SCE NON-RESIDENTIAL NEW CONSTRUCTION PERSISTENCE STUDY

FINAL REPORT

DECEMBER 15, 1998

Prepared For:

Southern California Edison

2131 Walnut Grove Ave.

Rosemead, CA 91770

Prepared By:

RLW Analytics

1055 Broadway, Suite G

Sonoma, CA 95476

SCE-04

Table of Contents

Executive Summary

Goals

The goal of the study was to estimate the persistence of savings and retention of measures installed in new construction and large remodeling applications for the two program years: 1994 and 1996. The key issues explored in this study were:

- Technical Degradation of Installed Measures, the reduction in the measure's efficiency due to age and wear
- Measure Retention, the continued use of the measure
- Survival Function of Measures, the mathematical expression used to calculate the percentage of the original savings remaining after a given amount of time
- Effective Useful Life of Equipment, the number of years from installation where the survival function equals 0.5

Methodology

We used a combination of telephone and on-site surveys to estimate the survival proportion of the savings and to estimate the effective useful life of installed measures. Our approach was designed to satisfy the requirements of the M&E Protocols issued by CADMAC and was reviewed by the appropriate CADMAC subcommittee. We used a whole-building approach in this evaluation which was consistent with the philosophy of the 1994 and 1996 impact evaluations and with the retroactive waiver filed by Edison.

The sample of NRNC program participants for this study was drawn from those buildings that were sampled in the 1994 or 1996 impact evaluation. We sampled 61 of the 133 impact study sample sites for this study. All results were weighted to extrapolate back to the total 1994 and 1996 populations of program participants, 131 sites in 1994 and 272 sites in 1996.

A phone survey was used to determine whether incented measures were still installed and working at the time of the survey. Follow-up on-site surveys were conducted for those sites where the telephone survey indicated changes that may have impacted the incented equipment such as turnover of occupants, renovation of space, or removal of equipment, or replacement of less efficient equipment. Changes that were not included were repairs, replacement with equally efficient equipment, and changes in operating schedules.

The on-site survey consisted of a walk-through of the building by a surveyor. During the on-site, the surveyor compared the data collected during the original survey to the observed equipment. Particular attention was given to the systems that had received incentives in the NRNC program.

DOE-2 models were constructed for all buildings surveyed in this study. In each sample building model, a technical degradation factor was applied to each category of equipment based on the degradation estimates developed by the Statewide Technical Degradation study. Affected building parameters for those buildings with on-site surveys were set to reflect as-found conditions.

The statistical analysis of the data consisted of calculating the current program savings, the current survival proportion, and the Effective Useful Life (EUL) for each program year. The current program savings were estimated from the data collected in the phone and on-site surveys for the current year together with the data collected for energy savings in the first-year evaluation study. The survival proportion was estimated as the ratio between the current energy or demand savings (reflecting degradation, persistence, or both) and the first-year savings of the program. All measures of savings were calculated from the DOE-2 models constructed for the sample sites.

The EUL was estimated from the survival proportion assuming an exponential survival function. This is equivalent to assuming that the probability of failure is constant over time. The EUL is defined to be the number of years after which the survival proportion for savings would equal 50%. For example, suppose that after four years, the survival proportion is equal to 85%, i.e., the savings have declined by 15% in four years. Then, under the exponential survival model, the EUL would be about 17 years.

The final objective of the present study was to compare the ex post estimate of the EUL (i.e., the estimate obtained in this study) to the ex anti estimate of the EUL that is currently assumed in the program. A statistical hypothesis test was carried out to determine whether the ex post estimate was significantly different than the ex anti assumed value.

Results

All building models were projected to the original program populations to obtain the total program results presented here. Table 1 summarizes the observed survival proportions for the total program savings in annual kWh energy. The table shows the survival proportions for both the 1994 and 1996 programs. Results are given considering technical degradation alone, persistence alone, and technical degradation and persistence taken together.

In the case of the 1994 program, considering only technical degradation, 98.6% of the energy savings of the 1994 program is still being achieved in 1998. In other words, only 1.4% of the original savings has been lost due to technical degradation. Considering only persistence, only 0.5% of the original savings has been lost. Considering both factors, only 1.9% of the original savings has been lost.

In the case of the 1996 program, only 2.0% of the savings has been lost to technical degradation. The technical degradation was slightly larger in PY96 than in PY94 because the mix of measures was somewhat different. In 1996 we saw no losses due to persistence.

	1994		***1996***	
Category	***Estimate***	***Rel Prec***	***Estimate***	***Rel Prec***
Degradation	98.6%	0.9%	98.0%	0.9%
Persistence	99.5%	0.7%	100.0%	0.0%
Both	98.1%	1.2%	98.0%	0.9%

Table 1: Survival Proportions for Annual Energy

Table 1 also shows the relative precision of each of its estimates. The relative precision reflects the variation due to sampling. For example the relative precision of the 1994 survival proportion due to degradation was 0.9%. This indicates that we would expect the result obtained from the sample to be within ±0.9% of the result that would have been obtained if the same engineering methodology had been applied to every project implemented in the 1994 program.

A 90% confidence interval for the true survival proportion can be calculated as 0.986 ± (0.009) (0.986). In words, considering the potential variation in the results due to sampling, we can be confident that the survival proportion is between 97.6% and 99.5%. Analogous confidence intervals are shown in Table 2 for each of the results reported in Table 1.

	Interval		***Interval***	
Category	***From***	***To***	***From***	***To***
Degradation	97.6%	99.5%	97.2%	98.9%
Persistence	98.8%	100.2%	100.0%	100.0%
Both	96.9%	99.3%	97.2%	98.9%

Table 2: Confidence Intervals for Annual Energy Survival Proportions

Table 3 summarizes our results for the Equivalent Useful Live (EUL). The top portion of Table 3 shows our ex post estimates for each case. These results were calculated from the lower estimate of the survival proportion (S) shown in Table 2. Therefore, this is a conservative estimate in terms of sampling variability. Obviously, these estimates cannot be taken literally. It is absurd to believe that half of the first-year savings will last 236 years. This is discussed further below.

The lower portion of Table 3 shows our ex anti estimates of the EUL of each of the two programs. We calculated this as an average of the assumptions made in the programs for each measure, using the net resource benefit of each measure as weight. For example, for the 1994 program we found the ex anti estimate of average EUL to be 12 years.

The table also shows the current survival proportion S that would be expected under the exponential survival function corresponding to the ex anti EUL. Considering the 1994 program as an example, we would expect to have seen 79.4% of the savings after four years, i.e., we would have expected to see a loss of 20.6% of the first-year savings if the actual EUL is 12 years.

A hypothesis test can be carried out to compare the ex post and ex anti estimates shown in Table 3. This test leads to the conclusion that the true EUL is greater than the ex anti value. Essentially, these results indicate that the losses due to technical degradation and persistence are currently very low – significantly lower than would be expected under the program assumptions.

	1994		***1996***	
Category	***S***	***EUL***	***S***	***EUL***
Degradation	97.6%	117	97.2%	48
Persistence	98.8%	236	100.0%	na
Both	96.9%	88	97.2%	48
Ex Anti Value	**79.4%**	**12**	**88.2%**	**11**

Table 3: Effective Useful Life for Annual Energy

The main report also reports results for peak demand. They were essentially the same as the results for annual energy that have been describe.

Discussion of the Results

It is important to consider the possibility that our methodology may be giving a biased result. We examined each of the key assumptions in the study and concluded that only one assumption is doubtful. In estimating the EUL and carrying out the related hypothesis test, we assumed an exponential survival function. This is equivalent to assuming that a fixed proportion of the current savings is lost each year. An assumption such as this is necessary to estimate the EUL since the survival proportions observed in the first two or four years must be extrapolated out into the future.

However, the implausibly long estimates of EUL indicate that the exponential survival model is wrong. It appears that savings fail at a higher rate as the buildings and equipment grows older. If this is true, it is simply too soon to estimate the rate of increase in the failure rate and the corresponding EUL. However we can be confident in concluding that up to the present time the savings are persisting at a higher rate than had been expected.

