Ninth Year Retention Study of the 1995 Southern California Gas Company Commercial New Construction Program

Study ID Number 718A

FINAL REPORT

February 24, 2005

Robert Mowris & Associates

P.O. Box 2141, Olympic Valley, California 96164 ■ (800) 786-4130 ■ FAX (530) 581-4970 ■ email: rmowris@earthlink.net

Table of Contents

Acknowledgments

Robert Mowris, P.E., Principal Consultant, and Kathleen Carlson, Esq., Principal Consultant, were responsible for project management, development of the retention survey instrument, and analyses of the results. Ean Jones, B.S., and Anne Blankenship, B.S., managed the onsite retention survey inspections performed by John Parker. Peter Pressley, Ph.D., developed the retention survey database, and Rachel Noack, managed and revised the database. Robert Mowris assisted with the database design and development. Robert Mowris and Rachel Noack performed the hazard function and survival and statistical analyses, developed the figures and tables, and wrote the report. Anne Blankenship, B.S., assisted with project management and development. Rob Rubin was the study manager for Southern California Gas Company.

1.0 Executive Summary

This report provides the results of a retention survey conducted by Robert Mowris & Associates (RMA) of measures installed under the Southern California Gas Company (SoCalGas) 1995 Commercial New Construction Program. The program primarily addressed cooking equipment in commercial kitchens. Other measures included in the retention survey include storage hot water (SHW), packaged gas/electric HVAC systems, and boilers. This retention study surveyed the following measures.

1) Oven
2) Fryer
3) Range
4) Griddle
5) Broiler
6) Packaged Gas/Electric HVAC Systems
7) Steamer
8) Hot Food Table
9) Kettle
10) Braising Pan
11) Other Cooking
12) Storage Hot Water
13) Boiler

RMA personnel collected data for this study between November 2004 and January 2005. A summary of measure retention rates for the surveyed sites is presented in **Table 1-1**. The average retention rate was 47%, including business failures and removals (column 2). However excluding business failures and removals (i.e., percentage of measures not failed), the average retention rate was 86% (column 5). Thus, the average retention rate based on measure performance was reasonably high after 10 years.

In addition to gathering site specific measure data, another key objective of the study was to estimate effective useful life (EUL) for each measure type and to determine if the estimated EULs are consistent with SoCalGas *ex ante* EULs. Direct estimation of hazard rates and survival functions from the collected data could not be made since retention rates for many of the measures were relatively high and observed removals of failures were relatively low. However, survival functions based on estimated hazard rates were estimated for four measures: ovens, fryers, ranges, and griddles. Hazard rates for the remaining measures could not be estimated due to the small number of failures or limited sample sizes.

Estimates of EULs for each measure are provided in **Table 1-2**. **Table 1-2** also includes the SoCalGas *ex ante* estimates of EULs. While sufficient data was available to estimate a median EUL for ovens, fryers, ranges, and griddles, the "null hypothesis" cannot be rejected for these measures.[1] As shown **Table 1-3**, the *ex ante* EUL is within the 80 percent effective useful life lower and upper bounds controlling for the important background variable of business failures. The "null hypothesis" regarding EULs cannot be rejected for the other measures examined in this study.

[1] The Fourth Year Retention Study found median EULs for ovens of 6.7 years and fryers of 6.4 years. These median estimates were based on five years of data.

Table 1-1. Retention Rates for SoCalGas 1995 Commercial New Construction Measures

Measure	Measures Removed or Failed within 10 Years	Measures Retained After 10 Years	Measures Failed within 10 Years	Measures not Failed After 10 Years
Oven	65%	35%	11%	89%
Fryer	64%	36%	22%	78%
Range	52%	48%	9%	91%
Griddle	35%	65%	14%	86%
Broiler	62%	38%	12%	88%
HVAC	11%	89%	11%	89%
Steamer	82%	18%	21%	79%
Hot Food Table	55%	46%	23%	77%
Kettle	60%	40%	20%	80%
Braising Pan	75%	25%	0%	100%
Other Cooking*	35%	65%	6%	94%
SHW	56%	44%	33%	67%
Boiler	0%	100%	0%	100%
Average (weighted)	**53%**	**47%**	**14%**	**86%**

Note: *Other Cooking includes pasta/rice cooker, rethermalizer, hot food plate, and broaster.

Table 1-2. Retention Study Estimate of Measure Effective Useful Life Compared to the SoCalGas Ex Ante Estimate for Effective Useful Life

Measure	Retention Study Median Effective Useful Life (years)	SoCalGas Ex Ante Estimate of Effective Useful Life (years)	Ex Ante Different From Ex Post at 80% Confidence Level ± 20 Percent?
Oven	10.1	12	No
Fryer	12.6	12	No
Range	13.3	12	No
Griddle	15.7	12	No
Broiler	*	12	No
HVAC	*	18	No
Steamer	*	12	No
Hot Food Table	*	12	No
Kettle	*	12	No
Braising Pan	*	12	No
Other Cooking	*	12	No
SHW	*	15	No
Boiler	*	15	No

Note: *Retention survey data regarding removals/failures is insufficient to estimate mean Effective Useful Life.

Table 1-3. Estimated Effective Useful Life (EUL) for Ovens, Fryers, Ranges, Griddles Excluding Failed Businesses

Measure	SoCalGas Ex Ante Estimated EUL	Estimated Median Life			Ex Ante Different From Ex Post at 80% Confidence Level ± 20 Percent?
		80% Lower bound	Estimated EUL	80% Upper Bound	
Ovens	12	5.6	10.1	23.7	No
Fryers	12	5.7	12.6	48.2	No
Ranges	12	5.3	13.3	74.2	No
Griddles	12	7.9	15.7	44.0	No

2.0 Introduction and Background

In the early 1990s Southern California Gas Company (SoCalGas) established the Commercial New Construction Program to encourage installation of high efficiency natural gas equipment in their service area. Financial incentives and technical assistance were provided to commercial establishments for high efficiency boilers, other high efficiency cooking equipment, and high-efficiency double-effect absorption chillers among other measures. The program was terminated after program activities were completed in 1995. Two years later, in 1997, an extensive evaluation of the program was undertaken to assess the program's impacts on natural gas consumption. The study, titled *An Evaluation of Southern California Gas Company's 1995 Commercial New Construction Program* (Evaluation Study), was completed in accordance with requirements established by the California DSM Advisory Committee (CADMAC).[2]

In addition to completing the initial Evaluation Study, CADMAC requires all utilities to complete a Measure Retention Study to be used in ongoing assessments of program costs and benefits. This requirement is defined in *Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-side Management Programs* adopted by the California Public Utilities Commission (CPUC).[3] This measure retention study meets the CADMAC Measure Retention Study requirements for the SoCalGas 1995 Commercial New Construction Program.

2.1 Project Objectives

The objectives of this measure retention study are as follows.

- Determine the appropriate target sample size for each measure based on the minimum number of removals or failures required to meet the 80 percent confidence level plus or minus 20 percent precision (as specified in the CADMAC Protocols).
- Locate energy conservation measures installed by participants in the SoCalGas 1995 Commercial New Construction Program
- Determine the fraction of measures that are installed and operational.
- Determine the rates of removal or failure, including survival functions.
- Determine reasons for early replacement and/or removal.
- Identify changes in business or service of measures.
- Estimate the measure effective useful life for each measure (if possible).
- Test the "null hypothesis" of "no difference between ex ante and *ex post* estimates of useful life." If the "null hypothesis" is rejected with 80 percent confidence (plus or minus 20 percent), then the retention study results can be used to establish new measure

---

[2] See, *An Evaluation of Southern California Gas Company's 1995 Commercial New Construction Program, Volume I*, prepared by Planergy, Inc., Equipoise Consulting, and Pacific Consulting Services, for Southern California Gas Company, Los Angeles, California, January 1998.

[3] See, *Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-side Management Programs,* as adopted by the California Public Utilities Commission Decision 93-05-063, Revised March 1998.

effective useful lives (EULs). Under the Protocols, effective useful life of a measure is defined as the median number of years that the measure is still in place and operable. In effect, the median age is the number of years that pass until 50% of the measures are no longer in place and operable.

Under the DSM Measurement Protocols adopted by the California Public Utilities Commission, measures are to be studied in retention studies that either make up 50 percent of the savings for their respective sectors or that account for the top 10 measures in a sector. Specific requirements for the Measure Retention Study are defined in Tables 8b and 9b of the *Protocols and Procedures* cited above. Based on these requirements, this retention study includes survey results and analysis of the following measures.

1) Oven
2) Fryer
3) Range
4) Griddle
5) Broiler
6) Packaged Gas/Electric HVAC Systems
7) Steamer
8) Hot Food Table
9) Kettle
10) Braising Pan
11) Other Cooking
12) Storage Hot Water
13) Boiler

The data for this study was gathered from onsite inspections of a random sample of commercial facilities that participated in the program. For the fourth year retention study, RMA personnel collected data from 150 onsite inspections in December 1999 and January 2000. For the ninth year retention study, another 163 inspections were completed between November 2004 and January 2005. A database with detailed inspection results of these companies is included with this report.

2.1 Organization of Report

This retention study is organized into the following sections.

- Chapter 1 provides the executive summary.
- Chapter 2 provides the project objectives.
- Chapter 3 discusses the retention survey database, sample design, survey instrument, data collection, and quality control.
- Chapter 4 discusses the results of the analyses.
- Appendix A contains a copy of the on-site data collection form.
- Appendix B contains the CADMAC Protocol Tables 6 and 7.
- Appendix C contains the Weibull Analysis of measure effective useful life.

3.0 Sample Design and Data Collection Methods

This chapter provides an overview of the survey sample design and data collection methodology used to collect onsite data for the measure retention study. **Section 3.1** discusses the retention survey database. **Section 3.2** discusses the sample design. **Section 3.3** discusses the retention survey instrument. **Section 3.4** discusses the data collection procedures. **Section 3.5** discusses quality control procedures.

3.1 Retention Survey Database

The Retention Survey Database (RSD) developed for this project is based upon the SoCalGas 1995 Commercial New Construction Program. This database contains a comprehensive record of all program participant information. The following data fields are contained in the RSD (partial list):

- Company identification number;
- Company name;
- Company contact name;
- Company address;
- Company telephone number;
- Measure number;
- 1995 New Construction Program identification number;
- Measure name;
- Measure description;
- Measure manufacturer model number;
- Measure rated input (MBtu/hr);
- Measure customer cost;
- Measure installation date;
- Measure energy savings;
- Retention Survey Codes (11 codes defined in Section 3.3, below);
- Retention Survey Date (indicating the removal or failure date); and
- Retention Survey Notes.

The information contained in the RSD was used to develop the sample of participants to survey as well as the retention survey forms. The RSD was also used to record the onsite survey results and to provide a quality control check on all retention survey data. Upon completion of the onsite inspections the RSD was redesigned to facilitate statistical analysis and improved reporting of relevant retention survey data used for this report. A copy of the database is included with this report.

3.2 Sample Design

The survey sample design was developed using survival analysis techniques. Survival analysis pertains to the analysis of data that correspond to a well-defined time origin until the occurrence of a particular event or end-point. For this study, the time origin event is defined by the installation of a measure under the SoCalGas 1995 New Construction Program. The failure or removal of the measure defines the end-point. Measure survival data were expected to have several features that warranted special treatment in preparing the sample design.

- The measure survival data would probably not be symmetrically distributed and cannot be reasonably represented by a normal distribution.
- The survival data would be *right-censored* in that the removal or failure end-points will not be observable for some of the installed measures.

A sample design for addressing these and other features of the data was developed using the following steps.

1) Determine the required number of removals or failures for each measure to test the “null hypothesis” with 80 percent confidence plus or minus 20 percent precision.
2) Determine the probability of removal or failure for each measure.
3) Determine the target sample size for each measure based on the number of removals or failures and the probability of failure for each measure.
4) Randomly select sample sites for each measure among facilities in the retention study database.

The first step in preparing the sample design was to arrive at quantitative estimates of the target sample size for each measure. The target sample size for each measure is based on the following exponential distribution assumed to represent the measure survival function.[4]

Eq. 1. $S(t) = \exp(-\lambda t)$

Where, λ = 1/mean effective useful life. The standard error is given by $1/(\lambda r)$, where *r* is the number of measure occurrences within a sample that have been removed or failed. With an exponential survival function, the standard error for the estimated mean from a sample depends on the number of observed removals/failures.

The goal in developing the sample design was to obtain results with an 80 percent confidence level plus or minus 20 percent. The following equation was used to calculate the sample size to meet the 80 percent confidence level.

Eq. 2. $r = \left[\frac{z}{0.2}\right]^2$

[4] *The Statistical Analysis of Failure Time Data*, Kalbfleisch, J.D., and Prentice, R.L., John Wiley and Sons, New York, NY 1980.

Where, z is defined as the critical value of the standard normal distribution at the desired level of confidence. For the 80 percent confidence level, $z = 1.28$. Thus, the number of removals/failures required to estimate mean measure life for a particular measure at the specified confidence level is $r = 41$. Wherever possible this value was used to establish the preliminary sample design size in order to meet the 80 percent confidence level.

As noted above, there would likely be right-censoring of the occurrences of a measure in the sample; not all of the occurrences would be observed until their life end-point. Accordingly, the number of measure occurrences brought into the sample had to be greater to accommodate this right censoring phenomenon. The sample size needed to provide the required number of removals was determined using the following equation.

Eq. 3. $$\text{Sample Size} = \frac{\text{Number of Required Removals/Failures}}{\text{Probability of Required Removals/Failures}}$$

As shown by Collett,[5] the probability of removal/failure with an assumed survivor function can be calculated with the following information.

1) Specified values for the survivor function;
2) Study accrual time (i.e., the period when measure occurrences take place); and
3) Study follow-up time (i.e., the period when occurrences are tracked to see whether they are removed or failed).

For this study, the accrual period was 12 months (assumed time activities for the SoCalGas 1995 New Construction Program), and the follow-up period is 108 months (the nine years 1996 through 2004 prior to onsite data collection). Ex ante mean values of measure life (for calculating the parameters of the assumed exponential survivor functions for the various types of measures) were taken from *DSM Measure Life Project: Master Tables of Measure Life Estimates and Final Report*[6], prepared for CADMAC by Energy Management Services.

Given the fixed length of the study, the probability of removal/failure was determined primarily by the expected mean life of a measure. Measures with shorter mean lifetimes have a higher probability of removal or failure. For example, the probability of removal or failure is 0.279 for measures with a mean life of 12 years, 0.231 for measures with a mean life of 15 years, and 0.197 for measures with a mean life of 18 years. With the required number of removals/failures for either type of measure being 41, the respective target sample sizes are 87 for 12-year life measures, 101 for 15-year life measures, and 116 for 18-year life measures. These target sample sizes were used to identify which measures might have a sufficient number of program installations to obtain the required 41 removals/failures and test the “null hypothesis.”

[5] *Modeling Survival Data in Medical Researc*h, Collett, D., Chapman & Hall, 1994, pp. 260-264.
[6] *DSM Measure Life Project: Master Tables of Measure Life Estimates and Final Report*, prepared by Energy Management Services, prepared for CADMAC, September 1993.

Total measures, target and actual sample sizes, and actual removals or failures for each measure are shown in **Table 3-1**. Of the 13 measures surveyed in the measure retention study, only the top four measures could be sampled sufficiently to test the “null hypothesis” and satisfy the 80 percent confidence/precision. The relatively small numbers of program installations limited sampling of the remaining measures.

Sampling frames for selecting the sample sites for the different measures were created by identifying information on program participant sites and measures from the initial program participant database. A measure-stratified random number was assigned to each site. Sample sites were then selected based on the numerical sorting of the random numbers assigned to each site. Adjustments were made to the required sample size based on retention survey results (i.e., comparing the number of removals or failures encountered during the onsite survey with the number of required removals or failures).

Table 3-1. Total Measures, Target Sample Size, Sample Size, and Removals or Failures

Measure	Total Measures in 1995 Program	Target Sample Size	Actual Retention Survey Sample Size in 2004	Actual Removals or Failures in 2004
Oven	489	147	125	81
Fryer	451	147	126	81
Range	361	147	139	72
Griddle	254	147	99	35
Broiler	207	147	77	48
HVAC	124	124	73	8
Steamer	83	35	28	23
Hot Food Table	65	20	22	12
Kettle	61	20	15	9
Braising Pan	23	10	8	6
Other Cooking	75	10	17	6
SHW	41	20	18	10
Boiler	34	10	6	0
Total	**2,268**	**570**	**753**	**391**

3.3 Survey Instrument

The survey instrument was designed by RMA to meet the CPUC data requirements defined in the *Protocols and Procedures* cited above. The sample survey instrument is provided in Appendix A. The survey instrument includes the following information:

- Company name and description;
- Contact name;
- Address;
- Telephone number;
- SoCalGas account number;
- Retention Survey Number;
- Retention Survey Date;

- Survey notes;
- Equipment category, description, manufacturer, model number, online date; and
- Retention survey codes (11 numbers used to define the retention survey status).

The eleven retention survey codes used in the survey are defined below.

1. Business and measure still in place.
2. New business, measure still in place.
3. Measure failed and replaced.
4. Measure failed, not replaced.
5. Business failed, unable to survey.
6. New business, measure removed.
7. Measure replaced, reason unknown.
8. Measure removed, did not fail, did not replace.
9. Removed, upgraded kitchen.
10. New business, measure replaced.
11. Measure not found.

Items 2 through 10 require a date in the "new date" column, if known. The date was used to keep track of changes to the measure or facility. For removals or failures, the date was used to estimate hazard rates and survival functions. The survey instrument was designed so that all data could be easily entered into the Microsoft Access Retention Survey Database (discussed above).

3.4 Data Collection

Data collection for the measure retention study was performed through onsite visits to customer facilities. A certified energy manager (CEM) implemented all onsite data collection efforts. The CEM was trained on how to use the retention survey instrument and retention survey database before beginning the data collection effort. The training session included a discussion of project objectives, review of the retention survey data collection instrument, and procedures to effectively collect data with minimum disruption to the customer. Before the CEM went to a facility, he reviewed information on the measures installed at that facility.

The retention survey database was used to produce electronic site-specific retention survey instruments. The survey instruments contained all the baseline measure information used by the CEM for the site visits. Observed changes from the baseline data were indicative of measure changes such as removals, replacements, failures, or other problems affecting the life of the measure. Observed changes were identified using the eleven Retention Survey Codes (noted above in Section 3.3) along with site notes that were later entered into the retention survey database. The CEM also obtained a Survey Code Date (if known) if the measure was removed or replaced as well as all relevant information to analyze the life of the measures.

In addition to the actual measure/facility data derived from the inspections several data elements were collected through interviews with the facility staff. For most facilities, a site manager was generally available who was familiar with the operation of the facility and its equipment. These interviews, conducted prior to the inspection, included a brief introduction regarding the purpose of the survey and follow-up questions. Facility staff members were asked questions directed at investigating inconsistencies in previous data as well as forming a basis for visual inspection of measures. Following the interview, the CEM visually inspected and verified measure installation and/or changes.

3.5 Quality Control

Quality control procedures were used to ensure that collected data was accurate and consistent with participant program information and study objectives. Among others, the following procedures were included.

- Reconciling discrepancies between baseline, interview, and visual inspection results prior to leaving a facility.
- Checking the SoCalGas program participant database and retention survey databases to ensure all customers included in the survey were in fact program participants.
- Creating onsite measure retention survey forms for each site using the retention survey database.
- Recording onsite data into the Microsoft Access Retention Survey Database.
- Exporting all data from the database into Microsoft Excel spreadsheets for validation and analysis.
- Double-checking all data entry (by the RMA project manager and the RMA database quality control engineer).

4.0 Analysis and Results

This chapter provides analysis and results of the onsite survey data collection. **Section 4.1** provides an overview of the measure retention survey results. **Section 4.2** provides a discussion of the impact of business failures on retention rates and effective useful life. **Section 4.3** provides the hazard and survival analysis for ovens and fryers.

4.1 Overview of Measure Retention Survey Results

Measure retention rates are presented in **Table 4-1** along with the total number of removed or failed measures, total measures in the 1995 program, and percent removed or failed measures. Survey results for all retention survey codes are provided in **Table 4-2**.

Table 4-1. Measure Retention Survey Results

Measure	Total Measures Failed or Removed	Total Measures in Survey	Measures Removed or Failed within 10 years	Measures Retained after 10 years
Oven	81	125	65%	35%
Fryer	81	126	64%	36%
Range	72	139	52%	48%
Griddle	35	99	35%	65%
Broiler	48	77	62%	38%
HVAC	8	73	11%	89%
Steamer	23	28	82%	18%
Hot Food Table	12	22	55%	45%
Kettle	9	15	60%	40%
Braising Pan	6	8	75%	25%
Other Cooking	6	17	35%	65%
SHW	10	18	56%	44%
Boiler	0	6	0%	100%
Total	391	753		

Table 4-2. Results of Measure Retention Survey by Survey Code

			Retention Survey Code										
			1	2	3	4	5	6	7	8	9	10	11
Measure	Total Measures in 1995 Program	Actual Retention Survey Sample	Business and measure still in place	New business measure still in place	Measure failed and replaced	Measure failed not replaced	Business failed and gone	New business measure removed	Measure replaced, reason unknown	Measure removed did not fail, not replaced	Removed upgraded kitchen	New business measure replaced	Measure not found
Oven	489	125	41	3	4	4	14	41	4	2	1	5	6
Fryer	451	126	41	4	24	1	11	20	2		3	17	3
Range	361	139	58	9	2	2	15	24	2		3	16	8
Griddle	254	99	62	2	9	1	7	3	2	2	3	4	4
Broiler	207	77	25	4	2	2	7	15	1		6	10	5
HVAC	124	73	62	3	3								5
Steamer	83	28	4	1	4		4	10				3	2
Hot Food Table	65	22	9	1	2		1	6					3
Kettle	61	15	6				1	4		1			3
Braising Pan	23	8	2				1	4				1	
Other Cooking	75	17	11				1	2				2	1
SHW	41	18	6	2	6		3					1	
Boilers	34	6	6										
Total	2,268	753	333	29	56	10	65	129	11	5	16	59	40

4.2 Impact of Business Failures

The most important background variable in this retention study is business failures. This background variable is important for two reasons.

1) Restaurants have a 28 percent higher business failure rate than other retail businesses.[7]
2) Restaurants accounted for 98 percent of equipment purchased with incentives through the 1995 Commercial New Construction Program.

