Final Report Study No. 720A

Ninth Year Measure Retention Study of the 1996 Commercial EEI Program

Submitted to
Southern California Gas Company

January 2006

Final Report Study No. 720A

Ninth Year Measure Retention Study of the 1996 Commercial EEI Program

Submitted to
Southern California Gas Company

January 2006

by

ASW Engineering Management Consultants, Inc.
David M. Wylie, P.E.
Chris Baginski, MSME
2512 Chambers Road, Suite 103
Tustin, CA 92780-6950
Telephone: 714/731-8193
Fax: 714/731-1921
E-Mail: cbaginski@aswengineering.com

and

Megdal & Associates
Dr. Lori Megdal
198 High Street
Acton, MA 01720-4218
Telephone: 978/461-3978
Fax: 978/461-3979
E-mail: megdal@verizon.net

Table of Contents

Executive Summary

Background and Study Method

The California Demand Side Management Advisory Committee (CADMAC) measurement and evaluation (M&E) Protocols require Retention Studies at specific retention years depending on the program. The purpose of the Retention Study is to collect data to determine the retention and effective useful life (EUL) for the primary measures in the program. This involves measuring the proportion of measures still in place, and operational. The retention information along with considerations of time since program participation provides the basis for development of the *ex post* EUL. The *ex post* EUL is then statistically compared with *the ex ante* EUL.

This study is the Measure Retention Study for the 1996 Commercial Energy Efficiency Incentives (EEI) Program operated by Southern California Gas Company (SoCalGas). This report includes the tables required by the M&E Protocols.

An examination of SoCalGas' Commercial EEI program database presented clear information with regard to what measures were to be examined in accordance with protocol requirements. Cooking measures alone constituted 75% of expected therm savings. This means that the only measures to be examined in the retention study are the cooking measures.

A stratified sampling plan was utilized to ensure the sample obtained the highest precision based upon expected energy savings while also ensuring representation of the smaller saving sites. The initial sampling was then winnowed to those sites that were part of the first year load impact study, as required by the Protocols for the SoCalGas retention studies. A total of 200 sites were visited to determine the retention of the cooking measures indicated in the program database.

The primary retention measurement is the proportion of measures that are in place and operational. This is derived from survey information by analyzing weighted proportions of the site visit data. The Effective Useful Life (EUL) analysis came from calculating the expected median from fitted values of a logistic function survival analysis based upon duration and duration-squared.

Findings

The savings weighted retention at nine years is 56.5 percent (adjusting for the stratified random sampling). Participants participated throughout 1996. Retention measurement occurred in August 2005. The 9-year retention rate is approaching the 50% level, the time period when half of the savings weighted equipment is remaining and operational.

The *ex post* EUL was estimated through fitted values on the results from a survival analysis with a logistic model based upon duration in years and duration-squared. This model controlled for right-censored data, had algorithm convergence, a good log likelihood statistics, very high Wald chi-square measurements for both independent variables, and produced reasonable fitted values. The EUL estimate and its confidence intervals are presented in Table ES.1.

Table ES.1[1] EUL Estimates

		Logistic Survival Analysis Model on Duration & Duration-Squared
9-year retention rate	56.5%	
Median value (fitted value where p=50%)		**9.5 years**
Confidence interval* (lower)		8.5 years
Confidence interval* (higher)		10.5 years

* Calculated at 80% confidence level.

The *ex ante* EUL for cooking measures for Southern California Gas Company's Commercial EEI Program is 12 years. Twelve (12) years does not fall within the confidence interval for the Effective Useful Life (EUL) estimate from this study. **The *ex post* EUL is a more accurate measure of the EUL of the 1996 Commercial EEI Program's cooking measures as of this 9th year retention**

[1] This table is the same as Table 3.2 and is further described in Section 3.

study. The *ex ante* EUL is 12 years but the 9-year retention measurement is 56.5% and the *ex post* EUL is 9.5 years.

1.0 Introduction

1.1 Project Background

Standardized protocols for demand-side management (DSM) evaluation were developed in California through the cooperative efforts of utility DSM evaluation experts, interested parties, regulatory staff, and outside consultants working through the California Demand Side Management Advisory Committee (CADMAC). These measurement and evaluation (M&E) protocols are the standardized expectations for DSM evaluation which serve as the basis for the measurement of *ex post* energy savings caused by energy efficiency programs, whose measurement determines the shareholder incentives to be received by the utility due to the utility's performance in obtaining these savings.

The M&E Protocols' require Retention Studies at a specified number of years after the program year depending on the program. This document presents the ninth (9^{th}) year retention study of Southern California Gas Company's (SoCalGas) Commercial Energy Efficiency Incentives (EEI) Program of 1996 as stipulated in Table 8B of the M&E Protocols.[2] The Retention Study must collect data to determine the proportion of measures that are in place and operational. This is derived from survey information by analyzing frequencies and means of the site visit data by measure. This involves measuring the proportion of measures still in place, operational, and effective. The retention information, along with considerations of time since program participation, provides the basis for development of the *ex post* EUL. The *ex post* EUL is then statistically compared with the *ex ante* EUL at an 80% confidence level.[3]

[2] *Protocols, and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management Programs*, 1999 Version, page 26.

[3] Ibid, Table 6, page 17.