Lessons Learned

The principle lessons from this study are:

- Persistence of savings is high in the first few years of these programs but it is too soon to determine reliably how long the savings will last.
- Therefore, it is important to continue to track persistence of savings over time.
- Persistence can be measured cost-effectively by building on the detailed engineering models and excellent customer relationships from the first-year evaluation studies.
- The statistical methodology of the present study seemed to work well.

Key Issues

This is the final report for the Southern California Edison 1994 and 1996 Non-Residential New Construction Program Persistence evaluation. This document summarizes the key issues in this study, presents the study methodology, and details the findings of the study.

This study can be thought of as having four phases. They are:

1. Study design
2. Data collection
3. Analysis
4. Reporting

Each phase of the project presented unique challenges.

The figure below shows the overall flow of the project from study design to final reporting. It also summarizes the key issues at each stage of the project. For the project to be successful, these issues had to be effectively addressed. The discussion below briefly describes how we addressed these issues. More complete discussion can be found later in the report.

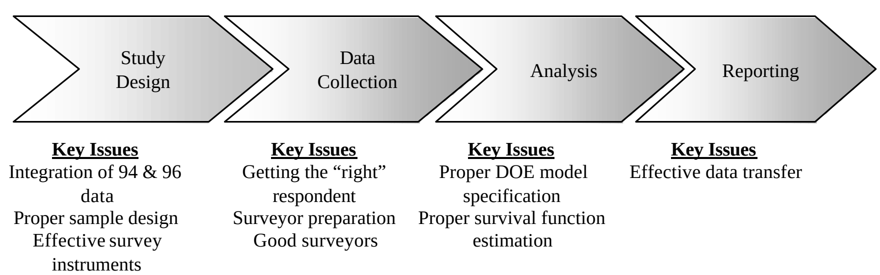

Figure 1: Overall Project Flow and Key Issues

Key Study Design Issues

Quality control steps that were taken in the early stages of the work profoundly effected the ultimate success of the project. To ensure that a solid foundation was set for the project, the senior staff at RLW Analytics and AEC personally performed the study design tasks.

The first key issue that had to be addressed was the integration of the 1994 and 1996 program data. The data collection for the 94 and 96 projects were slightly different and the data resided in databases with different formats. This had the potential to introduce errors into the study from the outset. Because we developed these databases for the original impact evaluations, we had an intimate understanding of the structures and the key differences. The same staff that originally developed each of the databases was responsible for merging the data. This use of the original database designers minimized the chances for data errors that could have propagated through the remainder of the analysis.

There were some slight differences in the data collected and modeling algorithms between the 94 and 96 studies. We updated the PY94 database to the PY96 database format and reran the PY94 models with the PY96 algorithms. The change in savings calculated from the new models was applied to the first year savings originally developed for the PY94 study.

The other key study-design issue was an effective sample design. The central issue was to choose a sample size large enough to comply with the Protocols and be defensible, but not excessively large. The sample design had to satisfy the hypothesis-testing framework for effective useful life on a whole-building basis. We believe this is new ground. Our analysis of the protocols indicated that as few as 28 sample sites would be adequate. Taking a conservative approach, we developed a sample of 61 sites from the two program years, 31 from 1996 and 30 from 1994. The approach behind the sample design is described in a later section.

Key Data Collection Issues

Perhaps the most critical data collection issue was ensuring that the proper respondent was contacted for the telephone survey. The proper respondent is the one who is most knowledgeable about construction and maintenance activity at the sample site. We started with the decision-maker survey respondent from the first year impact studies. We recognized that the appropriate decision-maker during the construction process was not necessarily the proper contact for maintenance issues. Therefore, we qualified the respondent and asked for a referral in the event that someone else was a more appropriate contact.

Once the appropriate person was identified, the next issue was accurate field data collection. The on-site staff used in the 1996 Edison NRNC impact evaluation was used to conduct the onsite surveys. Each surveyor was an experienced DOE-2 engineer, and was well-qualified to understand the data collection and modeling issues key to answering the research questions posed by this project. This issue proved to be a small one because of the very high persistence rate of the incented equipment.

Key Analysis Issues

Estimating savings persistence and measure retention involves two issues:

- Degradation in measure efficiency over time (persistence)
- Retention of installed equipment. (retention)

The application of the technical degradation factors to the simulation models was facilitated by the ModelIT automated modeling software. The issue of simulating impacts of measure retention and persistence at surveyed sites was expected to be more complicated. The key to obtaining meaningful measure retention and persistence results from the on-site survey and simulation exercise is to insure that the models respond only to observed building changes. Thus, once the on-site survey was completed, data entry and modeling needed to focus on these changes, while leaving other building attributes energy-neutral. This required knowledge of both the original modeling process and new modeling techniques necessary calculate the impact of the building changes. Because we were using the same on-site survey team that completed the 1996 NRNC evaluation, the group was intimately familiar with the process that created the original models. In hindsight, this turned out to be easier than expected, since only two sites required onsite visits.

The central issue of analysis was to carry out the survival analysis at the whole-building level, using a methodology that would yield unbiased estimates of program-level survival rates and effective useful life. The whole-building approach meant that standard statistical survival analysis was not applicable. Fortunately, the same statistical methods used in the original evaluation studies could be used to estimate aggregate program-level survival rates. These in turn could be used to estimate the aggregate program-level effective useful life results, and to carry out the required hypothesis tests. This methodology is described later in this report

Key Reporting Issues

The most important reporting issue is to ensure that the data and knowledge is effectively transferred to Edison at the conclusion of the project. The final report has been written by senior staff. There have been multiple iterations of review and revision before delivery of the draft to Edison.

The datasets to be delivered were assembled by senior database developers at AEC. The database structure conforms to common standards and has been documented such that anyone reasonably proficient with databases will easily understand the structure and be able to use the databases to perform additional analysis or reporting.

Detailed Methodology

Sample Design

Background

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

The ex-anti estimates were to be changed only if there is a significant difference between the ex-post and ex-anti estimates of measure life. At the planning stage, the 80% level of confidence was assumed although in the final analysis this was changed to the more conventional 90%. A two-sided test was assumed. We assumed that the sample size should be chosen so that the hypothesis test should have 80% probability of rejecting the null hypothesis under the assumption that the true value is 20% less than the ex-anti estimate.

We found that the preceding criterion requires a sample of 28 sites. We chose to apply this criterion to each of the two program years.

Technical Analysis

Our sample size planning was carried out in the following five steps:

1. Establish the procedure for estimating the survival proportion *S* of the measures in a set of buildings of a particular average age *t*. Specifically, consider a particular program year such as PY94 and assume an exponential survival function as specified in the work plan.
2. Establish the procedure for estimating the effective useful life *EUL* for a particular set of buildings, given an estimate of the survival proportion *S*.
3. Find the relationship between the sampling distributions for estimating survival and for estimating effective useful life. In particular, how is the coefficient of variation (*cv)* of the estimator of *EUL* related to the coefficient of variation of the estimator of *S?*
4. Find the required value of the coefficient of variation of the estimator of *EUL* to satisfy the hypothesis-testing framework of the proposed protocols.
5. Find the relationship between the required sample size *n* and the coefficient of variation of the estimator of *EUL*. Solve for the sample size *n*.

The methodology of steps 1 and 2 is discussed further under in a later section: Data Analysis.

For each of the two program years, we define the survival proportion *S* to be the current energy use of the corresponding population of program participants as a proportion of the gross first year savings found in the program evaluation. We use standard MBSS™ ratio estimation techniques similar to what was used in the first-year studies to estimate *S* from the information from the telephone and onsite surveys and the corresponding engineering models. This estimator may be denoted $\hat{S}$. The MBSS procedure gives the value of $\hat{S}$ and the corresponding standard error. Let *t* denote the age of the buildings, 2 years for PY96 and 4 years for PY94.