This influences the retention rates of most cooking measures as well as the failure and survival analysis for ovens, fryers, ranges, and griddles. The impact of failed businesses on retention rates is illustrated in **Table 4-3**. Some measures such as braising pan, SHW, and boiler were unaffected by excluding failed businesses. HVAC and other cooking measures were also relatively unaffected. However, ovens, fryers, broilers, steamers, hot food table, and kettles would have had at least a 10 percent higher retention rate if "removals/failures" at failed businesses were not included.

According to the CADMAC Protocols, effective useful life (EUL) of a measure is defined as "the median number of years that the measure is still in place and operable." Interpretation of the Protocol definition of EUL raises the following questions regarding measure retention.

1) Should a measure be defined as "removed or failed" if a business fails, but the measure is still in place and operable, or if the measure was sold to another business and at the new business the measure is still in place and operable?
2) Should the CADMAC definition of EUL be revised or expanded to include a methodology for consideration of background variables such as failed businesses?
3) Should background variables (e.g., business failures) be considered in retention rate calculations?

For the purpose of estimating retention rates, this study defines measures at sites with failed businesses as "removals or failures." However, for the purpose of drawing conclusions regarding the *ex post* measure EUL, this study controls for business failures by excluding these sites from the analysis. Further discussion regarding the impact of business failures and the failure and survival analysis of ovens and fryers is provided in **Section 4.3**. Business failures are included in the final retention rate estimates, but are not included in the final *ex post* measure EUL estimates (see **Appendix B**, **Table 6** per the CADMAC Protocols).

[7] The business failure rate for restaurants of 100 failures per 10,000 is 28 percent higher than the business failure rate for total retail trade of 78 failures per 10,000. See *Business Failure Record: A Comparative Statistical Analysis of Geographic and Industry Trends in Business Failures in the United States*, Neil DiBernardo, Editor, The Dun & Bradstreet Corporation, One Diamond Hill Road, Murray Hill, NJ 07974-1218, 1998.

Table 4-3. Impact of Failed Businesses on Retention Rates

Measure	All Sites Percent Measures Retained	Not Including Failed Businesses Percent Measures Retained	Impact on Retention Rate Not Including Failed Businesses (i.e., increase)
Oven	35%	46%	11%
Fryer	36%	51%	15%
Range	48%	56%	8%
Griddle	65%	73%	8%
Broiler	38%	49%	11%
HVAC	89%	96%	7%
Steamer	18%	32%	14%
Hot Food Table	45%	59%	14%
Kettle	40%	67%	27%
Braising Pan	25%	25%	none
Other Cooking	65%	71%	6%
SHW	44%	44%	none
Boiler	100%	100%	none

4.3 Weibull Analysis and Results

Weibull analysis is used to make predictions about the life of products in the population by "fitting" a statistical distribution of life data from a representative sample of units inspected in the field. The Weibull analysis is based on 10 years of removal or failure data rather than 9 years. This is due to the fact that some measures included within the SoCalGas 1995 Commercial New Construction Program were actually installed as early as December 1994. **Appendix C** provides more information about how to perform the Weibull analysis.

For measures evaluated in this study, the Weibull analysis uses a three-step process. First, hazard rates are developed using the available data. Second, hazard functions are developed based on the hazard rates.[8] Third, survival functions are developed using the estimated hazard function. The steps in the parametric procedure for estimating the survival functions are as follows:

1) Prepare data for calculation of hazard rate.
2) Estimate the hazard function.
3) Use hazard function to determine Weibull distribution survival function and measure effective useful life .

Estimating the hazard function is an essential component in this analytical procedure. Two of the distributions commonly used for survival analysis are the exponential distribution and the Weibull distribution. The probability density functions and associated hazard functions, as well as the survival functions for these distributions are shown in **Table 4-4**.

[8] Hazard functions are used to forecast the probability of removal or failure for a measure, given that the measure has survived to the present.

Table 4-4. Hazard and Survival Functions for Exponential and Weibull Distributions

Exponential Distribution	
Probability Density Function	$f(t) = \gamma\exp(-\gamma t)$
Hazard Function	$h(t) = \gamma$
Survival Function	$S(t) = \exp(-\gamma t)$
Weibull Distribution	
Probability Density Function	$f(t) = \alpha\beta t^{\beta-1}\exp(-\alpha t^{\beta})$
Hazard Function	$h(t) = \alpha\beta t^{\beta-1}$
Survival Function	$S(t) = \exp(-\alpha t^{\beta})$

As **Table 4-4** illustrated, the exponential distribution can be used to represent a hazard rate that is constant. The associated survival function is also exponential. However, the exponential distribution does not represent hazards that increase or decrease over time. If the hazard rate increases or decreases with age, the Weibull distribution can be used to represent the hazard function and the survival function. (Note the Weibull distribution, α is termed as the scale parameter, while β is termed as the shape parameter.)

Direct estimation of hazard rates and survival functions from collected data could not be made since retention rates for many of the measures were relatively high and observed removals of failures were relatively low. However, survival functions based on estimated hazard rates were estimated for four measures: ovens, fryers, ranges, and griddles. Hazard rates for the remaining measures could not be estimated due to the small number of failures or limited sample sizes.

The hazard rates are developed using removal or failure information about the lifetime of the measure. The lifetime of the measure is calculated as the difference between the date that the measure was removed or failed and the date the measure was installed. The date that the measure was installed was provided in the SoCalGas program participant baseline database. The date that the measure was removed or failed is identified in the retention survey database as the survey code date (i.e., month/day/year). Survey code dates were obtained from the facilities staff in interviews conducted during the onsite inspections. However, survey code dates were not provided for all measures due to the facilities staff not being aware of the date of removal or failure. Only 54 of 81 survey code dates were provided for the ovens, 58 of 81 survey code dates were provided for fryers, 56 of 72 survey code dates were provided for ranges, and 25 of 35 survey code dates were provided for griddles that were observed as removed or failed. This presented a problem in terms of developing the hazard rates. RMA used the following approach to solve this problem.

1) Calculate removals/failures for subset of measures where the facilities staff provided survey code dates. These calculations were performed for each year where data was available (e.g., year 1 through year 10).
2) Calculate removals/failures for all measures based on data from the subset of measures where the facilities staff provided survey code dates and lifetime information was available.

The removals/failures each year for all measures are calculated using the following equation.

Eq. 4. $$r(t) = r_s(t) \times \left[\frac{R}{R_s} \right]$$

Where, $r_s(t)$ = removals/failures in year, *t*, for subset of measures with survey code dates.
R_s = total removals/failures for subset of measures with survey code dates (R_s = 54).
R = total removals/failures for all measures in survey sample (R = 81).

Hazard rates are calculated using the following equation.

Eq. 5. $$h(t) = \frac{r(t)}{M(t)}$$

Where, $r(t)$ = removals/failures in year, *t*.
$M(t)$= Measures at start of year.

Hazard rates for all ovens (based on data from the subset of ovens with survey code dates) are shown in Table 4-5. An example calculation for ovens is shown in the following equation.

Eq. 4a. Calculated removals/failures for all ovens in the first year = $1 \times \left[\frac{81}{54} \right] = 1.5$.

Assuming 1.5 ovens in the survey sample were removed/failed in the first year implies a hazard rate of 1.2 percent (see **Table 4-5**). With 1.5 ovens removed/failed during the first year, there were 123.5 ovens “at risk” at the start of the second year. Assuming 1.5 ovens in the survey sample were removed/failed in the second year implies a hazard rate of 1.21 percent. Similar calculations provided hazard rate estimates for years 3 through 10. These calculated hazard rates for all ovens are plotted in **Figure 4-1**. Hazard rates for ovens not including failed businesses are provided in **Table 4-6** and **Figure 4-2**. Hazard rates for all fryers are provided in **Table 4-7** and **Figure 4-3**. Hazard rates for fryers not including failed businesses are provided in **Table 4-8** and **Figure 4-4**. Hazard rates for all ranges are provided in **Table 4-9** and **Figure 4-5**. Hazard rates for fryers not including failed businesses are provided in **Table 4-10** and **Figure 4-6**. Hazard rates for all griddles are provided in **Table 4-11** and **Figure 4-7**. Hazard rates for griddles not including failed businesses are provided in **Table 4-1**2 and **Figure 4-8**.

Table 4-5. Hazard Rates for All Ovens Based on Ovens with Survey Code Dates

All Ovens				Ovens with Survey Code Dates			
Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)	Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)
1	125.0	1.5	1.20%	1	98	1	1.02%
2	123.5	1.5	1.21%	2	97	1	1.03%
3	122.0	9.0	7.38%	3	96	6	6.25%
4	113.0	1.5	1.33%	4	90	1	1.11%
5	111.5	9.0	8.07%	5	89	6	6.74%
6	102.5	13.5	13.17%	6	83	9	10.84%
7	89.0	19.5	21.91%	7	74	13	17.57%
8	69.5	10.5	15.11%	8	61	7	11.48%
9	59.0	10.5	17.80%	9	54	7	12.96%
10	48.5	4.5	9.28%	10	47	3	6.38%
Total		81		Total		54	

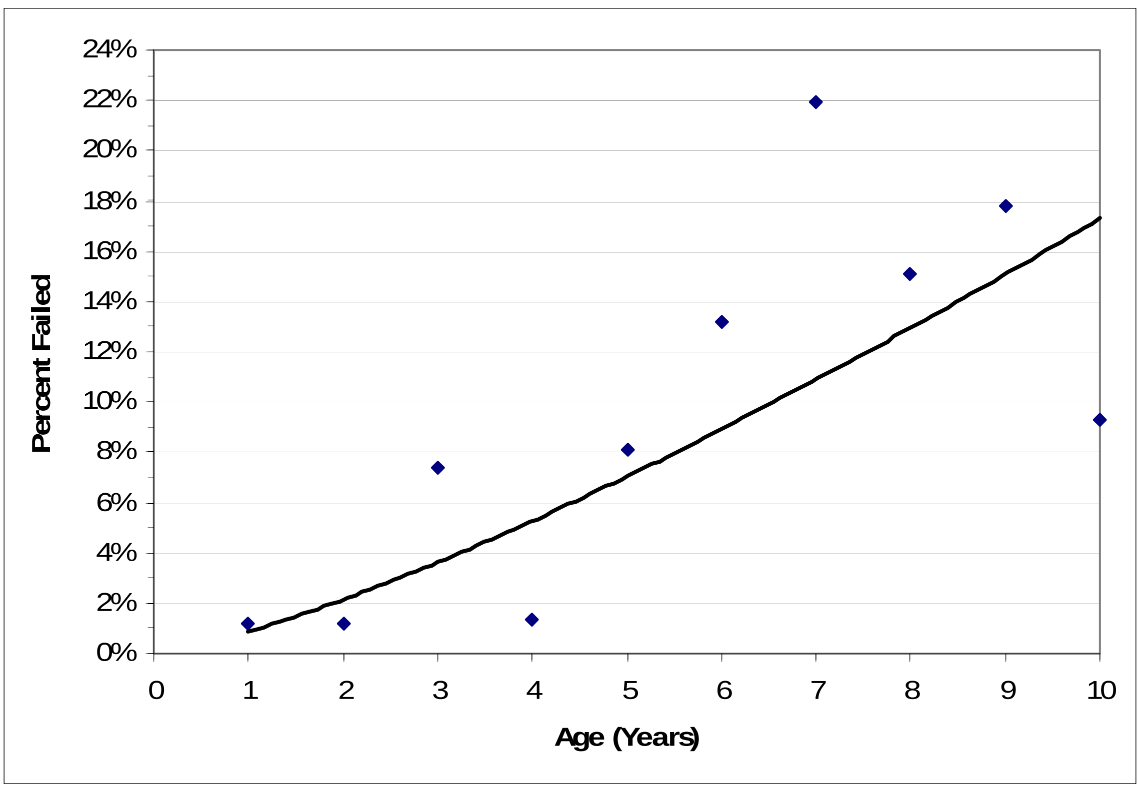


Figure 4-1. Plot of Hazard Rates for All Ovens with Survey Code Dates

Table 4-6. Hazard Rates for Ovens Not Including Failed Businesses.

All Ovens Not Including Failed Businesses				Ovens with Survey Code Dates			
Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)	Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)
1	111.0	1.3	1.14%	1	97	1	1.03%
2	109.7	1.3	1.15%	2	96	1	1.04%
3	108.5	7.6	6.99%	3	95	6	6.32%
4	100.9	1.3	1.25%	4	89	1	1.12%
5	99.6	7.6	7.61%	5	88	6	6.82%
6	92.0	10.1	10.99%	6	82	8	9.76%
7	81.9	16.4	20.06%	7	74	13	17.57%
8	65.5	8.8	13.51%	8	61	7	11.48%
9	56.6	8.8	15.62%	9	54	7	12.96%
10	47.8	3.8	7.94%	10	47	3	6.38%
Total		67		**Total**		53	

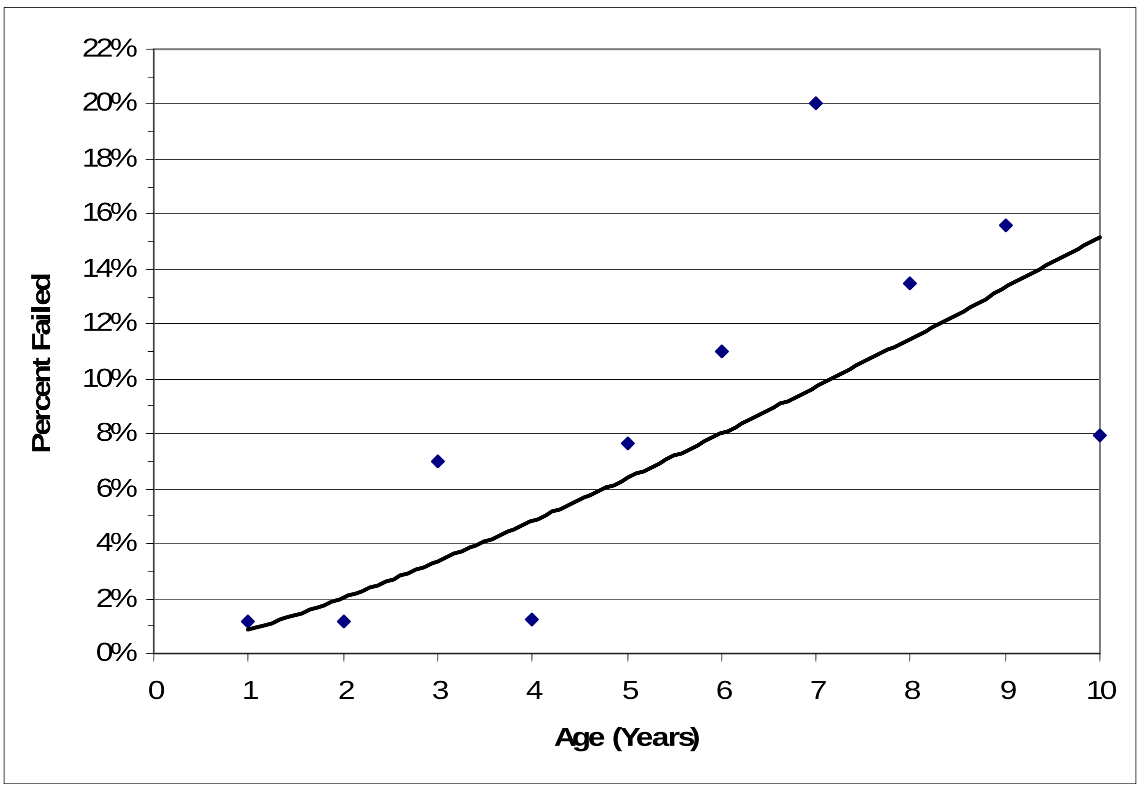


Figure 4-2. Plot of Hazard Rates for Ovens Not Including Failed Businesses

Table 4-7. Hazard Rates for All Fryers Based on Fryers with Survey Code Dates

All Fryers				Fryers With Survey Code Dates			
Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	124.0	1.4	1.13%	1	101	1	0.99%
2	122.6	1.4	1.14%	2	100	1	1.00%
3	121.2	1.4	1.15%	3	99	1	1.01%
4	119.8	7.0	5.83%	4	98	5	5.10%
5	112.8	5.6	4.95%	5	93	4	4.30%
6	107.2	18.2	16.93%	6	89	13	14.61%
7	89.1	4.2	4.70%	7	76	3	3.95%
8	84.9	8.4	9.87%	8	73	6	8.22%
9	76.5	32.1	41.98%	9	67	23	34.33%
10	44.4	1.4	3.15%	10	44	1	2.27%
Total		81		**Total**		58	

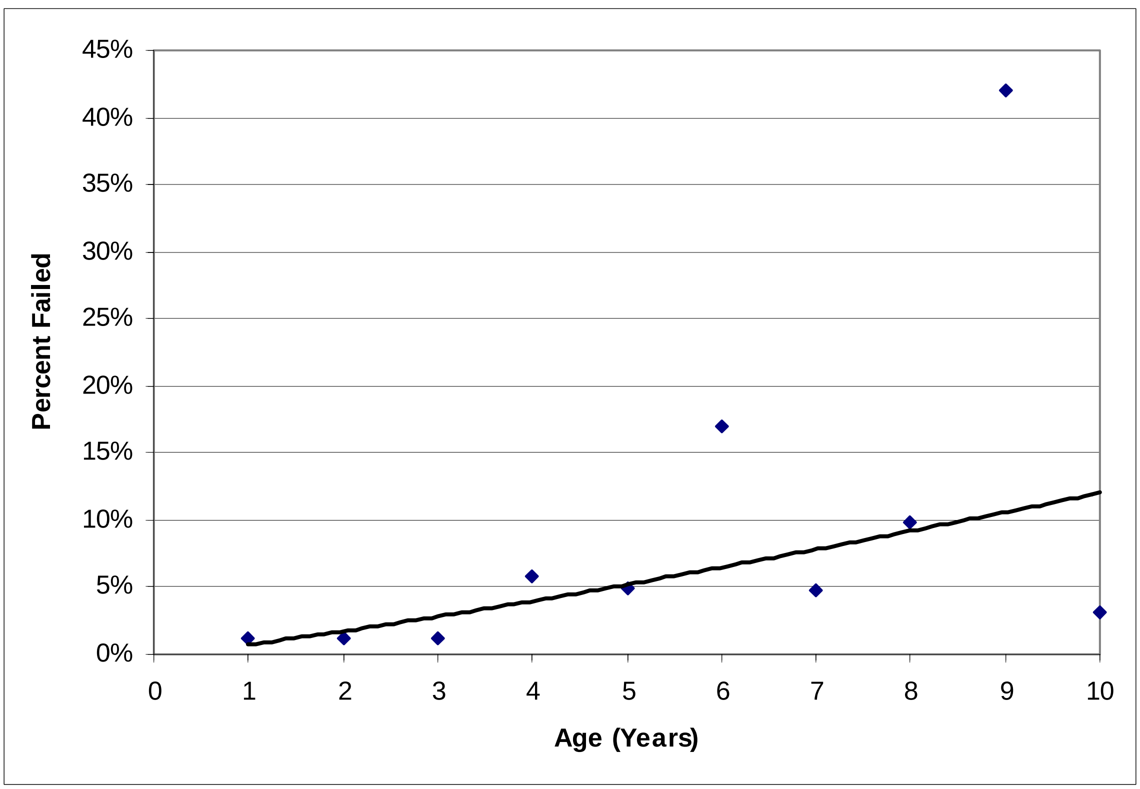


Figure 4-3. Plot of Hazard Rates for All Fryers with Survey Code Dates

Table 4-8. Hazard Rates for All Fryers Not Including Failed Businesses

All Fryers				Fryers With Survey Code Dates			
Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	113.0	1.2	1.08%	1	94	1	1.06%
2	111.8	1.2	1.09%	2	93	1	1.08%
3	110.6	1.2	1.10%	3	92	1	1.09%
4	109.4	6.1	5.56%	4	91	5	5.49%
5	103.3	2.4	2.35%	5	86	2	2.33%
6	100.8	12.2	12.06%	6	84	10	11.90%
7	88.7	3.6	4.11%	7	74	3	4.05%
8	85.0	7.3	8.58%	8	71	6	8.45%
9	77.7	25.5	32.84%	9	65	21	32.31%
10	52.2	1.2	2.33%	10	44	1	2.27%
Total		62		Total		51	

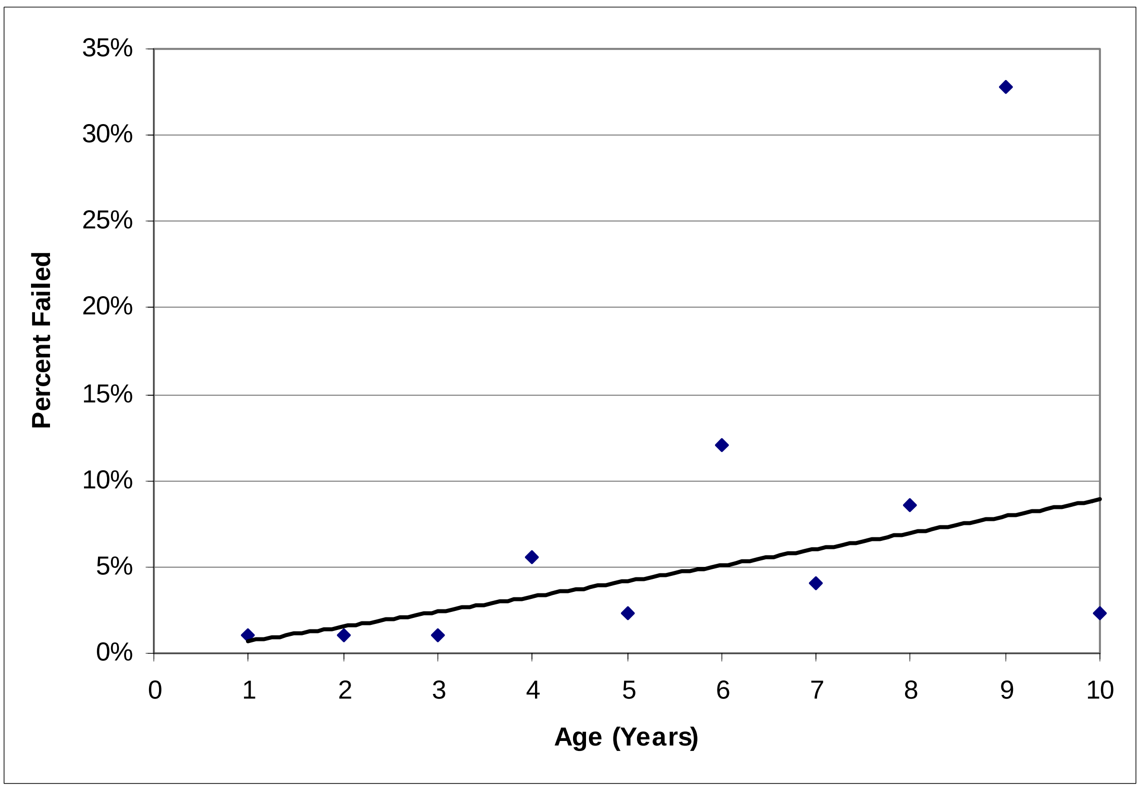


Figure 4-4. Plot of Hazard Rates for All Fryers Not Including Failed Businesses

Table 4-9. Hazard Rates for All Ranges with Survey Code Dates

All Ranges				Ranges With Survey Code Dates			
Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	139.0	5.1	3.70%	1	123	4	3.25%
2	133.9	1.3	0.96%	2	119	1	0.84%
3	132.6	1.3	0.97%	3	118	1	0.85%
4	131.3	3.9	2.94%	4	117	3	2.56%
5	127.4	2.6	2.02%	5	114	2	1.75%
6	124.9	12.9	10.30%	6	112	10	8.93%
7	112.0	5.1	4.59%	7	102	4	3.92%
8	106.9	9.0	8.42%	8	98	7	7.14%
9	97.9	27.0	27.59%	9	91	21	23.08%
10	70.9	3.9	5.44%	10	70	3	4.29%
Total		72		**Total**		56	

