SOUTHERN CALIFORNIA EDISON 1996 NON-RESIDENTIAL NEW CONSTRUCTION EVALUATION

FINAL REPORT

Study ID Number 543

February 18, 1998

Prepared for:
Andrea Horwatt
Southern California Edison
300 N. Lone Hill Ave
San Dimas, CA 91773

Prepared By:
RLW Analytics
1055 Broadway, Suite G
Sonoma, CA 95476

SCE-03

Table of Contents

Executive Summary

Introduction

This is the final report of the 1996 Non-Residential New Construction (NRNC) Program evaluation. The evaluation was conducted by RLW Analytics and Architectural Energy Corporation from May 1997 through December 1997.

This report details findings of energy and demand savings at the whole building level and for lighting, HVAC, and shell & daylighting end-uses. Both net and gross savings are presented.

The evaluation relied on the use of model-based statistical sampling, on-site engineering surveys, DOE –2.1 building simulation models, and econometric analysis to develop the findings presented. A sample of 73 participant buildings and 81 non-participant buildings were surveyed and modeled to estimate gross energy savings relative to a baseline level. An additional telephone survey was conducted with decision-makers to collect data to estimate free-ridership and spillover. Net savings were developed using logistic and linear regression modeling to predict efficiency choice in the absence of the program.

The 1996 evaluation benefited greatly from the project team's experience with the 1994 PG&E / SCE NRNC evaluation. Valuable lessons were learned during the 1994 evaluation that helped to refine the methodology used in this study. Four key refinements to the 1996 study were:

- An improved sample design stratified by the estimated energy savings of participants
- The use of DOE modelers to conduct the on-site surveys
- The development of the initial model shortly after the survey visit
- The introduction of scaled variables in the econometric analysis

A brief overview of the 1996 evaluation methodology appears below.

Study Design

The goal of this evaluation was to estimate the net and gross energy and demand savings of the 1996 nonresidential new construction program.

The primary deliverables of this evaluation were:

1. Gross savings estimates of annual energy and summer peak demand
2. Net savings estimates of annual energy and summer peak demand
3. Gross savings of lighting, HVAC, and shell / daylighting end-uses.

The RLW Analytics/AEC team used a methodology similar to the 1994 study, with important modifications to reflect what was learned from that study. The basic approach relied on engineering models to develop gross savings estimates and econometrics to determine the net-to-gross ratio. This methodology conforms to the CADMAC protocols with the important exception that statistical sampling was used in the place of an attempted census of program participants.

On August 20, 1997 CADMAC approved a waiver for this change in methodology.

The study was carried out in three phases – design, data collection, and data analysis – plus reporting. Each phase builds on the results of the previous phase. Figure 1 shows the major tasks for this project and their relationships.