The next step in our analysis was to obtain an estimate of EUL from $\hat{S}$. The ex-anti estimate of effective useful life was calculated on a whole-building basis for each program year. The starting point was the exponential survival function $S(t) = e^{-l t}$. Here the mean survival time is equal to $1/l$. We defined the EUL as the value of *t* that satisfies the equation: $S(t) = e^{-lt} = 0.5$. Solving for t = EUL, we obtained $EUL = -\frac{\ln(0.5)}{l}$. If we observe $\hat{S}$ in a sample with average measure age t, then we can solve the survival function for $\hat{l} = -\frac{\ln(\hat{S})}{t}$. If we substitute this equation in the preceding one, we obtain $E\hat{U}L = -\frac{t \ln(0.5)}{\ln(\hat{S})}$.

Thus, following the exponential failure model and the definition of EUL from the RFP, we used the estimated the EUL from the survival proportion using the equation:

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

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

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

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

This can be rewritten as

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

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

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

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

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

Final Sample Design

The following points summarize the steps that were taken to design the sample:

- The original participant sample of 133 sites from PY94 and PY96 data were weighted to reflect the population of 403 sites using the program estimate of savings determined by SCE.
- The persistence-study sample was designed using the weighted impact evaluation sample. The sample design for each of the program years appears in the tables below. The stratification of the sample was based on the actual measured savings from the impact evaluation.

The figure below shows how the data at each level is related to the other levels.

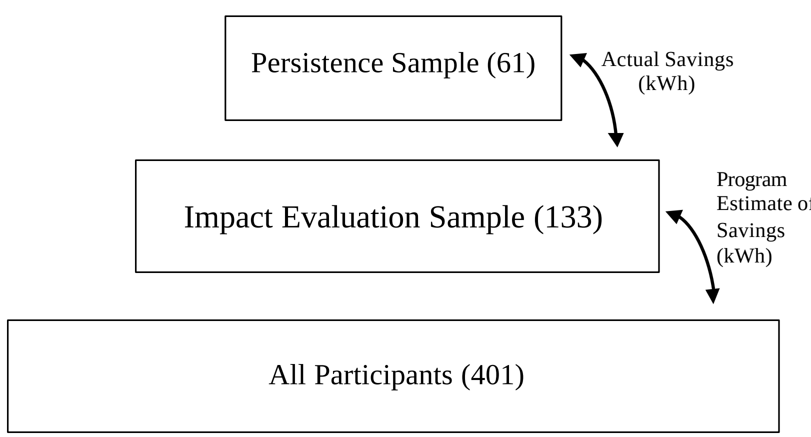

Figure 2: Persistence and Impact Samples

The sample frame used to design the final sample contained all 60 sampled sites from PY94 and all 73 sampled sites from the PY96 impact evaluations. A total sample of 61 sites satisfies the confidence and precision requirements of the Protocols. A gamma of 1.0 was used to determine the size of each stratum and the amount of necessary sample sites within each stratum. Table 4 shows the full set of parameters assumed to prepare the MBSS sample design.

Parameter	***Value***
Beta (β)	1.00
Gamma (γ)	1.00
Error Ratio	0.50

Table 4: MBSS Sample Design Parameters

The sample was stratified by program year and then by the actual savings calculated from the PY94 and PY96 impact evaluations. The PY1994 sample design contained five size strata, all of which are probability strata. According to the original sampling plan, six sites would be selected from each of the five strata. In practice, we were limited to the sites actually included in the first-year evaluation so the sample sizes were adjusted to reflect the available sites. Table 5 shows the strata cutpoints, population counts and final sample sizes for each strata.

Stratum Number	***Strata Cutpoint (kWh)***	***Weighted Building Population***	***Sample Size***
1	69,711	134	6
2	142,862	58	7
3	485,729	37	8
4	830,199	24	5
5	2,000,000	19	4
TOTAL		**272**	**30**

Table 5: Program Year 1994 Sampling Plan

The PY96 data was divided into 6 strata. The first 5 strata contained 6 sites to be sampled, and the last stratum contained the largest site in the population to be surveyed. In this case the sample could be implemented without further adjustment since the original sampling plan was also stratified on savings and gave an adequate number of sites in each stratum.

Stratum Number	***Strata Cutpoint (kWh)***	***Weighted Building Population***	***Sample Size***
1	77,702	61	6
2	203,279	28	6
3	472,942	18	6
4	943,337	12	6
5	1,752,276	10	6
6	7,000,000	1	1
TOTAL		**131**	**31**

Table 6: Program Year 1996 Sampling Plan

Telephone Survey Instrument

The goal of the telephone survey was to determine if the participating buildings are still in service and if there have been any significant changes to those buildings. The survey instrument addressed the following topics:

- Is the building, or portion of the building, which participated in the NRNC program still in service?
  - If no longer in service, when was it removed and why?
  - Is the building permanently out of service, or is it just temporarily vacant?
- Obtain information about any changes to the energy-consuming equipment in the building.
  - Is the incented equipment still in place and operable?
  - If not, was it removed, disconnected, broken, or damaged? Why / how?
  - When was the equipment removed or disconnected?
  - Was this part of a larger modification? What else happened?
  - What replaced the incented equipment?
  - Have other energy-consuming systems been removed or modified? Which systems?
- Determine if there is a new tenant in the building, and if so, determine if the type of business has changed.
  - Have there been any remodeling changes?
  - Were there any changes when the new occupant moved in?

The telephone survey instrument was written to function as a recruiting instrument for the onsite survey if the interviewer discovers any of the following:

- The facility has been removed from service
- A new tenant has moved into the facility, and changes have been made
- Any of the incented equipment has been removed or modified

The telephone survey contained a total of 17 questions. A draft instrument was pre-tested on ten customers and some refinements were made. The total time necessary to administer the survey was approximately 10 minutes. The draft of the survey instrument is contained in the appendix of this report.

On-Site Survey Training

In preparation for the PY96 NRNC evaluation, a detailed training course was developed and delivered to all surveyors. The course covered a range of issues, including program design and operation, targeted measures, customer relations and etiquette, measure identification, and surveying techniques. Since the surveyor used in this project was also involved in the PY96 evaluation, further detailed training was unnecessary. However, a refresher course was held at AEC to introduce the surveyor to the unique objectives of this project, and review key measure identification and customer relations issues from last year. The on-site survey conducted during the original evaluation was reviewed, and issues relative to the specific building surveyed were reviewed.

Telephone Surveys

The telephone surveys were conducted from RLW Analytics' Sonoma, CA office by two technically qualified surveyors. The flowchart below outlines the telephone survey process.