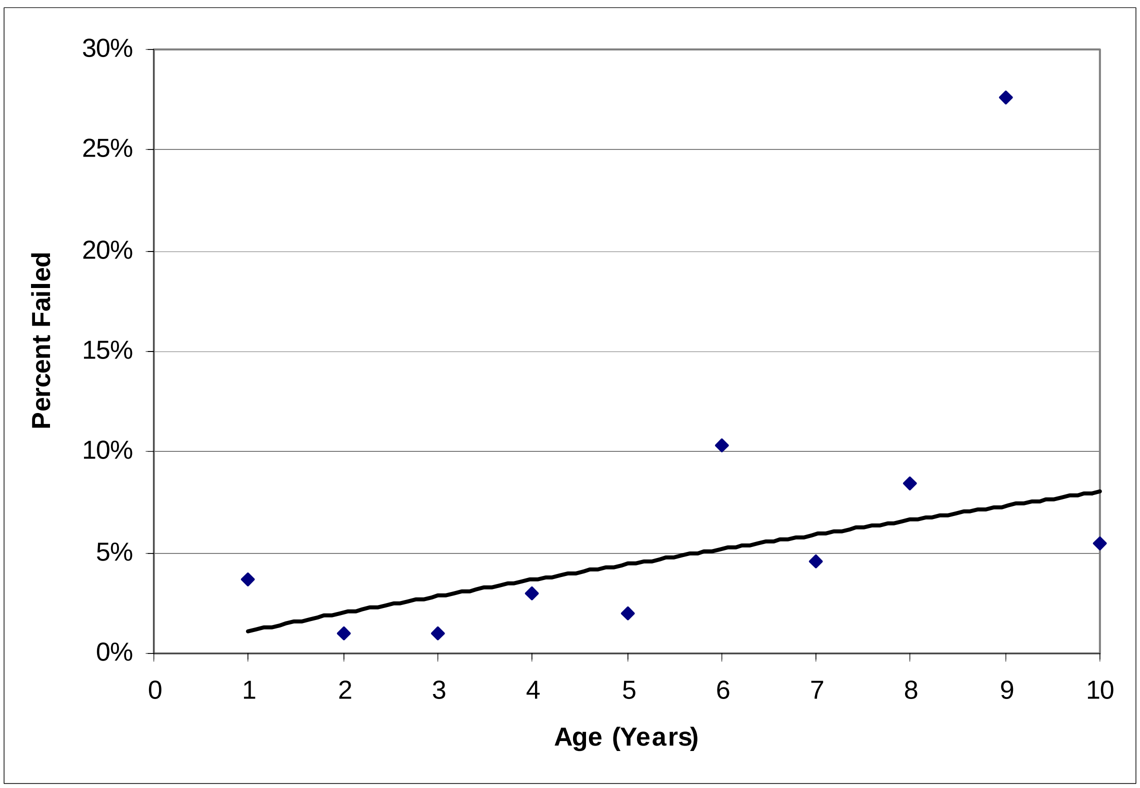


Figure 4-5. Plot of Hazard Rates for All Ranges with Survey Code Dates

Table 4-10. Hazard Rates for All Ranges Not Including Failed Businesses

All Ranges				Ranges With Survey Code Dates			
Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	124.0	5.3	4.28%	1	113	4	3.54%
2	118.7	1.3	1.12%	2	109	1	0.92%
3	117.4	1.3	1.13%	3	108	1	0.93%
4	116.0	4.0	3.43%	4	107	3	2.80%
5	112.1	1.3	1.18%	5	104	1	0.96%
6	110.7	8.0	7.18%	6	103	6	5.83%
7	102.8	5.3	5.16%	7	97	4	4.12%
8	97.5	9.3	9.52%	8	93	7	7.53%
9	88.2	21.2	24.06%	9	86	16	18.60%
10	67.0	4.0	5.94%	10	70	3	4.29%
Total		61		**Total**		46	

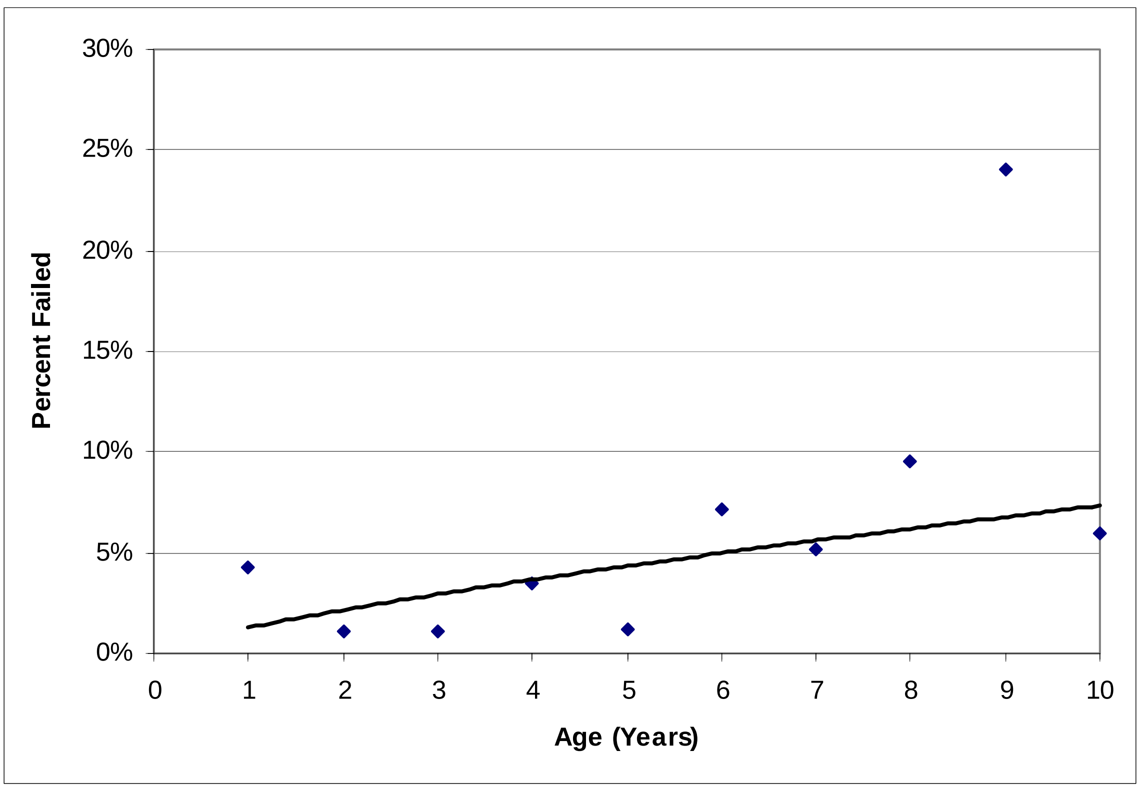


Figure 4-6. Plot of Hazard Rates for All Ranges Not Including Failed Businesses

Table 4-11. Hazard Rates for All Griddles with Survey Code Dates

All Griddles				Griddles With Survey Code Dates			
Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	99.0	1.4	1.41%	1	86	1	1.16%
2	97.6	1.4	1.43%	2	85	1	1.18%
3	96.2	1.4	1.46%	3	84	1	1.19%
4	94.8	1.4	1.48%	4	83	1	1.20%
5	93.4	4.2	4.50%	5	82	3	3.66%
6	89.2	1.4	1.57%	6	79	1	1.27%
7	87.8	4.2	4.78%	7	78	3	3.85%
8	83.6	4.2	5.02%	8	75	3	4.00%
9	79.4	12.6	15.87%	9	72	9	12.50%
10	66.8	2.8	4.19%	10	63	2	3.17%
Total		35		**Total**		25	

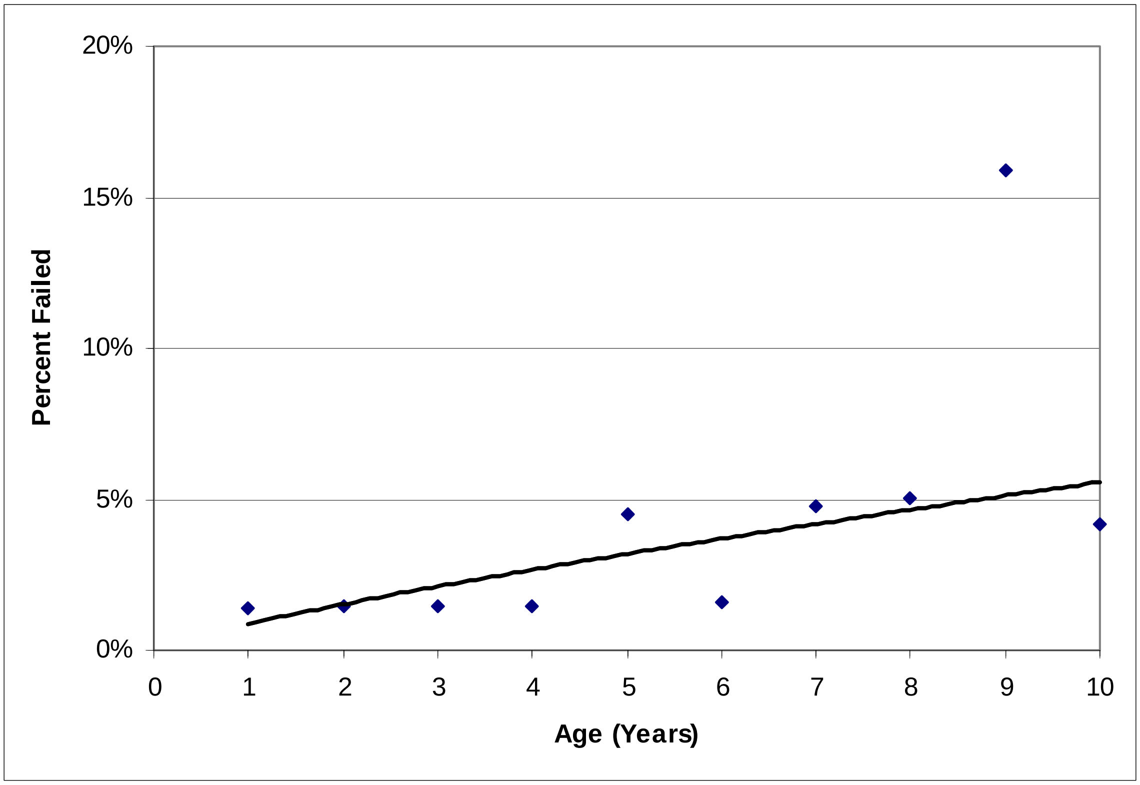


Figure 4-7. Plot of Hazard Rates for All Griddles with Survey Code Dates

Table 4-12. Hazard Rates for All Griddles Not Including Failed Businesses

All Griddles				Griddles With Survey Code Dates			
Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	92.0	1.3	1.47%	1	84	1	1.19%
2	90.7	1.3	1.49%	2	83	1	1.20%
3	89.3	1.3	1.51%	3	82	1	1.22%
4	88.0	1.3	1.53%	4	81	1	1.23%
5	86.6	1.3	1.56%	5	80	1	1.25%
6	85.3	1.3	1.58%	6	79	1	1.27%
7	83.9	4.0	4.82%	7	78	3	3.85%
8	79.9	4.0	5.06%	8	75	3	4.00%
9	75.8	12.1	16.00%	9	72	9	12.50%
10	63.7	2.7	4.23%	10	63	2	3.17%
Total		31		**Total**		23	

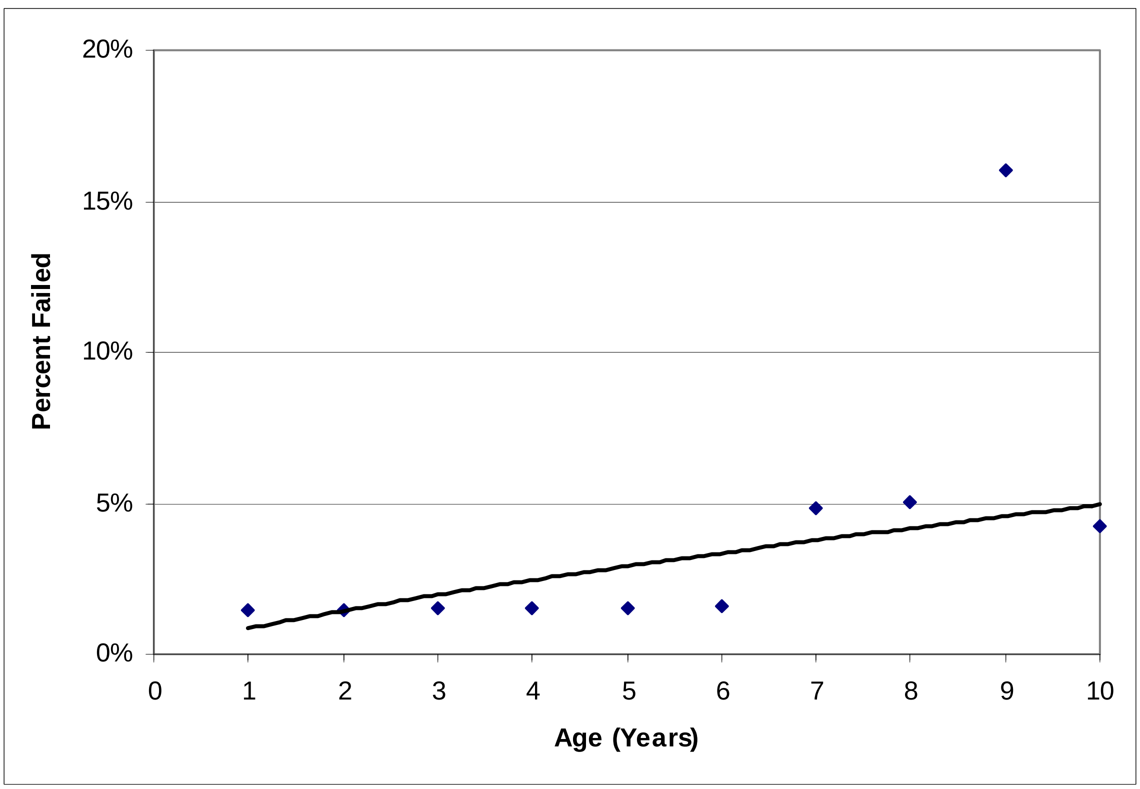


Figure 4-8. Plot of Hazard Rates for All Griddles Not Including Failed Businesses

Calculated hazard rates for ovens, fryers, ranges, and griddles vary over time as shown in **Tables 4-5** through **4-12**. This indicates that there is no reason to assume the survival function can be represented using the exponential distribution, since the hazard rate for an exponential survival function is constant. However, the Weibull distribution does allow for hazard rates that increase over time, and the Weibull-based hazard function is therefore used as the functional form for estimating the hazard functions.

Power curve fits to the hazard rate data in **Tables 4-5** through **4-12** provide estimates of the parameters for the Weibull distribution representation of the hazard rate function. The parameters estimated through power curve fits and the estimated scale and shape parameters of the Weibull function are reported in **Table 4-13**. Plots of the hazard rate functions for ovens, fryers, ranges, and griddles are illustrated in **Figures 4-9** through **4-16**.

With the hazard rate functions estimated and the associated survival functions derived, the effective useful life of ovens or fryers can be estimated as the median survival time. Median survival time is defined as that age where 50% of the appliances have been removed or failed. The resulting estimates of median survival lives are reported and compared to the SoCalGas *ex ante* estimates of effective useful lives in **Table 4-14**.

The M&E Protocols, noted earlier, require a statistical test of whether the *ex post* estimate of useful life is significantly different from the *ex ante* estimate. This test can be accomplished by constructing an 80% confidence interval around the *ex post* estimate and determining whether the *ex ante* estimate falls within this confidence interval. That is, if the *ex ante* estimate falls inside the constructed confidence interval, then the hypothesis of "no difference between the *ex ante* and *ex post* estimates" cannot be rejected. If the *ex ante* estimate falls outside the constructed confidence interval, then the hypothesis of "no difference between the *ex ante* and *ex post* estimates" can be rejected.

To estimate the useful lives of the measures addressed in this study, an 80% confidence interval for the estimated median life of a measure was calculated. This approach includes using the regression fit of the power curve coefficients to report the values of the estimated coefficients associated with the 80% confidence levels. Thus, the power curve regression analysis for each measure provided three sets of parameters for the Weibull hazard rate function: the "best" fit parameters and parameters for the upper and lower bounds of the 80% confidence interval for the estimated coefficients. In effect, the analysis provided an estimate of the "best" hazard function and survival function for a measure, plus estimates of the functions for the upper and lower bounds of the 80% confidence interval.

Figures 4-9 through **4-16** show the "best" fit survival function and the upper and lower bound survival functions associated with the 80% confidence level. The upper and lower bounds on the "best" fit survival function provide the confidence interval bounds for the estimated median useful life.

While sufficient data was available to estimate a median EUL of 10.1 years for ovens, 12.6 years for fryers, 13.3 years for ranges, and 16.7 years for griddles, the “null hypothesis” cannot be rejected for ovens, fryers, ranges, or griddles for the following reasons.

- For ovens, controlling for the important background variable of business failures, the *ex ante* EUL is within the 80 percent effective useful life lower and upper bounds of 5.6 and 23.7 years.
- For fryers controlling for business failures, the *ex ante* EUL is within the 80 percent effective useful life lower and upper bounds of 5.7 and 48.2 years.
- For ranges controlling for business failures, the *ex ante* EUL is within the 80 percent effective useful life lower and upper bounds of 5.3 and 74.2 years.
- For griddles controlling for business failures, the *ex ante* EUL is within the 80 percent effective useful life lower and upper bounds of 8.0 and 52.1 years.

Results are not reported for other measures in the study since there were not sufficient numbers of removals/failures for these measures to support the hazard function estimation.

