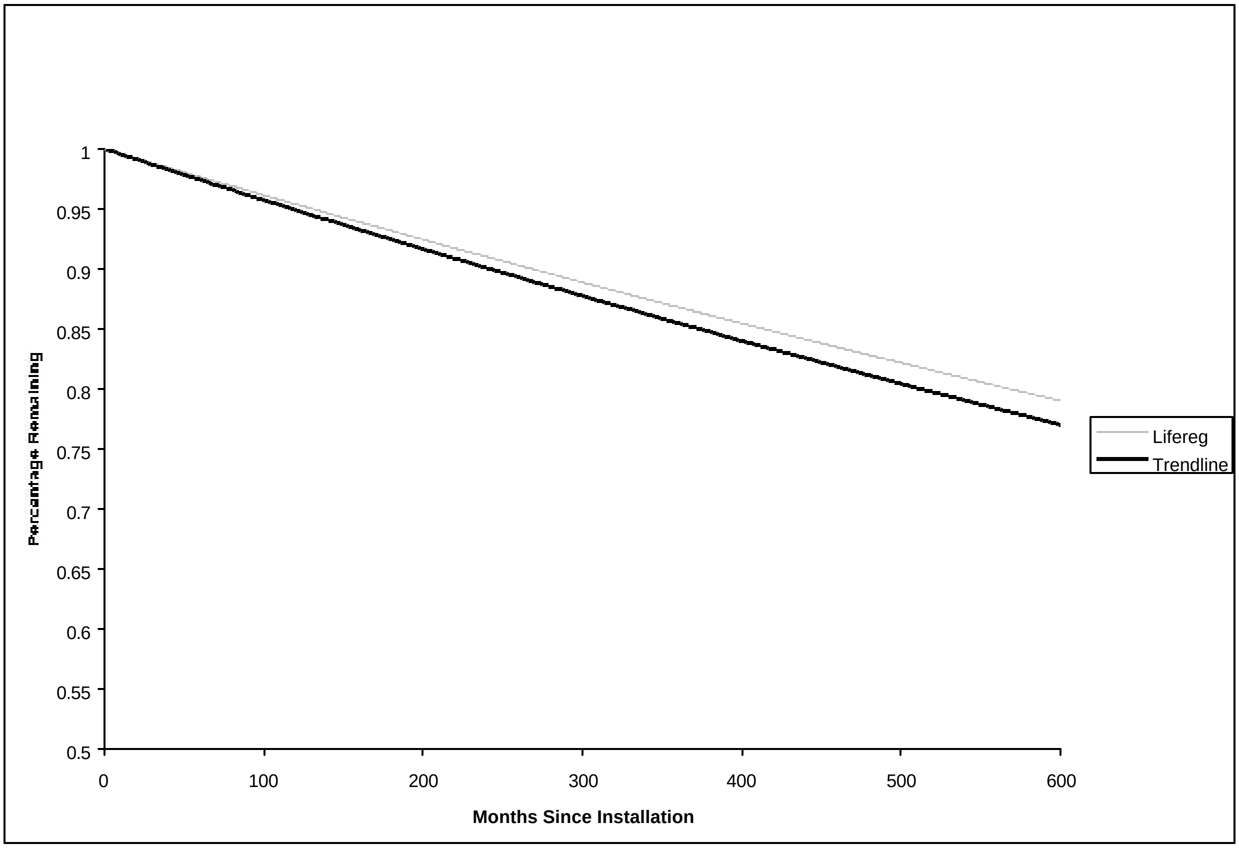


Exhibit 3-43 compares the survival functions based on the exponential, logistic, lognormal, Weibull and gamma distributions, estimated for the L81 HID 251-400W measure using the LIFEREG procedure with the empirical survival function, over the first 45 months of the measure's life. Exhibit 3-44 provides forecasts generated by these five survival functions over the first 600 months.

Exhibit 3-43
Comparison of Survival Functions
Exponential, Logistic, Lognormal, Weibull and Gamma versus Empirical Function
L81 HID 251-400W Measure

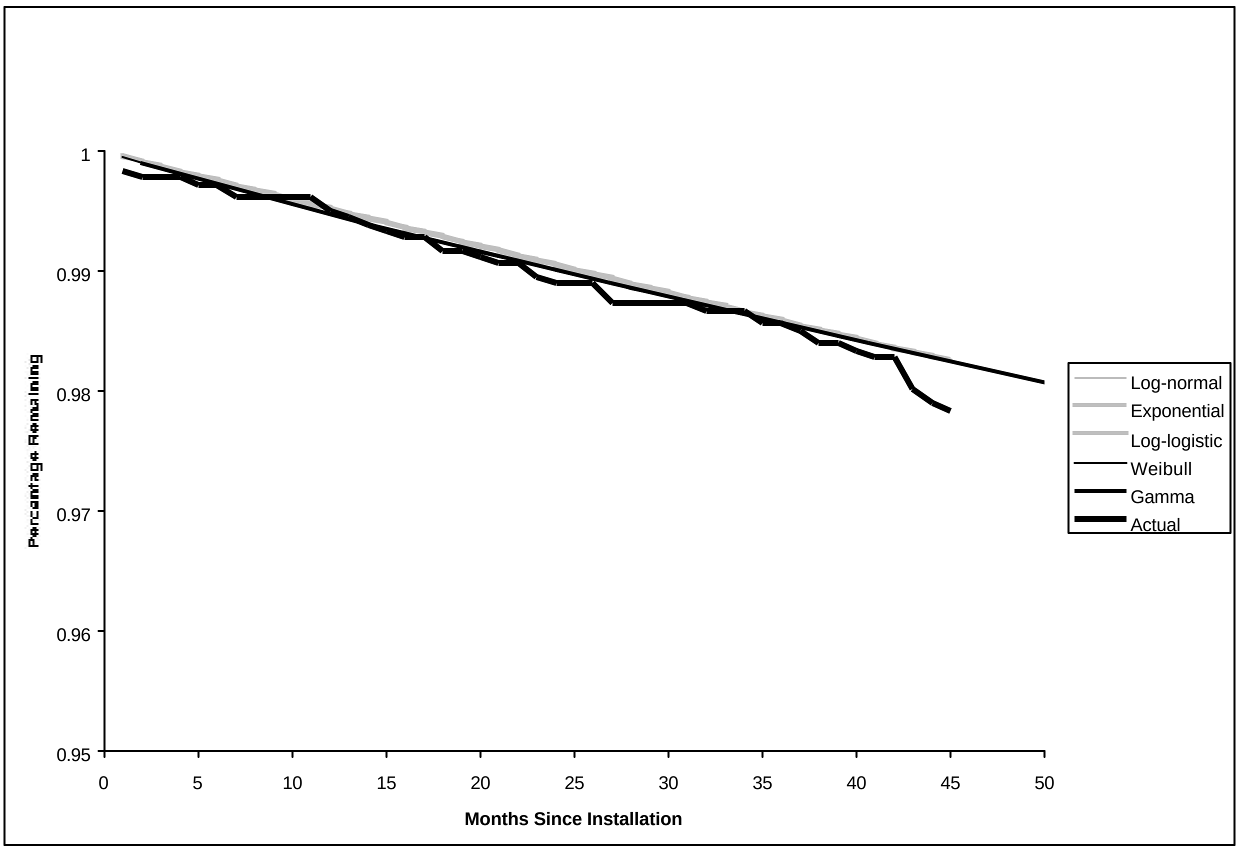


Exhibit 3-44
Exponential, Logistic, Lognormal, Weibull and Gamma Survival Functions
Based on LIFEREG Procedure
L81 HID 251-400W Measure

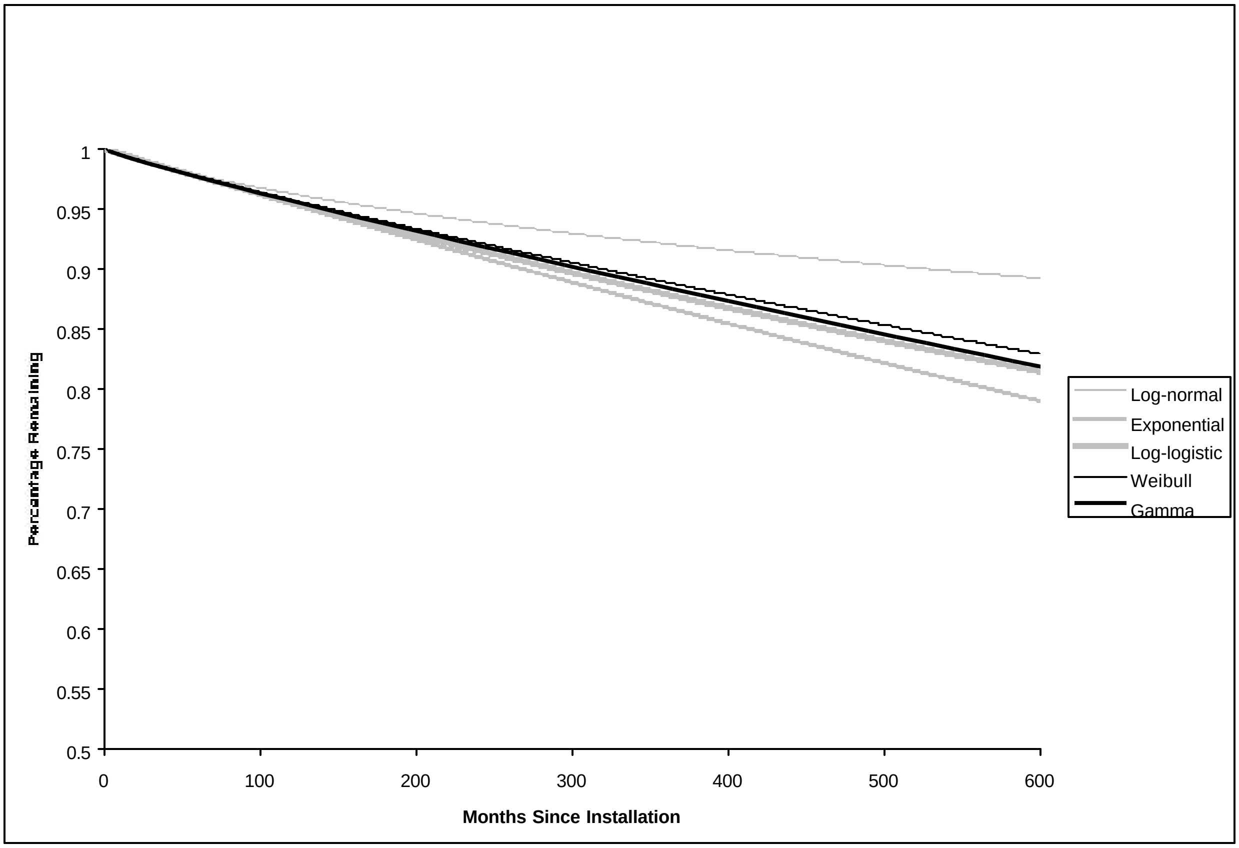


Exhibit 3-45 compares the empirical survival function for the S160 CAC measure, with both the LIFEREG estimate of the exponential survival function and the exponential trendline, over the first 45 months of the measure's life. Exhibit 3-46 provides the estimated exponential survival function, and compares it with exponential trendline that was estimated based on the empirical survival function discussed above.

Exhibit 3-45
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline versus Empirical Function
S160 CAC Measure

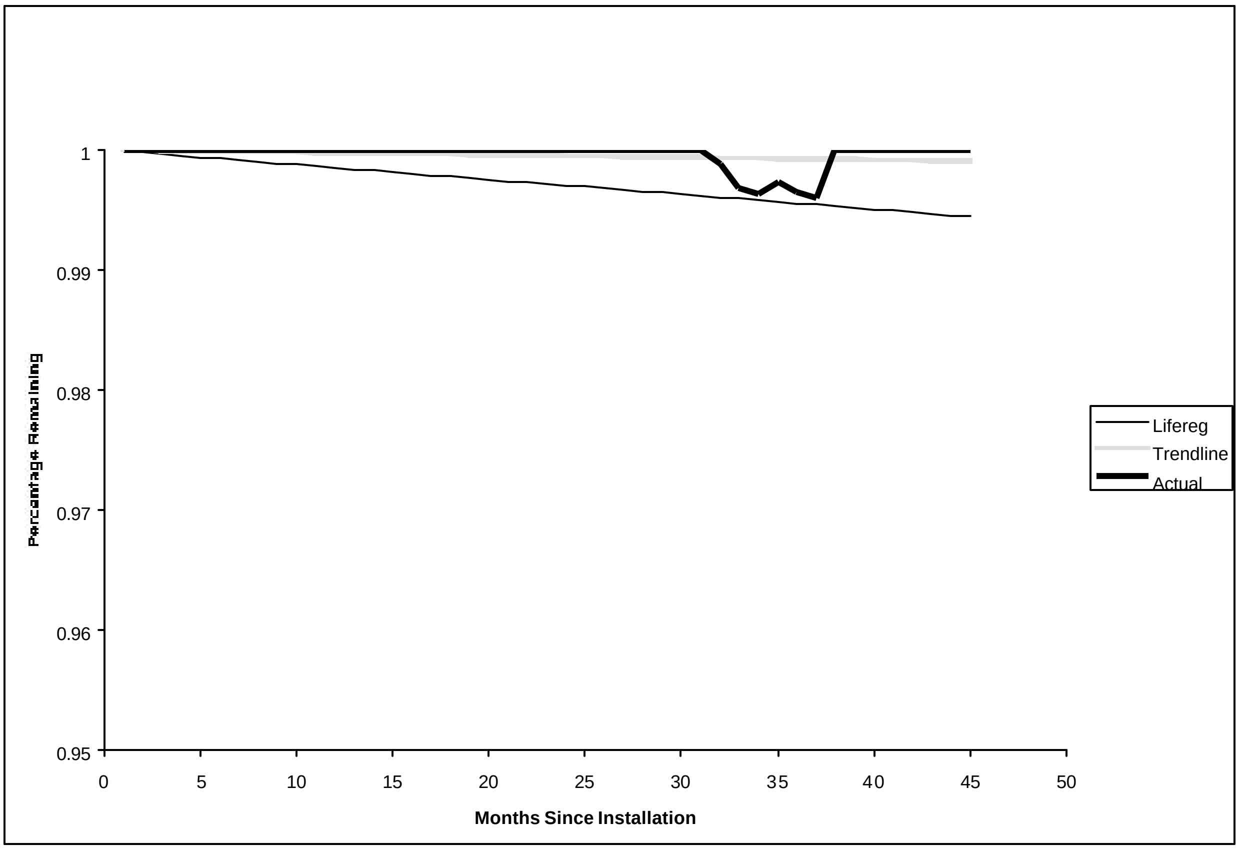


Exhibit 3-46
Comparison of Survival Functions
LIFEREG Exponential Model versus Exponential Trendline
S160 CAC Measure