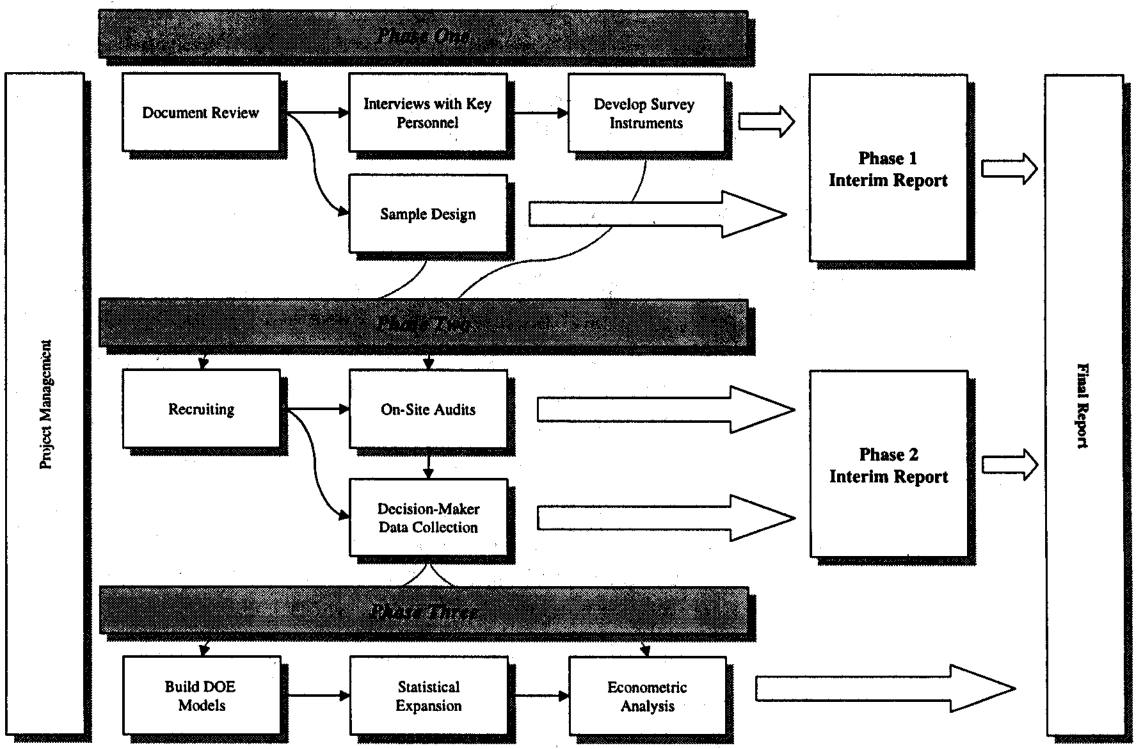


Figure 1: Study Flowchart

Data Collection

A major portion of this project was the collection of the building and decision-maker data necessary to determine the program impacts. Overall, the data collection process ran quite smoothly - no problems were encountered that had an adverse impact on the overall quality of the data. The data collection process used in this study yielded significantly better information than the process used in the 1994 NRNC study.

The data collection process was designed to collect the highest quality data in the most efficient manner possible. This process relied on several people working together to ensure a seamless information flow. Figure 2 shows a graphical representation of the data collection process.

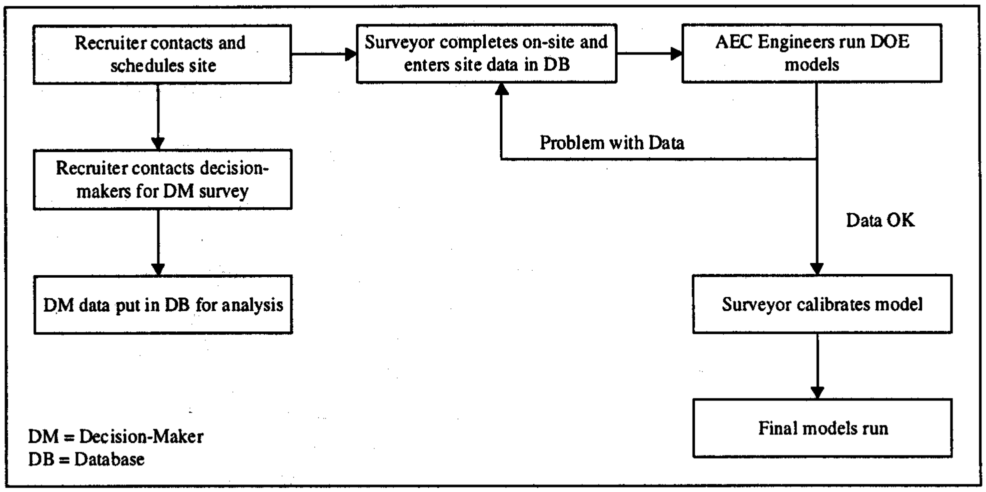


Figure 2: Data Collection Process

The recruiter was responsible for making contact with the site and securing its participation in the study. Once that was accomplished, the recruiter scheduled the on-site visit and provided the information to the field surveyors from RLW Analytics and AEC. The recruiter then completed the decision-maker survey with the initial site contact and any additional contacts that were necessary to answer the decision-maker questions.

The on-site surveyor collected building description and operation information from the site and entered the data into a database. Automated modeling software was used to create DOE-2 input files. The surveyors were responsible for checking the models created from the field data, and correcting the data if necessary. The on-site surveyor was also responsible for calibrating the model to billing data or short-term meter data, if available for the site. Senior staff engineers of AEC and RLW checked the final model results for reasonableness.

The calibrated models were delivered to AEC, who produced all of the required parametric runs of the engineering models.

Engineering Models

Engineering models were developed for each building in the on-site survey sample using the DOE-2.1E building simulation program. A series of models were developed for each sample site, including:

- A "baseline" model representing the building with minimally compliant equipment and envelope efficiencies.
- An as-built model representing the building as found by the surveyors.
- A series of parametric runs to isolate the impact of HVAC, lighting, and shell / daylighting end-uses.

The models were developed using an automated BDL[1] generator, developed by AEC and RLW Analytics. This method ensured that all of the models were consistent, thus eliminating a potential source of bias in the results.

Analysis Baseline and Gross Savings Calculations

The estimates of gross program savings were made by comparing the as-built simulated building energy consumption to a baseline level of energy consumption. The baseline energy consumption for all buildings was defined to be the energy consumption of the building as if all of the equipment was specified to be minimally compliant with Title 24 and the building was operated on the schedule found during the on-site survey. *Because the default Title 24 operating schedules were not used to develop the baseline and because the area category method was used for each building regardless of the Title 24 compliance path actually elected, the savings calculated relative to the baseline in this study cannot be interpreted as the degree of compliance with Title 24.*

A gross savings estimate was calculated for each building in the sample. The savings estimated were projected to the population of participants using model-based statistical sampling procedures. Gross savings estimates were developed for both the participant and the non-participant population.