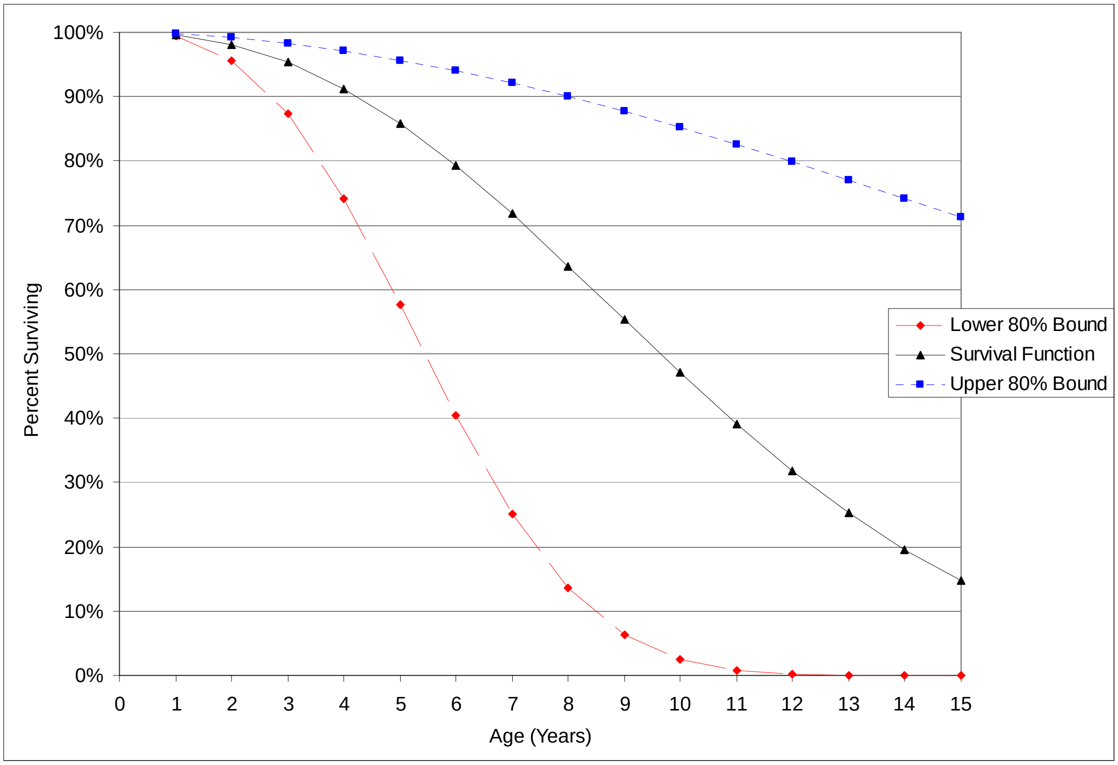


Figure 4-9. Survival Function Plot for All Ovens

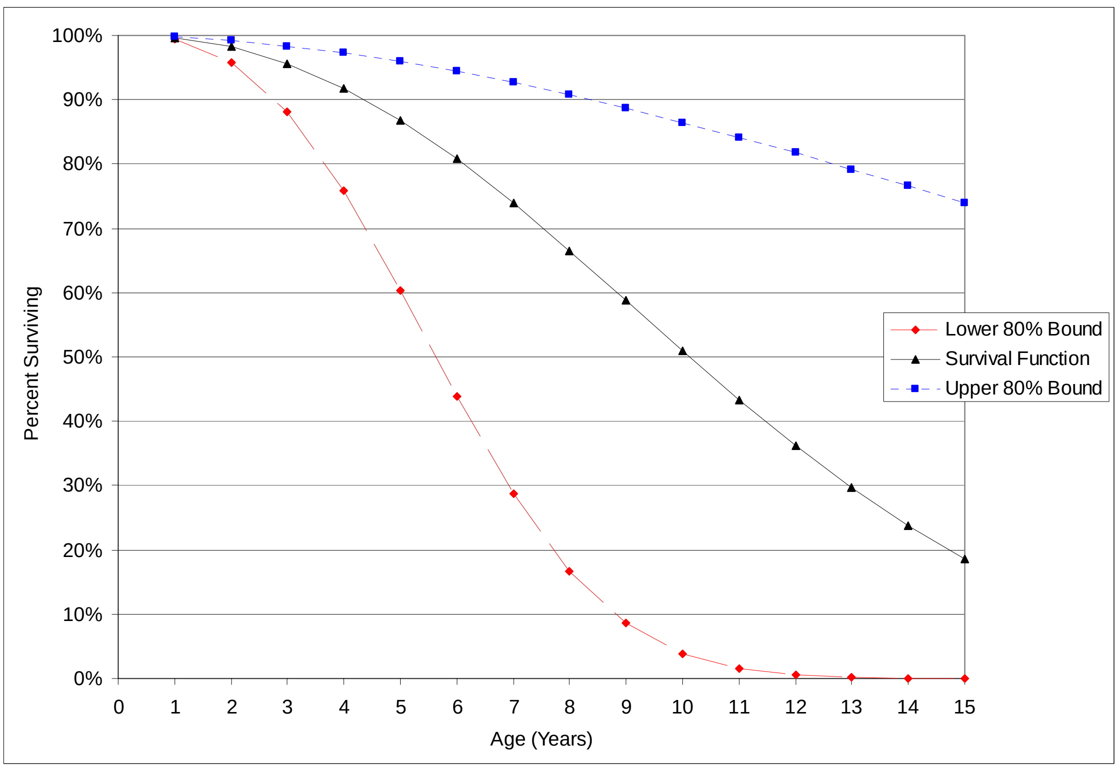


Figure 4-10. Survival Function Plot for Ovens Not Including Failed Businesses

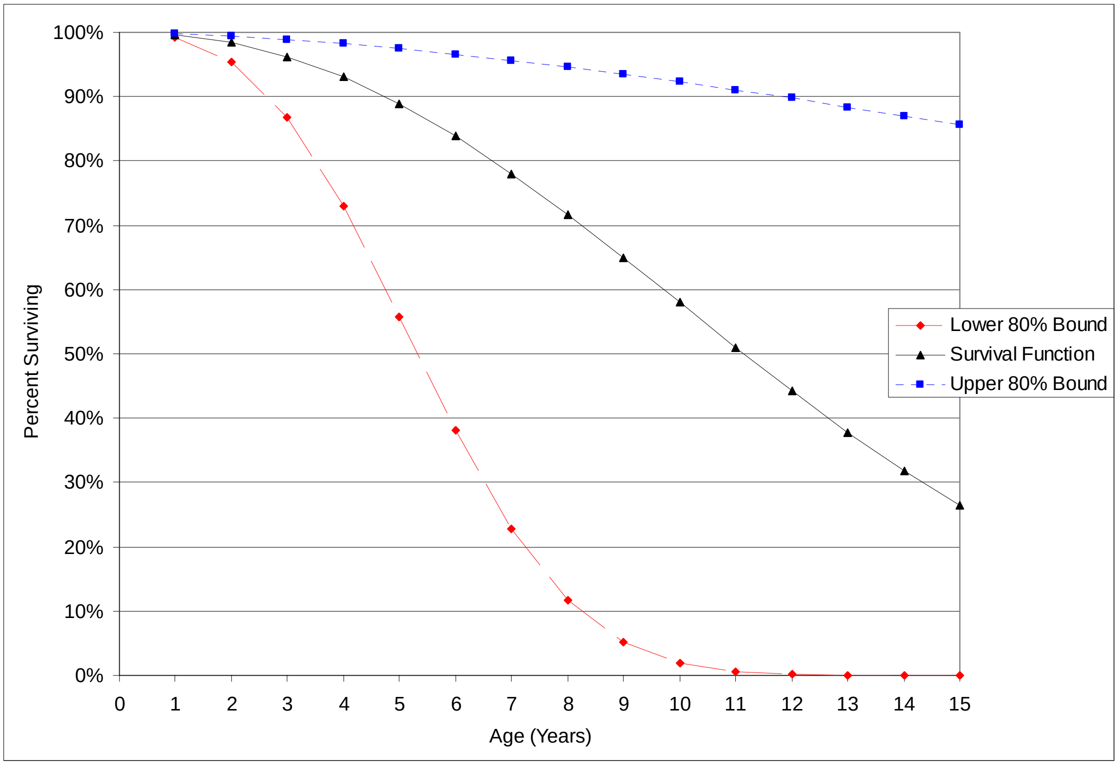


Figure 4-11. Survival Function Plot for All Fryers

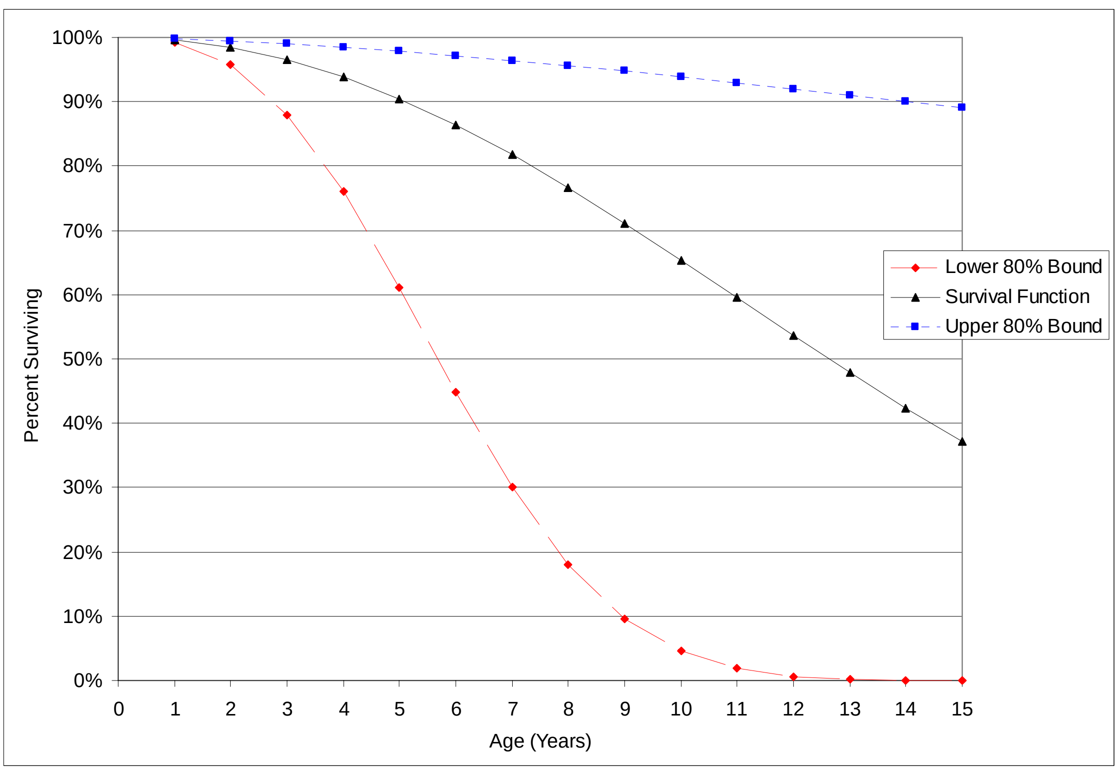


Figure 4-12. Survival Function Plot for Fryers Not Including Failed Businesses

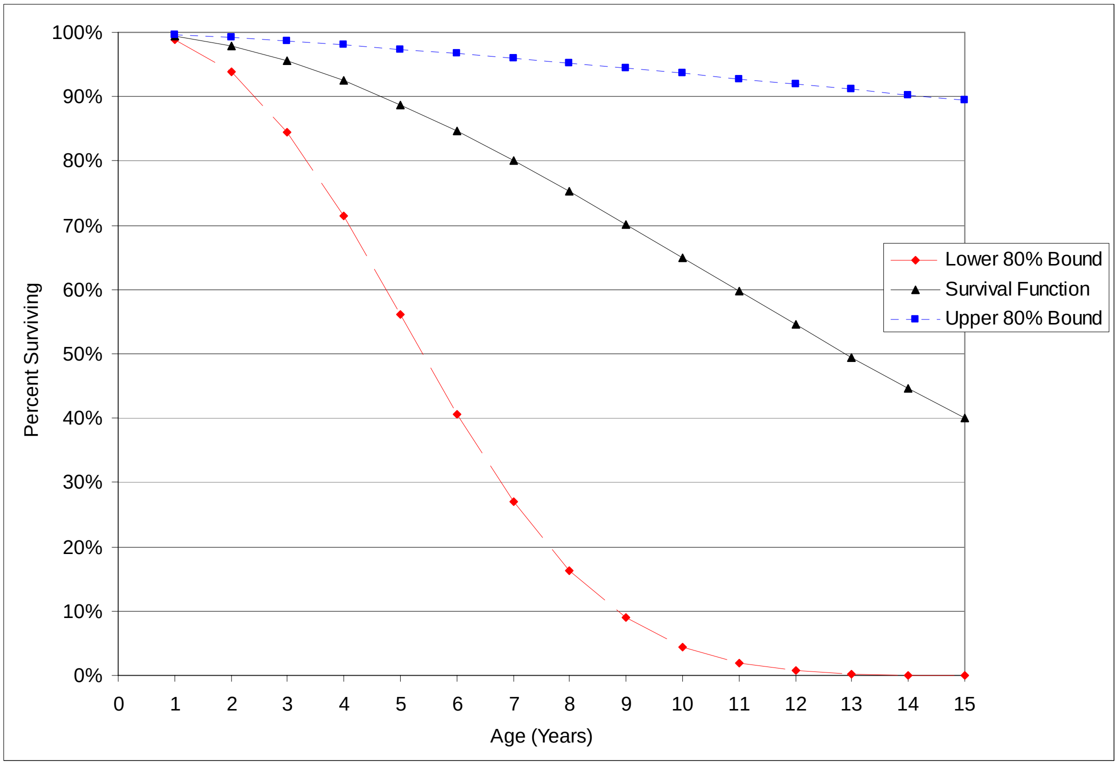


Figure 4-13. Survival Function Plot for All Ranges

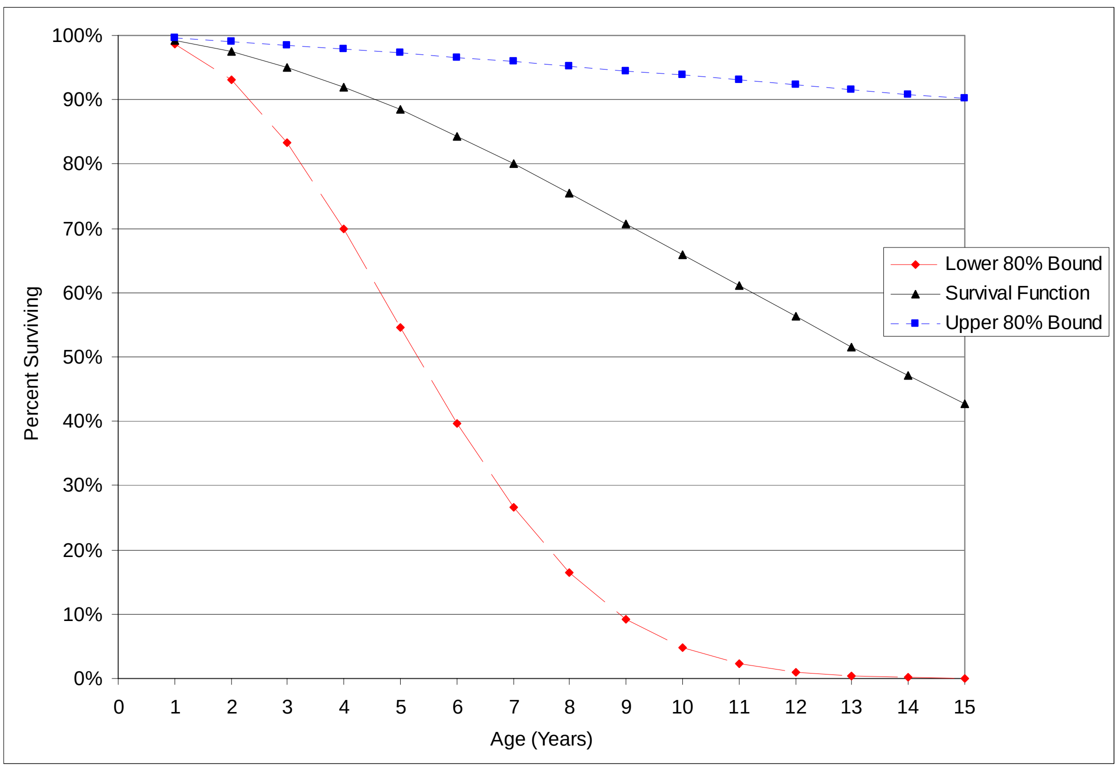


Figure 4-14. Survival Function Plot for Ranges Not Including Failed Businesses

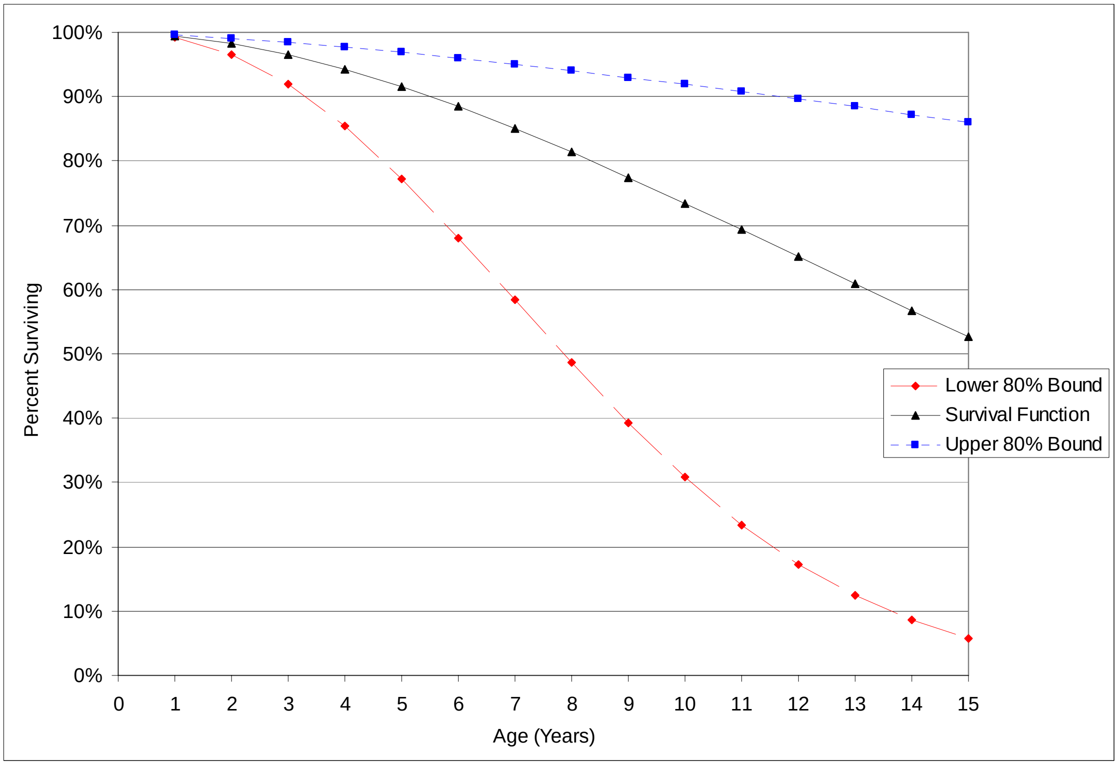


Figure 4-15. Survival Function Plot for All Griddles

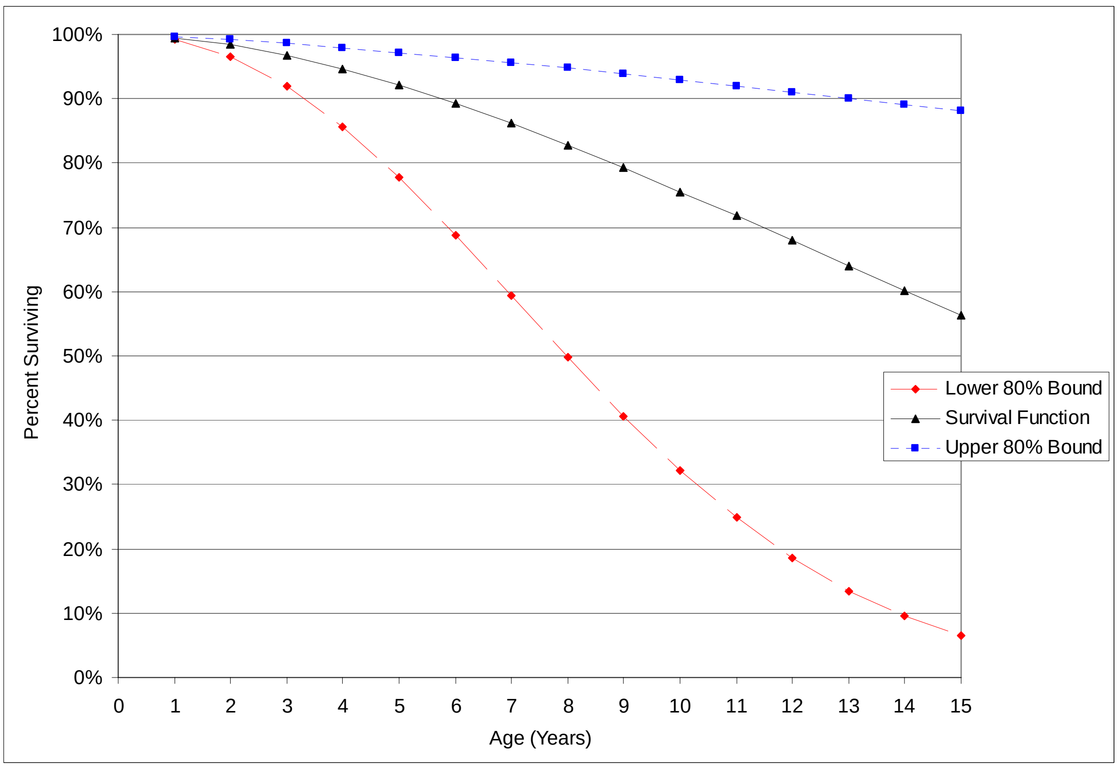


Figure 4-16. Survival Function Plot for Griddles Not Including Failed Businesses

Table 4-13. Hazard Rate Estimation for Fryers, Ovens, Ranges, and Griddles

Type of Measure	Power Curve Fit			Weibull Distribution Parameters	
	a	B	R-squared	α (Scale)	β (Shape)
Ovens All Ovens removed/failed	0.008735	1.296692	0.674295	0.003803	2.296692
Ovens removed/failed not including failed businesses	0.008517	1.25107	0.660800	0.003783	2.251070
Fryers All Fryers removed/failed	0.007449	1.208119	0.533880	0.003374	2.208119
Fryers removed/failed not including failed businesses	0.007249	1.087716	0.478374	0.003472	2.087716
Ranges All Ranges removed/failed	0.011084	0.859824	0.361053	0.005960	1.859824
Ranges removed/failed not including failed businesses	0.012760	0.759913	0.301183	0.007250	1.759913
Griddles All Griddles removed/failed	0.008745	0.80404	0.513689	0.004847	1.804040
Griddles removed/failed not including failed businesses	0.008486	0.765865	0.466741	0.004806	1.765865

Table 4-14. Estimated Effective Useful Life (EUL) for Fryers, Ovens, Ranges, and Griddles

Measure	SoCalGas Ex Ante Estimated EUL	Estimated Median Life			Ex Ante Different From Ex Post at 80% Confidence Level ± 20 Percent?
		80% Lower bound	Estimated EUL	80% Upper Bound	
Ovens All Ovens removed/failed	12	5.4	9.6	22.1	No
Ovens removed/failed not including failed businesses	12	5.6	10.1	23.7	No
Fryers All Fryers removed/failed	12	5.3	11.2	37.0	No
Fryers removed/failed not including failed businesses	12	5.7	12.6	48.2	No
Ranges All Ranges removed/failed	12	5.4	12.9	61.5	No
Ranges removed/failed not including failed businesses	12	5.3	13.3	74.2	No
Griddles All Griddles removed/failed	12	7.9	15.7	44.0	No
Griddles removed/failed not including failed businesses	12	8.0	16.7	52.1	No

Appendix A – Retention Survey Instrument

Retention Survey Instrument
SoCalGas 1995 Commercial New Construction Program

The SoCalGas 1995 Commercial New Construction Program Retention Survey Microsoft Access Database main menu dashboard, database screen to select companies, and retention survey instrument are shown in **Figures A.1**, **A.2**, and **A.3**. The database includes all participants from the original program. Data is entered using a customized database on wireless notebook personal computers and e-mailed to the office for archiving purposes.