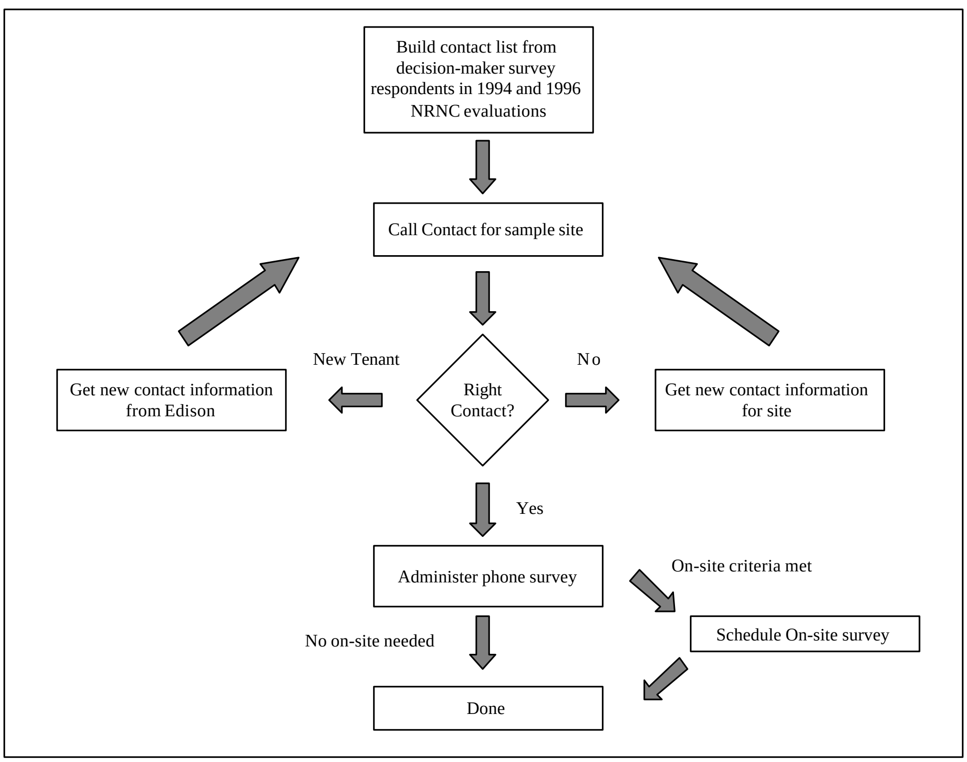

Figure 3: Flowchart of Telephone Survey Process

We began the telephone survey process by extracting the name and phone number of the primary respondent for the decision-maker (DM) survey for the 1994 or 1996 NRNC impact evaluation database. We were already in possession of these databases. They are part of our project records from the impact evaluations. The contact information was appended to the sample frame data for this study.

The telephone surveyors contacted the DM survey respondent and verify that they are still the appropriate respondent. The surveyors asked for a referral if the DM respondent was no longer the appropriate contact. The referral information was entered into the tracking database and the surveyors attempted to reach the new contact. If it was found that a new tenant has moved into a site, location information was provided to Edison so that the Edison project manager could provide the surveyors with contact information for the new customer.

A minimum of 5 attempts were made to contact each sample point before that point was deemed unreachable and replaced in the sample.

Once the appropriate person was reached, the survey was administered and, if appropriate, an on-site survey was scheduled. All contact and survey data were stored in a database for later analysis and delivery to Edison.

An on-site survey was triggered based on the telephone survey questions if:

1. the facility had been removed from service,
2. there had been a tenant change that included a tenant improvement, or
3. there had been a removal or modification of equipment installed through the NRNC Program

An on-site survey was not required if any of the following apply:

- The building was only temporarily vacant
- Only lamps, task lighting, or other process or plug loads were changed
- The measure was removed and replaced by a similar measure with the same or greater efficiency

The phone surveyors attempted, on average, 2.3 times to contact a site. The maximum number of attempts made to contact a site was 10. They spoke with an average of 1.2 contact persons per site. The phone surveyors were instructed to contact those individuals in charge of influencing or making the decisions on the installation choice and usage of the energy equipment at the site. Contact names from previous years were used as primary contacts for each site. The contact was the same person 40.0% of the time in PY 94 sites. PY 96 sites proved to have more of the same contact people at 54.8% of the sites.

On-Site Surveys

The original survey data that was collected in the first-year impact evaluations was the primary basis for the analysis. As explained in the prior section, new onsite surveys were only required if the telephone survey had indicated that the facility had been removed from service, and new tenant had moved in and made changes, or any of the incented equipment had been removed or modified.

Before going into the field, the surveyor examined the original data, and made changes only to data elements that are related to changes in the building or equipment affecting this study. PY94 data were collected using a paper on-site form, with numerous QC checks occurring after the original data were collected. PY96 data were entered directly into the SurveyIT on-site survey database.

The final “cleaned-up” PY94 data were merged with the PY96 data into a single database under a different project. Once these data were merged, a machine-generated on-site data report was generated for the surveyed site. During the on-site survey, changes to the building description data that relate to this study were “red-lined” onto the machine-generated report. This technique enforced a consistent modeling and data collection approach, while highlighting building changes that relate only to this study. Once the building changes were identified, data relating to these changes were entered into the SurveyIT database.

An important role of the on-site surveyor was to verify first hand the information given to the phone surveyor. Thus, phone survey responses relating to physical changes to the building were verified. The on-site survey began with an interview of the site contact, consisting of the following questions:

- Has the use of the participant building (or portion thereof) changed since the energy consuming systems were installed? If so, how?
- Have any of the rebated systems been removed? If so, why? What was installed in their place?
- Is energy-consuming equipment being used differently than it was originally? Has it been modified?
- Were any changes made since the building was occupied as a result of an Edison energy-efficiency retrofit program? If so, what equipment was affected? Was any equipment that was installed under the original program changed during a later retrofit program?
- Is there a maintenance schedule for the energy-efficient equipment?
- Are energy-consuming systems in a good state of repair?

An interview guide was developed to guide the surveyor through the interview process. The interview guide and the original on-site data were used as the data collection instruments for this study.

The overall process was:

1. If the phone survey indicated that an on-site visit was necessary, the site were recruited and scheduled.
2. Program records and previous on-site data forms and data were reviewed by the surveyor prior to the site visit.
3. The engineer responsible for the model collected the on-site data. As discussed above, data collection focused on changes to the building since the original survey.
4. The on-site surveyor entered the changes to the on-site survey data directly into the SurveyIT database.

5. As soon as the data were keyed into the program, the automated model building software created the DOE-2 model ***and*** calculated changes in energy savings for the surveyed site. The models were checked for reasonableness by the surveyor and AEC project manager..

Recruiting

Sites identified during the phone interview as potential survey sites were recruited at that time. The phone surveyor scheduled the site visit and confirmed the contact and building location information. A letter on Edison letterhead was sent out to each site recruited to verify the authenticity of the study and provide Edison contact information to the customer. A postage-paid customer-satisfaction post card was included in the letter to provide feedback to the project team on surveyor deportment and overall customer satisfaction with the survey. SCE account representatives had access to the survey schedule and could easily arrange to accompany the surveyor during the on-site survey.

On-Site Visits

The on-site visit at the surveyed site took about 4 hours to complete.. Areas of the building associated with changes identified during the phone survey and subsequent interviews with site personnel were surveyed. The surveyor also looked for evidence of other remodeling activity not reported by the site contact, but did not find any.

QC

After the data were collected, the changes were entered into a central database. Range checks at the data entry level were implemented, as in the PY96 impact evaluation. Once the data were entered, a DOE-2 model was automatically generated. The DOE-2 ouput reports were reviewed by the surveyor/modeler, and a senior RLW/AEC engineer.

The original data from the PY94 and PY96 evaluations were subject to a series of QC checks during the course of those studies. The original building description data were considered to be adequately validated for the purposes of this study. The two datasets were merged into a single database under another project. During the course of that project, additional QC checks were applied to the merged dataset, to identify errors introduced during the data merging process.

DOE-2 Simulations

DOE-2 models were developed using our automated modeling tool and the on-site survey database containing a merged set of the PY94 and PY96 buildings. Modeling algorithms and engineering assumptions from the PY96 study were used in this study. Revised models incorporating the Technical Degradation Factors developed from the CADMAC study were created for all sampled sites. Additional model changes relating to measure retention were implemented for surveyed sites. The key to obtaining meaningful results from the simulation exercise was to insure that the models respond only to observed changes in equipment performance or operation. Thus, once the survey was completed, data entry and modeling focused on these changes, while leaving other building attributes energy-neutral. Since the original savings estimates for the PY94 sites were made with an earlier version of the modeling software, a comparison of the savings estimates calculated using both versions of the software was made. Any sampled PY94 sites showing large differences were identified and reviewed. During this process, a number of database merge errors were identified and corrected.