Net Savings Methodologies

Net program savings estimates are the savings that directly result from program participation. Effects of free-ridership, or what the customer would have done anyway, have been factored out. Two net savings methodologies were used in this evaluation, a "difference of differences" approach and an econometric approach. Net-to-gross ratios from both methods are presented in this report.

Difference of Differences

A simple "difference of differences" estimation approach to net savings was done for this study. This method estimated net savings by comparing the savings of the participants in the sample to a "matched" sample of non-participants. The savings of the non-participant group is assumed to be the savings of the participants in the absence of the program. In this methodology, spillover among the non-participants is assumed to be offset by free-ridership among the participants but no attempt is made to measure either spillover or free-ridership

Econometric Modeling

An econometric approach to estimating net savings was also used in this study. The econometric approach appeared to provide a more unbiased and statistically reliable estimate of net savings than the difference of difference approach because it explicitly measured both free-ridership and spillover and controlled for self-selection and other decision-making factors affecting the efficiency choice of each sample site. Figure 3 shows the overall flow of data for the econometric modeling. In this methodology, a logistic regression was performed to create a participation model. This model estimated Mills' ratios for correcting

[1] BDL is DOE-2's Building Description Language

self-selection bias. A second model was built, a linear regression, to estimate the savings of participants in the absence of the program. The econometric approach also incorporated the relationship between SCE's influence on the design of projects and the energy efficiency of the current project. This component of the model was used to estimate the spillover effect, i.e., the effect of the program on non-participant savings.

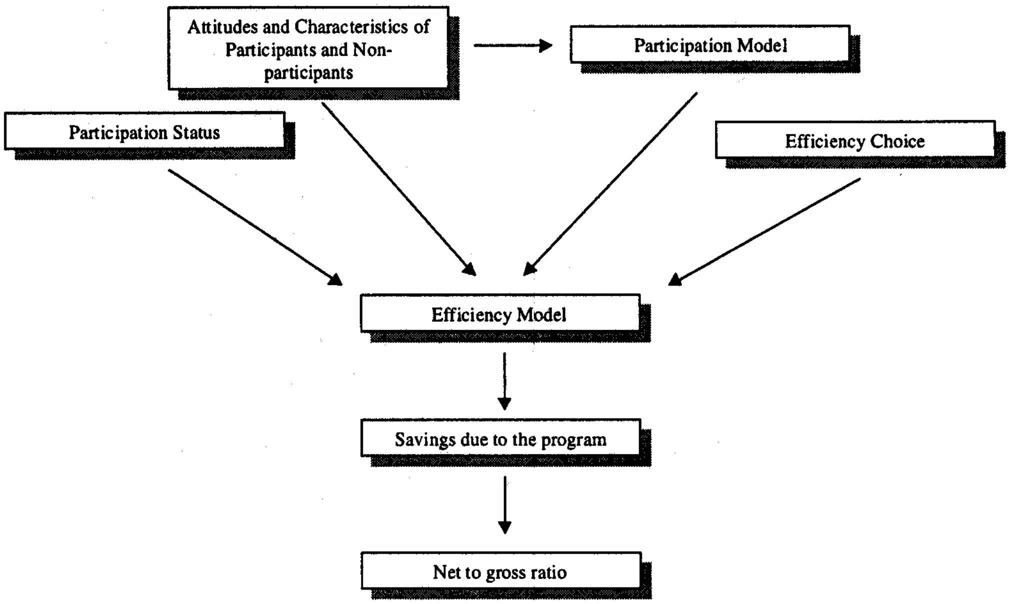


Figure 3: Econometric Modeling Overview

Findings

This section presents gross and net savings estimates for the population of program participants.

Gross Savings

Program participants saved 42,730 MWh of energy in their first year of operation. This is a realization rate of 116.4% of the verified savings estimate. The relative precision of the estimate is ±8.8% at the 90% confidence level, meaning that the gross program savings is estimated to be between 38,969 MWh and 46,490 MWh.

The summer on-peak demand savings is 10.13 MW. The realization rate is 115.0% of the verified program savings. The relative precision is ±11.0% at the 90% confidence level, meaning that the gross program demand savings is between 9.02 MW and 11.24 MW. Table 1 shows the energy and demand savings by SCE time-of-use period.