Figure A.1 Main Menu Dashboard for the SCG CNC Retention Survey Database

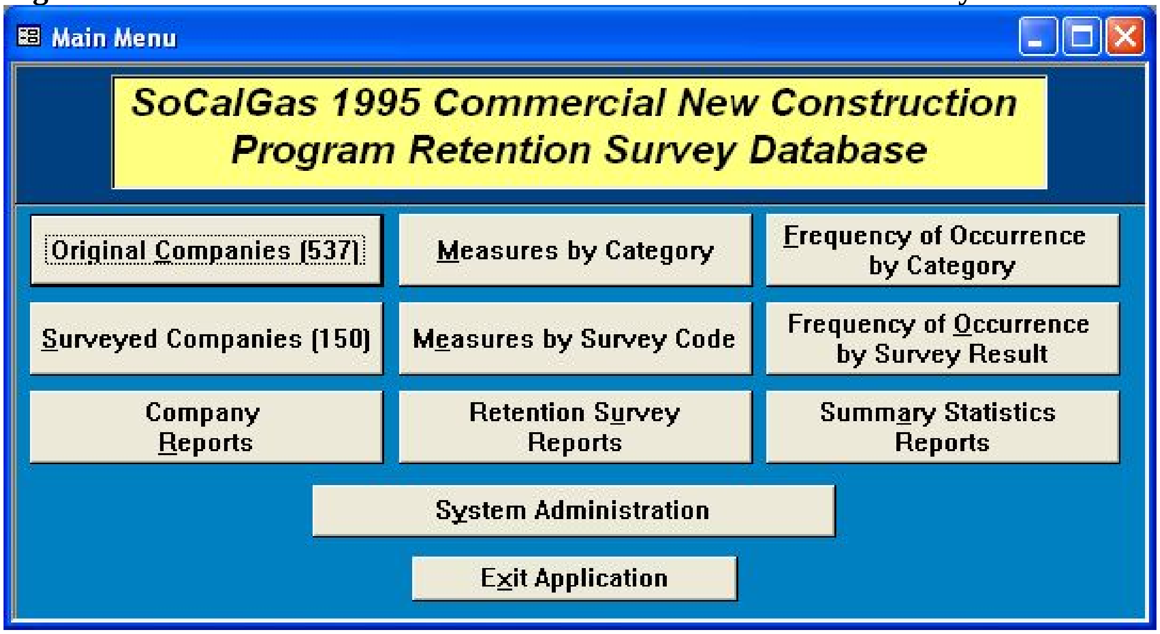


Figure A.2 Database Screen to Select Companies for the SCG CNC Retention Survey

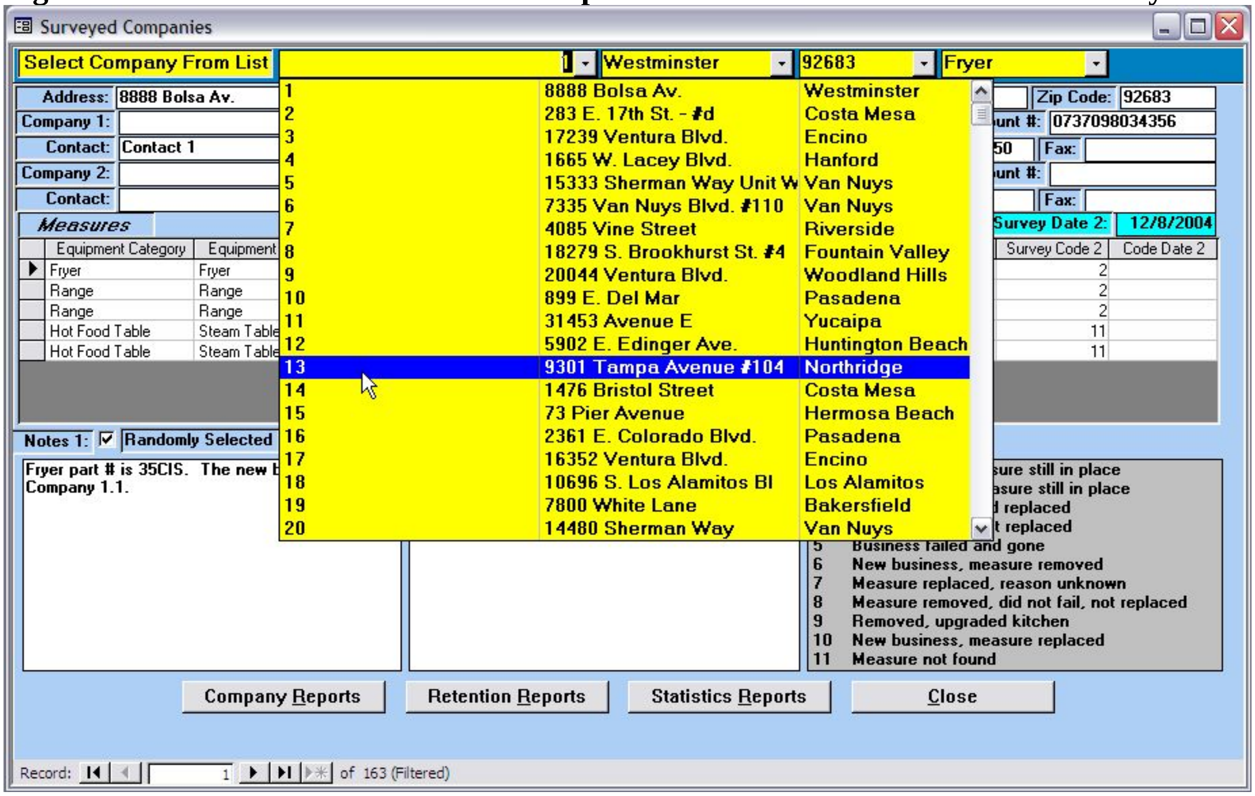
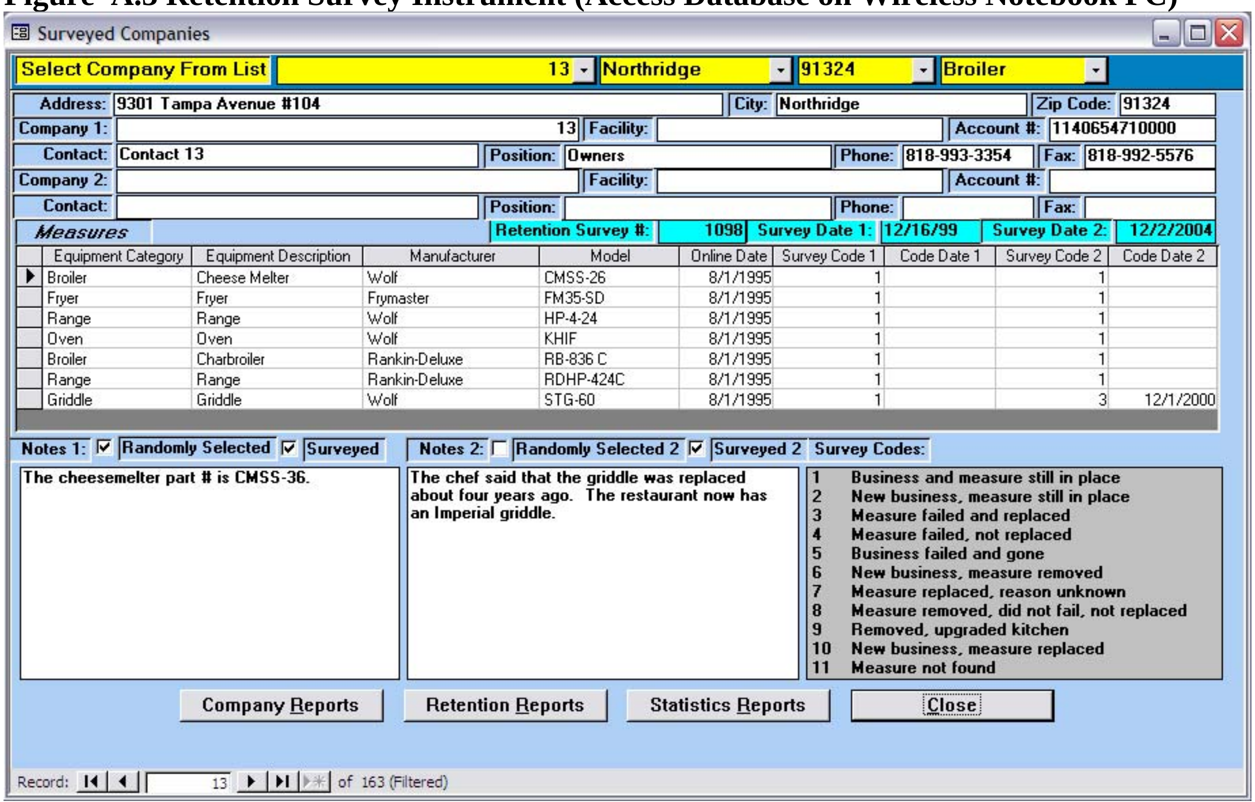


Figure A.3 Retention Survey Instrument (Access Database on Wireless Notebook PC)



Appendix B – CADMAC Protocol Tables 6 and 7

This appendix provides the information requested in Tables 6 and 7 of the CADMAC M&E Protocols.

B.1 Information Required per Table 6 of M&E Protocols

The information required per Table 6 of the M&E Protocols is reported in Table B-1.

1. Identify the studied measure and the end use it belongs to.
This information is provided in Columns (1) and (2) of Table B-1.

2. Identify the ex ante expected useful life and the source of the ex ante expected useful life.
This information is provided in Columns (3) and (4) of Table B-1.

3. Identify the ex post expected useful life estimated in the study.
This information is provided in Column (6) of Table B-1.

4. Identify the ex post expected useful life to be used by the utility in the third and fourth earnings claim.
This information is provided in Column (8) of Table B-1.

5. Identify the standard error associated with the ex post expected useful life.
The survival functions for the measures are not symmetric. Therefore, the standard error does not provide meaningful information on the spread around the estimated median life. Information on the spread around the estimated value is provided by the lower and upper bounds of the confidence interval, reported in Columns (5) and (7) of Table B-1.

6. Provide the 80% confidence interval associated with the ex post expected useful life.
This information is provided in Columns (5) and (7) of Table B-1.

7. Provide the p-value associated with the ex post expected useful life.
The p-value is 20%.

8. Provide the realization rate for the adopted ex post expected useful life. This is defined as the ratio of the adopted ex post expected useful life to the ex ante expected useful life.
This information is provided in Column (9) of Table B-1.

9. Identify all the "like" measures associated with the studied measure.
This information is provided in Column (10) of Table B-1.

Table B- 1. Required Information per Protocols Table 6

		SoCalGas Ex Ante Useful Life		Estimated Median Life					
Measure	End Use	Value	Source	80% Lower Bound	Estimated Mean Effective Useful Life	80% Upper Bound	Ex Post EUL for 3rd & 4th Year Earnings Claim	Realization Rate	"Like" Measures
Oven	Cooking	12	CADMAC Measure Life	5.6	10.1	23.7	12	1.00	None
Fryer	Cooking	12	CADMAC Measure Life	5.7	12.6	48.2	12	1.00	None
Range	Cooking	12	CADMAC Measure Life	5.3	13.3	74.2	12	1.00	None
Griddle	Cooking	12	CADMAC Measure Life	8.0	16.7	52.1	12	1.00	None
Broiler	Cooking	12	CADMAC Measure Life	**	*	**	12	1.00	None
HVAC	Space Conditioning	18	CADMAC Measure Life	**	*	**	18	1.00	None
Steamer	Cooking	12	CADMAC Measure Life	**	*	**	12	1.00	None
Hot Food Table	Cooking	12	CADMAC Measure Life	**	*	**	12	1.00	None
Kettle	Cooking	12	CADMAC Measure Life	**	*	**	12	1.00	None
Braising Pan	Cooking	12	CADMAC Measure Life	**	*	**	12	1.00	None
Other Cooking	Cooking	12	CADMAC Measure Life	**	*	**	12	1.00	None
SHW	Water Heating	15	CADMAC Measure Life	**	*	**	15	1.00	None
Boiler	Space Conditioning	15	CADMAC Measure Life	**	*	**	15	1.00	None
Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Column 10

B.2 Information Required per Table 7 of M&E Protocols

This section provides the information required per Table 7 of the M&E Protocols.

1.a. Study Title and Study ID Number.

Study title:

Southern California Gas Company 1995 Commercial New Construction Program Ninth Year Retention Study

Study ID Number:

Study ID Number 718

1.b. Program, Program years, and program description

Program and Program Year:

Commercial New Construction Program Fourth Year Program

Program Year is 1995

Program Description:

This study examined the retention rates and effective useful lives for measures installed by SoCalGas commercial customers who participated in the 1995 New Construction Program. These customers received financial incentives for installing eligible energy efficiency measures.

1.c. End Uses and Measures Covered:

The sectors, end uses and measures covered are listed in Table B-1.

1.d. Methods and Models Used: Describe the final model specification used for the study.

Where applicable, indicate the study location of the competing class or types of models that were estimated but were not selected. State why the final specification was chosen.

All data for the study, including whether or not installed measures were still in place and operable, was collected through onsite inspections of program participant sites. Collected data were used to determine the percent retention for each measure. Another objective of the study was to estimate effective useful life (EUL) for each measure and to determine if the estimated EULs were different from expected EULs. Direct estimation of survival functions from the collected data could not be made since retention rates for a number of measures were relatively high. Survival functions based on estimated hazard rates could only be estimated for four measures: ovens, fryers, ranges, and griddles. Hazard rates for the remaining measures could not be estimated due to the small number of failures and/or limited sample sizes.

1.e. Analysis Sample Size: Provide the number of customers, number of installations, number of measures (if different) and the number of observations in the analysis and time periods of data collection. If different for different units of analysis, a summary table should be provided.

Table B-2 shows the number of customers included in the study, number of installations and measures, and number of observations. The survey included a total of 163 customer sites and each site had more than one measure. For example, 125 ovens were observed at 70 different customer sites. All onsite inspections and data collection was performed during November 2004 and January 2005.

Table B- 2. Number of Customers, Program Installations, and Observations

Measure	Total Measures in 1995 Program	Observations Included in Retention Study	Customer Sites Included in Retention Study
Oven	489	125	70
Fryer	451	126	85
Range	361	139	92
Griddle	254	99	68
Broiler	207	77	57
HVAC	124	73	6
Steamer	83	28	17
Hot Food Table	65	22	17
Kettle	61	15	11
Braising Pan	23	8	5
Other	75	17	17
SHW	41	18	13
Boiler	34	6	3
Total	**2,268**	**753**	

2.a. Identify the specific data sources used for each data element.

Total program measures and participant customer information were obtained from the SoCalGas 1995 New Construction Program database. Data for the measure retention study were collected from 163 customer site inspections implemented over a three-month period in November 2004 and January 2005.

2.b. Diagram and describe the data attrition process commencing with the program database for participants. Specific numbers and decision points for inclusion and exclusion should be provided. Where different data sources are used (e.g., surveys and program records), appropriate attrition categories should be used (e.g., response rates for surveys).

The steps involved in preparing the various data sets used for the measure retention analysis are illustrated in Figure B-1.

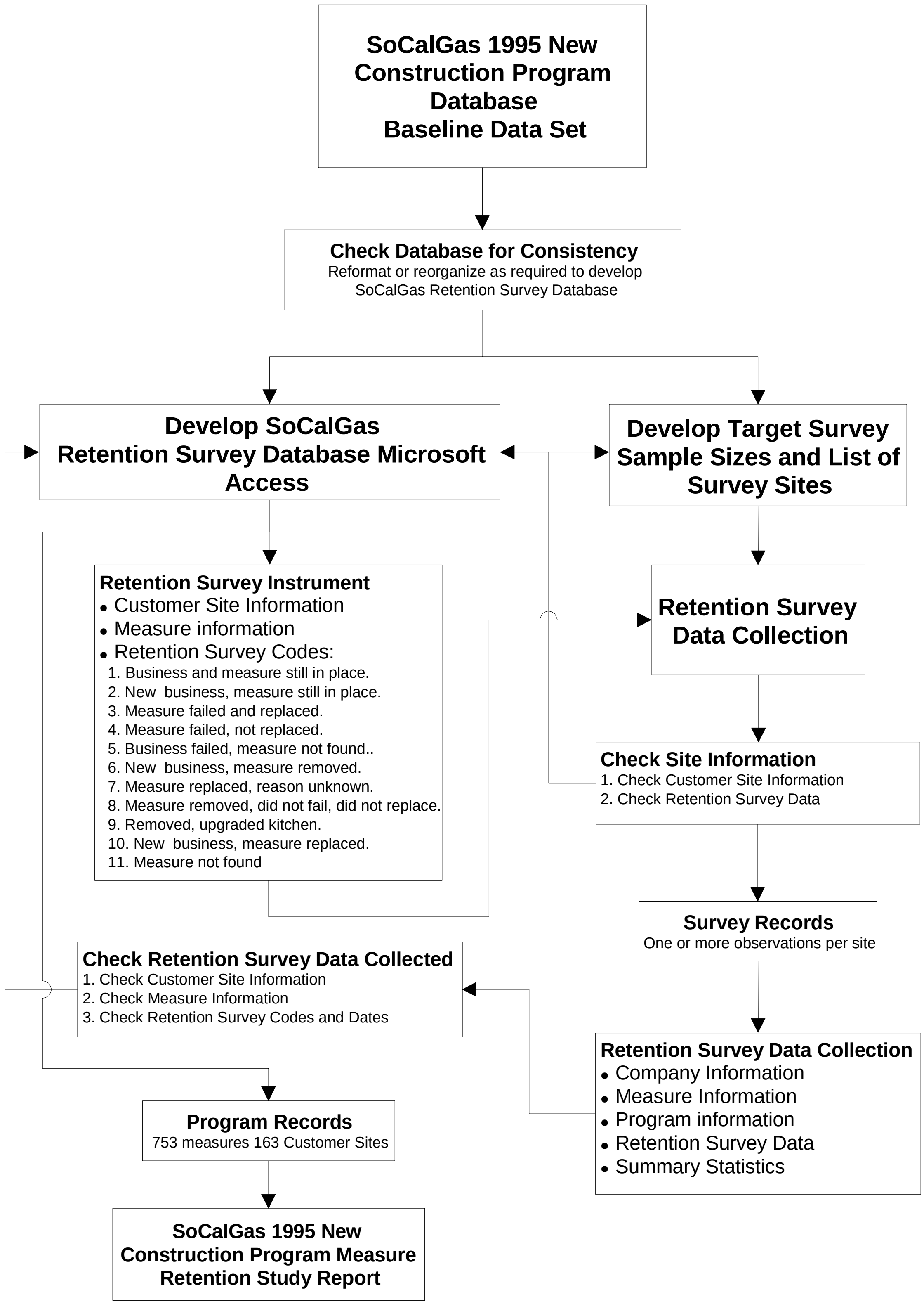


Figure B-1. Flow Chart of Baseline Data Development and Data Collection Process

2.c. Describe the internal/organizational data quality checks and data quality procedures used to match customers and surveys, participation records, and any other data used in the analysis.

Quality control procedures were used to ensure that collected data was accurate and consistent with participant program information and study objectives. Among others, the following procedures were included.

- Reconciling discrepancies between baseline, interview, and visual inspection results prior to leaving a facility.
- Checking the SoCalGas program participant database and retention survey databases to ensure all customers included in the survey were in fact program participants.
- Creating onsite measure retention survey forms for each site using the retention survey database.
- Recording onsite data into the Microsoft Access Retention Survey Database.
- Exporting all data from the database into Microsoft Excel spreadsheets for validation and analysis.
- Double-checking all data entry (by the RMA project manager and the RMA database quality control engineer).

2.d. Provide a summary of the data collected specifically for the analysis but not used, the reasons for them not being used, and a documentation of where those data reside.

The Retention Survey Instrument used for the onsite data collection is provided in Appendix A of the final report. The retention survey instrument shows all of the data that were collected for the analysis. The major items that were used for the analysis were the Retention Survey Codes (described in Chapter 3). All data collected during the on-site surveys were used in the analysis for this report.

3.a. Sampling procedures and protocols: Describe the sampling procedures and protocols used. Information provided should include the sampling frame (e.g., eligible population), sampling strategy (e.g., random, stratified, etc.), sampling basis (e.g., customers, installation, rebate issued), and stratification criteria (e.g., geographic, etc.). Specific data and formulas should be used to present sampling goals and achieved results.

The survey sample design was developed using survival analysis techniques. Survival analysis pertains to the analysis of data that correspond to a well-defined time origin until the occurrence of a particular event or end-point. For this study, the time origin event is defined by the installation of a measure under the SoCalGas 1995 New Construction Program. The failure or removal of the measure defines the end-point. Measure survival data were expected to have several features that warranted special treatment in preparing the sample design.

- The measure survival data would probably not be symmetrically distributed and cannot be reasonably represented by a normal distribution.
- The survival data would be *right-censored* in that the removal or failure end-points will not be observable for some of the installed measures.

A sample design for addressing these and other features of the data was developed using the following steps.

1) Determine the required number of removals or failures for each measure to test the "null hypothesis" with 80 percent confidence plus or minus 20 percent precision.
2) Determine the probability of removal or failure for each measure.
3) Determine the target sample size for each measure based on the number of removals or failures and the probability of failure for each measure.
4) Randomly select sample sites for each measure among facilities in the retention study database.

The first step in preparing the sample design was to arrive at quantitative estimates of the target sample size for each measure. The target sample size for each measure is based on the following exponential distribution assumed to represent the measure survival function.[9]

Eq. B-1. $S(t) = \exp(-\lambda t)$

Where, λ = 1/mean effective useful life. The standard error is given by $1/(\lambda r)$, where r is the number of measure occurrences within a sample that have been removed or failed. With an exponential survival function, the standard error for the estimated mean from a sample depends on the number of observed removals/failures.

The goal in developing the sample design was to obtain results with an 80 percent confidence level plus or minus 20 percent. The following equation was used to calculate the sample size to meet the 80 percent confidence level.

Eq. B-2. $r = \left[\dfrac{z}{0.2}\right]^2$

Where, z is defined as the critical value of the standard normal distribution at the desired level of confidence. For the 80 percent confidence level, z = 1.28. Thus, the number of removals/failures required to estimate mean measure life for a particular measure at the specified confidence level is r = 41. Wherever possible this value was used to establish the preliminary sample design size in order to meet the 80 percent confidence level.

As noted above, there would likely be right-censoring of the occurrences of a measure in the sample; not all of the occurrences would be observed until their life end-point. Accordingly, the number of measure occurrences brought into the sample had to be greater to accommodate this right censoring phenomenon. The sample size needed to provide the required number of removals was determined using the following equation.

Eq. B-3. $\text{Sample Size} = \dfrac{\text{Number of Required Removals/Failures}}{\text{Probability of Required Removals/Failures}}$

[9] *The Statistical Analysis of Failure Time Data*, Kalbfleisch, J.D., and Prentice, R.L., John Wiley and Sons, New York, NY 1980.

As shown by Collett,[10] the probability of removal/failure with an assumed survivor function can be calculated with the following information.

1) Specified values for the survivor function;
2) Study accrual time (i.e., the period when measure occurrences take place); and
3) Study follow-up time (i.e., the period when occurrences are tracked to see whether they are removed or fail).

For this study, the accrual period was 12 months (assumed time activities for the SoCalGas 1995 New Construction Program), and the follow-up period is 108 months (the nine years 1996 through 2004 prior to onsite data collection). Ex ante mean values of measure life (for calculating the parameters of the assumed exponential survivor functions for the various types of measures) were taken from *DSM Measure Life Project: Master Tables of Measure Life Estimates and Final Report*[11], prepared for CADMAC by Energy Management Services.

Given the fixed length of the study, the probability of removal/failure was determined primarily by the expected mean life of a measure. Measures with shorter mean lifetimes have a higher probability of removal or failure. For example, the probability of removal or failure is 0.279 for measures with a mean life of 12 years, 0.231 for measures with a mean life of 15 years, and 0.197 for measures with a mean life of 18 years. With the required number of removals/failures for either type of measure being 41, the respective target sample sizes are 87 for 12-year life measures, 101 for 15-year life measures, and 116 for 18-year life measures. These target sample sizes were used to identify which measures might have a sufficient number of program installations to obtain the required 41 removals/failures and test the “null hypothesis.”