1.2 Report Overview

Section 1 has provided an overview of the project. Section 2 presents the methodology of the study, to include the program information that determined the selection of the cooking measures as the measures to be examined in this study. The last section, Section 3, presents the study findings including information on the sample, the measure retention estimate, the effective useful life examination (EUL), and responses on why equipment was removed and satisfaction with the efficient cooking equipment. Section 3.3 presents a summary of the documentation protocols as required in Table 7, and the reporting protocols as required in Table 6 of the revised M&E Protocols. Section 3.4 presents the other findings on why cooking measures were removed and the satisfaction levels reported for the 1996 program-installed efficient cooking measures. Two appendices follow the body of the report. Appendix A contains a copy of the site visit instrument. Appendix B presents the datasets and documentation for the study (in accordance with the M&E Protocols).

2.0 Methodology

2.1 Measurement Issues

Criteria for Measure(s) Included in Retention Study

The M&E Protocols provide explicit direction as to the criteria to be used for determining what measures are to be examined in the required retention studies. This is as follows:

> *"The utility should select the top ten measures, excluding measures that have been identified as miscellaneous (per Table C-9), ranked by net resource value or the number of measures that constitutes the first 50% of the estimated resource value, whichever number of measures is less."*[4]

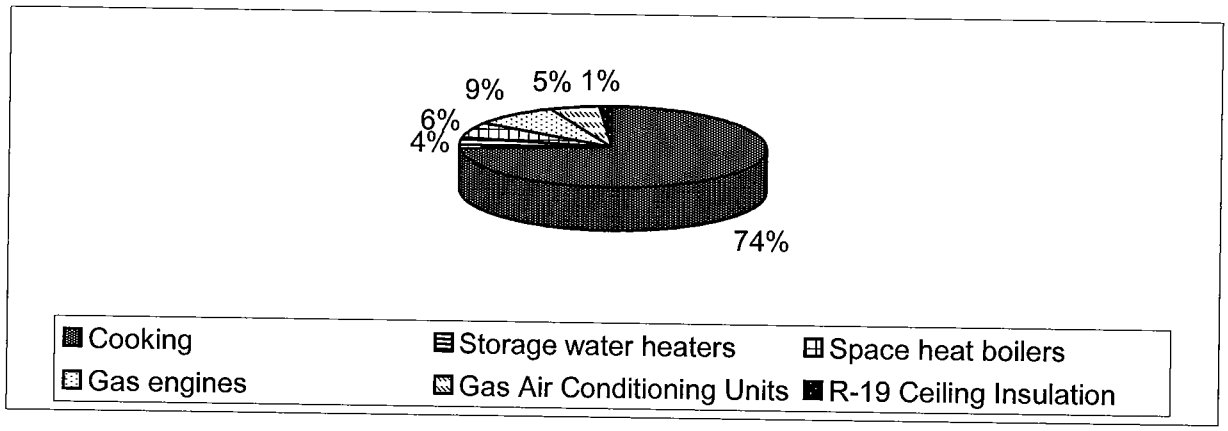
An examination of SoCalGas' Commercial EEI program database presents clear information with regard to these protocol requirements for the measures(s) to be studied. Cooking measures alone constitute 75% of expected therm savings, as shown in Table 2.1 and Figure 2.1. This means that the cooking measures are the only measures for which this retention study will be conducted.

Table 2.1 Distribution of Measures within SoCalGas' 1996 Commercial EEI Program

	% of therm savings
Cooking	74%
Storage water heaters	4%
Space heat boilers	6%
Gas engines	9%
Gas Air Conditioning Units	5%
R-19 Ceiling Insulation	1%
Total	100%

[4] Ibid, Table 9B, page 27, General comment section.

Figure 2.1 Expected Therm Savings Per Measure



The cooking equipment in SoCalGas' 1996 Commercial EEI Program included the following high efficiency gas cooking equipment:

- ranges,
- fryers,
- ovens,
- pizza ovens,
- griddles,
- Chinese ranges,
- Salamanders,
- steamers,
- rotisseries,
- broilers,
- pressure cookers,
- stockpots,
- donut fryers,
- steam tables, and
- cheesemelters.

Retention Measurement

One of the primary objectives of this study was to answer the questions: "Is the measure still in place?; Is it operational?; and Is it still effective?". This is in accordance with the M&E Protocols' definition of a Measure Retention Study:

> *"An assessment of (a) length of time the measure(s) installed during the program year are maintained in operating condition; and (b) the extent to which there has been a significant reduction in the effectiveness of the measure(s)."*[5]

The methodology selected was based upon these needs, understanding the differences between a measure retention study and a persistence study, and developing a workable methodology for conducting 200 site visits to gather the data to answer this question.

This study was designed only as a measure retention study and not a persistence study. Only a few practitioners with significant experience in conducting persistence studies understand the differences between these two types of studies. One of the primary differences after the studies are conducted lies in their acceptable uses. Given that this study is a measure retention study, the results should only be used as a measure retention study (unless further adjustments and examinations are made).

An example of an improper use of a measure retention study would be to use its results along with prior impact evaluation. This improper use of the retention results could yield a double-counting of losses. As an example, suppose a program database indicated that 100 low flow showerheads should have been installed. Then an impact evaluation is conducted one year post-participation. This impact evaluation finds 97 showerheads installed (or implicitly accounts for this loss in a lower realization rate in a billing analysis such as a 97% realization rate).

[5] Ibid. Measure Retention Study definition from page A-9.

Then suppose two years later a retention study is done and finds 90 showerheads in place and operational. If the study were conducted as a measure retention study only, using as its baseline the program database, the retention study would find a loss of 10 showerheads (100-90) or a 90% retention. This could be an accurate measure retention estimate.