Period	Energy Savings (MWh)	Energy Rel. precision	Demand Savings (MW)	Demand Rel. precision
Annual	42,730	± 8.8%	--	--
Summer On-Peak	4,196	± 9.6%	10.13	± 11.0%
Summer Mid-Peak	4,679	± 8.7%	9.73	± 11.1%
Summer Off-Peak	6,838	± 9.5%	9.53	± 11.1%
Winter Mid-Peak	14,220	± 9.4%	9.59	± 10.4%
Winter Off-Peak	12,800	± 11.2%	9.25	± 10.7%

Table 1: Participant Energy and Demand Gross Savings by Time-of-use period

To compare participants and non-participants, the savings of each group relative to their own baseline is plotted in Figure 4. The figure clearly shows much higher levels of energy efficiency among participants than among non-participants. The participants' energy use was 21.4% better than baseline, while the non-participants' energy use was only 8.1% better than baseline. "Better than baseline" means that the buildings are more energy efficient than the baseline efficiency levels established for this study. Numerically, a building that is 20% better than baseline uses 20% less energy than it would have used if built to baseline efficiency levels. For summer on-peak demand, the participant group was 20.2% better than baseline while the non-participant group was 9.7% better than baseline.

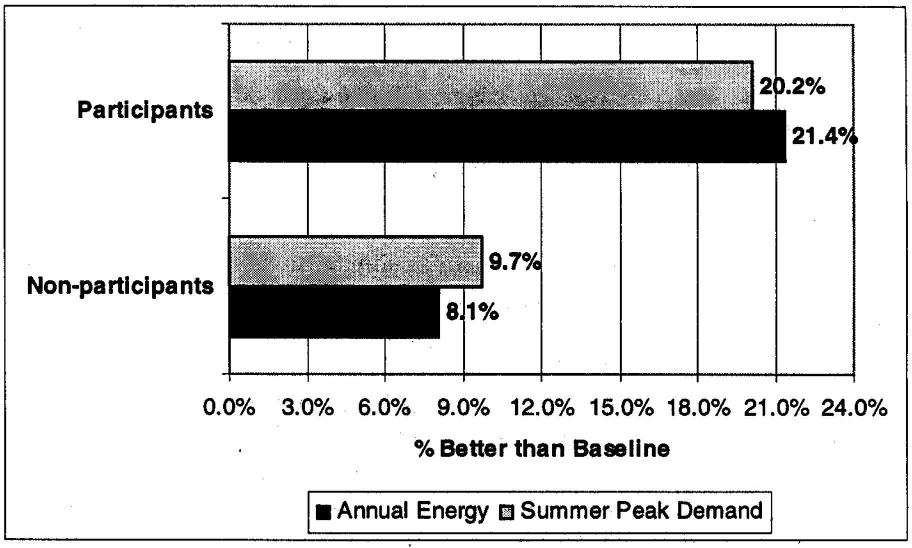


Figure 4: Gross Energy and Demand Savings Relative to Baseline

Energy and demand savings were also estimated for lighting, shell/daylighting, and HVAC end-uses. Figure 5 shows the composition of the annual energy savings and the summer on-peak demand savings for program participants.

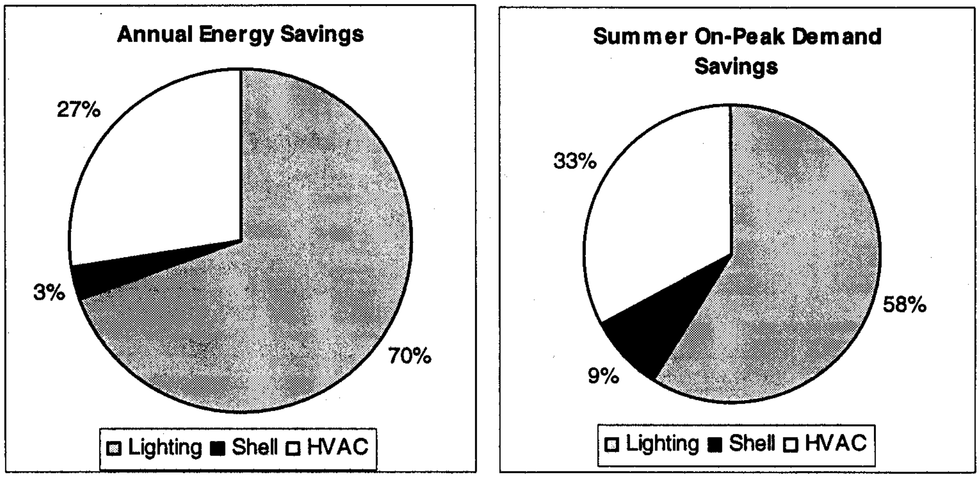


Figure 5: Composition of Gross Savings

Table 2 shows the energy savings by end-use for each of the time-of-use periods. Table 3 shows the summer on-peak demand savings for each end-use.