Total measures, target and actual sample sizes, and actual removals or failures for each measure are shown in Table B-3. Of the 13 measures surveyed in the measure retention study, only the top six measures could be sampled sufficiently to test the “null hypothesis” and satisfy the 80 percent confidence/precision. The relatively small numbers of program installations limited sampling of the remaining measures.

Sampling frames for selecting the sample sites for the different measures were created by identifying information on program participant sites and measures from the initial program participant database. A measure-stratified random number was assigned to each site. Sample sites were then selected based on the numerical sorting of the random numbers assigned to each site. Adjustments were made to the required sample size based on retention survey results (i.e., comparing the number of removals or failures encountered during the onsite survey with the number of required removals or failures).

[10] *Modeling Survival Data in Medical Researc*h, Collett, D., Chapman & Hall, 1994, pp. 260-264.

[11] *DSM Measure Life Project: Master Tables of Measure Life Estimates and Final Report*, prepared by Energy Management Services, prepared for CADMAC, September 1993.

Table B-3. Total Measures, Target Sample Size, Sample Size, and Removals or Failures

Measure	Total Measures in 1995 Program	Target Sample Size	Actual Retention Survey Sample Size	Actual Removals or Failures
Oven	489	147	125	81
Fryer	451	147	126	81
Range	361	147	139	72
Griddle	254	147	99	35
Broiler	207	147	77	48
HVAC	124	124	73	8
Steamer	83	35	28	23
Hot Food Table	65	20	22	12
Kettle	61	20	15	9
Braising Pan	23	10	8	6
Other Cooking	75	10	17	6
SHW	41	20	18	10
Boiler	34	10	6	0
Total	**2,268**	**570**	**753**	**391**

3.b. Survey information: Survey instruments should be provided. Response rates should be presented. Reasons for refusals should be presented in tabular form. Efforts to account for or test for non-response bias should be presented, as well as corrections to account for the bias.

The retention survey instrument used for the onsite data collection is provided in Appendix A. Customer data developed for this study included a random sample population of 200 sites from which 163 sites were selected for onsite inspection. There were no non-response sites. All sites cooperated with the onsite survey.

3.c. Statistical descriptions. For the key variables that were used in the final models, provide descriptive statistics for the participant group, and, when present, for the comparison group.

The key variable for the analysis of retention is the number of removal/failures that occur for a measure over a specified time period. The removal/failure rates over a ten-year period are summarized for the various measures in Table B-4. The average retention rate was 47%, including business failures and removals (column 2). However excluding business failures and removals (i.e., percentage of measures not failed), the average retention rate was 86% (column 5). Thus, the average retention rate based on measure performance was reasonably high after 10 years.

Table B-4. 10-Year Retention Rates for 1995 Commercial New Construction Measures

Measure	Measures Removed or Failed within 10 Years	Measures Retained After 10 Years	Measures Failed within 10 Years	Measures not Failed After 10 Years
Oven	65%	35%	11%	89%
Fryer	64%	36%	22%	78%
Range	52%	48%	9%	91%
Griddle	35%	65%	14%	86%
Broiler	62%	38%	12%	88%
HVAC	11%	89%	11%	89%
Steamer	82%	18%	21%	79%
Hot Food Table	55%	46%	23%	77%
Kettle	60%	40%	20%	80%
Braising Pan	75%	25%	0%	100%
Other Cooking*	35%	65%	6%	94%
SHW	56%	44%	33%	67%
Boiler	0%	100%	0%	100%
Average (weighted)	**53%**	**47%**	**14%**	**86%**

Note: Other Cooking includes pasta/rice cooker, rethermalizer, hot food plate, and broaster.

4.a. Describe procedures used for the treatment of outliers, and missing data points.

The information required for the analysis was whether a measure had failed or been removed within the time span of the study period. For an individual measure, a removal or failure is essentially a binary 0-1 decision for purposes of analysis. The problem of outliers would arise primarily at the aggregate level if there appeared to be a disproportionate percentage of removals or failures. The possibility of outlier percentages was examined on a measure-by-measure basis. No excessively high rates of removal/failure were detected.

The most important missing data points are the removal or failure dates. Removal or failure dates are identified in the retention survey database as the survey code date. Survey code dates were obtained from the facilities staff in interviews conducted during the onsite inspections. However, survey code dates were not provided for all measures due to the facilities staff not being aware of the date of removal or failure. Only 54 of 81 survey code dates were provided for the ovens, 58 of 81 survey code dates were provided for fryers, 56 of 72 survey code dates were provided for ranges, and 25 of 35 survey code dates were provided for griddles that were observed as removed or failed. This presented a problem in terms of developing the hazard rates. RMA used the following approach to solve this problem.

1) Calculate removals/failures for subset of measures where the facilities staff provided survey code dates. These calculations were performed for each year where data was available (e.g., year 1 through year 10).
2) Calculate removals/failures for all measures based on data from the subset of measures where the facilities staff provided survey code dates and lifetime information was available.

The removals/failures each year for all measures are calculated using the following equation.

Eq. B-4. $$r(t) = r_s(t) \times \left[\frac{R}{R_s} \right]$$

Where, $r_s(t)$ = removals/failures in year, t, for subset of measures with survey code dates.
R_s = total removals/failures for subset of measures with survey code dates (R_s = 54).
R = total removals/failures for all measures in survey sample (R = 81).

Hazard rates are calculated using the following equation.

Eq. B-5. $$h(t) = \frac{r(t)}{M(t)}$$

Where, $r(t)$ = removals/failures in year, t.
$M(t)$= Measures at start of year.

Hazard rates for all ovens (based on data from the subset of ovens with survey code dates) are shown in Table B-5. An example calculation for ovens is shown in the following equation.

Eq. B-4a. Calculated removals/failures for all ovens in the first year = $1 \times \left[\frac{81}{54}\right] = 1.5.$

Assuming 1.5 ovens in the survey sample were removed/failed in the first year implies a hazard rate of 1.2 percent. With 1.5 ovens removed/failed during the first year, there were 123.5 ovens "at risk" at the start of the second year. Assuming 1.5 ovens in the survey sample were removed/failed in the second year implies a hazard rate of 1.21 percent. Similar calculations provided hazard rate estimates for years 3 through 10. Hazard rates for all ovens, fryers, ranges and griddles are shown in **Tables B-5** through **B-8**.

Table B-5. Hazard Rates for All Ovens Based on Ovens with Survey Code Dates

All Ovens				Ovens with Survey Code Dates			
Year	Ovens at start of year	Year	Ovens at start of year	Year	Ovens at start of year	Year	Ovens at start of year
1	125.0	1	125.0	1	125.0	1	125.0
2	123.5	2	123.5	2	123.5	2	123.5
3	122.0	3	122.0	3	122.0	3	122.0
4	113.0	4	113.0	4	113.0	4	113.0
5	111.5	5	111.5	5	111.5	5	111.5
6	102.5	6	102.5	6	102.5	6	102.5
7	89.0	7	89.0	7	89.0	7	89.0
8	69.5	8	69.5	8	69.5	8	69.5
9	59.0	9	59.0	9	59.0	9	59.0
10	48.5	10	48.5	10	48.5	10	48.5
Total		Total		Total		Total	

Table B-6. Hazard Rates for All Fryers Based on Fryers with Survey Code Dates

All Fryers				Fryers With Survey Code Dates			
Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	124.0	1.4	1.13%	1	101	1	0.99%
2	122.6	1.4	1.14%	2	100	1	1.00%
3	121.2	1.4	1.15%	3	99	1	1.01%
4	119.8	7.0	5.83%	4	98	5	5.10%
5	112.8	5.6	4.95%	5	93	4	4.30%
6	107.2	18.2	16.93%	6	89	13	14.61%
7	89.1	4.2	4.70%	7	76	3	3.95%
8	84.9	8.4	9.87%	8	73	6	8.22%
9	76.5	32.1	41.98%	9	67	23	34.33%
10	44.4	1.4	3.15%	10	44	1	2.27%
Total		81		Total		58	

Table B-7. Hazard Rates for All Ranges with Survey Code Dates

All Ranges				Ranges With Survey Code Dates			
Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	139.0	5.1	3.70%	1	123	4	3.25%
2	133.9	1.3	0.96%	2	119	1	0.84%
3	132.6	1.3	0.97%	3	118	1	0.85%
4	131.3	3.9	2.94%	4	117	3	2.56%
5	127.4	2.6	2.02%	5	114	2	1.75%
6	124.9	12.9	10.30%	6	112	10	8.93%
7	112.0	5.1	4.59%	7	102	4	3.92%
8	106.9	9.0	8.42%	8	98	7	7.14%
9	97.9	27.0	27.59%	9	91	21	23.08%
10	70.9	3.9	5.44%	10	70	3	4.29%
Total		72		Total		56	

Table B-8. Hazard Rates for All Griddles with Survey Code Dates

All Griddles				Griddles With Survey Code Dates			
Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	99.0	1.4	1.41%	1	86	1	1.16%
2	97.6	1.4	1.43%	2	85	1	1.18%
3	96.2	1.4	1.46%	3	84	1	1.19%
4	94.8	1.4	1.48%	4	83	1	1.20%
5	93.4	4.2	4.50%	5	82	3	3.66%
6	89.2	1.4	1.57%	6	79	1	1.27%
7	87.8	4.2	4.78%	7	78	3	3.85%
8	83.6	4.2	5.02%	8	75	3	4.00%
9	79.4	12.6	15.87%	9	72	9	12.50%
10	66.8	2.8	4.19%	10	63	2	3.17%
Total		35		Total		25	

4.b. Describe what was done to control for the effects of background variables, such as economic, political activity, etc.

The most important background variable in this retention study is business failures. This background variable is important for two reasons.

1) Restaurants have a 28 percent higher business failure rate than other retail businesses.[12]
2) Restaurants accounted for 98 percent of equipment purchased with incentives through the 1995 Commercial New Construction Program.

This influences the retention rates of most cooking measures as well as the failure and survival analysis for ovens and fryers. The impact of failed businesses on retention rates is illustrated in **Table B-7**. Some measures such as braising pan, SHW, and boiler were unaffected by excluding failed businesses. HVAC and other cooking measures were also relatively unaffected. However, ovens, fryers, broilers, steamers, hot food table, and kettles would have had at least a 10 percent higher retention rate if "removals/failures" at failed businesses were not included.

Hazard rates for ovens, ranges, fryers, and griddles not including failed businesses are provided in **Tables B-9** through **B-12**. Survival functions for these two sets of data are illustrated in **Figure B-2** for ovens, **B-3** for fryers, **B-4** for ranges, and **B-5** for griddles. While sufficient data was available to estimate a median EUL for ovens, fryers, ranges, and griddles, the "null hypothesis" cannot be rejected for these measures. As shown **Table B-12**, the *ex ante* EUL is within the 80 percent effective useful life lower and upper bounds controlling for the important background variable of business failures. The "null hypothesis" regarding EULs cannot be rejected for any measures examined in this study.

According to the CADMAC Protocols effective useful life (EUL) of a measure is defined as "the median number of years that the measure is still in place and operable." Interpretation of the Protocol definition of EUL raises the following questions regarding measure retention.

1) Should background variables such as business failure be included or excluded in calculating retention rates?
2) Should a measure be defined as "removed or failed" if a business fails, but the measure is still in place and operable, or if the measure was sold to another business and at the new business the measure is still in place and operable?
3) Should the CADMAC definition of EUL be revised or expanded to include a methodology for consideration of background variables such as failed businesses?

For the purpose of estimating retention rates, this study defines measures at sites with failed businesses as "removals or failures." However, for the purpose of drawing conclusions regarding the *ex post* measure EUL, this study controls for business failures by excluding these sites from the analysis. Further discussion regarding the impact of business failures and the failure and survival analysis of ovens and fryers is provided in **Section 4.3**. Business failures are included in

[12] The business failure rate for restaurants of 100 failures per 10,000 is 28 percent higher than the business failure rate for total retail trade of 78 failures per 10,000. See *Business Failure Record: A Comparative Statistical Analysis of Geographic and Industry Trends in Business Failures in the United States,* Neil DiBernardo, Editor, The Dun & Bradstreet Corporation, One Diamond Hill Road, Murray Hill, NJ 07974-1218, 1998.

the final retention rate estimates, but are not included in the final *ex post* measure EUL estimates given in Table 6 per the CADMAC Protocols (see **Table B-1**, above).

Except for failed businesses, the overall measure attrition appears to be relatively low. This indicates there are no major economic or political events introducing bias into the data used for analysis of measure life.

Table B-7. Impact of Failed Businesses on Retention Rates

Measure	All Sites Percent Measures Retained	Not Including Failed Businesses Percent Measures Retained	Impact on Retention Rate Not Including Failed Businesses (i.e., increase)
Oven	35%	46%	11%
Fryer	36%	51%	15%
Range	48%	56%	8%
Griddle	65%	73%	8%
Broiler	38%	49%	11%
HVAC	89%	96%	7%
Steamer	18%	32%	14%
Hot Food Table	45%	59%	14%
Kettle	40%	67%	27%
Braising Pan	25%	25%	None
Other Cooking	65%	71%	6%
SHW	44%	44%	None
Boiler	100%	100%	None

Table B-8. Hazard Rates for Ovens Not Including Failed Businesses

All Ovens Not Including Failed Businesses				Ovens with Survey Code Dates			
Year	Ovens at start of year	Year	Ovens at start of year	Year	Ovens at start of year	Year	Ovens at start of year
1	111.0	1	111.0	1	111.0	1	111.0
2	109.7	2	109.7	2	109.7	2	109.7
3	108.5	3	108.5	3	108.5	3	108.5
4	100.9	4	100.9	4	100.9	4	100.9
5	99.6	5	99.6	5	99.6	5	99.6
6	92.0	6	92.0	6	92.0	6	92.0
7	81.9	7	81.9	7	81.9	7	81.9
8	65.5	8	65.5	8	65.5	8	65.5
9	56.6	9	56.6	9	56.6	9	56.6
10	47.8	10	47.8	10	47.8	10	47.8
Total		**Total**		**Total**		**Total**	

Table B-9. Hazard Rates for All Fryers Not Including Failed Businesses

All Fryers				Fryers With Survey Code Dates			
Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Fryers at start of year	Fryers Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	113.0	1.2	1.08%	1	94	1	1.06%
2	111.8	1.2	1.09%	2	93	1	1.08%
3	110.6	1.2	1.10%	3	92	1	1.09%
4	109.4	6.1	5.56%	4	91	5	5.49%
5	103.3	2.4	2.35%	5	86	2	2.33%
6	100.8	12.2	12.06%	6	84	10	11.90%
7	88.7	3.6	4.11%	7	74	3	4.05%
8	85.0	7.3	8.58%	8	71	6	8.45%
9	77.7	25.5	32.84%	9	65	21	32.31%
10	52.2	1.2	2.33%	10	44	1	2.27%
Total		62		**Total**		51	

Table B-10. Hazard Rates for All Ranges Not Including Failed Businesses

All Ranges				Ranges With Survey Code Dates			
Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Ranges at start of year	Ranges Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	124.0	5.3	4.28%	1	113	4	3.54%
2	118.7	1.3	1.12%	2	109	1	0.92%
3	117.4	1.3	1.13%	3	108	1	0.93%
4	116.0	4.0	3.43%	4	107	3	2.80%
5	112.1	1.3	1.18%	5	104	1	0.96%
6	110.7	8.0	7.18%	6	103	6	5.83%
7	102.8	5.3	5.16%	7	97	4	4.12%
8	97.5	9.3	9.52%	8	93	7	7.53%
9	88.2	21.2	24.06%	9	86	16	18.60%
10	67.0	4.0	5.94%	10	70	3	4.29%
Total		61		**Total**		46	

Table B-11. Hazard Rates for All Griddles Not Including Failed Businesses

All Griddles				Griddles With Survey Code Dates			
Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)	Year	Griddles at start of year	Griddles Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)
1	92.0	1.3	1.47%	1	84	1	1.19%
2	90.7	1.3	1.49%	2	83	1	1.20%
3	89.3	1.3	1.51%	3	82	1	1.22%
4	88.0	1.3	1.53%	4	81	1	1.23%
5	86.6	1.3	1.56%	5	80	1	1.25%
6	85.3	1.3	1.58%	6	79	1	1.27%
7	83.9	4.0	4.82%	7	78	3	3.85%
8	79.9	4.0	5.06%	8	75	3	4.00%
9	75.8	12.1	16.00%	9	72	9	12.50%
10	63.7	2.7	4.23%	10	63	2	3.17%
Total		31		**Total**		23	

Table B-12. Estimated Effective Useful Life for Ovens, Fryers, Ranges, and Griddles

Measure	SoCalGas Ex Ante Estimated EUL	Estimated Median Life			Ex Ante Different From Ex Post at 80% Confidence Level ± 20 Percent?
		80% Lower bound	Estimated EUL	80% Upper Bound	
Ovens					
All Ovens removed/failed	12	5.4	9.6	22.1	No
Ovens removed/failed not including failed businesses	12	5.6	10.1	23.7	No
Fryers					
All Fryers removed/failed	12	5.3	11.2	37.0	No
Fryers removed/failed not including failed businesses	12	5.7	12.6	48.2	No
Ranges					
All Ranges removed/failed	12	5.4	12.9	61.5	No
Ranges removed/failed not including failed businesses	12	5.3	13.3	74.2	No
Griddles					
All Griddles removed/failed	12	7.9	15.7	44.0	No
Griddles removed/failed not including failed businesses	12	8.0	16.7	52.1	No