However, if the retention study results were applied to the impact evaluation's savings to estimate savings still being achieved, there would be a double-count of the 3% loss. The persistence retention rate would need to be re-estimated as 93% (90/97) in order to be applied to the impact savings estimate. Of the 10 showerheads not in place at the time of the retention study, three are in the program database but were never actually installed and seven were the retention loss in the form of persistence from the impact evaluation.

Another alternative way to look at this issue can be seen in how this retention study is performed and how its information can be used. The M&E Protocols specifically state that for SoCalGas the measurement retention studies will be based upon sub-samples from the first year impact analyses and annual load impact studies.[6] This study was designed to meet this requirement, as discussed below in the section on sampling.

This means that this retention study is directly applicable to the first year impact assessment as it includes savings estimates (effectiveness). However, as retention is still measured from the original database, the double-count possibility still exists. (The first year impact study for SoCalGas' Commercial EEI Program was based upon a telephone survey and billing analysis. There were no on-site visits, metering or monitoring. First year retention estimates were not explicitly provided nor a database of retained measures after the first year.)

As this study is a measure retention study, and not a persistence study, it did not gather data on usage, as doing so could cause confusion to readers of the report.

[6] M&E Protocols (ibid), pg. 28.

Effective Useful Life Measurement

The second primary objective of this study is to estimate the *ex post* effective useful life (EUL) and compare this to the *ex ante* EUL. The M&E Protocols define effective useful life (EUL) as:

> *"An estimate of the median number of years that the measures installed under the program are still in place and operable."*[7]

This means we need to use the average retention rate at each follow-up point to estimate when half (50%) of the measures are still in place and operable and half are not.

2.2 Survey Instrument

The site visits and survey instruments were kept simple for easy and quick data collection, minimizing potential data errors, and ensuring high quality customer service. The instrument was drafted to allow each instrument to contain the specific customer information on the site and the number and types of cooking measures being examined for the retention study. As part of interacting with customers on-site, participants were also asked their satisfaction level with the cooking measures. The draft instrument was reviewed by the Utility Study Manager and then finalized. A copy of the site survey instrument is provided as Appendix A.

2.3 Sampling

There are a total of 735 sites in SoCalGas' 1996 Commercial EEI Program that received incentives for 1,309 pieces of cooking equipment. This resulted in an anticipated savings of 1,584,165 therms per year.

We conducted site surveys of 200 sites, 27% of the participating sites and collected data on the retention of all program-installed cooking measures at that site. A stratified sample for heavier expected savers was drawn (as further described below). Assessing all cooking measures installed at these 200 sites raises the retention sample size to 401.

[7] Ibid, pg. A-1.

In order to maximize the accuracy of the retention study, these 200 site visits were selected through a stratified random sampling process based upon estimated therm savings. The sampling strata were designed based upon the program site data and expected therm savings from installed cooking measures. The stratum information is provided in Table 2.2.

Table 2.2 Sampling Strata Design from Program Site Data of Cooking Measures

Stratum	Msr_Ids	Min.	Max	Avg.	Therms	% therms	% in Sampling Pool	Sampling Pool
1	42	5,111	16,817	8,459	355,279	22%	100%	42
2	212	2,012	4,892	2,937	622,578	39%	100%	212
3	222	1,001	1,987	1,395	309,774	19%	68%	151
4	634	38	998	493	312,316	20%	24%	152
	1,110							

Stratified random sampling was conducted based upon these strata. These initial sampling datasets from the program data were then compared to the list of sites (by premise identifier) included in the first year impact study. These constituted the sampling pool for the 4th year retention study. This same sampling pool was then used for this 9th year retention study.

2.4 Final Sample

Site visits were made by qualified ASW auditors for the selected sampling sites. Given the very short and easy nature of the site survey instrument once on-site, it was decided that the best customer service, quickest, and most cost-effective method was direct site visits (at non-meal times given we would be looking at cooking equipment often in restaurants) without prior scheduling (which would have taken the customer longer with headquarters contact and site scheduling than the actual site visit took). This procedure produces an almost 100% sampling hit rate and, therefore, eliminates potential sampling bias.

We conducted site surveys of 200 sites, 27% of the participating sites and collected data on the retention of all program-installed cooking measures at that site. Assessing all cooking measures installed at the 200 stratified sample sites raises the retention sample size to 401.

The final site sample distribution by the sampling strata is provided in Table 2.3. This table also provides the stratum weighting for the analysis given the final sampling distribution.

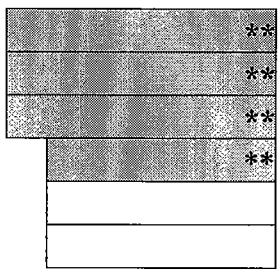
Table 2.3 Sampling Strata Design from Program Site Data of Cooking Measures

Stratum	Customers	Therms	% of Sample (a)	Strata % of therms for population (b)	Weights (b)/(a)
1	28	268,952.88	46%	22%	0.4784948
2	42	155,747.76	27%	39%	1.4647839
3	64	125,037.44	21%	19%	0.8888824
4	66	35,228.12	6%	20%	3.3210186
	200	584,966.20			

2.5 Analysis for Retention Estimates

Cooking equipment is a very straightforward measure for a retention study. Generally, the equipment can not be partially removed. It is either there and operational, there and non-operational, or removed. Therefore, the site survey collected this information for the cooking measures installed. This data was gathered as shown in Figure 2.2, an excerpt from the site survey instrument provided in Appendix A.

Figure 2.2 How Retention Data Was Collected



** Shaded areas provided by computerized read of sampling database.

2.6 Effective Useful Life Analysis

The purpose of the retention study is to develop an *ex post* Effective Useful Life (EUL) estimate and compare this with the *ex ante* EUL. According to the M&E Protocols, Table 10, this comparison is made at the 80% confidence level.