Technical Degradation

For all sampled sites, technical degradation factors (TDF) contained in the CADMAC report were applied to each measure identified in the building. The technical degradation factors were programmed into the modeling software, allowing efficient automated generation of new DOE-2 models. Each building attribute identified as a measure was modified according to the value of the degradation factor. The CADMAC Statewide Technical Degradation study covered a number of measures applicable to the NRNC program. The full list of measures considered by the CADMAC study, and their applicability to this study is shown in the Table below:

Measure	*Applies to DFE*	*Has Technical Degradation*
Residential Packaged Air-Conditioners		
Commercial Packaged Air Conditioners	✓	
Oversized evaporative condensers for grocery stores		✓
High-efficiency residential refrigerators		
Electronic ballasts	✓	
T-8 lamps and electronic ballasts	✓	
Reflector installation with de-lamping		
Metal halide lighting, 250-400 Watt	✓	✓
Occupancy sensors	✓	
High-efficiency motors	✓	
Adjustable speed drives for HVAC fans	✓	
Infra-red gas fryers		
Residential ceiling insulation		
LED exit signs	✓	
Adjustable speed drives for process pumping		
Adjustable speed drives for injection molding equipment		✓
Residential wall insulation		
Daylight dimming controls	✓	✓
Agricultural irrigation pumps		
VAV systems	✓	
Energy management systems	✓	✓
High-efficiency air compressors		✓
High-efficiency compressed air distribution		✓
Compact fluorescent downlights	✓	

Table 7: Measures where Technical Degradation Applies

Based on the results from the CADMAC study, the only measures that were applicable to the 1994 and 1996 Edison NRNC programs ***and*** shown to have technical degradation were metal halide lighting, daylight dimming controls, and energy management systems. Thus only buildings with these measures were resimulated. A total of 43 buildings out of the sample of 61 had one or more measure with technical degradation.

The TDFs defined in the CADMAC study were derived primarily from engineering studies on the physical causes of measure degradation. The TDF was defined as "a scalar amount to account for the time and use related change in the energy savings of a high efficiency measure or practice relative to a standard efficiency measure or practice." The TDFs are a series of yearly numbers which, when multiplied by the first year savings yield an estimate of the energy savings in years subsequent to the first year.

The savings reported for each participant were estimated on a whole-building level. Since the whole-building savings were made up of the net contributions of all conservation actions above Title 24, it was necessary to disaggregate the savings associated with the three affected measures, and apply the correct TDF to the savings from each of these measures. The process is further complicated by the interactions between measures, since savings of all affected measures taken together is likely to be different from the sum of the individual measure savings.

Due to the complexities of applying the TDFs to simulation ***results***, the general approach taken by the RLW/AEC team was to apply the TDFs to the simulation ***inputs***, and recalculate the savings. From an engineering perspective, this approach is more consistent with the engineering basis of the TDFs, and more straightforward to implement in the simulation model. The approach taken for each affected measure is outlined below:

Metal Halide Lighting. The TDFs for metal halide lighting fixtures were based on an engineering study of the stability of the fixture input power relative to a baseline mercury vapor fixture. The CADMAC study concluded that the input power to a metal halide fixture will increase at a rate of about 0.4% per 1000 hours over the 10,000 hour life of the lamp, while the input power to the baseline fixture will be stable. Based on this conclusion, the input power to a metal halide lamp increases an average of 2% for a lamp with an average age of 5,000 hours.

The CADMAC study reported TDFs in terms of savings rather than input power. The TDFs in the CADMAC report were developed for a specific set of conditions, where a 250 W metal halide fixture replaced a 400 W mercury vapor fixture. In this specific example, the savings degraded an average of 4% over the life of the lamp, thus a TDF of 0.96 was calculated for this technology. Note that the degradation in savings is a function both the increase in lamp watts and the original savings percentage.

In the first-year evaluation of DFE, savings were calculated against the Title 24 allowed lighting power density. The allowed LPD varied as a function of space occupancy type. The savings calculations didn't consider the baseline fixture type, only the difference between the installed and allowed lighting power densities. In the present study, we applied the CADMAC findings for the expected increase in input power rather than the degradation of saving. For both PY94 and PY96, we increased the input wattage of all metal halide fixtures by 2%, and recalculated the savings using the same baseline assumptions as those used in the original evaluation.

Daylighting Controls. In the CADMAC study, the TDFs for daylighting controls were calculated based on an engineering study of failure mechanisms for switched, stepped, and dimming controls. Switched and stepped controls were judged to have no technical degradation. TDFs were established for dimming controls to account for a portion of the controllers failing over time. The failure mechanisms identified for dimming controls caused were expected to cause uneven operation of the system, resulting in bypass of the controls by building occupants, and a reduction in the lighting connected load subject to daylight control.

Our simulation of energy savings from daylighting utilized a DOE-2 “function” to calculate the ratio of the exterior illuminance to the illuminance “seen” by the daylighting sensor. Standard DOE-2 algorithms were used to simulate the action of the control system in response to the interior illuminance levels calculated by the “function.” The fraction of the total lighting load in the daylit space connected to the control system was calculated directly from the onsite survey data.

To simulate the degradation in daylight dimming system performance, the daylighting TDFs were applied to the total lighting connected load controlled by dimming type daylighting controls. Electric lighting controlled by switched or stepped control systems was not affected. For PY96, we used a YDF of 0.73, i.e., we reduced the connected load controlled by dimming type daylighting controls by 27%. For PY94 the assumed TDF was 0.54.

Energy Management Systems*.* Energy management systems (EMS) perform a wide variety of functions, ranging from simple time clock control of HVAC equipment to sophisticated optimization of chiller plants. The original simulation of energy savings from EMS controls depended on the specific control actions taken, and whether or not these actions were required by Title 24. The savings calculations were further limited to control actions simulated by the DOE-2.1E program.

The CADMAC Study considered a wide range of possibilities for degradation in energy savings from EMSs, including sensor reliability, human operation, and maintenance. The TDFs developed for EMSs were not specific to any particular control action, nor were they expressed in any specific engineering terms. For this study, our approach was to simulate the energy savings of the building with and without the specified control actions to disaggregate the energy savings associated with the EMS, then apply the technical degradation factor to the simulated EMS-only savings. For PY96, we used a YDF of 0.8, i.e., we reduced the simulated EMS-only savings by 20%. For PY94 the assumed TDF was 0.4, and the reduction in EMS savings was 60%.

Measure Retention

For the surveyed building, changes in building characteristics identified at the site were implemented in the DOE-2 model. Our modeling software was used to generate revised DOE-2 models based on the observed changes in building characteristics.

Simulation Parametrics

The impacts of technical degradation and measure retention were studied in a series of parametric runs, as described below:

1. ***Impacts of technical degradation***. The first run incorporated the TDF to each measure in all sampled sites, and calculate the change in whole-building energy savings.
2. ***Impacts of building changes.*** The second run examined the impact of building changes on whole-building savings for the surveyed sites only.
3. ***Combined impacts.*** The third run incorporated the TDF into the second run, giving a combined impact of technical degradation and measure retention.

Data Analysis

The data analysis was carried out in the following four steps:

(a) Estimate the current program savings and survival proportion $\hat{S}$ for each of the two program years. Calculate the standard error of $\hat{S}$. Also estimate the current average age *t* of the measures for each program year.

(b) Calculate the ex post estimate of the effective useful life ($E\hat{U}L$) for each program year corresponding to the survival proportion $\hat{S}$.

(c) Calculate the ex anti estimate of effective useful life (EUL_0) for each program year as a weighted average of the assumed EUL of each measure.

(d) Test the null hypothesis that the EUL is equal to the ex-anti estimate EUL_0.

Each of these steps is discussed briefly below.

Estimating survival proportions is a standard application of MBSS™ ratio estimation. For each sample site, the data collection and engineering modeling yields a site-specific estimate of the current savings in energy after adjusting for degradation, persistence or both. This becomes the target y-variable of our analysis. We also know the evaluated first-year savings from our prior evaluation studies. This is the x-variable. The survival proportion is the ratio of the population sum of y to the population sum of x. In other words, the survival proportion is the total program savings after adjusting for degradation, persistence or both, relative to the total first-year evaluated savings.