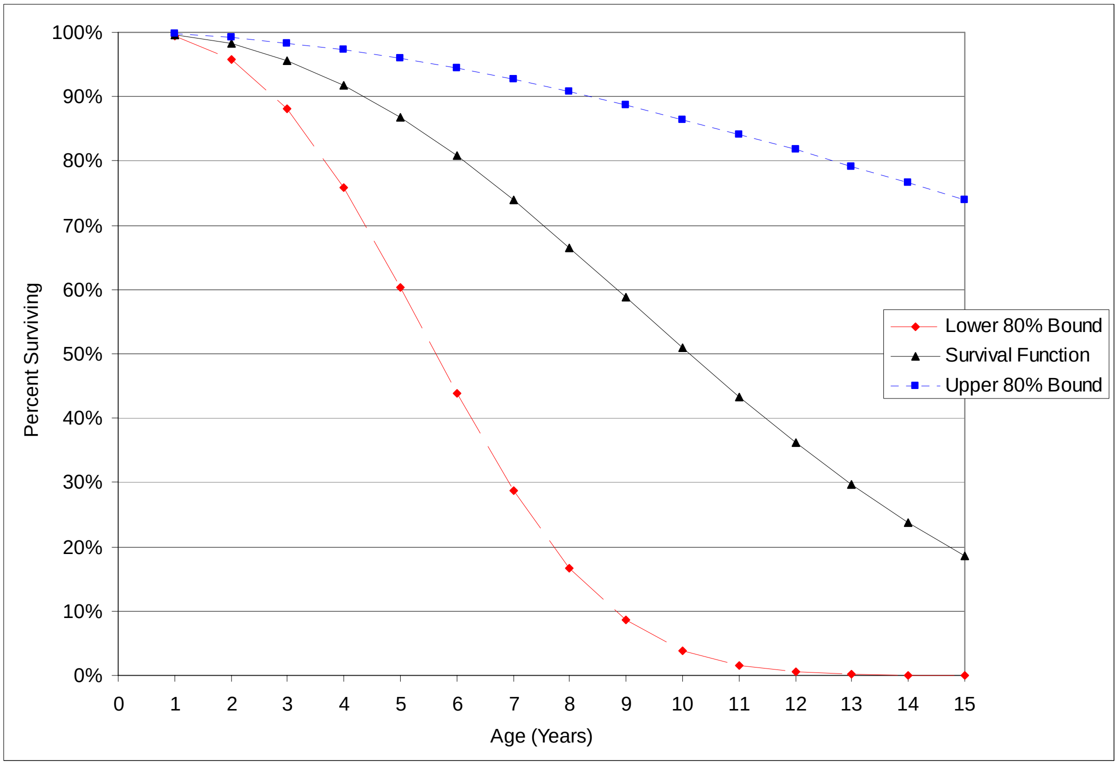


Figure B-2. Survival Function Plot for Ovens Not Including Failed Businesses

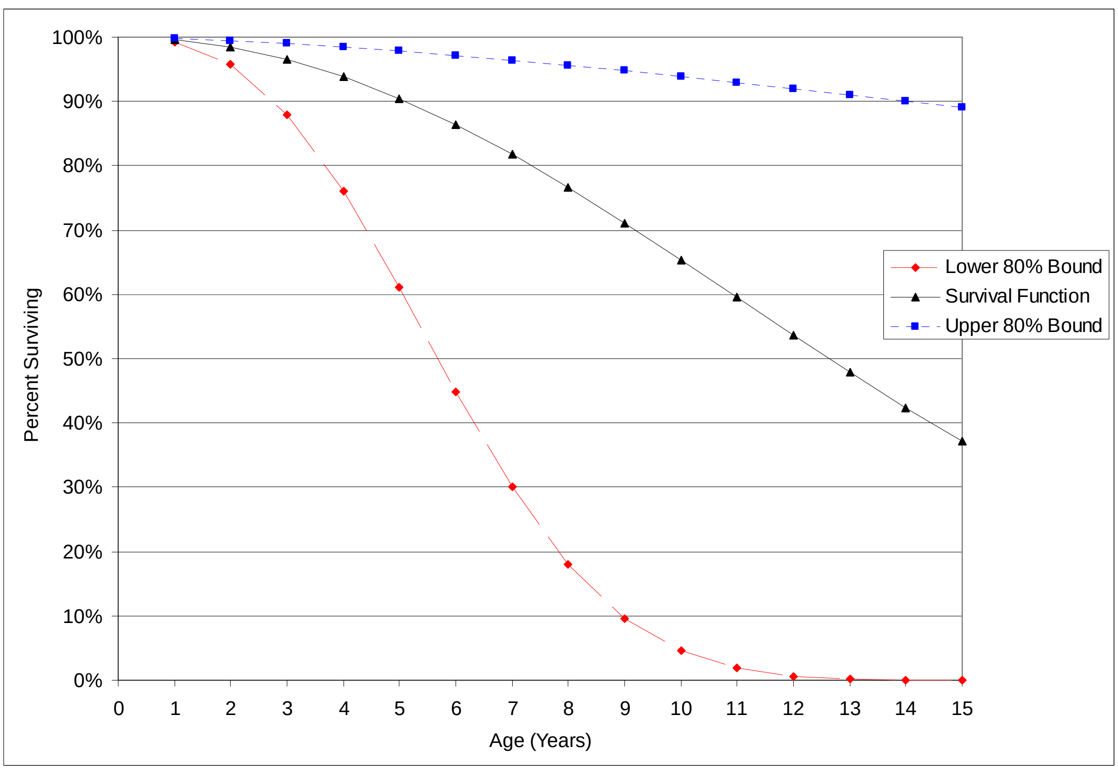


Figure B-3. Survival Function Plot for Fryers Not Including Failed Businesses

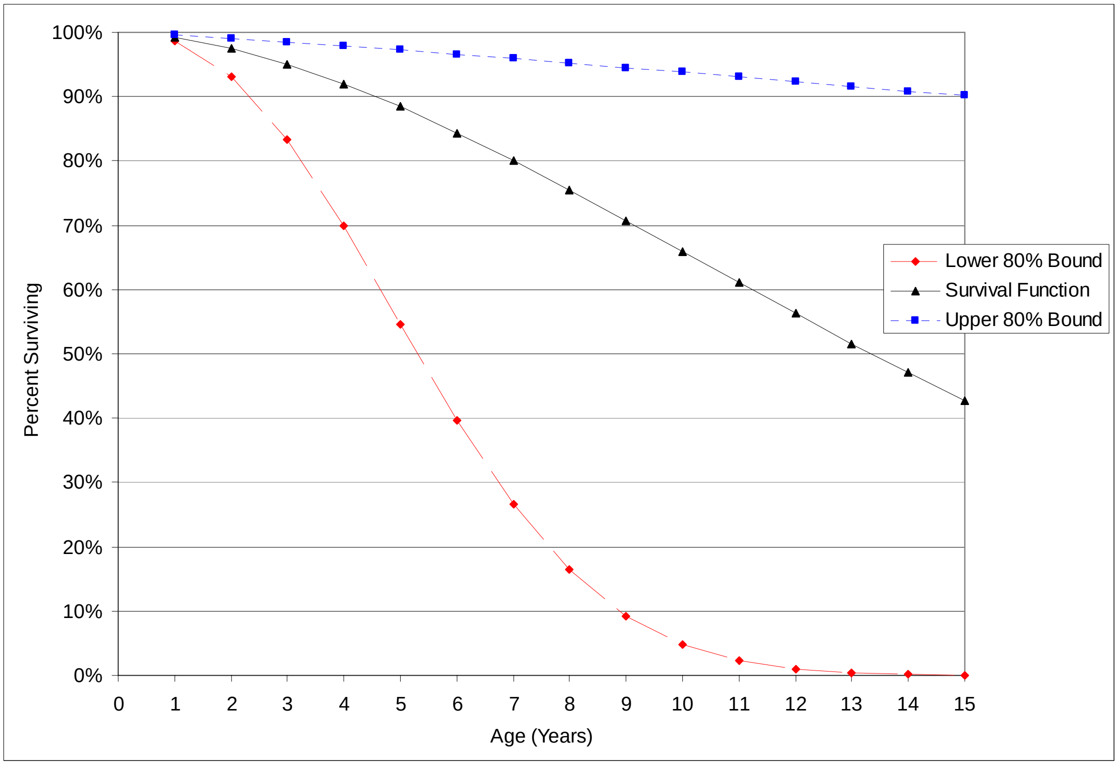


Figure B-4. Survival Function Plot for Ranges Not Including Failed Businesses

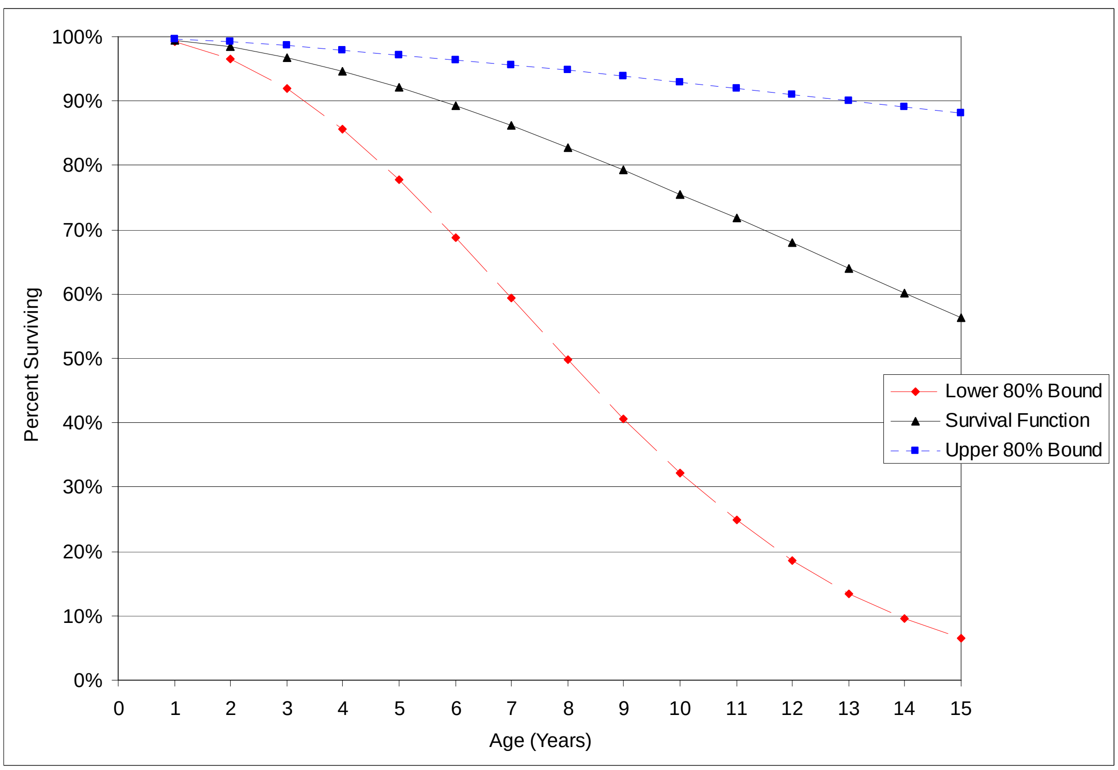


Figure B-5. Survival Function Plot for Griddles Not Including Failed Businesses

4.c. Describe procedures used to screen data for inclusion into the final analysis dataset. Show how many customers, installations or observations were eliminated with each screen.
No screens were used to eliminate customers, installations, or observations from the data set that was used for the analysis. The number of customer sites, program installations and observations are reported in **Table B-2** (above).

4.d. Model Statistics. For all final models, provide standard model statistics in a tabular form.
The final models used for estimating median useful lives for various measures were established by estimating hazard functions for each such measure. This was accomplished using power curve fits for a hazard function defined by a Weibull distribution. The summary statistics for the various models fitted are provided in **Table B-13**.

Table B-13. Hazard Rate Estimation for Fryers, Ovens, Ranges, and Griddles

Type of Measure	Power Curve Fit			Weibull Distribution Parameters	
	a	b	R-squared	α (Scale)	β (Shape)
Ovens All Ovens removed/failed	0.008735	1.296692	0.674295	0.003803	2.296692
Ovens removed/failed not including failed businesses	0.008517	1.25107	0.660800	0.003783	2.251070
Fryers All Fryers removed/failed	0.007449	1.208119	0.533880	0.003374	2.208119
Fryers removed/failed not including failed businesses	0.007249	1.087716	0.478374	0.003472	2.087716
Ranges All Ranges removed/failed	0.011084	0.859824	0.361053	0.005960	1.859824
Ranges removed/failed not including failed businesses	0.012760	0.759913	0.301183	0.007250	1.759913
Griddles All Griddles removed/failed	0.008745	0.80404	0.513689	0.004847	1.804040
Griddles removed/failed not including failed businesses	0.008486	0.765865	0.466741	0.004806	1.765865

4.e. Specification: Refer to the section(s) of the Study that present the initial and final model specifications that were used, the rationale for each, and the documentation for the major alternative models used. In addition, the presentation of the specification should address, at a minimum, the following:
1) Describe how the model specification and estimation procedures recognize and address heterogeneity of customers (i.e., cross-sectional variation)
2) Discuss the factors, and their associated measures, that are omitted from the analysis, and any tests, reasoning, or special circumstances that justify their omission.

The model specifications used for the study are presented and discussed in Chapters 3 and 4. An illustrative example is provided in Chapter 4. For all measures except ovens and fryers the numbers of removals/failures observed over the five-year period were too small or the random sample size was too small to support estimation of hazard functions.

4.f. Error in measuring variables: Describe whether and how this issue was addressed, and what was done to minimize the problem (e.g., response bias, measurement errors, etc.)
Because the removal/failure variable is binary, the issue of measurement error was not considered to affect the results of the analysis.

4.g. Influential data points. Describe the influential data diagnostics that were used, and how the identified outliers were treated.

For some measures, the hazard plots showed a sawtooth pattern over the five-year period of study (i.e., low, high, low, high). With this pattern, a low or a high point could move the fitted regression line. Further data collected over time will allow for better determination of the appropriate hazard function.

4.h. Missing data: Describe the methods used for handling missing data during the analysis phase of the study.

A discussion of survey code dates is provided above in 4.e.

4.i. Precision: Present the methods for the calculation of standard errors.

Because the survival functions for the measures studied are not symmetric, the standard error does not provide meaningful information on the spread around the estimated median life. The lower and upper bounds of the confidence interval provide sufficient information regarding the precision or “spread” around the estimated value. An 80% confidence interval for the estimated median life of a measure was calculated as follows. The regression fit of the power curve coefficients was used to report the values of the estimated coefficients associated with the 80% confidence levels. Thus, the power curve regression analysis for each measure provided three sets of parameters for the Weibull hazard rate function: the “best” fit parameters and parameters for the upper and lower bounds of the 80% confidence interval for the estimated coefficients. In effect, the analysis provided an estimate of the “best” hazard function and survival function for a measure, plus estimates of the functions for the upper and lower bounds of the 80% confidence interval.

Appendix C – Weilbull Analysis of Measure Effective Useful Life

C.1 Introduction

This appendix presents an overview of how Weibull analysis was used to evaluate effective useful measure life for the 1995 Southern California Gas Company Commercial New Construction Program using Microsoft® Excel. Weibull analysis is used to make predictions about the life of products in the population by "fitting" a statistical distribution of life data from a representative sample of units inspected in the field after 5, 10, 15 or more years.[13] The parameterized distribution for the data set can then be used to estimate important life characteristics of the product such as reliability or probability of failure at a specific time, the mean life for the product and failure rate. Weibull analysis requires the practitioner to: gather life data for the product, select a lifetime distribution that will fit the data and model the life of the product, estimate the parameters that will fit the distribution to the data, and generate plots and results that estimate the life characteristics such as reliability or mean life of the product.

For measures evaluated in the retention study for SoCalGas, the Weibull analyses uses a three-step process. First, hazard rates are developed using the available data. Second, hazard functions are developed based on the hazard rates.[14] Third, Weibull distribution survival functions are developed using the estimated hazard function. The steps in the parametric procedure for estimating the survival functions are as follows:

1) Prepare data for calculation of the hazard rate.
2) Estimate the hazard function.
3) Use hazard function to determine Weibull distribution survival function and measure effective useful life.

C.2 Data Preparation

The first step is to prepare data for the calculation of hazard rates. To accomplish this, measure information is exported from the Access database and organized in an Excel spreadsheet. Each measure is listed with the following information: company name, online date, survey date, code date, survey code, equipment category, and equipment description. The measures are sorted first by equipment category and then by survey code. Calculations are performed on the data of each equipment category. See attached CD-ROM for the 1995 Southern California Gas Company Commercial New Construction Program spreadsheet data.

Hazard rates are developed using removal or failure information about the lifetime of the measure. For each failed or replaced measure (survey codes 3-11), its life is calculated as the difference between the date that the measure was removed or failed (code date) and the date the measure was installed (online date). The date that the measure was installed was provided in the SoCalGas program participant baseline database. Survey code dates were obtained from the facilities staff in interviews conducted during the onsite inspections. However, survey code dates

[13] More information is available at: http://www.itl.nist.gov/div898/handbook/eda/section3/eda3668.htm http://www.mathpages.com/home/kmath122/kmath122.htm; http://www.weibull.com/basics/lifedata.htm.
[14] Hazard functions are used to forecast the probability of removal or failure for a measure, given that the measure has survived to the present.

were not provided for all measures due to the facilities staff not being aware of the date of removal or failure. To get around this issue, the following approach is followed:

1) Calculate replacements/failures for a subset of measures where the facilities staff provided survey code dates. These calculations are performed for each year where data is available (e.g., year 1 through year 10).
2) Calculate replacements/failures for *all* measures based on the results from step 1.

When working with measures where survey code dates were provided (step 1), the number of measures that were removed each year is calculated twice: once to include survey codes 3 through 11 and a second time to include survey codes 3 through 11 excluding survey code 5. The first analysis determines the total measures that were removed each year, whereas the second analysis excludes measures from failed businesses.

When calculating replacements/failures for *all* measures (step 2), the number of measures that were removed each year is again calculated twice: once to include all measures that were removed and a second time to exclude measures from failed businesses. The values are calculated using **Eq. C-1**.

Eq. C-1. $r(t) = r_s(t) \times \left[\frac{R}{R_s} \right]$

Where, $r_s(t)$ = removals/failures in year, t, for subset of measures with survey code dates.
R_s = total removals/failures for subset of measures with survey code dates.
R = total removals/failures for all measures in survey sample.

For example, of the 81 ovens removed/failed in the survey, only 54 had code dates. Therefore, R_s = 54 and R = 81. Excluding measures from failed businesses, 53 of the 67 removed/failed measures had code dates. For this analysis, R_s = 53 and R = 67. Of the ovens in the subset with survey code dates, only one oven was removed/failed in the first year, so $r_s(t)$ = 1.

Eq. C-1a. Removals/failures for all ovens in the first year = $1 \times \left[\frac{81}{54} \right] = 1.5$.

Eq. C-1b. Removals/failures excluding failed businesses = $1 \times \left[\frac{67}{53} \right] = 1.26$.

Hazard rates are calculated using **Eq. C-2**.

Eq. C-2. $h(t) = \frac{r(t)}{M(t)}$

Where, $r(t)$ = removals/failures in year, t.
$M(t)$= Measures at start of year.

For example, there were a total of 125 ovens at the beginning of year one. In **Eq. C-1a.** we calculated that 1.5 ovens were removed/failed within the first year. Therefore, **Eq. C-2a**. demonstrates a hazard rate of 1.20%. Excluding failed businesses, there were a total of 111 ovens at the beginning of year one. In **Eq. C-1b.** we calculated that 1.26 ovens were removed/failed within the first year. Therefore, **Eq. C-2a**. demonstrates a hazard rate of 1.14% for non-failed businesses in year one.

Eq. C-2a. Hazard rate for all ovens in the first year = $\left[\frac{1.5}{125}\right] = 0.012.$

Eq. C-2b. Hazard rate excluding failed businesses in year one = $\left[\frac{1.26}{111}\right] = 0.0114.$

The number of measures removed/failed and the corresponding hazard rate are calculated for each of the 10 years. See **Table C-1** for all ovens. See **Table C-2** for ovens excluding failed businesses.

Table C-1. Hazard Rates for All Ovens Based on Ovens with Survey Code Dates

All Ovens				Ovens with Survey Code Dates			
Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)	Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)
1	125.0	1.5	1.20%	1	98	1	1.02%
2	123.5	1.5	1.21%	2	97	1	1.03%
3	122.0	9.0	7.38%	3	96	6	6.25%
4	113.0	1.5	1.33%	4	90	1	1.11%
5	111.5	9.0	8.07%	5	89	6	6.74%
6	102.5	13.5	13.17%	6	83	9	10.84%
7	89.0	19.5	21.91%	7	74	13	17.57%
8	69.5	10.5	15.11%	8	61	7	11.48%
9	59.0	10.5	17.80%	9	54	7	12.96%
10	48.5	4.5	9.28%	10	47	3	6.38%
Total		81		Total		54	

Table C-2. Hazard Rates for Ovens Not Including Failed Businesses.

All Ovens Not Including Failed Businesses				Ovens with Survey Code Dates			
Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rates Used in the Analysis (Rate of Removal/Failure)	Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rate (Rate of Removal/Failure)
1	111.0	1.3	1.14%	1	97	1	1.03%
2	109.7	1.3	1.15%	2	96	1	1.04%
3	108.5	7.6	6.99%	3	95	6	6.32%
4	100.9	1.3	1.25%	4	89	1	1.12%
5	99.6	7.6	7.61%	5	88	6	6.82%
6	92.0	10.1	10.99%	6	82	8	9.76%
7	81.9	16.4	20.06%	7	74	13	17.57%
8	65.5	8.8	13.51%	8	61	7	11.48%
9	56.6	8.8	15.62%	9	54	7	12.96%
10	47.8	3.8	7.94%	10	47	3	6.38%
Total		67		Total		53	

Once this table information has been entered into an Excel spreadsheet, we can create a Hazard Chart of the data.

C.3 Estimation of Hazard Functions

In order to chart the hazard data, you must have the Excel Data Analysis Toolpak installed on your computer. Then, select Tools from the top menu bar and choose Data Analysis.

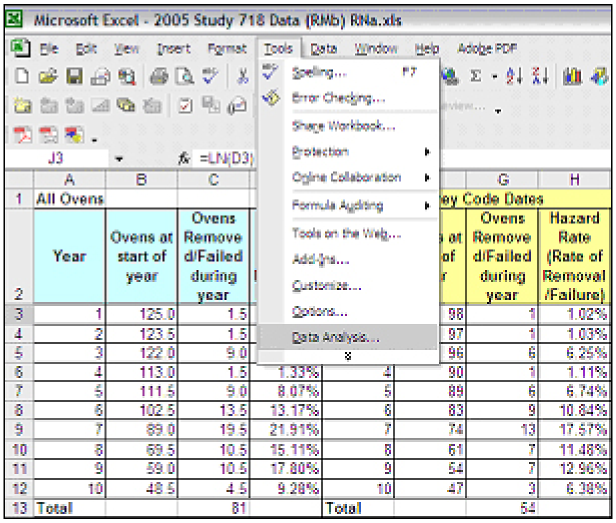


Figure C-1. Select Data Analysis

Select Regression and click OK.

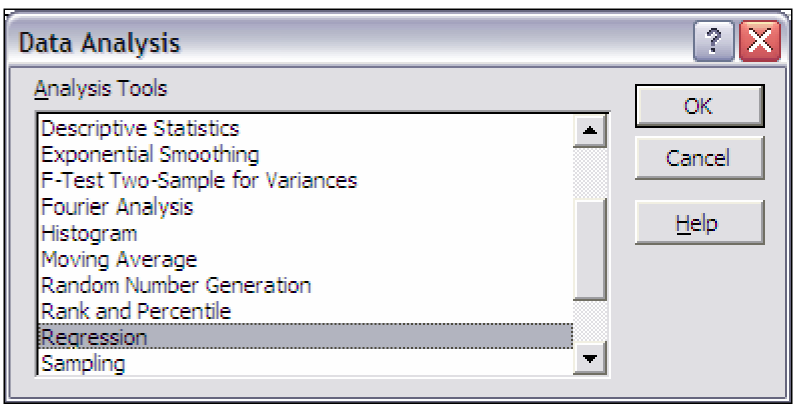


Figure C-2. Analysis Tools

For the Input Y Range, select cells D2 through D12 (column heading and hazard rate for years 1-10). For the Input X Range, select cells A2 through A12 (column heading and years 1-10). Then, check the Labels box, Confidence is Zero box, and the Confidence Level box. Set the Confidence Level to 80%. In the Residuals section, check the Residuals box and the Line Fits Plots box. See **Figure C-3**. Click OK to produce the hazard chart and output information.

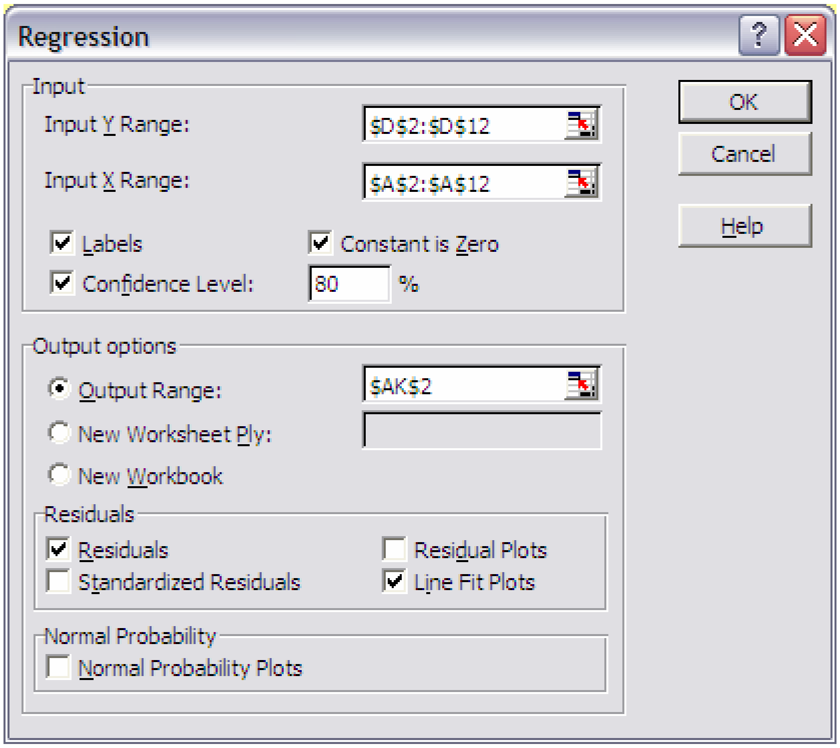


Figure C-3. Regression Options

A hazard rate function chart similar to **Figure C-4** will be produced.