"1) For measures selected for 3rd/4th year retention studies: The estimated ex post measure effective useful life that results from the retention study will be compared to the ex ante (i.e., forecast) effective useful life estimates. Hypothesis testing procedures will be used to determine if the estimated ex post measure effective useful life is statistically significantly different from the ex ante measure effective useful life. If the estimated ex post measure effective useful life is significantly different than the ex ante measure effective useful life, the estimated ex post measure effective useful life will be used to recalculate the Resource Benefits, net. Otherwise, the Resource Benefits, net estimate will continue to use the ex ante measure effective useful life. Hypothesis testing will be conducted at the 20% significance level.

An equivalent representation is to construct 80% confidence intervals around the estimated ex post measure effective useful life. If the ex ante measure effective useful life estimate is within the constructed confidence interval, then the Resource Benefits, net calculation will continue to use the ex ante measure effective useful life. If the ex ante measure effective useful life estimate is outside the constructed confidence interval, the estimated ex post measure effective useful life will be used to recalculate the Resource Benefits, net."[8]

Using the post-impact data and the site visit instrument provides at least two points of data. All equipment is installed and operational in 1996 that is in the sample. Then the 9th year site visit data provides which of this equipment is still in place and operational.

The experienced site auditors also talked to the customers about why equipment had been removed. Within this discussion, the auditors made notes of when equipment had been removed. This provided more specific retention dates for these cases. The analysis dataset was then also compared with the 2001 4th year retention study. This allowed annotation of whether equipment had been removed in 2001. Between this data and the discussions the auditors had with customers, some of the sites had data of whether the equipment was removed earlier or later than 2001. All of this information was used to create a dataset where a year of removal estimate was derived. This helps provide more than the two time points of data that would have been provided by only using the original and 9th year site visit inspection.

[8] M&E Protocols, Table 10, pg. 31.

The *ex ante* EULs were developed using engineering experience and information. Engineering observations were that energy efficiency measures generally lasted a certain length of time and then rapid losses would occur as the measures reached their expected lifetime. A few measures then continue to last significantly beyond their expected lifetime.

An initial approximation for most types of forecasting efforts includes a linear estimate. This involves trying to fit a line to the observed data and use this to predict future estimates. Yet, the engineering experience for efficiency measures suggests that a linear model would not well represent the survival function of energy efficiency measures.

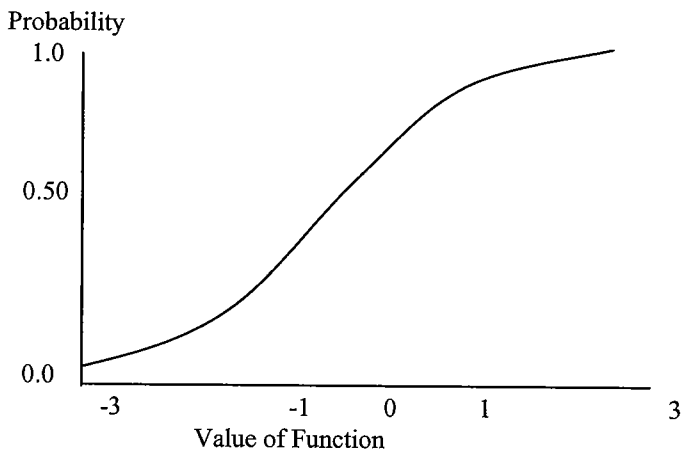
Common alternative models include logistic and exponential models. A variation of the logistic function can be used to describe the pattern of little loss in the early years and increasing loss (hazard probability) as the measure approaches its expected life, with a flattening hazard occurring thereafter.

The standard cumulative logistic probability function is:

$$Pi = F(Z_i) = F(\alpha + \beta X_i) = 1/(1 + e^{-(\alpha + \beta X_i)})$$

The logistic model is generally used to measure and predict probabilities that an event will occur. This model limits the end points to zero and one. The cumulative logistic, the logit model, looks like the curve shown in Figure 2.3.

Figure 2.3 Cumulative Logistic Function

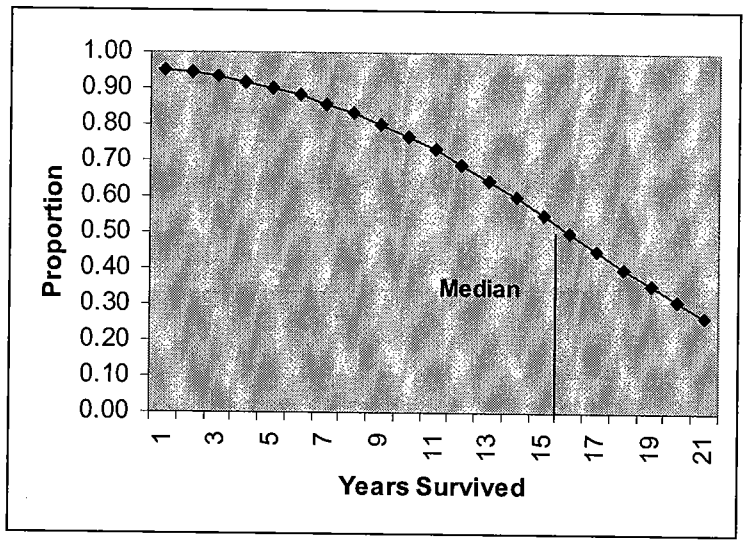


The logistic function that best fits the engineering observations described above relies upon a logistic function of time and the EUL. This is:

$$F(Z_i) = 1-[1/(1 + e^{-(t + EUL)\,b})$$

With the survival function as in Figure 2.4.