	Lighting	Shell / Daylighting	HVAC
Annual	29,580	1,378	11,720
Summer on-peak	2,746	251	1,191
Summer mid-peak	3,136	252	1,282
Summer off-peak	4,539	284	2,007
Winter mid-peak	10,290	271	3,635
Winter off-peak	8,865	319	3,602

Table 2: End-Use Gross Energy Savings by Time-of-use period (MWh)

	Lighting	Shell / Daylighting	HVAC
Summer on-peak	5.94	0.86	3.30
Summer mid-peak	5.48	0.90	3.33
Summer off-peak	5.53	0.77	3.21
Winter mid-peak	5.51	0.87	3.19
Winter off-peak	5.32	0.68	3.23

Table 3: End-Use Gross Demand Savings by Time-of-use period (MW)

Net Savings

As discussed in a prior section, two different methodologies were followed in the net-to-gross analysis: a relatively simple difference of differences approach and a more complex econometric approach. In the difference of differences methodology, the net-to-gross ratio was calculated by comparing (a) the gross savings relative to baseline of the program participants and (b) the gross savings relative to baseline of the non-participants. In the econometric approach, the net-to-gross ratio was calculated by using regression modeling techniques to estimate the net savings due to the program for each of the program participants and non-participants.

Difference of Differences Net-to-Gross

The following table summarizes the findings from the difference of differences analysis. The table shows the estimated net-to-gross ratio for both annual energy and summer peak demand savings.

	Net Savings	Net-to-gross Ratio	Relative Precision
Annual Energy	26,621 MWh	62.3%	± 22.0%
Summer Peak Demand	5.27 MW	52.0%	± 24.9%

Table 4: Difference of Differences Net-to-gross Ratio

The table also shows the relative precision of each estimate.[2] For example, in the case of annual energy, the net-to-gross ratio was estimated to be 62.3% with a relative precision of ±22.0%. The error bound for the 90% confidence interval for the true net-to-gross ratio is equal to 22.0% of the estimate, i.e. to ±13.7%. The 90% confidence interval for the true net-to-gross ratio can be calculated using the equation:

$$0.623 \pm 0.623 \times 0.220 = (0.486, 0.760)$$

We can be quite confident that this interval contains the net-to-gross ratio that would have been obtained by developing onsite surveys and building engineering simulation models for all program participants and a very large sample of non-participants using the methodology of this study and then analyzing the resulting data using the difference of differences methodology. The confidence interval reflects sampling variability and random measurement error but does not reflect any possible systematic measurement error that might be repeated throughout the data collection and engineering simulation or that might arise by neglecting explicit estimation of free ridership and spillover.

Econometric Net-to-Gross

The following table summarizes the findings from the econometric analysis. The table shows the estimated savings and net-to-gross ratio for both annual energy and summer peak demand savings.

[2] Some definitions: The standard error reflects the standard deviation of an estimate in repeated sampling. The error bound at the 90% level of confidence is 1.645 times the standard error. The confidence interval is the estimate plus or minus the error bound. The relative precision is the error bound divided by the estimate itself.

	Savings	Net to Gross	Rel Prec
Annual Energy	41,005 MWh	96.0%	±30.5%
Summer Peak Demand	10.818 MW	106.8%	±21.5%

Table 5: Econometric Net-to-Gross Ratios

The table also shows the relative precision of each estimate. For example, in the case of annual energy, the savings was estimated to be 41,005 MWh and the net-to-gross ratio was estimated to be 96.0%, both with a relative precision of ±30.5%.

Table 6 shows conservative estimates of the saving and net to gross ratios, obtained by discounting the point estimates to reflect their relative precision.[3] There is a 90% probability that the conservative estimates are less than or equal to the true values that would have been obtained by developing onsite surveys and building engineering simulation models for all program participants and a very large sample of non-participants using the methodology of this study and then analyzing the resulting data using the econometric methodology. The conservative estimates reflect sampling variability, random measurement error, and explicit estimation of free-ridership and spillover. But these estimates do not reflect any possible systematic measurement error that might be repeated throughout the data collection and engineering simulation or any possible bias arising from inaccuracy in the assumed econometric model.

	Savings	Net to Gross
Annual Energy	31,273 MWh	73.2%
Summer Peak Demand	9.008 MW	88.9%

Table 6: Conservative Estimates

Spillover

Statistically significant non-participant spillover was found in the econometric analysis. The analysis indicated that for non-participants in the current program, the degree of influence of SCE on the projects was significantly related to the efficiency of their projects. In other words, decision makers in the non-participant sample who reported being strongly influenced by SCE tended to have significantly more efficient projects than those who were not influenced. For participants, on the other hand, there was no statistically significant relationship, perhaps because of less variation on the amount of influence reported by the participants.

[3] The conservative estimate is equal to the point estimate – 1.28 * the standard error of the point estimate. This is the lower bound of a one-sided confidence interval at the 90% level of confidence.