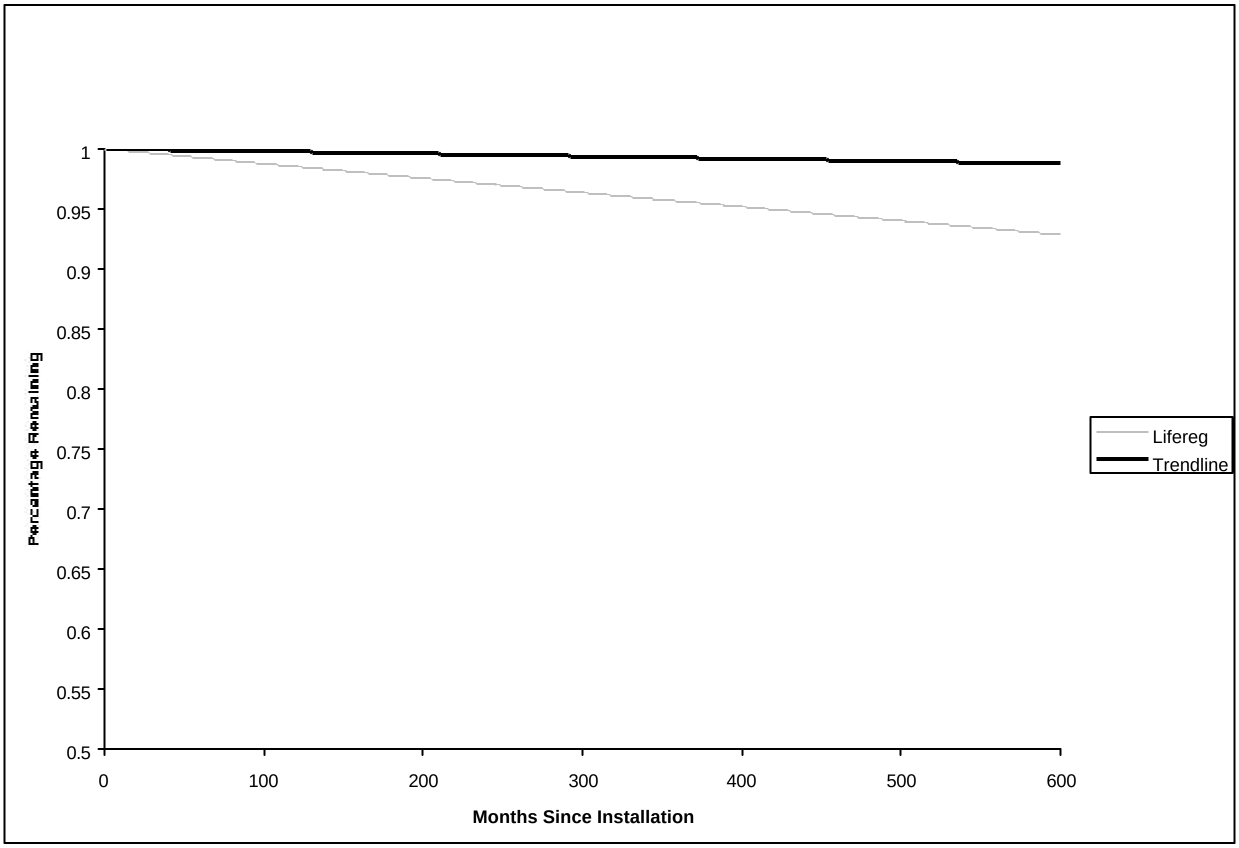
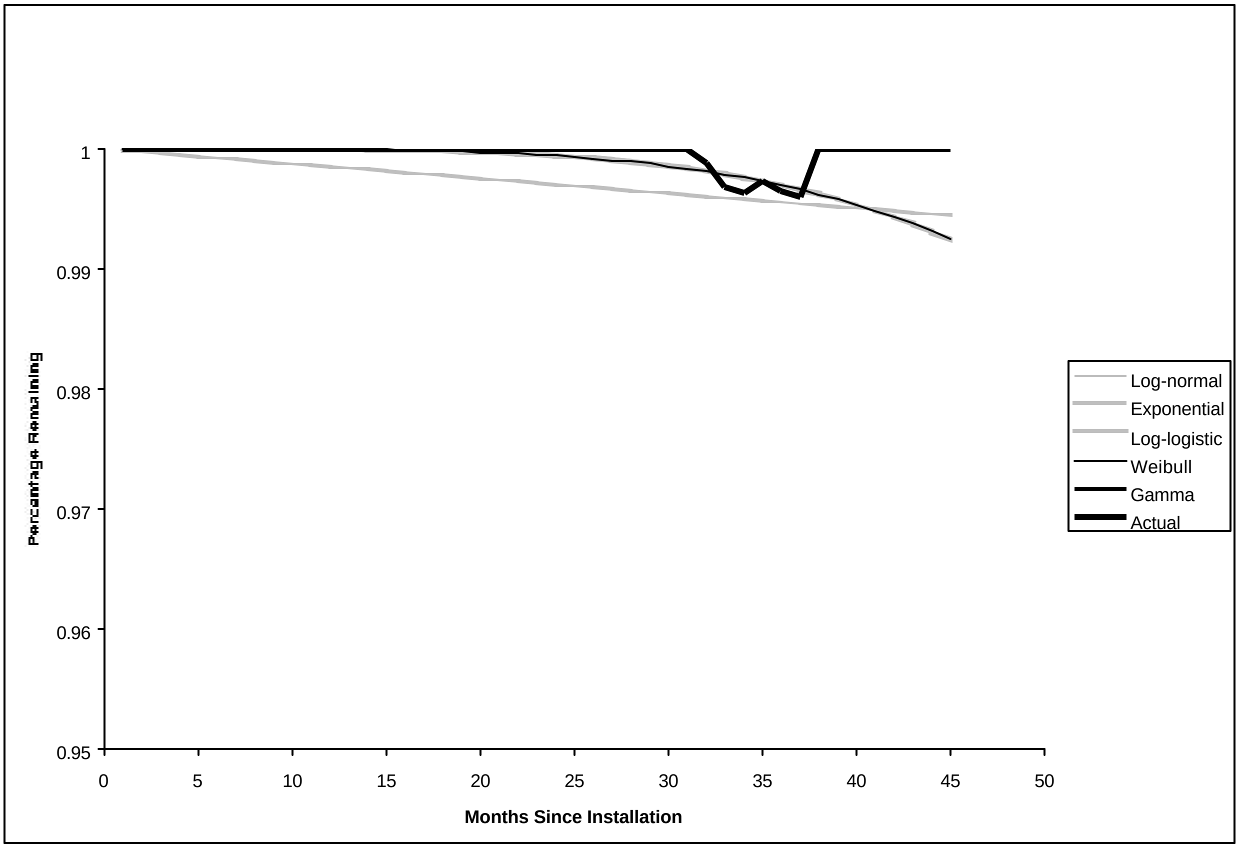


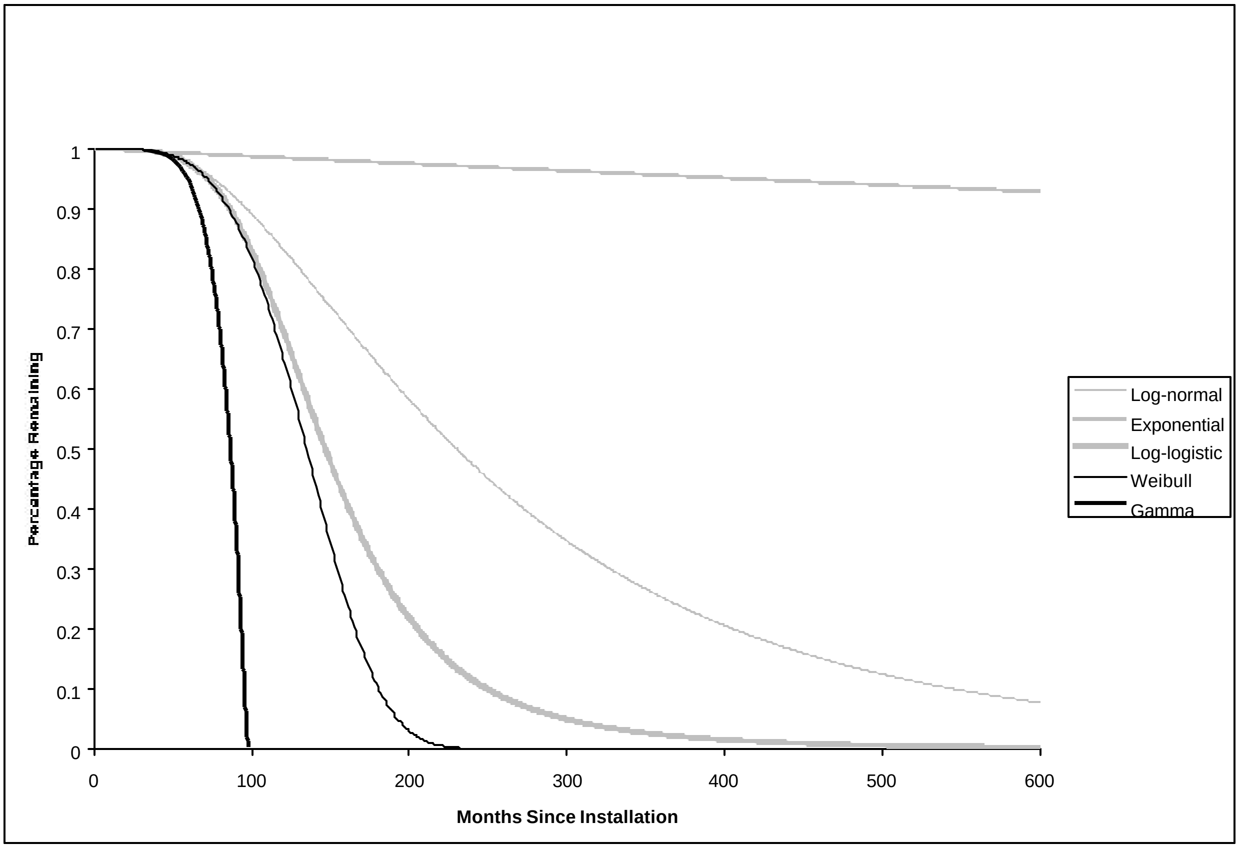
Exhibit 3-47 compares the survival functions based on the exponential, logistic, lognormal, Weibull and gamma distributions, estimated for the S160 CAC measure using the LIFEREG procedure, with the empirical survival function over the first 45 months of the measures's life.

Exhibit 3-47
Comparison of Survival Functions
Exponential, Logistic, Lognormal, Weibull and Gamma versus Empirical Function
S160 CAC Measure



It is clear in this exhibit that the model sees the few failures (3 air conditioners representing 12 tons of cooling out of over 2,500 total tons) at the end of the study period. This causes the model to forecast the distribution with a sharply increasing hazard rate, leading to a very short EUL. This is illustrated below in Exhibit 3-48, which provides forecasts generated by these five survival functions over the first 600 months.

Exhibit 3-48
Exponential, Logistic, Lognormal, Weibull and Gamma Survival Functions
Based on LIFEREG Procedure
S160 CAC Measure

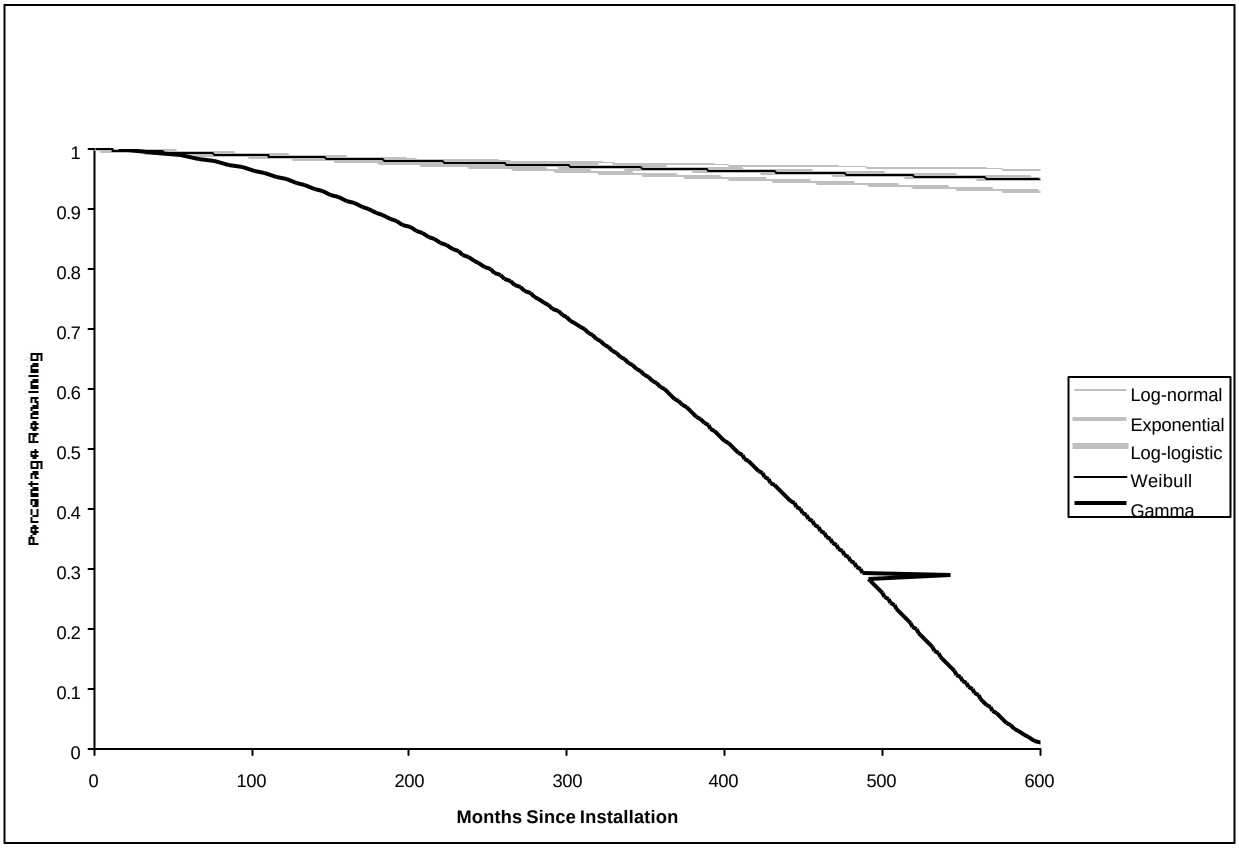


The distributions displayed in Exhibit 3-48 predict a very short EUL for all but the Exponential distribution. At first glance this may not seem unusual, but a closer look at the data reveals that the distributions may be subject to a bias. We do not think that this is an accurate representation of the distribution for the S160 CAC measure for several reasons. First of all, it is highly unlikely that new energy efficient air conditioners would exhibit this type of survival function without older, standard efficiency air conditioners having similar survival functions. Furthermore, this extremely high hazard function would not have gone unnoticed by the industry or consumers.