Figure 2.4 Logistic Survival Function with EUL=15 and b=0.2



Many energy efficiency retention studies examine energy efficiency equipment as being either there or not. This dichotomous scale allows the possibility of using classical survival analysis techniques. These techniques originated in the medical field where the concern was for mortality or whether someone contracted the studied disease. These outcomes are dichotomous, they either occur or not and can be measured as zero or one events.

Classic survival analysis is generally considered the "gold-standard" in measuring survival functions. This form of regression analysis is specifically designed to account for the fact that occurrence of non-survival might take place some time after when data is measured. These cases are those that are currently retained and are missing measurement for when they would no longer be operational. These cases are said to be "right censored" as we do not have the longer time period data that would say when they would fail.

In these cases, given enough data, many model forms can be tested with available survival analysis statistical programs. The regression techniques available allow consideration of censored data and can handle continuous

time data, discrete time data, and time-dependent covariates. These models, techniques, and SAS procedures for this modeling are well described in *Survival Analysis Using the SAS® System: A Practical Guide* by Dr. Paul D. Allison, SAS® Institute, 1995.

These techniques often require at least a minimum number of failures. Many of the functional forms also require measurement to include a significant proportion of the time between installation and the expected median life to be part of the dataset in order to achieve an accurate model. This is because significant changes in the slope occur around this point.

Early retention studies in energy efficiency programs have seldom had enough failure data to obtain a solving model (obtaining convergence) or reasonable estimates. The 4th year retention study for this program tested classical survival analysis but was unable to obtain model convergence.

The *ex ante* estimate for SoCalGas' cooking measures in the Commercial EEI Program is 12 years. This is the 9th year retention study on an expected 12 year measure. The 9-year site visits found a retention of 56.5%. This combined with greater time variance provided with the comparison with the 4th year retention data and the inquiries of year of removal by auditors have provided a dataset that did allow model convergence for a survival analysis model that obtained reasonable results that matched well to the data obtained and in a defensible model form and function that fits well with retention theory. This study was then able to meet the "gold standard" of using classical survival analysis.

Simpler analysis of the data, a linear model, and an exponential simple model were also examined. These estimates provided additional evidence that the results from the model form and function for the classical survival analysis method selected are as would be expected. This adds additional support for the findings obtained.

3.0 Findings and Results

3.1 Retention Findings

The first step for final sample was removing the customer identifiers and removing unnecessary variables to create the final sample for documentation (retaining customer confidentiality). It is from this point on that all the datasets and analysis efforts (spreadsheets and SAS® programs) have been provided within the documentation and datasets for this project. Appendix B contains all the datasets, documentation, and elements of this analysis as required by the M&E Protocols.

The SAS customer dataset was transformed in SAS to a measure dataset, with one observation per measure type rather than one observation per customer with data for multiple measure types.

A weighted average retention rate for cooking measures was obtained. (The weighting ensured that the analysis of the sample included customer types in the same proportion as in the program population.) The 9^{th} year retention is 56.5%.

3.2 Effective Useful Life (EUL) Analyses

The retention analysis dataset still had multiple pieces of equipment per observation. For example, if a customer received two cheesemelters through the program they would have one observation saying they received two and then reporting how many were still in place and operational during the site visit and approximately what year removal occurred, why, and their satisfaction with the efficient cheesemelters.

Survival analysis is performed on each piece of equipment. The analysis dataset, therefore, needs to be for each piece of equipment. It is also needs to be a dichotomous variable if there and operational or not for each time period in the study.

Cross-tabulations for the number installed and the number operational for each measure for each estimate of retention year were obtained through SAS analysis. This provided the information to create the 0-1 dataset for each possible retention year (1 through 9) for each piece of equipment installed. This dataset was created in Excel and imported into SAS as the survival dataset.

All of the survival analysis models were tested using the SAS procedure of Lifereg. Several model functional forms were tested. These were: logistic, logistic with duration squared (to fit expected pattern of inflection point slowing of retention losses), log normal, exponential, Weibull, and gamma. All of these tests achieved algorithm convergence. (No model achieved convergence within the 4th year retention study.)

The logistic model is that which best fits the theory of EUL for energy efficiency. The log likelihood measurement for goodness-of-fit for the models was the best for the logistic models, verifying the validity of this functional form and model. The Wald chi-square test statistics for duration were not statistically significant, or barely so, for the non-logistic functional forms. This statistic for the duration variable was very highly significant for the logistic model. This provided further strong evidence supporting the logistic survival analysis conducted.

In general, the logistic model can be viewed as an odds ratio where:

Logit (p) = log (p/(1-p)) = regression model = a + ∃ x + $\exists_2$y +… = eq.

The odds ratio is then p/(1-p) where the log(odds) equal the regression equation.

The odds are then: exp(eq). If z = exp(eq) then p = z/(1+z). As duration varies the resulting probability of retention varies. Fitted values can be derived by plugging in the different retention years. The estimate for median lifetime (EUL) is the fitted value where the probability of retention equals 50%.

The logistic model of duration and duration squared allows for the non-linear retention function based upon duration, as would be expected. The results and statistics from the Lifereg logistic procedure for this model are presented in Table 3.1. The log likelihood for goodness-of-fit is quite good and the Wald chi-square statistics provide a very high probability that the effect of duration is not random chance.