Table 7 shows the added efficiency due to the program implied by the econometric analysis. The variable across the top of the table represents the level of influence of SCE on the design and equipment specified for the project. The values in the table show the increase in expected efficiency due to the program for non-participants as a function of the degree of influence, holding other factors fixed. For example, if a non-participant reported strong influence from SCE, the building tended to be 10.2 percent points more efficient than a similar non-participant who reported low influence. By contrast, if the site was a participant, then it tended to be 13.2 percent points more efficient than a similar non-participant who reported low influence. These results are for annual energy, but the results for peak summer demand were similar.

These results indicates that the program had two impacts. First the program had a direct net impact on the participants. Second, the program had an indirect or spillover impact on the non-participants.

Current Influence	1 (low)	2	3	4	5	6	7 (High)
Energy Efficiency	0.0%	1.7%	3.4%	5.1%	6.8%	8.5%	10.2%

Table 7: Annual Energy Efficiency of Non-Participants

Sample Design

Introduction

The key to effective sample design is to take advantage of the association between the target variables to be measured in the study and any supporting variables already known from the sampling frame. For example, the savings of each program participant measured in this project can be associated with the estimate of savings recorded in the program tracking system. Stratified sampling is used to ensure that the sample has the best mix of small and large sites. Ratio estimation is used to expand the sample data to the target population, taking advantage of the supporting information. Both stratified sampling and ratio estimation are well known and widely used in load research and DSM evaluation.

The principal questions addressed in sample design are:

- How big should the sample be, both overall and within different subsets of the target population?
- How much statistical precision can we expect from the sample?
- How should the sample be stratified to get the best statistical precision?

The usual approach is to estimate the variance of the estimated savings in the program tracking system. This approach is not appropriate for stratified ratio estimation since the statistical precision depends not on the variance of estimated savings but on the strength of the association between the measured savings and the tracking estimate of savings. The Model-Based Statistical Sampling (MBSS) approach is to develop a statistical model describing the relationship between these variables, and then use the parameters of this model to develop the sample design. In this project the parameters of the MBSS model were estimated in our prior evaluation of the 1994 program.

Using this approach, RLW Analytics designed the participant sample to achieve ±10 percent precision at the 90 percent confidence level for the participants' annual measured energy savings. This analysis indicated that the participant sample size should be 72 sites, stratified by the tracking estimate of savings. The non-participant sample was matched to the participant population in terms of square footage and building type. A sample of 80 non-participant sites was selected from F.W. Dodge New Construction data.

Participants

RLW Analytics used the sites that received incentive checks dated in 1996 as a participant sample frame. A sample of 72 sites was drawn from a population of 133. The sample was stratified into 5 sampling strata and one certainty strata for a total of 6 strata by estimated annual energy savings. Sample size, population size, and stratum cutpoints are indicated in the table below.

Stratum	Maximum Energy Savings (kWh)	Population Size	Population Energy Savings (kWh)	Sample Size
1	46,582.	45.	801,079.	10.
2	85,095.	24.	1,636,352.	10.
3	163,928.	18.	2,261,959.	10.
4	314,316.	13.	3,316,179.	10.
5	490,821.	11.	4,376,334.	10.
6	6,000,000.	22.	28,099,069.	22.
Total		**133.**	**40,490,972.**	**72.**

Table 8: Stratified Sampling Plan for Participants

The total tracking savings for the 133 program participants was 40,490 MWh[4]. The anticipated precision from this sample design was ± 9.7 percent at 90 percent confidence. The estimated precision for participants was based on the model parameters used in the sample design, which are shown in Table 9.

Model Parameter	Value
error ratio	1.02
γ	0.44

Table 9: Model-Based Sampling Parameters for Participant Sample

The error ratio and γ were taken from the actual model parameters found in the 1994 NRNC study. The analysis variable is the actual energy saved and the explanatory variable is the tracking estimate of energy saved. The error ratio is a measure of the spread of the data around the trendline. It is analogous to the coefficient of variation. γ is a measure of the heteroskedastisity of the data. Heteroskedastisity is the tendency for the variation around the trendline to increase as the value of the stratification variable increases.

Non-participants

For the non-participant sample design, the *participant* population was re-stratified on building type and square footage. This two-way stratification defined the cells in the sample design, which was then filled with non-participant sites from the Dodge database. This procedure ensured that the non-participant sample could be well matched to the participant sample. Later in this section, a comparison between the participant and non-participant population is shown.

The sample frame for the non-participants was taken from the F.W. Dodge new construction database. The database was screened to eliminate out-of-scope and out-of-territory projects. The Dodge project was considered in scope if the building type was eligible for NRNC incentives.