A brief discussion of the properties of the model is warranted here. In the case of the S160 CAC measure, there are no failures until just before the end of the study period. The model accounts for these late failures as a sharply increasing hazard rate. With no future events to stabilize the model, it continues to forecast with this increasing hazard rate through the EUL point. With only one half of one percent of the surveyed equipment failed, removed, or replaced, there simply is not enough information to support this case.

To test the sensitivity of the model to the date of the failure, the same analysis was performed for the S160 CAC measure using event dates that were randomly distributed during the survey length. Exhibit 3-49 presents the resulting distribution.

Exhibit 3-49
Exponential, Logistic, Lognormal, Weibull and Gamma Survival Functions
Based on LIFEREG Procedure and Random Event Date Distribution
S160 CAC Measure



As shown in the previous exhibit, once the date of the events are dispersed more uniformly the model predicts much larger EUL's (as opposed to occurring near the end of the survey length). Even the Gamma distribution predicts an EUL of 34 years under these conditions, compared to 7.2 years using the actual dates. Although we are not suggesting to use these modified results over the actual results, the differences in the final EUL are worth noting.

Exhibit 3-50 below summarizes all of the results of the LIFEREG models. Shown for each model are the parameter estimates and standard errors for every variable included in the model specification. Furthermore, the resulting EUL and its associated standard error are provided.

It should be noted that the standard errors that were directly output by SAS were adjusted to account for the correlation problem discussed earlier in Section 3.1. Recall that the failure and removal rates associated with measures installed at the same site are correlated. For example, when a removal occurs, it is likely that many measures are removed at once. To a lesser extent, failures are correlated since they may all come from the same manufacturing lot, they are all likely to be installed under the same circumstances, and they are also used in a similar manner. Attachment 3, Protocol Table 7B, discusses the development of standard errors in more detail.

Exhibit 3-50
Comparison of Survival Model Results
Exponential, Logistic, Lognormal, Weibull and Gamma Models
L23 T8, L19 Delamping, L81 HID, and S160 CAC Measures

			Variable			Resulting
Measure	**Model**		**Intercept**	**Scale**	**Ophours**	**EUL**
L23	Exponential	Parameter Estimate	8.01	1.00	-0.00018	**90.0**
		Standard Error	1.69	0.00	0.00042	22.49
	Logistic	Parameter Estimate	5.92	0.53	-0.00007	**24.1**
		Standard Error	1.18	0.17	0.00023	13.92
	Log-Normal	Parameter Estimate	6.81	1.37	-0.00011	**50.5**
		Standard Error	1.35	0.40	0.00025	38.95
	Weibull	Parameter Estimate	7.08	0.79	-0.00013	**46.2**
		Standard Error	1.59	0.19	0.00034	28.28
	Gamma	Estimate	6.92	0.34	-0.00013	**36.8**
		Standard Error	5.24	15.58	0.00034	1,598.94
L19	Exponential	Parameter Estimate	17.32	1.00	-0.00204	**364.7**
		Standard Error	9.19	0.00	0.00213	252.80
	Logistic	Parameter Estimate	17.94	1.04	-0.00214	**649.6**
		Standard Error	13.18	0.70	0.00260	2,262.21
	Log-Normal	Parameter Estimate	20.31	2.95	-0.00221	**5,031.5**
		Standard Error	14.50	1.86	0.00277	23,344.88
	Weibull	Parameter Estimate	18.79	1.11	-0.00226	**609.4**
		Standard Error	13.73	0.71	0.00272	2,021.88
	Gamma	Estimate	17.29	0.37	-0.00204	**928.6**
		Standard Error	12.03	0.22	0.00244	3,290.07
L81	Exponential	Parameter Estimate	8.27	1.00	-0.00011	**147.2**
		Standard Error	5.17	0.00	0.00126	69.86
	Logistic	Parameter Estimate	8.36	1.02	-0.00012	**225.3**
		Standard Error	5.88	0.52	0.00130	477.50
	Log-Normal	Parameter Estimate	10.88	2.96	-0.00022	**1,818.3**
		Standard Error	7.41	1.34	0.00162	5,298.42
	Weibull	Parameter Estimate	8.74	1.10	-0.00013	**210.8**
		Standard Error	6.20	0.52	0.00139	408.00
	Gamma	Estimate	8.54	0.85	-0.00013	**171.0**
		Standard Error	11.50	0.00	0.00138	2,312.94
S160	Exponential	Parameter Estimate	9.00	1.00	-	**467.0**
		Standard Error	0.94	0.00	-	437.44
	Logistic	Parameter Estimate	4.99	0.24	-	**12.3**
		Standard Error	1.00	0.19	-	12.21
	Log-Normal	Parameter Estimate	5.44	0.67	-	**19.2**
		Standard Error	1.34	0.51	-	25.74
	Weibull	Parameter Estimate	4.99	0.24	-	**11.3**
		Standard Error	1.00	0.19	-	10.49
	Gamma	Estimate	4.55	0.04	-	**7.2**
		Standard Error	0.39	0.02	-	2.52

Keep in mind that, as discussed above, the S160 CAC measure did not exhibit enough events during the study period to accurately model the survival function. Again, this is not an issue of sample size, but instead it is an issue of not enough elapsed time between installation and retention study.

3.7 COMPETING RISKS MODELS

The final analysis step, as described in Section 3.2 above, was to develop competing risks models to account for multiple events influencing the survival distribution. The first task in developing competing risks models was to calculate hazard functions for all events individually. The hazard rate at each time step is simply the derivative of the survival function, or the number of events occurring over that time step divided by the remaining population at that time.

The next task is to create the competing risks model. This is accomplished by combining hazard rates from both failures and removals into one joint probability function.

Three different sets of output were generated from this model. The first output contains the best-fitting distribution for each event based on the log-likelihood estimate, which is a parameter output by SAS used to judge how well the model fits the actual data. The second output provides the minimum EUL estimate, and the third output provides the maximum EUL estimate. A summary of the different distributions that were chosen for each of the models is presented in Exhibit 3-51. As shown in the exhibit, the only measure to exhibit all three events (failures, removals, and replacements) was the L23 T8 measure. The L19 Delamping measure only experienced one event type during this study, excluding it from the competing risks models.

Exhibit 3-51
Comparison of Distributions used in the Competing Risks Model

		Distribution		
Measure	**Method**	**Failures**	**Removals**	**Replacements**
L23 T8	Best Fit	Logistic	Weibull	Log-Normal
	Min EUL	Gamma	Gamma	Logistic
	Max EUL	Exponential	Exponential	Exponential
L81 HID	Best Fit	Logistic	-	Logistic
	Min EUL	Weibull	-	Exponential
	Max EUL	Exponential	-	Log-Normal
S160 CAC	Best Fit	-	Logistic	Log-Normal
	Min EUL	-	Gamma	Gamma
	Max EUL	-	Exponential	Exponential

The resulting survival functions for the L23 T8 measure are provided in Exhibit 3-52. For the best-fitting model, the Logistic distribution was selected for failures, the Weibull distribution was selected for removals, and the Lognormal distribution was selected for replacements.

Exhibit 3-52
Resulting Survival Functions from the Competing Risks Model
L23 T8 Measure

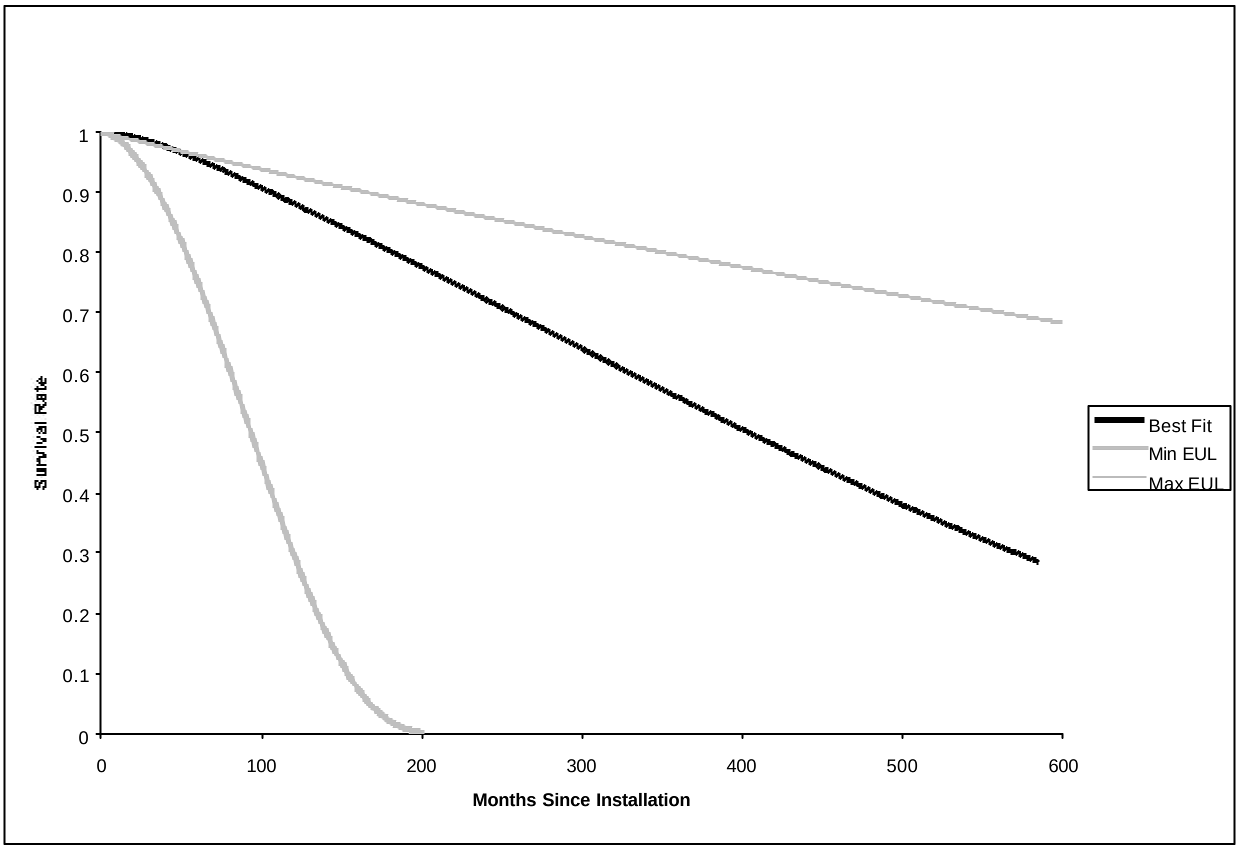


Exhibit 3-53 presents the results from the competing risks models in tabular format for the L23 T8 measure. For each case, the competing risks model EUL prediction is given along with its associated standard error. The properties for the event distributions (from the LIFEREG procedure in SAS) used to construct each competing risks model are also provided.

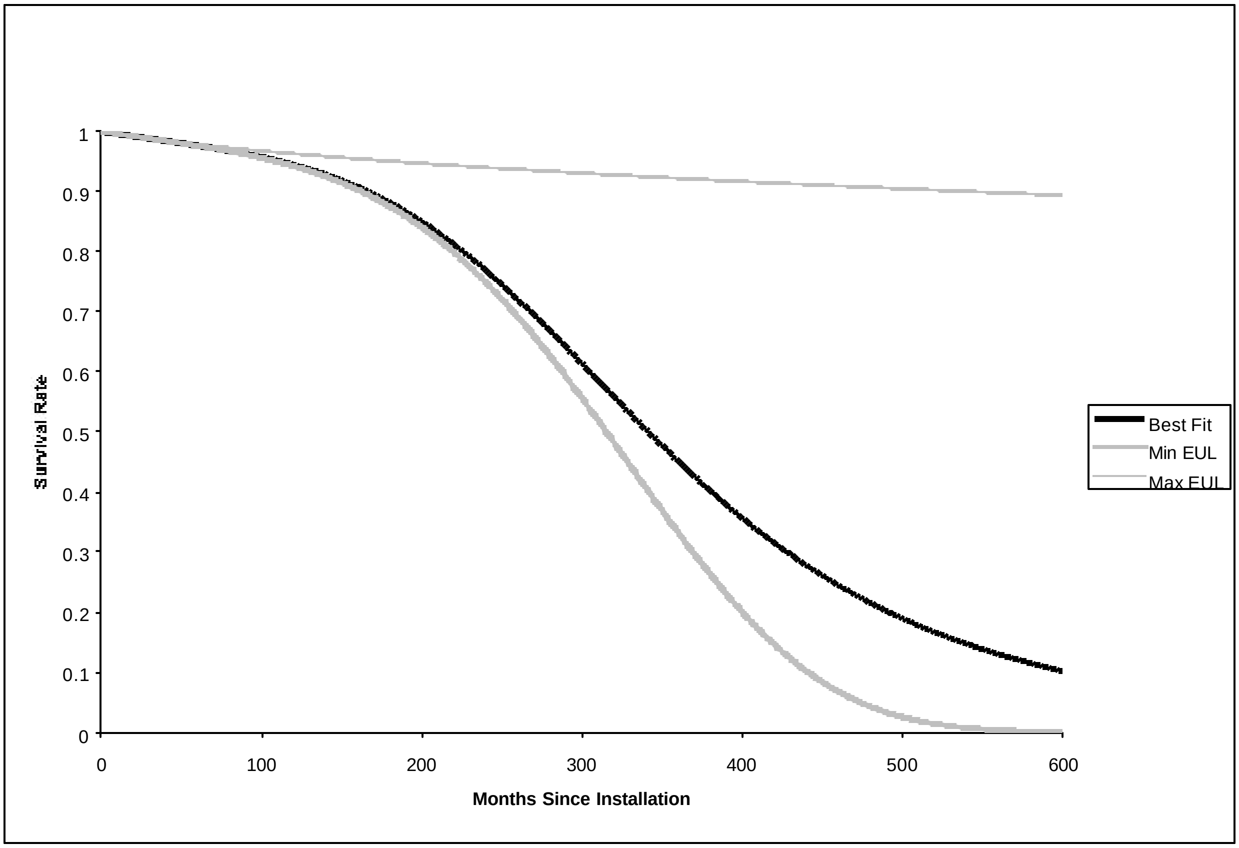
Exhibit 3-53
Results from Competing Risks Models
L23 T8 Measure