We estimated this ratio by calculating the weighted sum of y observed in the sample to the weighted sum of x observed in the sample. The weights were calculated to reflect the population of program participants relative to the current sample. MBSS provides the standard error of this estimate, calculated to reflect the weights and the strength of the correlation between current energy savings and first-year energy savings. In most circumstances, we can expect a strong association, so a small sample can be expected to yield a rather precision estimate of the survival proportion. See the discussion of the sample design.

The ex post estimate of effective useful life was calculated from program tracking data. We calculated a weighted average of the assumed measure lives, using the net resource benefit as the weights applied to each category of measure.

The ex-anti estimate of effective useful life was calculated on a whole-building basis for each program year. The starting point was the exponential survival function $S(t) = e^{-lt}$. Here the mean survival time is equal to $1/l$. We defined the EUL as the value of *t* that satisfies the equation: $S(t) = e^{-lt} = 0.5$. Solving for t=EUL, we obtained $EUL = -\frac{\ln(0.5)}{l}$

If we observe $\hat{S}$ in a sample with average measure age t, then we can solve the survival function for $\hat{l} = -\frac{\ln(\hat{S})}{t}$. By substituting this equation in the preceding one, we obtained $E\hat{U}L = -\frac{t\ln(0.5)}{\ln(\hat{S})}$.

This established the relationship between the estimated survival proportion and the estimated EUL. It will be noted from the preceding equation, that the relationship is monotonic increasing. In other words, higher survival proportions give higher values of EUL. This means that hypothesis tests can be carried out either using EUL or S. In the final analysis, we tested the hypothesis comparing the ex anti EUL to the ex post EUL in terms of the associated values of the survival proportion.

Specifically we wanted to test the null hypothesis that the true EUL is equal to the ex post estimate of EUL. We did this by finding the ex post value of S corresponding to the ex post estimate of EUL. Then we tested the hypothesis that the ex anti value of S is equal to the ex post value of S.

It should be emphasized that this approach was on a whole-building basis. All sample results were expanded to the program populations using standard MBSS methodology. Therefore the results were adjusted for the following factors:

1. The stratified sample designs used in the original evaluations
2. The stratified sample design used in the current project
3. Any attrition experienced in the present study
4. Any technical interaction between measures

Findings

Of all 61 sites in the sample, only one met the on-site criteria. One additional building in the 1994 program had been demolished. No on-site was needed to assess the level of savings persistence at that site.

Figure 4 shows how each site responded to the question about modifications to the building. Not surprisingly, PY94 sites had undergone more changes than PY96 buildings, although the vast majority of buildings from both years had not been modified at all.

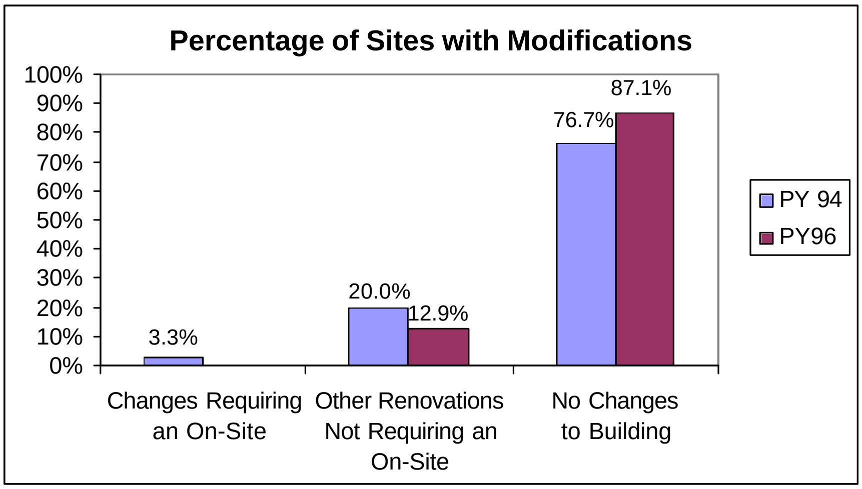

Figure 4 – Percentage of Sites with Modifications

Results for Annual Energy

Table 8 summarizes the observed survival proportions for the savings in annual kWh energy. The table shows the survival proportions for both the 1994 and 1996 programs. Results are given considering technical degradation alone, persistence alone, and technical degradation and persistence taken together.

In the case of the 1994 program, considering only technical degradation, 98.6% of the energy savings of the 1994 program is still being achieved in 1998. In other words, only 1.4% of the original savings has been lost due to technical degradation. Considering only persistence, only 0.5% of the original savings has been lost. Considering both factors, only 1.9% of the original savings has been lost. In the case of the 1996 program, only 2.0% of the savings has been lost to technical degradation. The technical degradation was slightly lower in PY96 than in PY94 because the mix of measures was slightly different. In 1996 we saw no losses due to persistence.

	1994		*1996*	
Category	***Estimate***	***Rel Prec***	***Estimate***	***Rel Prec***
Degradation	98.6%	0.9%	98.0%	0.9%
Persistence	99.5%	0.7%	100.0%	0.0%
Both	98.1%	1.2%	98.0%	0.9%

Table 8: Survival Proportions for Annual Energy

We also calculated degradation and persistence factors by costing period. These results showed the same high levels found in Table 8.

Table 8 also shows the relative precision of each of its estimates. The relative precision reflects the variation due to sampling. For example the relative precision of the 1994 survival proportion due to degradation was 0.9%. This indicates that we would expect the result obtained from the sample to be within ±0.9% of the result that would have been obtained if the same engineering methodology had been applied to every project implemented in the 1994 program.

A 90% confidence interval for the true survival proportion can be calculated as 0.986 ± (0.009) (0.986). In words, considering the potential variation in the results due to sampling, we can be confident that the survival proportion is between 97.6% and 99.5%. Analogous confidence intervals are shown in Table 9 for each of the results reported in Table 8.

	PY94		***PY96***	
Category	***From***	***To***	***From***	***To***
Degradation	97.6%	99.5%	97.2%	98.9%
Persistence	98.8%	100.2%	100.0%	100.0%
Both	96.9%	99.3%	97.2%	98.9%

Table 9: Confidence Intervals for Annual Energy Survival Proportions

The next step in the analysis was to calculate the ex anti estimate of EUL on a whole-building basis. Table 10 shows the calculations for the 1994 program year. The table shows the assumed EUL for each measure as well as the net resource base (NRB) for the program. Edison provided the assumed EUL and the NRB. The NRB was used to calculate the weighted average of the EUL across all categories.

Measure Category	***EUL***	***NRB***
Window Treatment	20	1,176
Chillers	20	3,080
HVAC	15	1,361
Daylighting Controls	10	1,323
Occupancy Sensors	10	441
LPD Reduction	10	4,587
Electronic Ballasts	10	2,236
High Efficiency Motors	15	351
Variable Speed Drives	10	3,896
Energy Management Systems	10	1,953
Performance	10	6,550
Average EUL	**12**	

Table 10: Calculating the Ex Anti EUL for PY94

Table 11 shows the similar calculations for the 1996 program. In this case we have combined measures from both the DFE and NRNC programs.

Measure Category	***EUL***	***NRB***
Window Treatment	20	0
Chillers	20	949
HVAC	15	5
High Efficiency Motors	15	13
Variable Speed Drives	10	667
Daylighting Controls	10	0
Occupancy Sensors	10	0
LPD Reduction	10	0
Electronic Ballasts	10	7
Energy Management Systems	10	0
Performance	10	1,179
Low Shading Coefficent Glazing	10	21
Heat Pump	10	0
HVAC	10	44
HVAC: Energy Reduction	10	927
High Efficiency Motors	15	0
Daylighting Controls	10	742
LPD Reductions	10	1,547
Skylighting w/controls	10	302
T8 w/ Electronic Ballasts	10	478
Design Assistance Incentives	10	279
Average EUL	**11**	

Table 11: Calculating the Ex Anti EUL for PY96

The top portion of Table 12 shows our ex post estimates of the Equivalent Useful Live (EUL) for each case. These results were calculated from the lower estimate of the survival proportion (S) shown in Table 9. Therefore, this is a conservative estimate in terms of sampling variability. Obviously, these estimates cannot be taken literally. It is absurd to believe that half of the first-year savings will last 236 years. This is discussed further below.