[4] The list of 133 program participants with 40,490 MWh of savings was used for the sample design. Later review indicated that 2 sites should not have been in the dataset. All savings estimates are based on the corrected dataset and verified savings estimates.

The non-participant sample was developed using the method outlined in the flowchart below. This led to a non-participant sampling frame of 2,438 sites.

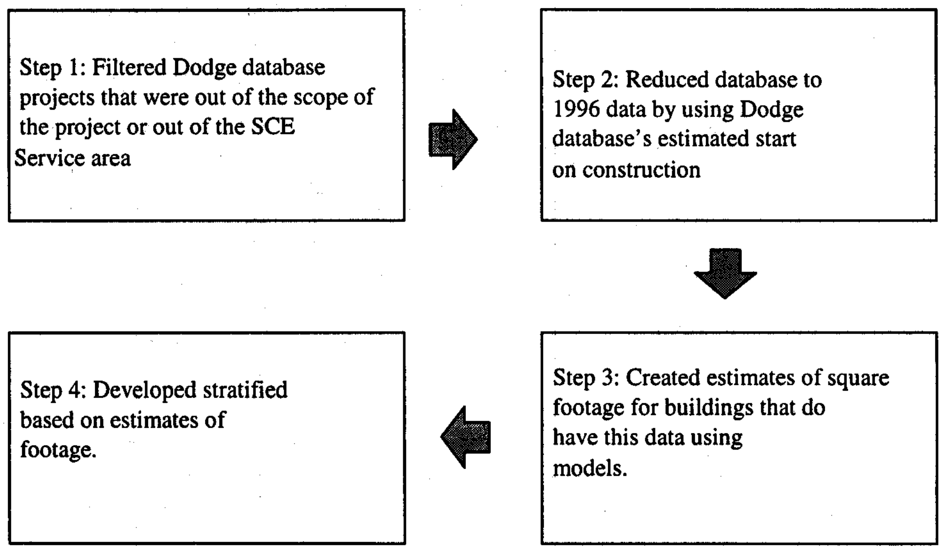


Figure 6: Non-Participant Sample Frame Development

The non-participant sample size was chosen to be 80 sites to approximately match the participant sample size. The non-participant sample was stratified by building type and by square footage. Table 10 below summarizes the sample design used to select the 80 non-participants. For example, in the case of food stores, 2 sites were selected from a single size stratum. This is equivalent to no stratification by size. By contrast, 21 offices were selected from three size strata. The number of sizes from each building type and the allocation of the sample to the size strata was selected to match the participant population. In Table 10 and Table 11, a dash in the cell indicates that the data element is not applicable to that building type. For example, there was only 1 food store stratum, therefore there was no strata 2 or strata 3 sample (Table 10) and there were no strata cutpoints (Table 11).

Building Type	Stratum 1	Stratum 2	Stratum 3	Total
College	2	2	--	4
Food Store	2	--	--	2
Hospital	2	--	--	2
Medical	2		--	2
Manufacturing	2	2	2	6
Miscellaneous	4	3	3	10
Office	7	7	7	21
Restaurant	2	2	2	6
Retail	7	7	7	21
School	2	2	--	4
Warehouse	2	--	--	2
Total	**34**	**26**	**20**	**80**

Table 10: Stratified Sampling Plan for Non-Participants

The square footage cutpoints for the non-participant strata are shown in Table 11. For example, in the college category, stratum 1 consists of sites with square footage less than 73,000 square feet, and stratum 2 of larger sites. Warehouses greater than 225,000 square feet were excluded from the non-participant sample because there were no warehouses in the participant population of that size.

Building Type	Stratum 1 Max. Square Footage	Stratum 2 Max. Square Footage
College	73,000	--
Food Store	--	--
Hospital	--	--
Medical	--	--
Manufacturing	165,500	258,000
Miscellaneous	45,590	102,728
Office	33,216	121,500
Restaurant	47,217	104,000
Retail	90,800	133,998
School	73,000	--
Warehouse	--	--

Table 11: Strata Cutpoints

Sample Design vs. Actual Sample

Table 12 shows the participant sample design and the actual participant sample. As the table shows, fewer than desired large customers (higher strata numbers) were successfully recruited.

Stratum	Design	Actual
1	10	15
2	10	12
3	10	10
4	10	11
5	10	8
6	22	17
Total	**72**	**73**

Table 12: Participant Sample Design and Actual Sample

There was no stratification of the participant sample by building type. Table 13 shows the expected distribution of the participant sample by building type and the actual distribution of the participant sample. The distributions have been weighted by their inclusion probability, to reflect the fact that a particular participant building type would have tended to be in larger or smaller savings strata.