					Variable			Resulting
Measure	**Method**	**Model**			**Intercept**	**Scale**	**Ophours**	**EUL**
L23	Best Fit	Combined		Parameter Estimate	-	-	-	**33.8**
				Standard Error	-	-	-	87.33
		Failures	Logistic	Parameter Estimate	11.91	0.35	-0.0013	**112.1**
				Standard Error	14.24	0.56	0.0026	495.70
		Removals	Weibull	Parameter Estimate	7.34	0.43	-0.0002	**51.7**
				Standard Error	4.95	0.38	0.0009	115.54
		Replacements	Log-Normal	Parameter Estimate	6.97	1.46	-0.0001	**63.6**
				Standard Error	1.50	0.47	0.0003	58.12
L23	Min EUL	Combined		Parameter Estimate	-	-	-	**7.8**
				Standard Error	-	-	-	59.39
		Failures	Gamma	Parameter Estimate	11.87	0.16	-0.0013	**89.8**
				Standard Error	14.28	0.26	0.0026	452.0
		Removals	Gamma	Parameter Estimate	6.06	0.11	-0.0001	**21.0**
				Standard Error	2.89	0.07	0.0006	22.72
		Replacements	Logistic	Parameter Estimate	5.98	0.55	0.0000	**27.4**
				Standard Error	1.28	0.19	0.0002	18.52
L23	Max EUL	Combined		Parameter Estimate	-	-	-	**91.1**
				Standard Error	-	-	-	193.72
		Failures	Exponential	Parameter Estimate	24.82	1.00	-0.0032	**7,111.5**
				Standard Error	25.50	0.00	0.0057	18,510.3
		Removals	Exponential	Parameter Estimate	12.11	1.00	-0.0005	**1,276.4**
				Standard Error	8.09	0.00	0.0020	1,234.4
		Replacements	Exponential	Parameter Estimate	7.91	1.00	-0.0001	**96.8**
				Standard Error	1.72	0.00	0.0004	25.4

As the exhibit shows, there is a wide variation in the EUL from the minimum to the maximum. The actual range is 8 years for the minimum EUL and 91 years for the maximum EUL, with the best fit having an EUL of 34 years.

Exhibit 3-54 provides the competing risks results for the L81 HID measure. This exhibit clearly illustrates the wide range in values. The maximum EUL model did not reach the median point within 100 years, so the model was stopped. An EUL of 2,218 years was given to this model, which represents the minimum EUL from the Exponential failure and Log-Normal replacement models.

Exhibit 3-54
Resulting Survival Functions from the Competing Risks Model
L81 HID Measure



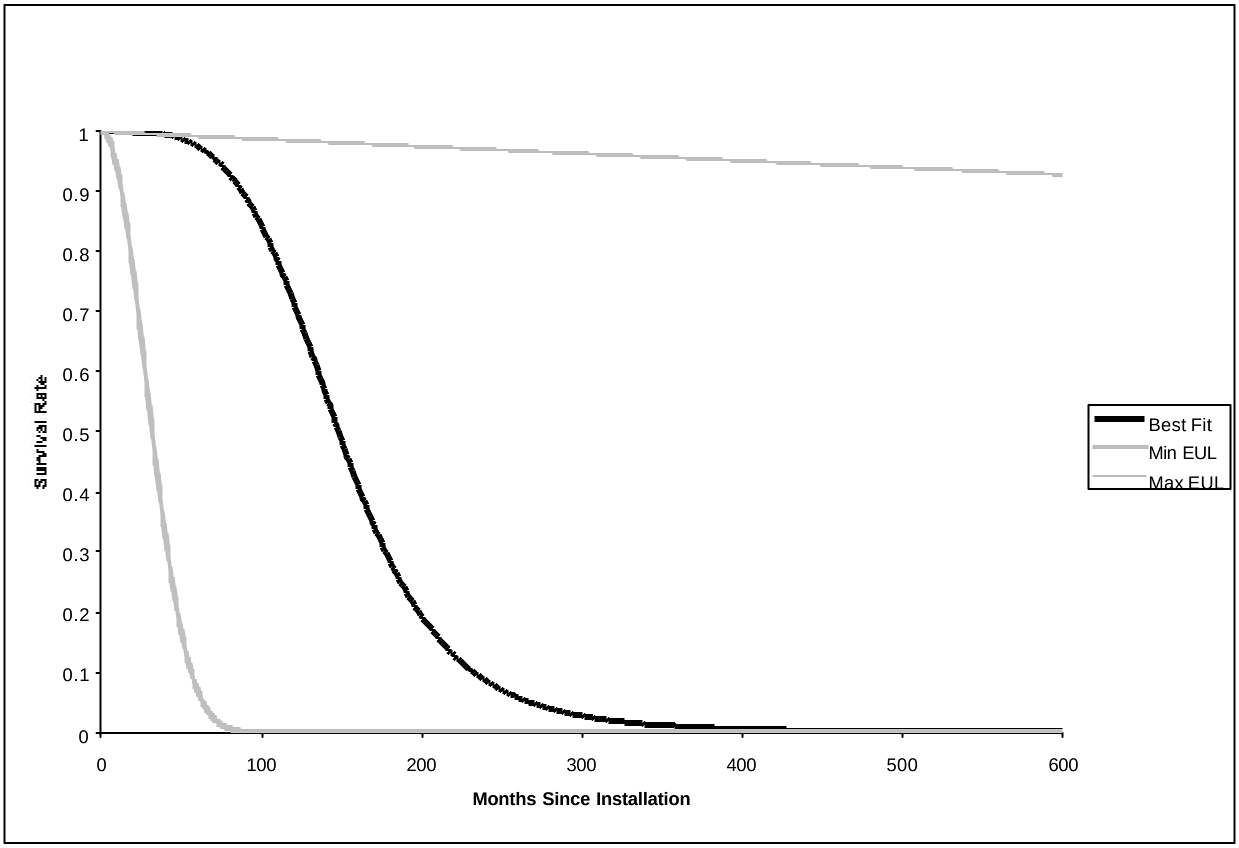
The detailed results from the competing risks models for the L81 HID measure are presented in Exhibit 3-55. Again, the competing risks model EUL prediction is provided along with the underlying assumptions.

Exhibit 3-55
Results from Competing Risks Models
L81 HID Measure

					Variable			Resulting
Measure	**Method**	**Model**			**Intercept**	**Scale**	**Ophours**	**EUL**
L81	Best Fit	Combined		Parameter Estimate	-	-	-	**28.5**
				Standard Error	-	-	-	346.1
		Failures	Logistic	Parameter Estimate	7.03	0.26	-0.0003	**30.3**
				Standard Error	11.62	0.67	0.0019	155.5
		Replacements	Logistic	Estimate	8.34	1.05	-0.0001	**257.2**
				Standard Error	6.04	0.54	0.0013	572.1
L81	Min EUL	Combined		Parameter Estimate	-	-	-	**26.3**
				Standard Error	-	-	-	315.69
		Failures	Weibull	Parameter Estimate	7.02	0.26	-0.0003	**27.6**
				Standard Error	11.61	0.67	0.0019	134.8
		Replacements	Exponential	Parameter Estimate	8.12	1.00	-0.0001	**149.7**
				Standard Error	5.20	0.00	0.0013	72.1
L81	Max EUL	Combined		Parameter Estimate	-	-	-	**2,218.1**
				Standard Error	-	-	-	6,623.47
		Failures	Exponential	Parameter Estimate	17.55	1.00	-0.0015	**5516.6**
				Standard Error	26.70	0.00	0.0062	16718.5
		Replacements	Log-Normal	Parameter Estimate	10.88	3.03	-0.0002	**2218.1**
				Standard Error	7.60	1.40	0.0017	6774.8

Similarly, Exhibit 3-56 provides the resulting survival functions for the S160 CAC measure. This measure had no observed failures during the study period, so the competing risks models are based upon removals and replacements only. It is interesting to note the impact that a very small number of events occurring at the end of the study period has on the resulting EUL. It is no surprise that the exponential distribution predicts the maximum EUL due to its constant hazard rate property. Although the best fitting model comes close to predicting the ex ante EUL, the result is not statistically significant.

Exhibit 3-56
Resulting Survival Functions from the Competing Risks Model
S160 CAC Measure



Finally, Exhibit 3-57 presents the detailed results for the S160 CAC competing risks model. As shown in the exhibit, the resulting EUL prediction is not statistically significant due to the large standard error. This point is further illustrated by the large variation among EUL predictions.

Exhibit 3-57
Results from Competing Risks Models
S160 CAC Measure

Measure	Method	Model			Variable: Intercept	Variable: Scale	Variable: Ophours	Resulting EUL
S160	Best Fit	Combined		Parameter Estimate	-	-	-	**12.4**
				Standard Error	-	-	-	22.40
		Removals	Logistic	Parameter Estimate	5.08	0.21	-	**13.50**
				Standard Error	1.76	0.28	-	23.80
		Replacements	Log-Normal	Parameter Estimate	5.74	0.75	-	**25.83**
				Standard Error	1.96	0.71	-	50.57
	Min EUL	Combined		Parameter Estimate	-	-	-	**2.7**
				Standard Error	-	-	-	0.89
		Removals	Gamma	Parameter Estimate	4.31	0.02	-	**5.89**
				Standard Error	0.27	0.01	-	1.41
		Replacements	Gamma	Estimate	4.68	0.06	-	**8.21**
				Standard Error	0.59	0.04	-	4.40
	Max EUL	Combined		Parameter Estimate	-	-	-	**467.0**
				Standard Error	-	-	-	239.78
		Removals	Exponential	Parameter Estimate	10.10	1.00	-	**1401.20**
				Standard Error	1.62	0.00	-	2273.22
		Replacements	Exponential	Estimate	9.40	1.00	-	**700.58**
				Standard Error	1.15	0.00	-	803.68

Section 4 provides the recommended results by studied measure, and summarizes all of the results developed in this section.

4. RESULTS

This section presents the final results of the 1996 and 1997 CEEI Retention Study. As discussed in detail in Section 3, the overall approach consists of five analysis steps that were used to estimate each of the studied measures' EULs:

1. ***Compile summary statistics*** on the raw retention data.
2. ***Visually inspect*** the retention data.
3. ***Develop a trend line*** from the survival plots.
4. ***Develop a survival function*** using classical survival techniques.
5. ***Develop competing risks models*** that incorporate different distributions for failures, removals, and replacements.

4.1 COMPILE SUMMARY STATISTICS

For all studied measures, it was clear that such a small percentage of failures and removals had occurred, that it would be nearly impossible to model the equipment's survival function.

Exhibit 4-1 presents the percentage of measures that were found to have failed or been removed over the study period. From this percentage, an EUL was estimated, assuming a constant failure rate over the life of the measure.

Exhibit 4-1
Summary Statistics on Raw Retention Data

End Use	Technology	Measure	Percent Failed, Removed, Replaced	Annualized Failure, Removal, Replacement Rate^	Mean Life*	Median Life*	Ex Ante EUL
Lighting	T8 Lamps and Electronic Ballasts	L23	2.56%	0.85%	117	81	16
	Optical Reflectors w/ Fluor. Delamp	L19	0.76%	0.25%	396	275	16
	High Intensity Discharge	L81	1.88%	0.63%	159	110	16
HVAC	CAC	S160	0.48%	0.16%	627	434	15

^ Assuming a percentage of failed, removed, replaced occurs over three years.
* Assuming a constant failure rate over time.

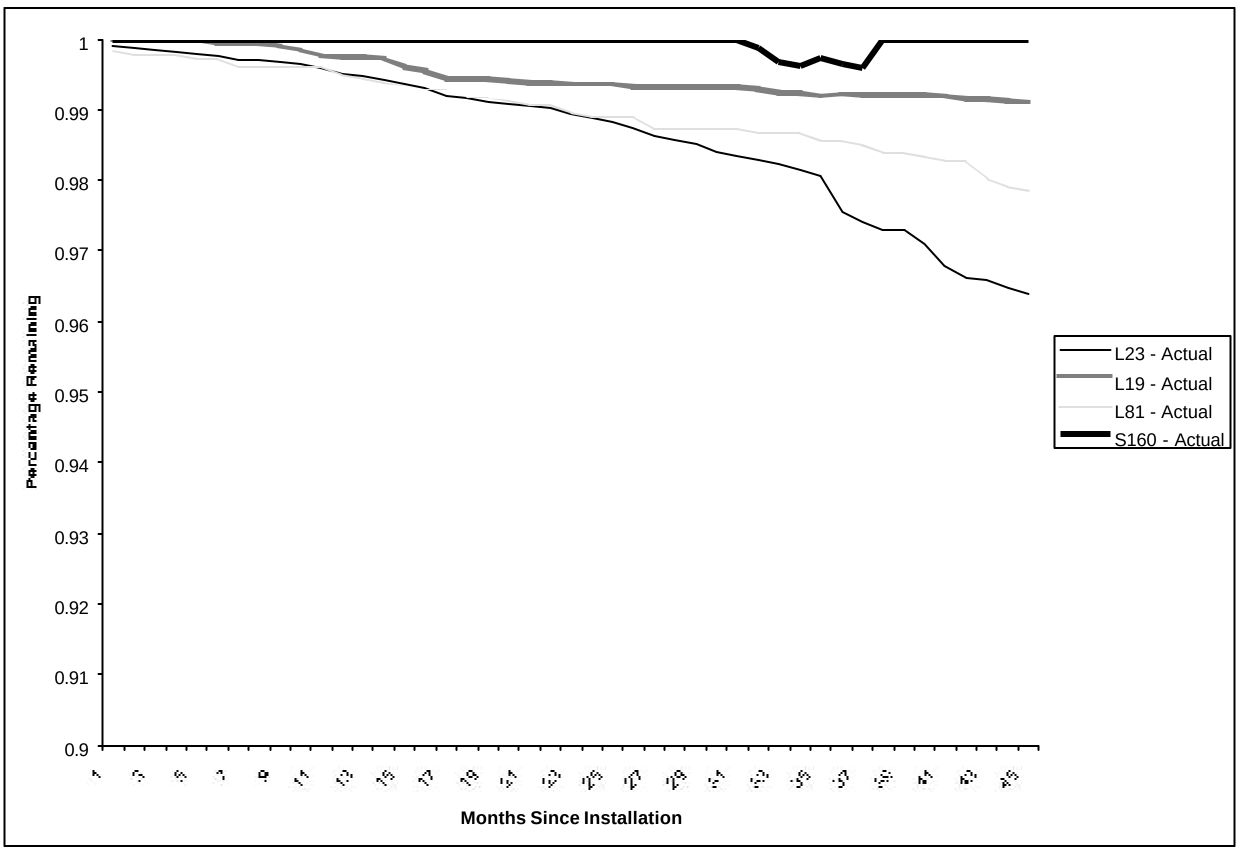
Exhibit 4-1 clearly demonstrates that for the S160 CAC measure, it will be difficult to develop a survival function or an ex post EUL estimate, since only a few events occurred during the study period. With such limited data on failures, a reliable survival function cannot be developed nor can an ex post EUL estimate.