Table 3.1 Lifereg Logistic Survival Analysis Results

Observations	5,022
Noncensored values	4,464
Right censored values	558
Log Likelihood	-6,994.70

	Wald Chi-Square	Probability
Duration (in years)	898.76	<0.0001
Duration-squared (years)	518.98	<0.0001
	Estimate	**Standard error**
Duration (in years)	0.4421	0.0147
Duration-squared (years)	-0.0459	0.002

The *ex post* EUL was estimated through fitted values on the results from a survival analysis with this logistic survival analysis model based upon duration in years and duration-squared. The EUL estimate and its confidence intervals are presented in Table 3.2.

Table 3.2[9] EUL Estimates

		Logistic Survival Analysis Model on Duration & Duration-Squared
9-year retention rate	56.5%	
Median value (fitted value where p=50%)		**9.5 years**
Confidence interval* (lower)		8.5 years
Confidence interval* (higher)		10.5 years

* Calculated at 80% confidence level.

A few simplistic exams were made with the retention year estimates and using the current 9-year retention estimate. A simple linear estimate using the overall loss divided by nine for an annual rate provides an estimated EUL of 10. An exponential model based on the 9-year retention provides an estimate of 10.6 years for the median. Both of these methods do not include the non-linear form that would be expected and do not include the right censored data that occurs. Because of these, we would expect these estimates to be higher than the true estimate. These provide further

[9] This table is the same as Table ES.1.

verification as to the reasonableness of the logistic survival analysis estimate.

The *ex ante* EUL for cooking measures for Southern California Gas Company's Commercial EEI Program is 12 years. Twelve (12) years does not fall within the confidence interval for the Effective Useful Life (EUL) estimate from this study. **The *ex post* EUL is a more accurate measure of the EUL of the 1996 Commercial EEI Program's cooking measures as of this 9th year retention study. The *ex ante* EUL is 12 years but the 9-year retention measurement is 56.5% and the *ex post* EUL is 9.5 years.**

3.3 Required Protocol Tables

This subsection provides the summary tables as required in the M&E Protocols.

Table 3.3 provides the summary documentation for data quality and processing as required in Table 7 of the M&E Protocols.

Table 3.3 Data Quality and Processing Documentation
Protocol Table 7B

Protocol Table Item #	
	Overview Information
1a. Study Title & ID	Measure Retention Study for SoCalGas Company's 1996 Commercial EEI Program, Study 720A
1b. Program, years, & descrip.	1996 Commercial Energy Efficiency Incentives Program Assistance provided for high efficiency gas measures for commercial customers.
1c. End uses & measures	End Use: Cooking Study measures: Cooking (e.g., ranges, boilers, fryers, melters)
1d. Methods & models	Site survey analysis produced retention counts. Given a stratified sample, weighted average retention calculated. Retention year estimates from site visits, prior 4th year retention findings, and interview comments received by auditors. Survival analysis performed using SAS Lifereg with a logistic model of duration and duration squared.
1e. Analysis sample sizes	200 sites for a total retention study of 401 measures.
	Data collection: August 2005.
	Database Management
2a. Data sources	Program tracking database provided information for initial stratified sampling pool. Final sampling pool was selected sites from the 4th year retention study that had been matched to be within the first year impact study.
2b. Data attrition	Unscheduled on-sites had no scheduling attrition. Only 5 back-up sample sites were required.
2c. Data quality checks	All program data pulled along with initial sampling. Each customer in sampling pool was assigned a tracking number that was used throughout surveying, data entry and verification, and initial analysis checks within study. Protocols established for site visits, utility interactions, and data entry.
2d. Collected data not used	None

	Sampling
3a. Sampling procedures	Stratified random sample by expected site therm savings for cooking measures.
3b. Survey information	Survey instrument provided in Appendix A.
3c. Statistical descrip.	Retention findings based on weighted mean of measure type retention estimate. Survival analysis performed using SAS Lifereg with a logistic model of duration and duration squared.
	Data Screening and Analysis
4a. Outliers	No outliers identified or treated.
4b. Background var.	None.
4c. Screened data	No screening, all data utilized.
4d. Model statistics	Log likelihood for model of -6,997.7. The Wald chi-square statistics for the duration and duration-squared terms are 898.76 and 518.98, respectively. These have a random probability of less than 0.0001. The estimates are 0.4421 for duration and -0.0459 for duration-squared. The standard errors are, respectively, 0.0147 and 0.002. The fitted values provide that at p=50% the EUL is 9.5 years. With the above standard errors, the 80% confidence interval is 8.5 years to 10.5 years. See Section 3.2 of report for further information and discussion.
4e. Specification	Survival = duration and duration-squared.
4e1 Heterogeneity	All cooking measures with no heterogeneity considered.
4e2 Omitted Factors	No omissions.
4f Error	Simple survey with no scheduling attrition. There were 200 site visits.
4g Influential data points	No outliers identified.
4h Missing data	None.
4i Precision	Confidence levels computed on retention rates. Standard errors for variable estimates in the logistic model used with fitted values to derive 80% confidence interval.

Table 3.4 provides a reporting summary of the study results as required in Table 6 of the M&E Protocols.