The lower portion of Table 12 shows our ex anti estimates of the EUL of each of the two programs. We calculated this as an average of the assumptions made in the programs for each measure, using the net resource benefit of each measure as weight. For example, for the 1994 program we found the ex anti estimate of average EUL to be 12 years.

The table also shows the current survival proportion S that would be expected under the exponential survival function corresponding to the ex anti EUL. Considering the 1994 program as an example, we would expect to have seen 79.4% of the savings after four years, i.e., we would have expected to see a loss of 20.6% of the first-year savings if the actual EUL is 12 years.

A formal hypothesis test can be carried out simply by comparing the ex post and ex anti estimates shown in Table 12. Recall that the ex post estimates shown in Table 12 are the lower bounds of the corresponding 90% confidence intervals. Consider the null hypothesis that the true EUL is less than or equal to the ex anti value. We can reject the null hypothesis at the 5% level of significance as long as the ex anti value is below the lower bound of the ex post confidence interval.

This test leads to the conclusion that the true EUL is greater than the ex anti value. Essentially, these results indicate that the losses due to technical degradation and persistence are currently very low – significantly lower than would be expected under the program assumptions.[1]

	1994		***1996***	
Category	***S***	***EUL***	***S***	***EUL***
Degradation	97.6%	117	97.2%	48
Persistence	98.8%	236	100.0%	na
Both	96.9%	88	97.2%	48
Ex Anti Value	**79.4%**	**12**	**88.2%**	**11**

Table 12: Effective Useful Life for Annual Energy

Results for Peak Demand

Table 13 summarizes the observed survival proportions for the savings in peak kW demand. The table shows the survival proportions for both the 1994 and 1996 programs. Results are given considering technical degradation alone, persistence alone, and technical degradation and persistence taken together.

Take the case of the 1994 program. Considering only technical degradation, 99.5% of the demand savings of the 1994 program were still being achieved in 1998. In other words, only 0.5% of the original savings has been lost due to technical degradation. Considering only persistence, only 0.4% of the original savings has been lost. Considering both factors, only 0.9% of the original savings has been lost. In the 1996 program, only 0.8% of the original savings has been lost to technical degradation. In 1996, we saw no losses due to persistence.

	1994		***1996***	
Category	***Estimate***	***Rel Prec***	***Estimate***	***Rel Prec***
Degradation	99.5%	0.4%	99.2%	0.3%
Persistence	99.6%	0.6%	100.0%	0.0%
Both	99.1%	0.8%	99.2%	0.3%

Table 13: Survival Proportions for Peak Demand

[1] The difference between the 1994 and 1996 estimates of EUL shown in Table 12 is an artifact of the exponential survival model. In fact, most of the loss of savings is due to technical degradation of metal halide lighting fixtures which was assumed to be a one-time lose of 2%.

Table 13 also shows the relative precision of each of its estimates. The relative precision reflects the variation due to sampling as discussed above. The corresponding confidence intervals are shown in Table 14.

	Interval		***Interval***	
Category	***From***	***To***	***From***	***To***
Degradation	99.1%	100.0%	98.9%	99.5%
Persistence	99.0%	100.2%	100.0%	100.0%
Both	98.4%	99.9%	98.9%	99.5%

Table 14: Confidence Intervals for Peak Demand Survival Proportions

Table 15 shows our estimates of the Equivalent Useful Live for each case. These results were calculated from the lower estimate of the survival proportion shown in Table 14. Therefore, this is a conservative estimate in terms of sampling variability. As previously pointed out, these estimates cannot be taken literally. However, as before, these results indicate that the losses due to technical degradation and persistence are currently very low – significantly lower than would be expected under the program assumptions.

	1994		***1996***	
Category	***S***	***EUL***	***S***	***EUL***
Degradation	99.1%	302	98.9%	126
Persistence	99.0%	283	100.0%	na
Both	98.4%	170	98.9%	126

Table 15: Effective Useful Life for Peak Demand

Discussion of the Results

It is important to consider the possibility that our methodology may be giving a biased result. What were the key underlying assumptions in the study? Could these assumptions be false?

The following are the principal assumptions:

1. **Sampling Bias:** We assumed that our sample of 61 sites was a proper, stratified sample of the program participants.
2. **Response Bias:** We assumed that the removal of buildings or measures was accurately described in our telephone survey, i.e., that our respondents provided accurate information.
3. **Technical Degradation:** We restricted our analysis of technical degradation to factors that were considered in the Statewide Technical Degradation Study.
4. **Engineering Modeling:** We used engineering models to determine the impact of technical degradation and removal for the sample sites.

5. **Exponential Survival Function:** We assumed that the survival function is exponential.

We will discuss each assumption in turn.

Sampling Bias. Sampling bias can arise in this type of study if the sample is not actually selected following the planned sample design. For example, it may be necessary to choose substitutes if some of the specified sample refuse to participate in the survey. Bias can arise if there is a systematic difference between those who cooperate and those who refuse.

Fortunately, we were able to follow the sample design very closely in this study. The specified sample was randomly chosen from among the projects that we had studied in the first-year evaluations. Therefore we had good information about who to interview in the present telephone survey, and often an established relationship. Moreover the survey was very short and easy. So there were no refusals. Bias could have arisen indirectly from refusals in the sampling for the original first-year impact-evaluation study, but the refusal rates were also low in the first year studies.

One adjustment was required in the 1994 portion of the present sample. The current sample design reflected the complete population of 1994 program participants. When the current sample was selected, we found that the first-year sample contained fewer projects in the higher strata than called for by the sample design. To complete the sample, additional projects were randomly selected from lower strata. Therefore the final sample had fewer large projects than specified in the original sample.

In the analysis, new weights were calculated to adjust for the final sampling distribution. Because the selection was still random, this revision of the sample design is not believed to have biased the sample. This problem was not encountered with the 1996 sample since the original sample was stratified by savings.

Response Bias: In analyzing the sample data we assumed that the removal of buildings or measures was accurately described in our telephone survey, i.e., that our respondents gave us accurate information. As already indicated, in most cases we surveyed the respondents to the prior study, and we had established relationships with many of the respondents. Therefore we received excellent cooperation. Moreover the survey asked for simple, factual information that would be easily known by the respondent. The survey was carried out by the same experienced person who did the 1996 survey. Where appropriate, she used followup questions or probes to resolve an answer that was incomplete or unclear. Finally, we know of no motivation for the respondent to distort the truth. Therefore we see no reason to expect that the information was biased.

Technical Degradation: The accuracy of our results for technical degradation rests on the accuracy and completeness of the Statewide Study as well as the relevance of the Statewide Study to our sample projects. We have no reason to doubt the validity of the Statewide Study.

Following the evaluation guidelines established by CADMAC, we restricted our analysis of technical degradation to those factors that were considered in the Statewide Technical Degradation Study. In applying the Statewide Study, we took into account the expected degradation after two years for the 1996 program and after four years for the 1994 program. In the case of measures that were not considered in the Statewide Technical Degradation Study, no degradation was assumed.

Engineering Modeling: The DOE-2 simulation models of each sample site that had been used in the first-year evaluation were modified to reflect technical degradation and any removal of measures found in the onsite surveys. Each DOE-2 model was run with and without the technical degradation and compared to the baseline model. This gave an estimate of both the original first-year savings and the current annual savings for each sample site considering technical degradation, removals, or both. Given the policy of relying on the Statewide Technical Degradation Study, we believe that this approach is sound and appropriate.

Exponential Survival Function: We assumed that the survival function is exponential. This is equivalent to assuming that a fixed proportion of the current savings is lost each year. An assumption such as this is necessary to estimate the EUL since the survival proportions observed in the first two or four years must be extrapolated out into the future.