4.2 VISUAL INSPECTION

Using the raw retention data, we developed empirical distributions of the survival function for each of the studied measures. This step clearly illustrated that for each studied measure, there was not enough data over time to support an accurate estimate of the survival function. For this study, the vast majority of measures were in place less than five years (few were installed prior to 1996, and follow-up data collection was conducted no later than the end of 2000). Because the ex ante EUL is 15-16 years for the measures, our data were not capable of accurately estimating the survival function of failures and removals.

Exhibit 4-2 provides the empirical survival function for the four studied measures.

Exhibit 4-2
Empirical Survival Functions
L23 T8, L19 Delamping, L81 HID 251-400W and S160 CAC Measures



4.3 DEVELOP A TREND LINE

Using the empirical functions developed above, a trend line was estimated using standard linear regression techniques. We modeled the trend as a linear and an exponential function (by taking the log of the percentage operable). In each case, we plotted the resulting trend line and visually compared it to the empirical survival function developed above.

The results of the trendline regressions are provided in Exhibit 4-3 for each of the four measures. Also provided in Exhibit 4-3 is the estimated EUL for each measure. Clearly, the results of the linear and exponential trendline estimate indicate that the ex post EUL estimates are significantly larger than the ex ante estimates (which are 15 years for the S160 CAC measure and 16 years for all studied lighting measures). Each of these results would easily reject the ex ante estimate at the 80 percent confidence level.

Exhibit 4-3
Regression Results of Linear and Exponential Trendlines and Resulting Ex Post EUL Estimates

Measure	Measure Description	Intercept	t-Statistic	Slope	t-Statistic	EUL
	Linear Distribution					
L23	FIXTURE: MODIFICATION / REPLACE LAMPS & BLST, 4 FT FIXTURE	1.00	1,103	-0.0008	-22.53	54
L19	REFLECTORS WITH DELAMPING, 4 FT LAMP REMOVED	1.00	3,554	-0.0002	-20.43	191
L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	1.00	4,141	-0.0004	-45.71	100
S160	A / C: CENTRAL, < 65 KBTU / HR, AIR-COOLED, SPLIT-SYS / SNGL PKG	1.00	3,135	0.0000	-2.38	1,450
	Exponential Distribution					
L23	FIXTURE: MODIFICATION / REPLACE LAMPS & BLST, 4 FT FIXTURE	-	-	0.0006	30.00	90
L19	REFLECTORS WITH DELAMPING, 4 FT LAMP REMOVED	-	-	0.0002	42.68	261
L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	-	-	0.0004	93.04	133
S160	A / C: CENTRAL, < 65 KBTU / HR, AIR-COOLED, SPLIT-SYS / SNGL PKG	-	-	0.0000	3.45	2,827

4.4 DEVELOP A SURVIVAL FUNCTION

Using classical survival techniques, we modeled the survival function assuming five of the most common survival distributions: exponential, logistic, lognormal, Weibull and gamma. In each case, we plotted the resulting distribution and visually compared it to the survival plot developed above. Furthermore, we used the resulting survival function to estimate the EUL.

Exhibit 4-4 provides the results of the classical survival analysis. Shown are the model results for each measure, and for each type of distribution modeled. Furthermore, the resulting EUL estimates are provided.

Exhibit 4-4
Comparison of Survival Model Results
Exponential, Logistic, Lognormal, Weibull and Gamma Models
L23 T8, L19 Delamping, L81 HID, and S160 CAC Measures

			Variable			Resulting
Measure	**Model**		**Intercept**	**Scale**	**Ophours**	**EUL**
L23	Exponential	Parameter Estimate	8.01	1.00	-0.00018	**90.0**
		Standard Error	1.69	0.00	0.00042	22.49
	Logistic	Parameter Estimate	5.92	0.53	-0.00007	**24.1**
		Standard Error	1.18	0.17	0.00023	13.92
	Log-Normal	Parameter Estimate	6.81	1.37	-0.00011	**50.5**
		Standard Error	1.35	0.40	0.00025	38.95
	Weibull	Parameter Estimate	7.08	0.79	-0.00013	**46.2**
		Standard Error	1.59	0.19	0.00034	28.28
	Gamma	Estimate	6.92	0.34	-0.00013	**36.8**
		Standard Error	5.24	15.58	0.00034	1,598.94
L19	Exponential	Parameter Estimate	17.32	1.00	-0.00204	**364.7**
		Standard Error	9.19	0.00	0.00213	252.80
	Logistic	Parameter Estimate	17.94	1.04	-0.00214	**649.6**
		Standard Error	13.18	0.70	0.00260	2,262.21
	Log-Normal	Parameter Estimate	20.31	2.95	-0.00221	**5,031.5**
		Standard Error	14.50	1.86	0.00277	23,344.88
	Weibull	Parameter Estimate	18.79	1.11	-0.00226	**609.4**
		Standard Error	13.73	0.71	0.00272	2,021.88
	Gamma	Estimate	17.29	0.37	-0.00204	**928.6**
		Standard Error	12.03	0.22	0.00244	3,290.07
L81	Exponential	Parameter Estimate	8.27	1.00	-0.00011	**147.2**
		Standard Error	5.17	0.00	0.00126	69.86
	Logistic	Parameter Estimate	8.36	1.02	-0.00012	**225.3**
		Standard Error	5.88	0.52	0.00130	477.50
	Log-Normal	Parameter Estimate	10.88	2.96	-0.00022	**1,818.3**
		Standard Error	7.41	1.34	0.00162	5,298.42
	Weibull	Parameter Estimate	8.74	1.10	-0.00013	**210.8**
		Standard Error	6.20	0.52	0.00139	408.00
	Gamma	Estimate	8.54	0.85	-0.00013	**171.0**
		Standard Error	11.50	0.00	0.00138	2,312.94
S160	Exponential	Parameter Estimate	9.00	1.00	-	**467.0**
		Standard Error	0.94	0.00	-	437.44
	Logistic	Parameter Estimate	4.99	0.24	-	**12.3**
		Standard Error	1.00	0.19	-	12.21
	Log-Normal	Parameter Estimate	5.44	0.67	-	**19.2**
		Standard Error	1.34	0.51	-	25.74
	Weibull	Parameter Estimate	4.99	0.24	-	**11.3**
		Standard Error	1.00	0.19	-	10.49
	Gamma	Estimate	4.55	0.04	-	**7.2**
		Standard Error	0.39	0.02	-	2.52

To repeat discussions presented in Section 3 for the S160 CAC measure, the timing of the events had a very strong influence on the model predictions for all but the exponential model. A separate analysis that randomly distributed the 12 tons that either failed or were removed over the study period (rather than at the end of the study period, as indicated by empirical data) produced drastically different results. Therefore, accepting LIFEREG model results for the ex post EUL is strongly discouraged.

4.5 DEVELOP COMPETING RISKS MODELS

Competing risks models were developed to account for different events having different underlying distributions. Models were developed for all measures except the L19 Delamping measure, where there was only one event type observed during the study period. Results from the best-fitting competing risks models are provided in Exhibit 4-5.

Exhibit 4-5
Competing Risks Model Results
L23 T8, L81 HID, and S160 CAC Measures

Measure	Method	EUL	Standard Error
L23	Best Fit	**33.8**	87.33
	Min EUL	**7.8**	59.39
	Max EUL	**91.1**	193.72
L81	Best Fit	**28.5**	346.11
	Min EUL	**26.3**	315.69
	Max EUL	**2,218.1**	6,623.47
S160	Best Fit	**12.4**	22.40
	Min EUL	**2.7**	0.89
	Max EUL	**467.0**	239.78

Although many of the EUL values seem reasonable under the best fit scenario, all have standard errors that indicate that the EUL is not statistically significantly different from both zero and the ex ante EUL. In addition, the S160 CAC measure has a minimum EUL value that appears to be statistically significant. It should be noted that this scenario models the absolute minimum model results, no matter how well the model fits the empirical data. Also, the minimum EUL for the S160 measure is not borne out by empirical data, since at the time of this study more than 2.7 years has elapsed, and the median was not observed as only 0.48% of the total tons surveyed are inoperable.

4.6 FINAL RESULTS

Exhibit 4-6 summarizes the estimated EULs for each studied measure for each approach and corresponding model. The median EULs are provided, along with the upper and lower confidence bounds, based on the 80 percent confidence interval.

Exhibit 4-6
Comparison of Survival Model Results
Summary Statistics, Trendlines, LIFEREG, and Competing Risks Models
L23 T8, L19 Delamping, L81 HID, and S160 CAC Measures[1]

			Measures			
Approach	**Model**		**L23**	**L19**	**L81**	**S160**
Summary	Exponential	Median EUL	81	275	110	434
Statistics		Upper Bound	-	-	-	-
		Lower Bound	-	-	-	-
Trendlines	Linear	Median EUL	54	191	100	1,450
		Upper Bound	57	203	102	2,231
		Lower Bound	51	179	97	669
	Exponential	Median EUL	90	261	133	2,827
		Upper Bound	94	268	134	3,878
		Lower Bound	86	253	131	1,777
LIFEREG	Exponential	Median EUL	90	365	147	467
		Upper Bound	119	689	237	1,028
		Lower Bound	61	41	58	-94
	Logistic	Median EUL	24	650	225	12
		Upper Bound	42	3,550	837	28
		Lower Bound	6	-2,251	-387	-3
	Log-Normal	Median EUL	51	5,031	1,818	19
		Upper Bound	100	34,960	8,611	52
		Lower Bound	1	-24,897	-4,974	-14
	Weibull	Median EUL	46	609	211	11
		Upper Bound	82	3,201	734	25
		Lower Bound	10	-1,983	-312	-2
	Gamma	Median EUL	37	929	171	7
		Upper Bound	2,087	5,146	3,136	10
		Lower Bound	-2013	-3289	-2794	4
Competing Risks	Best Fit	Median EUL	34	-	29	12
		Upper Bound	121	-	375	35
		Lower Bound	-54	-	-318	-10
	Min EUL	Median EUL	8	-	26	3
		Upper Bound	67	-	342	4
		Lower Bound	-52	-	-289	2
	Max EUL	Median EUL	91	-	2218	467
		Upper Bound	285	-	8842	707
		Lower Bound	-103	-	-4405	227

[1] Although negative EUL values are a physical impossibility, the values are presented so that the reader may understand the magnitude of the standard error.

Before recommending a methodology to estimate the ex post EUL, it is first important to consider the definition of a confidence interval. Most people mistakenly interpret an 80 percent confidence interval, for example, to mean that there is an 80 percent probability that the true median EUL is contained within the interval provided. This is ***not*** true. The correct interpretation of an 80 percent confidence interval is that if a given experiment is repeated a large enough number of times (say 30 or more), the median obtained from the same model will be contained in the confidence interval 80 percent of the time.

Take for example the exponential distribution modeled for the L23 T8 measure, using the LIFEREG procedure. If we were to repeat our experiment and create a retention panel of 131 sites with 10,450 units originally installed (as was done for this study), there would be an 80 percent probability that the resulting median EUL using the exponential LIFEREG model would result in a value between 61 and 119 years.

Therefore, the results presented above should not be interpreted as data intervals that have an 80 percent probability of containing the true median EUL. One common use of confidence intervals is to identify models that provide results that are not statistically significantly different than zero. As we can see above, many of our model results are not statistically significantly different than zero when measured at the 80 percent confidence level. In fact, the only model from the LIFEREG procedure that produces a statistically significant result for all measures is the gamma distribution.

We point this all out, because based on our extensive analysis of the retention data, we believe that there is insufficient data to provide reliable model results. There may be sufficient sample sizes to produce statistically significant results, but there clearly is not enough data over time to reliably estimate the median EUL. This can be illustrated by the sensitivity in the model results.

Take, for example, the five model results based on the LIFEREG procedure for the L23 T8 measure. The median EUL based on the exponential distribution was 90 years, versus only 24 years using the logistic distribution. If we had a sufficient amount of data over time, such that the retention data actually covered the true median, we would expect the median result for the two models to be extremely close! Recall that only about 45 months of valid data was collected for this measure, and that the ex ante EUL is 192 months. After 45 months, the logistic distribution actually estimated fewer failure/removals than the exponential distribution, as shown below in Exhibit 4-7.