Table 3.4 Data Quality and Processing Documentation
Protocol Table 6

Protocol Table Item #	
	Overview Information
1. Studied measure & end-use	Cooking
2. *Ex ante* EUL	12 years.
3. *Ex post* EUL	9.5 years
4. *Ex Post* to be used	9.5 years
5. EUL Standard Error	Standard error for duration=0.0147 and 0.002 for duration-squared. Confidence interval calculated using these intervals with the coefficient estimates and fitted values.
6. 80% Confidence Interval	8.5 years to 10.5 years
7. p-Value	20%
8. Realization Rate	1.0
9. Like measures	None

3.4 Other Findings

The greatest responses for equipment removal were the reason wasn't known at the time of the 9-year retention site visit. New tenants replacing the equipment was the reason given for equipment removal 20% of the time. The next highest response is that the equipment was replaced by the previous tenant (17%). Old, worn-out or broken equipment that was replaced occurred for 11% of the equipment removed. The building was removed for 6% of the reasons provided. Other lower response rates were a different type of business (4%), that the building is vacant and the equipment was removed (3%), equipment was removed because the tenant did not use it (3%), and equipment burned up (2%). Table 3.5 presents these equipment removal reasons and their response portions.

Table 3.5 Reasons for Equipment Removal

Reason	Response % for Equipment Removed
Do not know	34%
New tenant replaced equipment	20%
Previous tenant replaced	17%
Old, worn out, broken—replaced	11%
Building removed	6%
Different type of business	4%
Building vacant, equipment removed	3%
Removed by tenant because not used	3%
Burned up	2%

Overall, customers were satisfied (78%) with the efficient cooking equipment. Most were very satisfied (62%), while 16% were somewhat satisfied. There were 19% that were neither satisfied nor dissatisfied. Only two percent were somewhat dissatisfied and two percent were very dissatisfied.

As would be expected, those with retained equipment are more satisfied than those with removed equipment. Table 3.6 presents the equipment satisfaction ratings for those the retained equipment and the removed equipment. Satisfaction runs quite high for the retained equipment with 79% very satisfied and 16% somewhat satisfied with this equipment. No one is very dissatisfied with retained equipment and only five percent (5%) of retained equipment received a somewhat dissatisfied or neutral rating.

Table 3.6 Reasons for Equipment Removal

Satisfaction Rating for Efficient Equipment	Equipment has been Retained	Equipment had been Removed
Very Satisfied	79%	14%
Somewhat satisfied	16%	14%
Neither satisfied nor dissatisfied	3%	62%
Somewhat dissatisfied	2%	4%
Very Dissatisfied	0%	6%

Appendix A: Site Survey Instrument

Introduction
Southern California Gas Company
1996 Commercial EEI 9-Year Retention Study

Original Customer

Hello, my name is _____. I work for ASW Engineering and we have been hired by Southern California Gas Company to do a very quick survey at this site. Our records show that your firm purchased some energy efficient cooking equipment with the help of SoCalGas back in 1996. We are performing a study to see how much of this equipment has since failed or been removed. This is just a study and there is no penalty or consequences if the equipment is no longer here. It should take me less than five minutes to see the equipment and turn on to test that they work.

CONDUCT SURVEY

Thank you very much for your assistance.

New Customer

Hello, my name is _____. I work for ASW Engineering and we have been hired by Southern California Gas Company to do a very quick survey at this site. Our records show that a firm that was located here in 1996 purchased some energy efficient cooking equipment with the help of SoCalGas. This equipment may have remained in the facility as your firm moved in. We are performing a study to see how much of this equipment has since failed or been removed. We would like to see if this cooking equipment is still on this site. This is just a study and there is no penalty or consequences to any customer if the equipment is no longer here. It should take me less than five minutes to look through your cooking equipment to find these pieces and turn them on to test that they work.

CONDUCT SURVEY

Thank you very much for your assistance.

Site Survey Data Collection Instrument

Site Visit Survey for the Retention Study
of the Commercial EEI Program -- SoCalGas

Site Tracking#		Surveyor Initials	Survey Date

Company Name	
Prev. Contact Name	
Prev. Contact Position	
Prev. Contact Phone	
Customer Name	
Street Address	
City	
Zip	
# of Types	

		Yes	No
Q1	Is customer name different at this address?		

Q2 (IF YES) _____
New customer name

Measure Type #1		**Measure Type #2**	
	Mfg.		Mfg.
	Model		Model
	Number		Number
Q3	# Still in place	Q7	# Still in place
Q4	# also operational	Q8	# also operational
Q5	If some have been removed or replaced: Why?	Q9	If some have been removed or replaced: Why?

Q6 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Q10 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Site | Tracking # 0 | Page 2

Measure Type #3

Mfg.
Model
Number
Q11 # Still in place
Q12 # also operational
Q13 If some have been removed or replaced: Why?

Q14 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Measure Type #4

Mfg.
Model
Number
Q15 # Still in place
Q16 # also operational
Q17 If some have been removed or replaced: Why?

Q18 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Measure Type #5

Mfg.
Model
Number
Q19 # Still in place
Q20 # also operational
Q21 If some have been removed or replaced: Why?

Q22 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Measure Type #6

Mfg.
Model
Number #N/A
Q23 # Still in place
Q24 # also operational
Q25 If some have been removed or replaced: Why?

Q26 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Measure Type #7

Mfg.
Model
Number
Q27 # Still in place
Q28 # also operational
Q29 If some have been removed or replaced: Why?

Q30 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Measure Type #8

Mfg.
Model
Number
Q31 # Still in place
Q32 # also operational
Q33 If some have been removed or replaced: Why?

Q34 How satisfied are you with this equipment?

Very satisfied	Somewhat Satisfied	Neither satisfied nor dissatisfied	Somewhat Dissatisfied	Very dissatisfied

Appendix B. Datasets and Documentation

This study was specifically designed to be as simple and straight forward as possible. As the analysis progressed, the steps and programs were continually refined in order to accomplish this goal. The result was the development of small set of concise data analysis steps. The use of these steps and copies of the programs are provided in this Appendix. The datasets, SAS© programs, and Excel© spreadsheets are provided on diskette at the end of this Appendix. Following the description contained below, the work should be easily replicable.