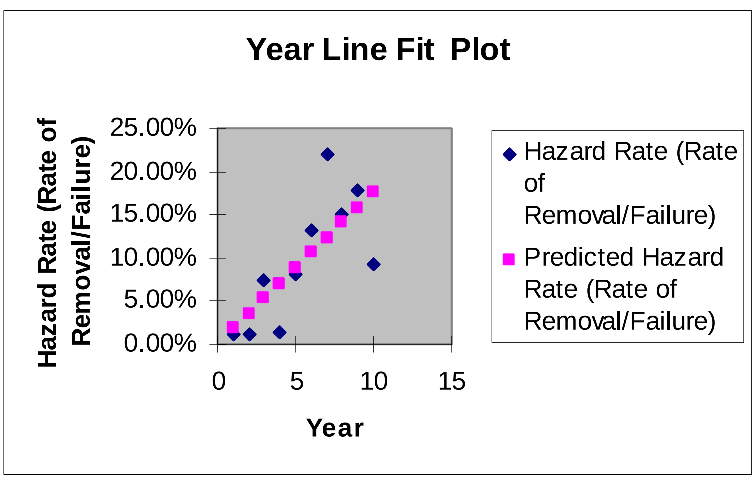


Figure C-4. Hazard Rate Function Chart

To reformat the hazard chart, perform the following steps:

- Right-click on the chart and select Chart Type.
  - Choose XY (Scatter) and click OK.
- Right-click on the chart and select Source Data.
  - Click on the Series tab.

    - In the series scroll-box, select Predicted Hazard Rate and click on the Remove button.
    - Click OK.
- Right-click on the chart and select Chart Options.
    - Click on the Gridlines tab.
    - Place a mark in the Major Gridlines box.
    - Click on the Legends tab.
    - De-select the Show legend box if necessary.
    - Click OK.
- Right-click on the x-axis of the chart and select Format Axis.
    - Click on the Scale tab.
    - Set Minimum to 0.
    - Set Maximum to 10.
    - Set Major Unit to 1.
    - Set Minor Unit to 1.
    - Set Value (Y) axis Crosses at to 0.
    - Click OK.
- Right-click on the y-axis of the chart and select Format Axis to reformat.
- Right-click on a data point on the chart and select Add Trendline.
    - Click on the Type tab.
    - Select Power.
    - Click on the Options tab.
    - Place a mark in the Display Equation on Chart box.
    - Place a mark in the Display R-Squared Value on Chart box.
    - Click OK.

The hazard rate function chart should now look similar to **Figure C-5**.

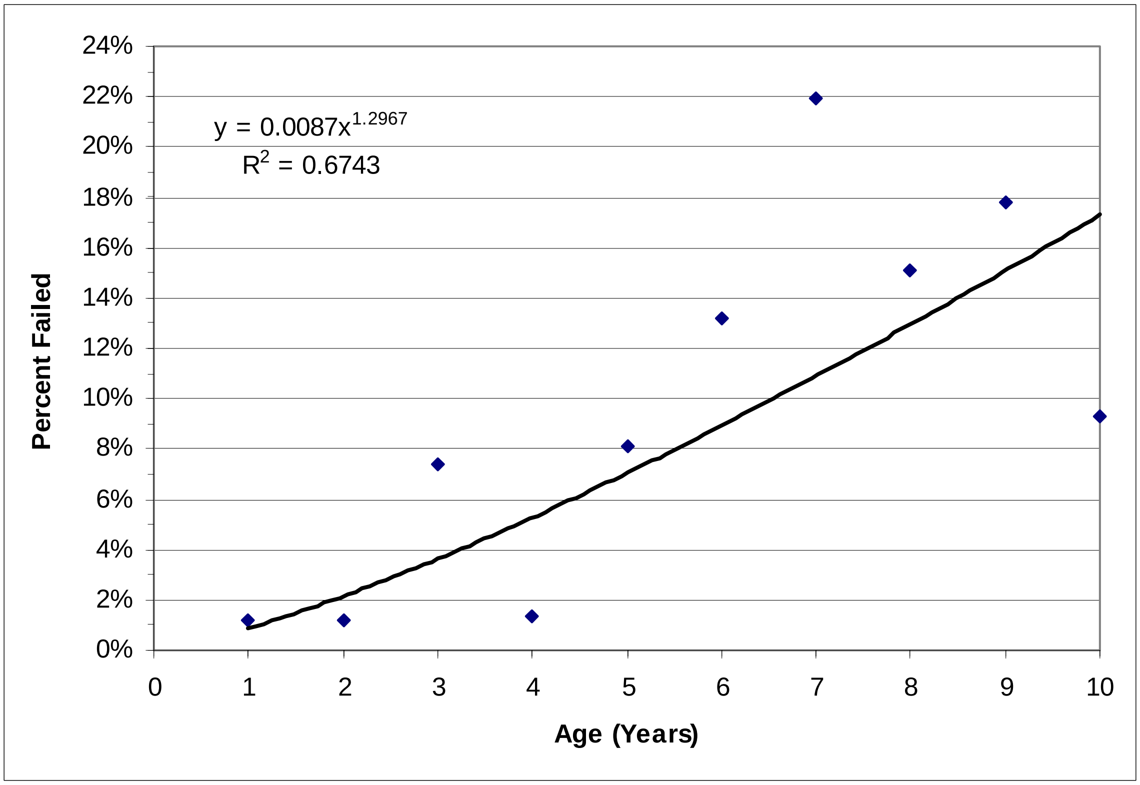


Figure C-5. Hazard Rate Function Chart

The equation and R-squared value displayed on the hazard rate function chart represent the power curve fit. For the oven example above a = 0.0087, B = 1.2967, and R-squared = 0.6743. Power curve fits to the hazard rate data provide estimates of the parameters for the Weibull distribution representation of the hazard rate function.

C.4 Determination of Weibull Distribution Functions and Effective Useful Life

Calculated hazard rates for measures vary over time. This indicates that there is no reason to assume the survival function can be represented using the exponential distribution, since the hazard rate for an exponential survival function is constant. However, the Weibull distribution does allow for hazard rates that increase over time, and the Weibull-based hazard function is therefore used as the functional form for estimating the hazard functions.

Before we can produce a Weibull distribution, we must first calculate the logarithmic functions of the hazard rate and year. The values are calculated using **Eq. C-3** and **Eq. C-4**.

Eq. C-3. $f(h(t)) = LN(h(t))$

Where, h(t) = hazard rate in year, t.

Eq. C-4. $f(t) = LN(t)$

Where, t = year number (1 through 10).

For example, to analyze the hazard rate logarithmic function of year one we calculate LN(1.20%). Likewise, we use LN(1) to calculate the logarithmic function of the year for year one.

	A	B	C	D	E	F	G	H	I	J	K
1	All Ovens				Ovens with Survey Code Dates						
2	Year	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rate (Rate of Removal /Failure)	Age (Years)	Ovens at start of year	Ovens Removed/Failed during year	Hazard Rate (Rate of Removal /Failure)		LnHR	LnYears
3	1	125.0	1.5	1.20%	1	98	1	1.02%		-4.42285	0
4	2	123.5	1.5	1.21%	2	97	1	1.03%		-4.41078	0.693147
5	3	122.0	9.0	7.38%	3	96	6	6.25%		-2.6068	1.098612
6	4	113.0	1.5	1.33%	4	90	1	1.11%		-4.32192	1.386294
7	5	111.5	9.0	8.07%	5	89	6	6.74%		-2.5168	1.609438
8	6	102.5	13.5	13.17%	6	83	9	10.84%		-2.02717	1.791759
9	7	89.0	19.5	21.91%	7	74	13	17.57%		-1.51822	1.94591
10	8	69.5	10.5	15.11%	8	61	7	11.48%		-1.88995	2.079442
11	9	59.0	10.5	17.80%	9	54	7	12.96%		-1.72616	2.197225
12	10	48.5	4.5	9.28%	10	47	3	6.38%		-2.37749	2.302585
13	Total		81		Total		54				

Figure C-6. Logarithmic Functions

Again, in order to produce the Weibull distribution, you must have the Excel Data Analysis Toolpak installed on your computer. Then, select Tools from the top menu bar and choose Data Analysis. Select Regression and click OK.

For the Input Y Range, select cells J2 through J12 (column heading and logarithmic function of hazard rate for years 1-10). For the Input X Range, select cells K2 through K12 (column heading and logarithmic function of years 1-10). Then, check the Labels box and the Confidence Level box. Set the Confidence Level to 80%. In the Residuals section, check the Residuals box and the Line Fits Plots box. See **Figure C-7**. Click OK to produce the regression output information used to calculate the Weibull distribution.

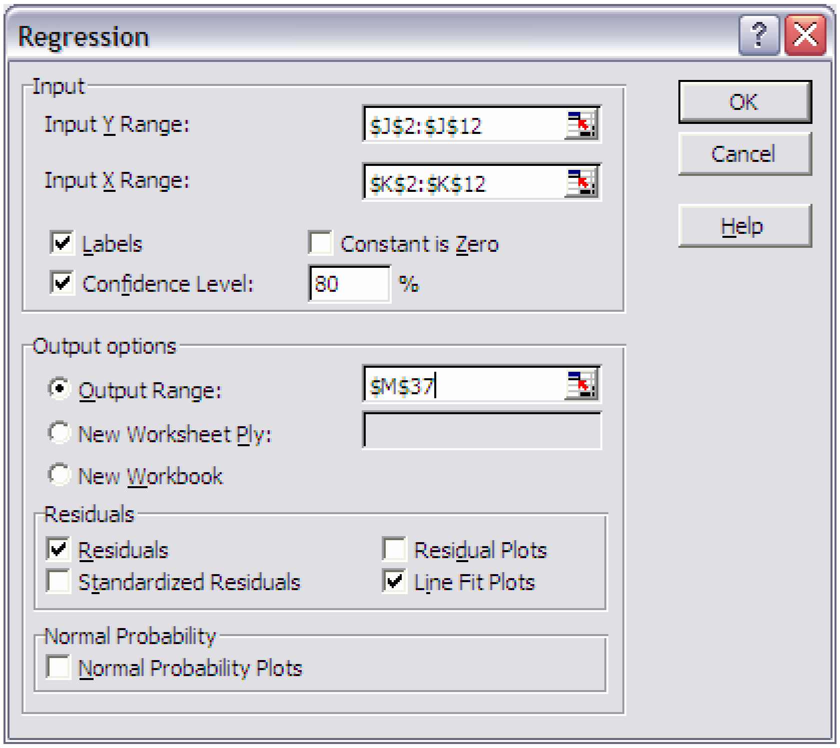


Figure C-7. Regression Options

A regression similar to **Figure C-8** will be produced.

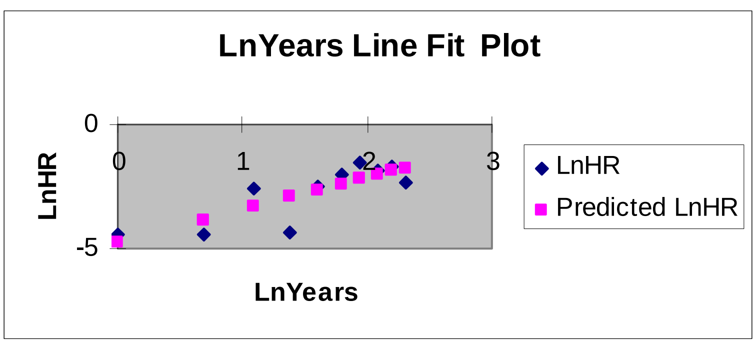


Figure C-8. Regression Chart

We will reformat this regression into Weibull distribution survival functions after performing more calculations.

Based on the Anova (ANalysis Of VAriance) output calculated when we produced the regression, we can determine the Weibull distribution parameters alpha and beta. Beta (β), the estimated shape parameter, is calculated by adding 1 to the LnYears Coefficient. Alpha (α), the estimated scale parameter, is calculated by taking e to the power of the Intercept Coefficient, divided by β. Similarly, we calculate the upper bound β value by adding 1 to the LnYears Lower 80% value. We calculate the lower bound β value by adding 1 to the LnYears Upper 80% value. Likewise, we calculate the upper bound α value by taking e to the power of the

Lower 80% Intercept, divided by the upper bound β. We calculate the lower bound α value by taking e to the power of the Upper 80% Intercept, divided by the lower bound β. **Table C-3** shows the alpha and beta calculations for ovens.

Table C-3. Weibull Distribution Alpha (Scale) and Beta (Shape) for Ovens

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 80.0%*	*Upper 80.0%*
Intercept	-4.740391268	0.529820364	-8.947166984	1.93505E-05	-5.480452621	-4.000329916
LnYears	1.296692183	0.318624498	4.069656261	0.003585128	0.851632512	1.741751854
Alpha (Scale)	0.003803395				0.002250686	0.006678065
Beta (Shape)	2.296692183				1.851632512	2.741751854

To calculate the 80% Lower Bound, Survival Function, and 80% Upper Bound plots for the Weibull distribution charts, use **Eq. C-5**.

Eq. C-5. $v^{\wedge} = EXP(-\alpha * t^{\beta})$

Where, α = alpha, the estimated scale parameter.
t = time, years 1-15.
β = beta, the estimated shape parameter.

The **Table C-4** shows the Weibull survival function for ovens with lower and upper bounds.

Table C-4. Weibull Distribution Survival Function for All Ovens

Time	Lower 80% Bound	Survival Function	Upper 80% Bound
1	0.9933	0.9978	0.9962
2	0.9563	0.9919	0.9815
3	0.8730	0.9829	0.9537
4	0.7417	0.9711	0.9123
5	0.5764	0.9567	0.8579
6	0.4033	0.9398	0.7922
7	0.2501	0.9207	0.7175
8	0.1355	0.8996	0.6369
9	0.0633	0.8767	0.5536
10	0.0251	0.8522	0.4709
11	0.0084	0.8263	0.3916
12	0.0023	0.7992	0.3183
13	0.0005	0.7711	0.2526
14	0.0001	0.7421	0.1957
15	0.0000	0.7126	0.1479

To reformat the regression figure (**Figure C-8**), perform the following steps:

- Right-click on the chart and select Chart Type.
  - Choose XY (Scatter) and click OK.
- Right-click on the chart and select Source Data.
  - Click on the Series tab.
  - In the series scroll-box, select LnHR and click on the Remove button.

- In the series scroll-box, select Predicted LnHR and click on the Remove button.
- Click on the Add button to add the Upper 80% Bound data series we just produced.
  - Click in the Name box, and then click on the table heading Upper 80% Bound in the spreadsheet.
  - Click in the X Values box, and then select the 15 table values under the Time heading in the spreadsheet.
  - Click in the Y Values box, and then select the 15 table values under the Upper 80% Bound heading in the spreadsheet.
- Click on the Add button to add the Survival Function data series we just produced.
  - Click in the Name box, and then click on the table heading Survival Function in the spreadsheet.
  - Click in the X Values box, and then select the 15 table values under the Time heading in the spreadsheet.
  - Click in the Y Values box, and then select the 15 table values under the Survival Function heading in the spreadsheet.
- Click on the Add button to add the Lower 80% Bound data series we just produced.
  - Click in the Name box, and then click on the table heading Lower 80% Bound in the spreadsheet.
  - Click in the X Values box, and then select the 15 table values under the Time heading in the spreadsheet.
  - Click in the Y Values box, and then select the 15 table values under the Lower 80% Bound heading in the spreadsheet.
- Click OK.

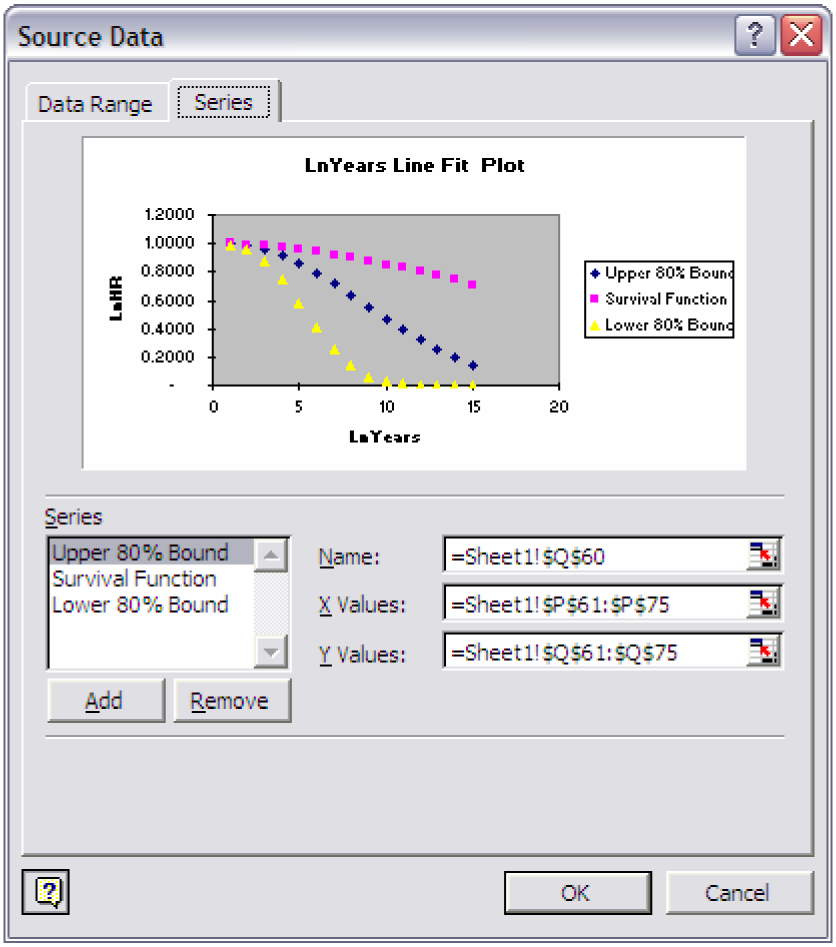


Figure C-9. Source Data

- Right-click on the chart and select Chart Options.
  - Click on the Gridlines tab.
  - Place a mark in the Major Gridlines box.
  - Click OK.
- Right-click on the x-axis of the chart and select Format Axis.
  - Click on the Scale tab.
    - Set Minimum to 0.
    - Set Maximum to 15.
    - Set Major Unit to 1.
    - Set Minor Unit to 1.
    - Set Value (Y) axis Crosses at to 0.
    - Click OK.
- Right-click on the y-axis of the chart and select Format Axis.
  - Click on the Number tab and select Percentage.
  - Click on the Scale tab.
    - Set Minimum to 0.
    - Set Maximum to 1.
    - Set Major Unit to 0.1.
    - Set Minor Unit to 0.1.

- Set Value (Y) axis Crosses at to 0.
- Click OK.

The resulting Weibull distribution survival function plot is shown in **Figure C-10**.

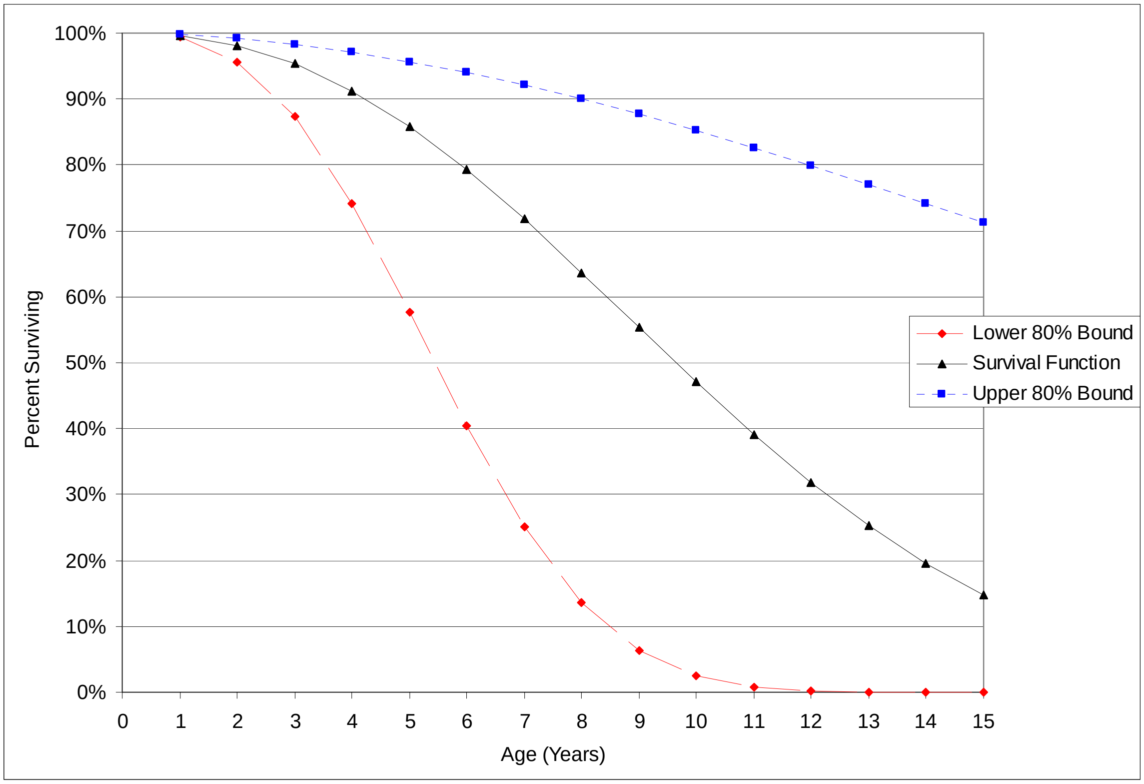


Figure C-10. Weibull Distribution Survival Function Plot for All Ovens

Next we calculate the estimated median life. Median survival time is defined as that age where 50% of the appliances have been removed or failed. Again, we calculate the 80% Lower Bound, Estimated EUL, and 80% Upper Bound using **Eq. C-6**.

Eq. C-6. $\breve{T} = ((1/\alpha) * LN(2))^{(1/\beta)}$

Where, α = alpha, the estimated scale parameter.
β = beta, the estimated shape parameter.

For example, the following calculations are used to determine the 80% Lower Bound, Estimated EUL, and 80% Upper Bound for ovens:

Eq. C-6a. $80\%\ \text{Lower Bound Median Survival Time} = ((1/0.0067) * LN(2))^{(1/2.7418)} = 5.4$

Eq. C-6b. $\text{Estimated EUL Median Survival Time} = ((1/0.0038) * LN(2))^{(1/2.2967)} = 9.6$

Eq. C-6c. $80\%\ \text{Upper Bound Median Survival Time} = ((1/0.0023) * LN(2))^{(1/1.8516)} = 22.1$

These equations are used to estimate the measure effective useful life for ovens in **Table C-5**. While sufficient data was available to estimate a median EUL of 10.1 years for ovens, the "null hypothesis" cannot be rejected for ovens since the *ex ante* EUL is within the 80 percent effective useful life lower and upper bounds of 5.4 and 22.1 years.

Table C-5. Estimated Effective Useful Life (EUL) for Ovens

Measure	SoCalGas Ex Ante Estimated EUL	Estimated Median Life			Ex Ante Different From Ex Post at 80% Confidence Level ± 20 Percent?
		80% Lower bound	Estimated EUL	80% Upper Bound	
Ovens					
All Ovens removed/failed	12	5.4	9.6	22.1	No