Exhibit 4-7
Comparison of Survival Functions
Exponential and Logistic versus Empirical Function
L23 T8 Measure

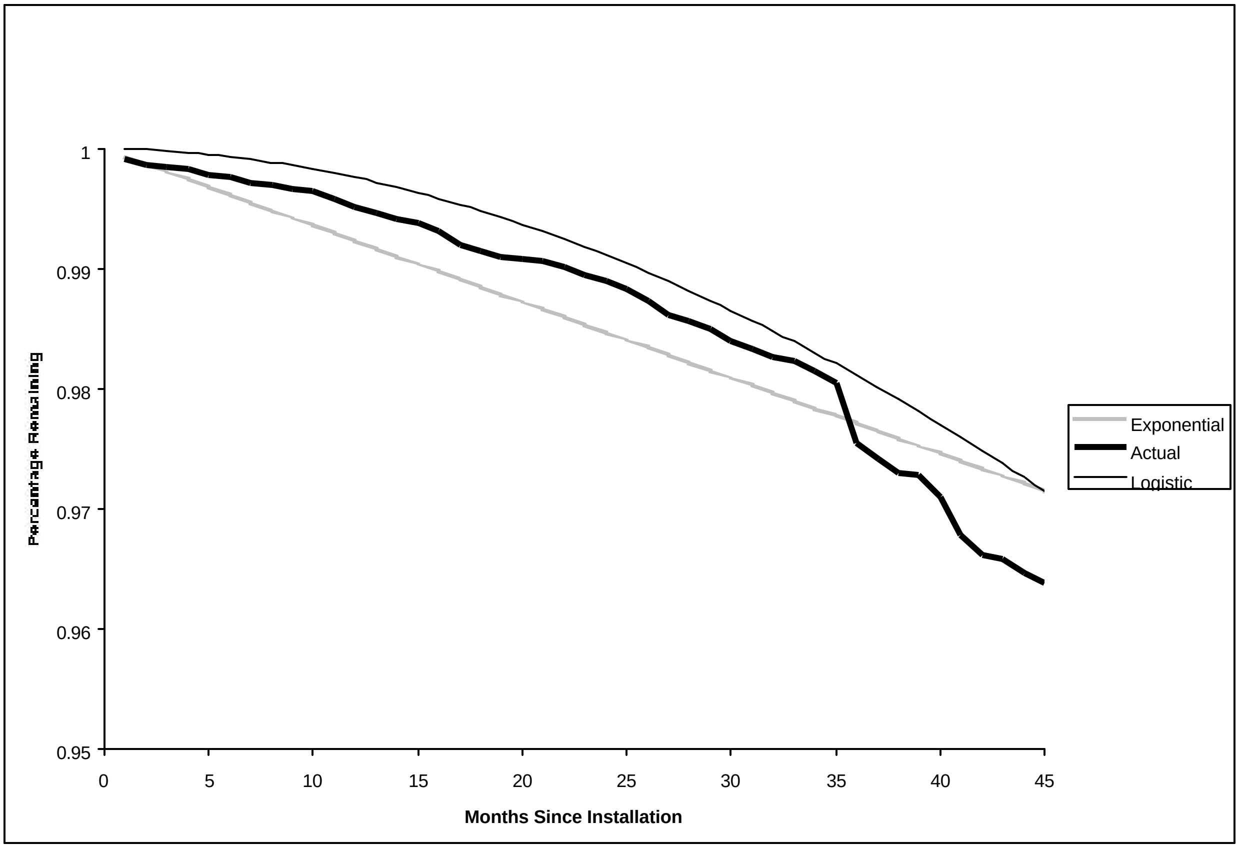
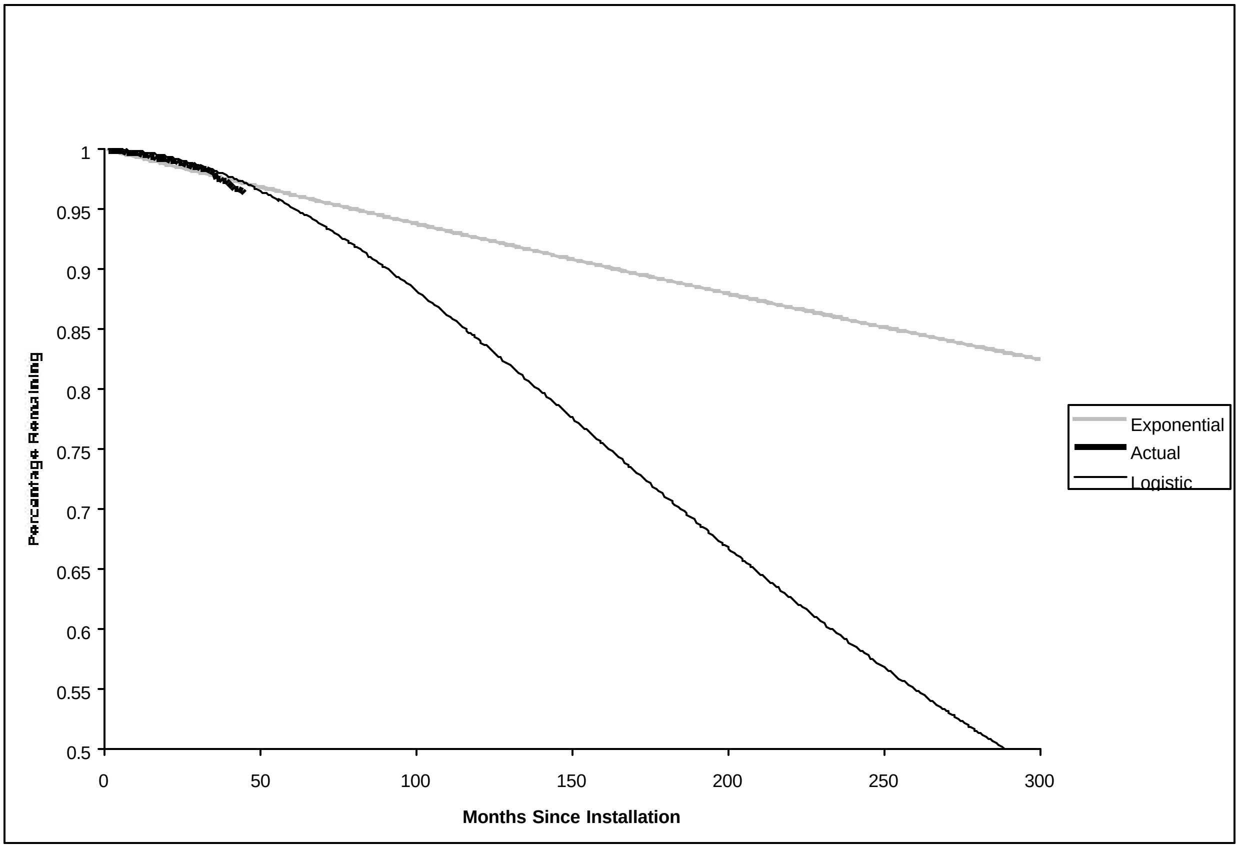


Exhibit 4-6 further illustrates how close the two models estimate the empirical survival function, and how close the two models are to each other. Beyond the 45 months, however, there is little data for the model to structure the remaining survival function. Consider what happens over the next 300 months, up to the 25th year. As shown in Exhibit 4-8, in year 24, the logistic model has reached its median point; whereas the exponential distribution still predicts that 83 percent of the measures are in place and operable. Which model result is better?

Clearly at this point in the measure's life it is not possible to state with much certainty, which model result is superior to the other. Yet, the Protocols require an ex post estimate of the EUL during the fourth year study, and if the ex post estimate is statistically significantly different than the ex ante estimate at the 80 percent confidence level, then we accept the ex post estimate. Under this guideline, one could select the exponential results, which are statistically significantly different than the ex ante estimate at the 80 percent confidence level for three measures, and provide EULs that are as much as 22 times larger than ex ante. Conversely, we could select the lognormal or Weibull results, which are not statistically significantly different than the ex ante results for any measure.

Exhibit 4-8
Comparison of Survival Functions Over 25 Years
Exponential and Logistic versus Empirical Function
L23 T8 Measure



Now take into account the results obtained from the competing risks models. The problems realized with the LIFEREG procedure are present in this method as well. When the events are separated and modeled independently, the number of events observed becomes so small that a large standard error results and the EUL is not statistically significant for all of the measures modeled under the best fit scenario. Although there is evidence to support the theory that different events have unique distributions, there simply is not enough data to precisely predict the EUL at this time.

Our recommendation would be to discard all of the model results on the basis that there is insufficient data over the life of the measures. We want to stress that we believe the sample sizes are sufficient. It is only that we have not observed the sample over a long enough period of time. However, because we are required by the Protocols to report a study result, we will select one of the approaches as our recommended result.

All approaches discussed in Section 3 were implemented for all four measures with the exception of the competing risks model for the L19 Delamping measure. The results based on the summary statistics are not recommended, as they based solely on the overall failure/removal rate observed during the study period. In addition, the results based on the trendlines are not recommended, as they are based on a number of assumptions, as discussed earlier. The results from LIFEREG are also not recommended as they do not take into account the different distributions that occur due to failures vs. removals vs. replacements.

Therefore, the recommended results are based on the competing risks models built from classical survival analysis using the LIFEREG procedure. Of the three models constructed for each measure, the best fit model is the model of choice. Because the best fit model is based upon the fit of the distribution to all of the actual data, we believe that in time the competing risks model will produce the most reliable results. The minimum and maximum EUL methods are not recommended because they seek to minimize/maximize the EUL at the expense of goodness of fit. The L19 Delamping measure study results are based upon the LIFEREG procedure because there was only one event type observed during the study period. The lognormal distribution was chosen because it provided the largest log-likelihood estimate.

Exhibit 4-9 presents the recommended ex post estimates of the EUL. Because the competing risks models did not provide results that were statistically significantly different from the ex ante results, measured at the 80 percent confidence interval, all of the ex post EULs are based on the ex ante estimates. The ex post estimates are compared to the favored study results and the corresponding upper and lower 80 percent confidence interval, when available. Finally, the program realization rates are provided, which are the ratios of the ex ante and ex post estimates. For all measures, the realization rate is one.

Exhibit 4-9

Final Ex Post EUL Estimates[2]

End Use	Technology	Measure	Ex Ante	Study Results Upper	Study Results Median	Study Results Lower	Ex Post	Realization Rate
Lighting	T8 Lamps and Electronic Ballasts	L23	16	121	34	-54	16	100%
	Optical Reflectors w/ Fluor. Delamp	L19	16	28,376	5,031	-18,313	16	100%
	High Intensity Discharge	L81	16	375	29	-318	16	100%
HVAC	CAC	S160	15	35	12	-10	15	100%

2 Although negative EUL values are a physical impossibility, the values are presented so that the reader may understand the magnitude of the standard error.

Attachments

Attachment 1

Survival Data Collection Instrument

Appointment Date: ___________ Appointment Time(s) ________________

Contact Name:
Contact Number:

Company Name:

Multi ID:

QC Site ID

STRATA:

☐ Phone ☐ OnSite

Actual Name:
Actual Number:

Site Address:
Site City, Zip:

YEAR:

Count As Complete?
☐ Yes ☐ No

Measure Code: **Measure Description:** **Make 1** **Make 2**
Model 1 **Model 2**

Retention Quantity: **Quantity Units:** **Technology/ Location Description**

Qty in Operation **Given Units** **Alternative Units** **Printed Response re: Equipment Verification**

Qty Failed, Removed or Replaced **Given Units** **Alternative Units**

UTD = Unable to determine
NA=not applicable

Qty Failed **Est. Date Failed**

UTD = Unable to determine
NA=not applicable

(mm/dd/yy)

Description of Failure (check one)

☐ 1 = Manufacturing defect
☐ 2 = Improper Installation
☐ 3 = Improper Maintenance
☐ 4 = Accident/human error
☐ 5 = Other (Print Reason)
☐ 99 = Unable to determine

Failure +/or Other comments

Qty Removed **Est. Date Removed**

UTD = Unable to determine
NA=not applicable

(mm/dd/yy)

Reason for Removal (check one)

☐ 1 = Unsatisfactory Performance
☐ 2 = Savings not worth the effort
☐ 3 = Remodeling disabled the installation
☐ 4 = Type of business changed
☐ 5 = Moved
☐ 6 = Equipment Upgraded
☐ 7 = Other (Print Reason)
☐ 99 = Unable to determine

Removal +/or Other comments

Qty Replaced **Est. Date Replaced**

UTD = Unable to determine
NA=not applicable

(mm/dd/yy)

Reason for Replacement (check one)

☐ 1 = Unsatisfactory Performance
☐ 2 = Savings not worth the effort
☐ 3 = Remodeling disabled the installation
☐ 4 = Type of business changed
☐ 5 = Moved
☐ 6 = Equipment Upgraded
☐ 7 = Other (Print Reason)
☐ 99 = Unable to determine

Number Warrantied, Replacement +/or Other comments

Replaced w/ Equivalent Technology? (check one)

☐ 1: Higher Efficiency ☐ 2: Equivalent Efficiency ☐ 3: Base line Efficiency ☐ 4: Other or UTD

Record Replacement Tech

Appointment Date: ___________ Appointment Time(s) ________________

Company Name:

Site Address:

Site City, Zip:

Contact Name:

Contact Number

Actual Name:

Actual Number:

Multi ID: **QC Site ID:**

STRATA:

YEAR:

☐ Phone ☐ OnSite

Count As Complete?