Flow of Datasets and Analysis Programs

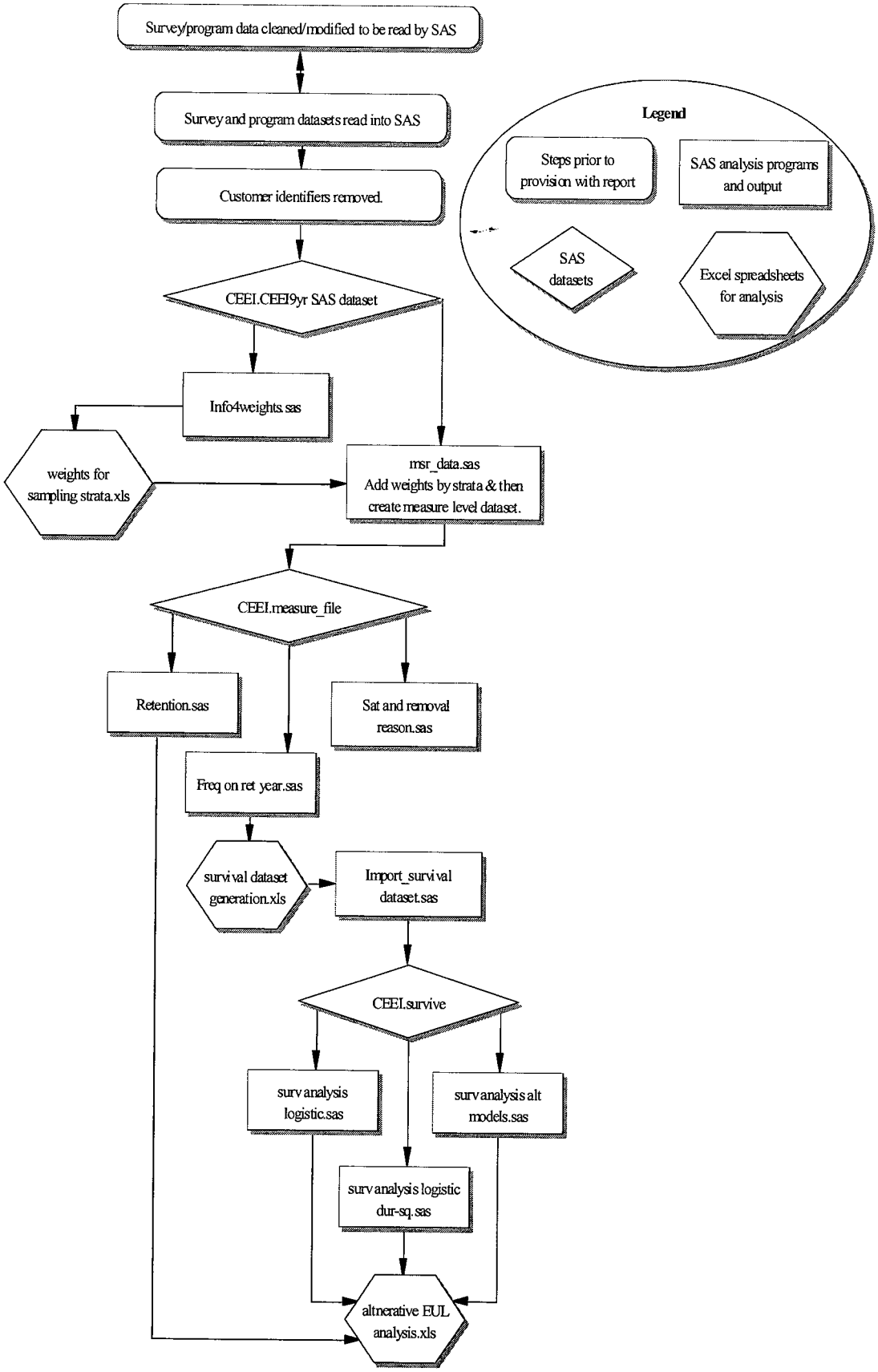
A step-by-step schematic of the use of datasets and analysis programs is presented in Figure B.1. This diagram also indicates the complete flow of the material provided and the type of material (dataset and type, program and type). This diagram can be used with the datasets and programs provided on diskette to replicate all of the results discussed in this report.

Printed copies of each of the SAS© programs and Excel© spreadsheets are provided in the pages following the flow chart. They are provided in the order that they are used.

Set-Up Reminders for Replication

The SAS© and Excel© programs are the exact ones used for this study. A few minor changes will need to be made to replicate the work.

SAS© programs contain LIBNAME statements and FILENAME statements in the beginning of the programs to tell the program where to find datasets and where to place datasets. These will need to be changed to reflect the folder set-up being used in the replication.


Survey/program data cleaned/modified to be read by SAS
Survey and program datasets read into SAS
Customer identifiers removed.
CEEI.CEEI9yr SAS dataset
Info4weights.sas
weights for sampling strata.xls
msr_data.sas
Add weights by strata & then create measure level dataset.
CEEI.measure_file
Retention.sas
Sat and removal reason.sas
Freq on ret year.sas
survival dataset generation.xls
Import_survival dataset.sas
CEEI.survive
surv analysis logistic.sas
surv analysis alt models.sas
surv analysis logistic dur-sq.sas
altnerative EUL analysis.xls
Legend
Steps prior to provision with report
SAS analysis programs and output
SAS datasets
Excel spreadsheets for analysis