In actuality, the assumption might be wrong. The proportion of the savings lost each year may accelerate as the buildings and equipment grow older. In fact, the estimates of EUL given by the exponential survival model are unreasonably long. This suggests that the exponential model may not be correct. *However it is simply too soon to identify a more accurate survival model than the exponential model.*

As we have discussed, we can conclude with statistical confidence that the EUL is longer than the program assumptions. However, this conclusion is conditional on the validity of the exponential survival model. Suppose for example, that after 5 years the measures begin to fail at a very high rate. Under these circumstances our conclusion that the EUL is longer than 12 years would be in error. However, the present study shows no reason to expect that this might happen.

To summarize, there is always the possibility of flaws in any study. However, we know of no special circumstances in this study that might be of concern. On the contrary, the study was unusually free of problems and we are confident in concluding that the survival rates are higher than the ex anti program estimates.

Methodological Lessons Learned

There has been some uncertainty about the most appropriate methodology for measuring persistence in nonresidential new construction. Since this type of study is relatively new, somewhat different methodologies have been suggested.

Some have suggested the application of the conventional statistical survival analysis developed for analyzing clinical trials in the medical field. However it is difficult to apply these techniques to buildings since whole-building savings do not usually 'die' but rather they gradually degrade. While it might be postulated that individual measures or 'bundles' of savings do die, it is difficult to justify the underlying assumption that the failures are statistically independent, especially when they are in the same building. Moreover, in nonresidential new construction it is best to measure savings on a whole-building basis to reflect interaction between measures. So conventional statistical survival analysis does not seem to be a fruitful approach in nonresidential new construction.

Others might conclude that the sample size in the present study was too small for definitive results. Certainly, one can never predict absolutely what would have happened with a larger sample. However, we have measured and reported the sampling variability of our estimates. The estimated relative precision was excellent, and all findings were consistent. This indicates that the findings were not subject to excess sampling variation. In other words, we would have expected essentially the same results if we had studied every project in each of the two programs. Instead, we would emphasize that it is too early to tell what might eventually happen in these programs.

It has also been pointed out that the initial persistence studies must establish whatever foundation is necessary for future persistence studies. In carrying out the present study, we took great advantage of the detailed onsite audit data and engineering models that had been developed for the 133 sites in the first-year evaluation studies. In future studies, we would not hesitate to include all 133 sites if it seemed to be necessary. However, it would be extremely difficult to increase the sample beyond 133 sites since it would be necessary to carry out a detailed onsite audit for each added site.

Some might have concern about revisiting the present sample of 61 sites in future studies. On the contrary, our experience suggests that we would receive good cooperation from the sites in the current study as well as the remaining sites in the original sample. We get excellent cooperation with the telephone survey as long as the interviewer is intelligent and knowledgeable, and the survey is short and asks for information that is known to the respondent.

In conclusion, we believe that the methodology we followed establishes a sound foundation for future persistence studies. In essence, the first-year evaluation studies created a panel of sites with extensive technical characteristics, detailed simulation models, and well-identified, cooperative respondents. Our combination of site-specific, semi-automated engineering simulation and statistical extrapolation provides a straightforward, cost-effective and defensible set of tools for mining this information in future studies.

Appendix A

Table 6. Protocols for Reporting of Results of Required Studies

1. **Identify the studied measures and the end use it belongs to**: All measures that received rebates through the program were studied. The impact of the measures was studied on a whole-building basis.
2. **Identify the ex-ante expected useful live and the source**. The ex-ante expected useful life was found to be 12 years for PY94 and 11 years for PY96. These were calculated on a whole-building basis, by weighting the assumed EUL of each measure by the net resource benefit. See Tables 10-11 of the present report.
3. **Identify the** ex-post expected useful life estimated in the study. Considering both degradation and persistence we found EULs of 144 years for PY94 and 70 years for PY96. These are the point estimates calculated assuming the exponential survival model. These results differ from the results in Table 12
4. **Identify the ex-post expected useful life to be used by the utility in the third and fourth earnings claims.** Since there was no statistically significant lose of savings, the third and fourth earnings claims will be based on the ex-ante EULs by measure category.
5. **Identify the standard error of the ex-post EUL**. Assuming the exponential survival model, the relative precision for EUL is approximately the same as the relative precision for the survival proportion. Considering both technical degradation and persistence, the estimated EUL is shown in the following table together with the relative precision. This reflects both technical degradation and persistence and is based on the exponential survival model applied to the survival proportions shown in Table 8. We calculated the standard errors shown below by dividing the relative precision by 1.645 (reflecting the 90% level of confidence) and multiplying the result by the EUL Estimate.

Program	EUL	Rel Prec	Standard	80% Conf Inter.	
Year	Estimate	at 90%	Error	From	To
1994	144	1.2%	1.08	143	146
1996	70	0.9%	0.37	69	70

6. **Provide the 80% confidence interval for the ex-post EUL**. We multiplied the standard errors by 1.28 reflecting the 80% level of confidence interval, and calculated the corresponding confidence interval shown above.
7. **Provide the associate p-value**. The z-value for 1994 is (144 – 12)/1.08 = 122. The associate p-value is 0.0000. For 1996, the z-value is 159 and the p-value is 0.0000.

8. **Provide the realization rate for the adopted ex-post expected useful life, i.e., the ratio between the adopted ex-post EUL and the ex-ante EUL**. Since there was no statistically significant lose of savings, the third and fourth earnings claims will be based on the ex-ante EULs by measure category. Therefore the realization rate is one.
9. **Identify all "like" measures associated with the studied measure**. Since the evaluation was on a whole-building basis, this is not applicable.

Table 7. Documentation Protocols for Data Quality and Processing

Overall Information

a) **Study Title and Study ID No**.

b) **Program and program year**: Nonresidential new construction, 1994 and 1996 program years.

c) **End Uses and Measures**: Whole building, all rebated measures.

d) **Methods and Models Used**: Standard stratified ratio estimation was used to estimate the survival proportion of savings for each program year. The exponential survival model was used to estimate the effective useful life from the observed survival proportion. No other model was tried.

e) **Analysis sample size**: 30 sites from PY94, 31 sites from PY96. The survival proportions were estimated from collected during 1998 and from the first-year evaluation studies done in 1995 and 1997.

Database Management

a) Initial observations of persistence were determined from a telephone survey, supported by an onsite audit if the respondent indicated that installed measures had been removed or disabled. The estimates of technical degradation were determined by the technical degradation factors identified in the CADMAC state-wide study, and were applied to the engineering models developed in the first-year impact evaluation studies.

b) No data points were excluded.

c) We used the files from the first-year evaluation studies to identify the sites and respondents for the survey.

d) All data collected in this study was used in the analysis.

Sampling

a) The sampling procedure is described in detail in the report. In summary, a model-based analysis was used to choose the sample size, and model-based methods were used to stratify the sample by first-year savings. The sampling frame was the sites evaluated in the first year impact evaluation, weighted up to the full population of program participants. The sample was designed to achieve ± 20% relative precision at 80% confidence, assuming 80% persistence of savings. In fact, the persistence of savings was virtually 100% and the relative precision was about ± 1% at 90% confidence.

b) The survey instrument is in the appendix. The response rate was 100% so there was no response bias.

c) Not applicable since there was no statistical modeling.

Data Screening and Analysis

a) Not applicable since there was no statistical modeling.

b) The engineering models held weather, hours of occupancy and operating schedules fixed to the levels assumed in the first-year impact evaluations.

c) There was no screening.

d) Not applicable since there was no statistical modeling.

e) The only specification was the exponential survival model. This model was not statistically tested since we lacked data over enough time to identify the true survival model.

f) Considerable care was taken to make the engineering simulations as accurate as possible and to reflect the technical degradation following appropriate engineering principles.

g) Not applicable since there was no statistical modeling.

h) No missing data.

i) The standard errors for the survival proportions were calculated using standard stratified ratio estimation methods. Using a Taylor's Theorem expansion, the relative precision of the EUL was shown to be approximately equal to the relative precision of the survival proportion.