☐ Yes ☐ No

Measure Code: **Measure Description** **Tech Type** **Watts/ Lamp** **Lamps/ Fixture**

Retention Quantity: **Quantity Units** **Location Description**

Qty in Operation **Given Units** **Alternative Units**

Printed Response re: Equipment Verification

Qty BALLASTS Failed, Removed or Replaced

UTD = Unable to determine
NA=not applicable

#Lamps/ Ballast **Total # of Ballasts**

 # BALLASTS Failed

UTD = Unable to determine
NA=not applicable

Est. Date Failed

(mm/dd/yy)

Description of Failure (check one)

☐ 1 = Manufacturing defect
☐ 2 = Improper Installation
☐ 3 = Improper Maintenance
☐ 4 = Accident/human error
☐ 5 = Other (Print Reason)
☐ 99 = Unable to determine

Failure +/or Other comments

 # BALLASTS Removed

UTD = Unable to determine
NA=not applicable

Est. Date Removed

(mm/dd/yy)

Reason for Removal (check one)

☐ 1 = Unsatisfactory Performance
☐ 2 = Savings not worth the effort
☐ 3 = Remodeling disabled the installation
☐ 4 = Type of business changed
☐ 5 = Moved
☐ 6 = Equipment Upgraded
☐ 7 = Other (Print Reason)
☐ 99 = Unable to determine

Removal +/or Other comments

 # BALLASTS Replaced

UTD = Unable to determine
NA=not applicable

Est. Date Replaced

(mm/dd/yy)

Reason for Replacement (check one)

☐ 1 = Unsatisfactory Performance
☐ 2 = Savings not worth the effort
☐ 3 = Remodeling disabled the installation
☐ 4 = Manufacturing Defect
☐ 5 = Moved
☐ 6 = Equipment Upgraded
☐ 7 = Other (Print Reason)
☐ 99 = Unable to determine

Replacement +/or Other comments

 # BALLASTS Warrantied

Record Replacement Tech

Replaced w/ Equivalent Technology? (Check one)

☐ 1: Higher Efficiency ☐ 2: Equivalent Efficiency ☐ 3: Base line Efficiency ☐ 4: Other or UTD

Attachment 2

Sample Design Memos Submitted to
the CADMAC Subcommittee on Persistence

RLW Statistical Methodology

Task 3: Sample Design

Background

Our preliminary calculation of the required sample size was based on the hypothesis-testing approach described in the RFP, following the proposed changes to the Protocols. The null hypothesis is that the ex-anti estimates of measure life still reflect the current population. For this purpose, the ex-anti estimate of measure life will be calculated as a weighted average of the individual measure lives, using the net resource benefit as the weights applied to each category of measure.

The ex-anti estimates will be changed only if there is a significant difference between the ex-post and ex-anti estimates of measure life at the 80% level of confidence. Unless agreed otherwise, a two-sided test will be used. We have assumed that the sample size should be chosen so that the hypothesis test should have 80% probability of rejecting the null hypothesis under the assumption that the true value is 20% less than the ex-anti estimate.

We found that the preceding criterion requires a sample of 28 sites. We chose to apply this criterion to each of the two program years. Our sample size planning was carried out in the following five steps:

1. Establish the procedure for estimating the survival proportion S of the measures in a set of buildings of a particular average age t. Specifically, consider a particular program year such as PY94 and assume an exponential survival function as specified in the RFP.

2. Establish the procedure for estimating the effective useful life *EUL* for a particular set of buildings, given an estimate of the survival proportion S.

3. Find the relationship between the sampling distributions for estimating survival and for estimating effective useful life. In particular, how is the coefficient of variation (*cv)* of the estimator of *EUL* related to the coefficient of variation of the estimator of S?

4. Find the required value of the coefficient of variation of the estimator of *EUL* to satisfy the hypothesis-testing framework of the proposed protocols.

5. Find the relationship between the required sample size n and the coefficient of variation of the estimator of *EUL.* Solve for the sample size n.

The results of steps 1 and 2 are discussed under Task 9 – Analyze Data.

For each of the two program years, we will define the survival proportion S to be the current energy use of the corresponding population of program participants as a proportion of the gross first year savings found in the program evaluation. We will use standard MBSS™ ratio estimation techniques to estimate S from the information from the telephone and onsite surveys and the corresponding engineering models. This estimator may be denoted $\hat{S}$. The MBSS procedure will give the value of $\hat{S}$ and the corresponding standard error. We will also determine the savings-weighted average age of the buildings, denoted *t*.

The next step in our analysis was to obtain an estimate of EUL from $\hat{S}$. Following the exponential failure model and the definition of EUL from the RFP, we will use the estimator

$$E\hat{U}L = \frac{t \ln(.5)}{\ln(\hat{S})}$$

The third step was to find the relationship between the sampling distributions for estimating survival and for estimating effective useful life. Using a standard tailor's series expansion of the preceding equation, we found that the coefficient of variation of $E\hat{U}L = \frac{t \ln(.5)}{\ln(\hat{S})}$ is approximately equal to the coefficient of variation of $\hat{S}$ itself.

The fourth step was to find the coefficient of variation (*cv)* of the estimator of *EUL* to satisfy the hypothesis-testing framework of the proposed protocols. Using the Central Limit Theorem, we assumed that $E\hat{U}L$ is normally distributed with unknown expected value m and standard deviation s. We specified the null hypothesis $H_0 : m = m_0$ based on the ex anti estimate of measure life. The decision rule was to reject the null hypothesis if $|z| > z_0 = 1.28$ where z is the usual test statistic. Assuming that $m = m_1 = 0.8m_0$, we want the probability of rejecting the null hypothesis to be 0.8. From the normal distribution we defined $z_1 = 0.84$ and determined the design equation satisfying the preceding requirement:

$$.8m_0 + z_1 s = m_0 - z_0 s$$

This can be rewritten as

$$cv = \frac{s}{m_0} = \frac{.2}{z_0 + z_1} = .0943$$

This implies that the study will satisfy the protocols if the coefficient of variation of the estimator of the EUL is equal to .0943.

The final task was to determine the relationship between the required sample size *n* and the desired coefficient of variation and then to solve for the sample size *n*. For this purpose we assumed that each site satisfies a binary failure model. We assumed that the current savings of each site was either the measured first-year saving, with probability *p* = 0.8, or zero otherwise. Under MBSS analysis, it can be shown that if each site is selected with probability proportional to savings, then the coefficient of variation of the estimated survival is approximately

$$cv = \sqrt{\frac{1-p}{np}} = \sqrt{\frac{1}{4n}}$$

From step 2, this is also the coefficient of variation of the estimated EUL. Solving the preceding two equations, we found n = 28.

Memo

To: Valerie Richardson

From: Richard Ridge

CC: Mike Baker, John Cavalli, Tim Caulfield, Roger Wright,

Date: 02/12/01

Re: Retention Methods

Per our agreement at the kickoff meeting on 7/21/98, I have determined the various approaches that different contractors are planing to use to estimate effective useful lives (EULs). I have focused on the differences among PG&E contractors and across contractors for PG&E's and SCE's commercial new construction retention studies. In addition, I have estimated the number of failures that we need to see in the sample of 150 sites in order to achieve the Protocol-required level of precision.

Differences Across Consultants

I have spoken with John Cavalli of Quantum and Tim Caulfield of Equipoise in order to determine how they are approaching the estimation of EULs. John Cavalli also indicated that Lisa Skumatz is using the same techniques as Quantum. There appear to be four basic approaches. The first is what I call classic survival analysis (CSA) which involves the analysis of data that correspond to the time from a well-defined time origin until the occurrence of some particular event or end-point (Collett, 1994). Regression refers to the familiar estimation of an ordinary least squares (OLS) regression that estimates the relationship between time and the percent of savings remaining at a site or the percent of equipment still present and operable (Maddala, 1992). The third approach involves assuming a function form (AFF) such as the logistic or exponential, conducting a survey at a given point in time after the installation, and using the data in conjunction with the adopted functional form to estimate the EUL. A fourth technique, time series, refers to an analysis of a single variable over time. Such methods include Box-Jenkins and exponential smoothing (Goodrich, 1992).

In Table 1, I indicate that a consultant is using one of these four approaches by placing an "X" in a cell.

Table 1. Analysis Techniques by Proposed by PG&E Consultants

	SBW	Quantum	Equipoise	Skumatz	RLW
CSA	X	X		X	
OLS		X	X		
AFF					X
Time Series					

Table 1 provides some useful information. First, one can see that there is a fair amount of consistency across the PG&E consultants with three of the four planning to use the CSA approach. The primary reason why I, and presumably the other two PG&E consultants, have chosen the CSA approach is that it is specifically designed to address problems having to do with persistence and retention. Also, using CSA, one can test various functional forms rather than assuming one. Note that Equipoise, using the OLS approach, is not assuming a functional form either. Second, only Quantum has specifically listed a backup method if the CSA approach does not perform well. However, I, and presumably the other two PG&E consultants, are prepared to try other approaches if our primary approach fails. Third, with respect to SCE's Commercial New Construction Retention Study, RLW will be using the AFF approach. Having reviewed the various methods proposed by me and the other consultants, I want to emphasize that all of the methodological choices are legitimate, are within the spirit of the Protocols, and are based on their expectations regarding the quantity and quality of the data that are available.

As we all know, aside from the reporting requirements in Tables 6 and 7, the Protocols have virtually nothing to say about retention study issues such as sample sizes, the kinds of statistical models that should be used, and the size of the difference between the *ex ante* EUL and the *ex post* EUL that our statistical models should be designed to detect. The Protocols only state that the confidence level should be set at 80 percent. This provides utilities with a fair amount of latitude.

Required Samples Sizes

For the PG&E Commercial New Construction Retention Study, I have attempted to estimate the number of failures required for the CSA approach to achieve the required level of precision. To perform this calculation, one must make a number of other assumptions in addition to the confidence level. For example, how big a difference between the *ex ante* and the *ex post* EULs (the so-called effect size) should the statistical test be able to detect as significant?[1] This is a particularly critical factor since the sample size is to a large extent a function of the effect size. Assuming a large effect size allows one to reduce the sample size accordingly. Because, the Protocols say nothing about effect size, utilities have a fair amount

[1] The effect size, the size of the sample, and the confidence level can be used to determine the *power* of the test (Cohen, 1988). Alternatively, the desired power of the test, the expected effect size, and the confidence level can be used to determine the size of the sample.

of latitude regarding the size of their retention samples. Simply setting the desired level of confidence at 80 percent does not lead one to the desired sample size.

For our purposes, I have assumed a logistic functional form, a power[2] of .8, an alpha of .20 (i.e., 80 percent confidence level), an *ex ante* EUL of 16 years, and an expected difference between the *ex ante* and *ex post* EULs set first at 20 percent and then at 30 percent (i.e., the savings expected to survive until the 16th year were set at 26 percent and 17 percent respectively). At an effect size of 20 percent, the required number of failures is 24 while at an effect size of 30 percent the required number of failures is 17 (recall that the larger the effect the smaller the required number of failures). Note that RLW chose the 20 percent effect size and a power of .8, both of which are reasonable. However, the Protocols do not prohibit assuming a larger effect size or a lower power.

While we plan to survey 150 sites, we are assuming, for this calculation, that we will actually visit no more than 30 sites. Lets also assume that the kWh savings at each site can be divided into ten bundles bringing the total number of bundles to 300 (10 x 30). If we choose an effect size of 20 percent, we must observe failures in at least 24 or 8 percent of the 300 bundles. If we chose an effect size of 30 percent, we must observe failures in at least 17 bundles or 5.7 percent. At this time, both of these numbers (24 and 17) seem like reasonable expectations.

If the number of expected bundle failures is not observed during the on sites, we will adopt one of the alternative methods described earlier.

References

1. D. Collett. *Modeling Survival Data in Medical Research*. New York: Chapman & Hall, 1994.

2. Maddala, G. S. *Introduction to Econometrics*. Englewood Cliffs, NJ: Prentice Hall, 1992.

3. Goodrich, Robert L. *Applied Statistical Forecasting*. Belmont, MA: Business Forecast Systems, 1992.

[2] The power of a statistical test of a null hypothesis is defined as the probability that it will lead to a rejection of the null hypothesis when it is false.

Attachment 3

Protocol Tables 6 and 7

PROTOCOL TABLES 6B AND 7B

FOURTH YEAR RETENTION STUDY FOR THE 1996 & 1997 COMMERCIAL EEI PROGRAM LIGHTING AND HVAC TECHNOLOGIES

PG&E STUDY ID #s 349R1 & 351R1

This Attachment presents Tables 6B and 7B for the above referenced study as required under the "Protocols and Procedures for the Verification of Cost, Benefits, and Shareholder Earnings from Demand Side Management Programs" (the Protocols), as adopted by the California Public Utility Commission (CPUC) Decision 93-05-063, Revised March 1998 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.

The Table 7B synopsis of analytical methods applied follows Protocol Table 6B.

Protocol Table 6.B
Results of Retention Study
PG&E 1996 & 1997 Commercial Energy Efficiency Incentives Program
Study ID #s 349R1 & 351R1

Item 1			Item 2		Item 3	Item 4	Item 5	Item 6		Item 7	Item 8	Item 9
PG&E Measure Code	**Studied Measure Description**	**End Use**	**Ex Ante EUL**	**Source of Ex Ante EUL**	**Ex post EUL from Study**	**Ex Post EUL to be used in Claim**	**Ex Post EUL Standard Error**	**80% Conf. Interval Lower Bound**	**80% Conf. Interval Upper Bound**	**p-Value for Ex Post EUL**	**EUL Realizat'n Rate (ex post/ex ante)**	**"Like" Measures Associated with Studied Measure (by measure code)**
L19	FIXTURE: MODIFICATION/LAMP REMOVAL, 4 FT LAMP REMOVED	Lighting	16	Advice Filing & MDSS	5031	16	23345	-18313	28376	0.999	100%	L17, L18, L20, L76 - L77
L23	FIXTURE: MODIFICATION/REPLACE LAMPS & BLST, 4 FT FIXTURE	Lighting	16	Advice Filing & MDSS	34	16	87	-54	121	0.990	100%	L9 - L12, L21, L22, L24, L69 - L75, L117 - L124, L160, L13, L112
L81	HID FIXTURE: INTERIOR, 251-400 WATTS LAMP	Lighting	16	Advice Filing & MDSS	29	16	346	-318	375	0.999	100%	L25, L78 - L80, L26, L27
S160	A/C: CENTRAL, < 65 KBTU/HR, AIR-COOLED, SPLIT-SYS/SNGL PKG	HVAC	15	Advice Filing & MDSS	12	15	22	-10	35	0.939	100%	S1, S2, S4, S160 - S163

PROTOCOL TABLE 7B

1996 & 1997 COMMERCIAL EEI PROGRAM
FOURTH YEAR RETENTION STUDY
PG&E STUDY ID #349R1 AND #351R1

The purpose of this section is to provide the documentation for data quality and processing as required in Table 7B of the California Public Utility Commission (CPUC) Evaluation and Measurement Protocols (the Protocols). The major topics covered in this section are organized and presented in the same order as they are listed in Table 7B for ease of reference and review. For items discussed in detail elsewhere in the report, only a brief summary will be given in this section to avoid redundancy.

1. OVERVIEW INFORMATION

A. Study Title and Study ID Number

Study Title: Fourth Year Retention Study of PG&E's 1996 & 1997 Commercial EEI Program.

Study ID Numbers: 349R1 and 351R1

B. Program, Program Year and Program Description

Program: PG&E Commercial EEI Program.

Program Year: Rebates Received in the 1996 & 1997 Calendar Year.

Program Description:

The Commercial Energy Efficiency Incentives Program for lighting and HVAC technologies offered by PG&E has three components: the Retrofit Express (RE) Program, the Retrofit Efficiency Options (REO) Program and the Customized Incentives (CI) Program.

The RE Program

The RE program offered fixed rebates to customers who installed specific electric energy-efficient equipment. The program covered the most common energy saving measures and spans lighting, air conditioning, refrigeration, motors, and food service. Customers were required to submit proof of purchase with these applications in order to receive rebates. The program was marketed to small- and medium-sized commercial, industrial, and agricultural customers. The maximum rebate amount, including all measure types, was $300,000 per account. No minimum amount was required to qualify for a rebate.

The REO Program

The REO program targeted commercial, industrial, agricultural, and multi-family market segments most likely to benefit from these selected measures. Customers were required to submit calculations for the projected first-year energy savings along with their application prior to installation of the high efficiency equipment. PG&E representatives worked with customers to identify cost-effective improvements, with special emphasis on operational and maintenance measures at the customers' facilities. Marketing efforts were coordinated amongst PG&E's divisions, emphasizing local planning areas with high marginal electric costs to maximum the program's benefits.

The Customized Incentives Program

The Customized Incentives program offered financial incentives to CIA customers who undertook large or complex projects that save gas or electricity. Customers may also participate under the APOS program. These customers were required to submit calculations for projected first-year energy impacts with their applications prior to installation of the project. The maximum incentive amount for the Customized Incentives program was $500,000 per account, and the minimum qualifying incentive was $2,500 per project. The total incentive payment for kW, kWh, and therm savings was limited to 50 percent of direct project cost for retrofit of existing systems. Since the program also applied to expansion projects, the new systems incentive was limited to 100 percent of the incremental cost to make new processes or added systems energy efficient. Customers were paid 4¢ per kWh and 20¢ per therm for first-year annual energy impacts. A $200 per peak kW incentive for peak demand impacts required that savings be achieved during the hours PG&E experiences high power demand.

Due to the significant documentation and analysis involved in Customized Incentives program measures, however, rebates for a number of 1992 through 1995 measures were delayed for payment until 1996. This evaluation covers those measures where rebates were paid in 1996 and 1997.

As a result of program design, the measures installed were similar to or the same as those for the RE program, but were installed in larger and more complex projects.

C. End Uses and/or Measures Covered

End Use Covered: Indoor Lighting and HVAC Technologies.

Measures Covered: For the list of measures covered in this evaluation, see *Exhibit 2-3*.

D. Methods and Models Used

Our overall approach consists of five analysis steps that were used to estimate each of the studied measures' EULs:

1. ***Compile summary statistics*** on the raw retention data. Upon review of the summary statistics, it became clear that such a small percentage of failures and removals had occurred, that it would be difficult to model the equipment's survival function.

2. ***Visually inspect*** the retention data, by simply calculating the cumulative percentage of equipment that had failed in a given month, and plotting the percentage over time. This step clearly illustrated that for each studied measure, there was not enough data over time to support an accurate estimate of the survival function.

3. ***Develop a trend line*** from the survival plots. Using the plots developed in (2) above, a trend line was estimated using standard linear regression techniques. We modeled the trend as a linear and an exponential function. In each case, we used the resulting trend line to estimate the EUL, which was statistically significantly larger than the ex ante estimate.

4. ***Develop a survival function*** using classical survival techniques. We modeled the survival function assuming five of the most common survival distributions: exponential, logistic, lognormal, Weibull and gamma. In each case, we used the resulting survival function to estimate the EUL. In nearly every case, the resulting EUL was either statistically significantly larger than the ex ante EUL, or was not statistically significantly different than the EUL. In only 1 out of 20 cases was the resulting EUL statistically significantly less than the ex ante EUL. In this case the failure events observed during the study period clearly do not provide adequate information for a reliable estimate.

5. ***Develop competing risks models*** that incorporate different distributions for failures, removals, and replacements. Using the LIFEREG procedure in SAS from step 4 above, separate output was generated for failures, removals, and replacements. Then, the best fitting distributions for each event were combined to form one combined survival function. This additional analysis step provided valuable results that have not been previously utilized in retention studies.

The details surrounding each of these steps is provided in *Section 3.*

E. Analysis Sample Size

Exhibit 3-2 provides the final sample disposition used in the study analysis.

2. DATABASE MANAGEMENT

A. Key Data Elements and Sources

The MDSS, the original retention panels and the follow-up survey data were the only data sources used for this analysis.

B. Data Attrition Process

All data points that had follow-up survey data were utilized in the analysis. As discussed in *Section 3,* the SAS analysis procedures we implemented were able to handle interval censored data, in the cases when failure/removal dates were not obtainable.

C. Internal Data Quality Procedures

The Evaluation contractor of this project, Quantum Consulting Inc. (QC), has performed extensive data quality control on all retention and follow-up survey data. QC's data quality procedures are consistent with PG&E's internal database guidelines and the guidelines established in the Protocols.

Throughout every step of this project, numerous data quality assurance procedures were in place to ensure that all data used in analysis and all survey data collected was of the highest quality. All data entry was performed using blind double-key data entry. On questionable responses follow-up phone calls or site visits were made.

D. Unused Data Elements

Without exception, all data collected specifically for the Evaluation were utilized in the analysis.

3. SAMPLING

A. Sampling Procedures and Protocols

Section 3.1 describes the sample procedures and protocols.

B. Survey Information

The data collection instrument is presented in the *Attachment 1. Exhibit 3-2* provides the final sample disposition, which contains the number of sites and units that were in the sample frame, and the number surveyed.

C. Statistical Descriptions

Statistics variables that were used in the survival models are also presented in *Section 3.*

4. DATA SCREENING AND ANALYSIS

A. Procedures for Treating Outliers and Missing Data

All data points that had follow-up survey data were utilized in the analysis. As discussed in *Section 3,* the SAS analysis procedures we implemented were able to handle interval censored data, in the cases when failure/removal dates were not obtainable.

B. Background Variables

Due to the nature of this analysis (survival analysis), background variables, such as interest rates, unemployment rates and other economic factors, were not considered to be a necessary component of the analysis.

C. Data Screen Process

Again, all data points that had follow-up survey data were utilized in the analysis.

D. Regression Statistics

The regression statistics for the models implemented are provided in *Section 3*.

E. Model Specification

The model specifications are presented in *Section 3*.

F. Measurement Errors

For the survival analysis, the main source of measurement errors is the survey data. Our approach has been to proactively stop the problem before it happens so that statistical corrections are kept to a minimum.

Measurement errors are a combination of random and non-random error components that plague all survey data. The non-random error frequently takes the form of systematic bias, which includes, but is not limited to, ill-formed or misleading questions and mis-coded study variables. In this project, we implemented several controls to reduce systematic bias in the data. These steps include: (1) thorough auditor/coder training; (2) instrument pretest; and (3) cross-validation between on-site audit data and telephone survey responses.

The random measurement error, such as data entry error, has no impact on estimating mean values because the errors are typically unbiased. For the measures that were modeled in the survival analysis, the impact of random unbiased measurement errors was accounted for as part of the overall standard variance in the parameter estimate.

G. Influential Data Points

No diagnostics were used to identify outliers.

H. Missing Data

As discussed in *Section 3*, the SAS analysis procedures we implemented were able to handle interval censored data, in the cases when failure/removal dates were missing. There were no other missing data points, other than failure/removal dates.

I. Precision

The SAS output provided the standard errors for the 50th percentile (or median). Because the analysis was conducted on the unit of measure (e.g., a ballast) and not a site, the standard errors from SAS were grossly underestimated. SAS treats each observation in the dataset as independent. However, it is likely that there is significant correlation in the observations that are common to a single site (especially in the event that a removal occurs.) For example, when a removal occurs, it is likely that many measures are removed at once. To a lesser extent, failures are correlated since they may all come from the same manufacturing lot, they are all likely to be installed under the same circumstances, and they are also used in a similar manner.

If we believed that there was 100 percent correlation of failure/removal for all measures with a site, we could simply multiply the standard error calculated from SAS by the square root of the ratio of the number of units to sites. Therefore, if there were an average of 100 units installed per measure, we would multiply by 10.

We felt, however, that there were two components to our error: one caused by variation across sites, and another caused by variation across measures. The errors calculated by SAS correspond only to the error across measures.

To estimate the standard error associated with failures and removals, we first took the SAS output and backed out a standard deviation. This was achieved by multiplying the standard error from SAS by the square root of the sample size (in units.) We then assumed that this standard deviation was associated with the joint probability density function of failures and removals.

(1) $StdErr_{SAS} * \sqrt{N_{Units}} = StdDev_{Failures,\mathrm{Re}movals}$

Where,

$StdErr_{SAS}$ is the standard error around the median EUL projected with the SAS System;

$\sqrt{N_{Units}}$ is the square root of the number of sites that contributed to the regression model;

$StdDev_{Failures,\mathrm{Re}movals}$ is the standard deviation associated with the median EUL of failures and removals.

We then assumed that failures were independent of removals (Which is of course not true, since a high failure rate may cause a customer to decide to make removal. But we felt this was reasonable overall.) Therefore, the variance of removals and failures is equal to the variance of removals plus the variance of failures:

(2) $$\begin{aligned} StdDev_{Failures,\mathrm{Re}movals}{}^2 &= Var_{Failures,\mathrm{Re}movals} \\ &= Var_{Failures} + Var_{\mathrm{Re}movals} \end{aligned}$$

Where,

$StdDev_{Failures,\mathrm{Re}movals}{}^2$ is the square of the standard deviation associated with the median EUL of failures and removals;

$Var_{Failures,\mathrm{Re}movals}$ is the variance which is equivalent to the square of the standard deviation.

If we assume that failures are independent across units, and removals are independent across sites, then the standard error can be calculated as:

$$(3)\quad \begin{aligned} StdErr_{Failures,\mathrm{Re}\,movals} &= \sqrt{StdErr_{Failures}{}^{2} + StdErr_{\mathrm{Re}\,movals}{}^{2}} \\ &= \sqrt{\frac{Var_{Failures}}{N_{Units}} + \frac{Var_{\mathrm{Re}\,movals}}{N_{Sites}}} \end{aligned}$$

Where,

$StdErr_{Failures,\mathrm{Re}\,movals}$ is the standard deviation associated with the median EUL of failures and removals;

N_{Units} is the number of units used for the regression models;

N_{Sites} is the total number of sites having those units.

Furthermore, if we assume that the underlying standard deviation of failures and removals are equivalent, then:

$$(4)\quad \begin{aligned} StdDev_{Failures,\mathrm{Re}\,movals}{}^{2} &= Var_{Failures,\mathrm{Re}\,movals} \\ &= Var_{Failures} + Var_{\mathrm{Re}\,movals} \\ &= 2Var_{Failures,or\,\mathrm{Re}\,movals} \end{aligned}$$

So,

$$(5)\quad \begin{aligned} Var_{Failures,or\,\mathrm{Re}\,movals} &= 0.5 * (StdDev_{Failures,\mathrm{Re}\,movals})^{2} \\ &= 0.5 * (StdErr_{SAS})^{2} * N_{Units} \end{aligned}$$

Therefore, substituting equation (5) in equation (3), we get

$$(6)\quad \begin{aligned} StdErr_{Failures,\mathrm{Re}\,movals} &= \sqrt{\frac{0.5 * (StdErr_{SAS})^{2} * N_{Units}}{N_{Units}} + \frac{0.5 * (StdErr_{SAS})^{2} * N_{Units}}{N_{Sites}}} \\ &= StdErr_{SAS} * \sqrt{0.5 + 0.5 * \frac{N_{Units}}{N_{Sites}}} \end{aligned}$$

It is interesting to note that if there was only one unit per site, the standard error would equal the standard error calculated in SAS. Our resulting standard error is somewhere between the standard error found in SAS, and the standard error from SAS multiplied by the square root of the ratio of the number of units to sites (the method discussed as the beginning of this section.)

Skinner and Kish[1] both offer a more theoretical approach to solving the problem of estimating a standard error when the data are not identical and independently distributed (IID). They define this problem as a design effect, which is the case when the sample is not a simple random sample that is IID, but rather is a cluster sample such as ours. In our case, each site contains a cluster of sample points.

Skinner developed a design effect factor, Deff, that can be used to adjust the standard error obtained from SAS to estimate the true standard error:

$$(7)\quad Deff = \frac{StdErr_{TRUE}{}^{2}}{StdErr_{SAS}{}^{2}}$$

Where,

$StdErr_{TRUE}$ is the actual standard error associated with the median EUL;

$StdErr_{SAS}$ is the standard error associated with the median EUL obtained from SAS;

Skinner estimated the design effect factor as:

$$(8)\quad Deff = 1 + (n - 1) * t$$

Where,

n = the average number of sample points per cluster (or, in our case, per site)

$$= \frac{N_{Units}}{N_{Sites}}$$

t = the intra-cluster correlation

[1] Skinner, C. J., "Analysis of Complex Surveys," John Wiley & Sons, 1989, pp. 23-46.
Kish, L., "Survey Sampling," John Wiley & Sons, 1965, pp. 162.

Skinner's design effect factor can be compare directly to the factor we developed in equation (6):

$$(9)\quad Deff(Eq.6) = 0.5 + 0.5 * (\frac{N_{Units}}{N_{Sites}}) = 1 + (n-1) * 0.5$$

Our method discussed above is identical to that developed by Skinner, with an intra-cluster correlation equal to 0.5. As discussed above, we believe that there are two types of events: removals and failures. Our assumption above was that removals are perfectly correlated and failures are totally uncorrelated. Therefore, an intra-cluster correlation of 0.5 is not unreasonable.

To calculate the intra-cluster correlation, it would require knowing the time of failure or removal for all units in our analysis. The intra-cluster correlation measures how correlated the failure/removal times are across all units within a site. Because our analysis is being conducted in such an early stage of the measures life, it is not possible to accurately estimate the correlation. However, given that (1) it is likely that removals are highly correlated, and failures are relatively uncorrelated; and (2) removals are expected to be as prevalent as failures over the life of the measure; then an intra-cluster correlation of 0.5 is a reasonable approximation.

Finally, relative precision estimated at the 80 percent confidence interval was calculated using the following equation:

$$RP = \frac{1.282 * StdErr}{EUL}$$

Where,

$StdErr$ = the standard error calculated using Equation 6, above.