Participant Sample Actual vs. Expected		
Category	**Expected %**	**Actual %**
College	9%	11%
Food store	1%	1%
Hospital	3%	1%
Medical	2%	1%
Mfg	9%	7%
Miscellaneous	19%	18%
Office	15%	15%
Restaurant	4%	4%
Retail	29%	29%
School	5%	8%
Warehouse	3%	4%
Total	100%	100%

Table 13: Expected Vs. Actual Participant Sample by Building Type

Table 14 shows the sample design and the actual non-participant sample by building type and square footage strata. To be read, the first number in a cell is the achieved sample in the cell. The second number is the designed sample. So, 3 stratum 1 food stores were surveyed and the sample design called for 2 to be surveyed. In Table 14, stratum 1 consists of the smallest buildings, stratum 3 consists of the largest buildings, in square footage terms.

Filling some of the cells proved difficult due to differences in the participant and non-participant population. The largest projects were generally program participants, thus making it difficult or impossible to find equally large non-participants for the comparison sample. As a result, the smaller size strata were typically overfilled and the larger size strata were under-filled.

Category	Stratum 1	Stratum 2	Stratum 3	Total
College	0 of 2	1 of 2	--	1 of 4
Food Store	3 of 2	--	--	3 of 2
Hospital	2 of 2	--	--	2 of 2
Medical	2 of 2	--	--	2 of 2
Mfg	5 of 2	0 of 2	1 of 2	6 of 6
Miscellaneous	7 of 4	2 of 3	2 of 3	11 of 10
Office	6 of 7	12 of 7	2 of 7	20 of 21
Restaurant	7 of 2	0 of 2	0 of 2	7 of 6
Retail	12 of 7	1 of 7	6 of 7	19 of 21
School	7 of 2	1 of 2	--	8 of 4
Warehouse	2 of 2	--	--	2 of 2
Total				**81 of 80**

Table 14: Non-participant Sample by Building Type and Size Strata

Figure 7 shows the non-participant sample design and the actual non-participant sample by building type. The figure shows that the non-participant design was fairly well filled with respect to building type.

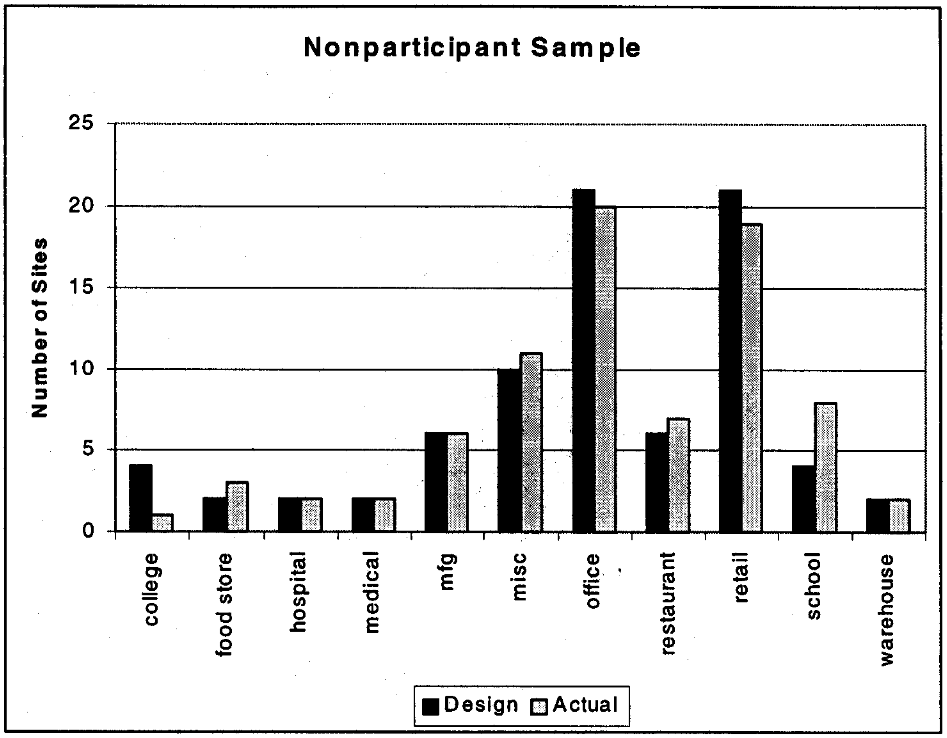


Figure 7: Non-participant Sample by Building Type

Differences Between the Participant and Non-participant Populations

The non-participant sample was chosen to be representative of the participant population to facilitate comparisons between the participant and non-participant groups. The figure below shows the true distribution by building type of the participant and non-participant populations. This figure suggests that the distribution of the participant population is not the same as the non-participant population. Because the non-participant sample was designed to be representative of the participant population, the non-participant sample is ***not*** representative of the non-participant population and therefore ***should not be used to draw general inferences about the non-participant population***.

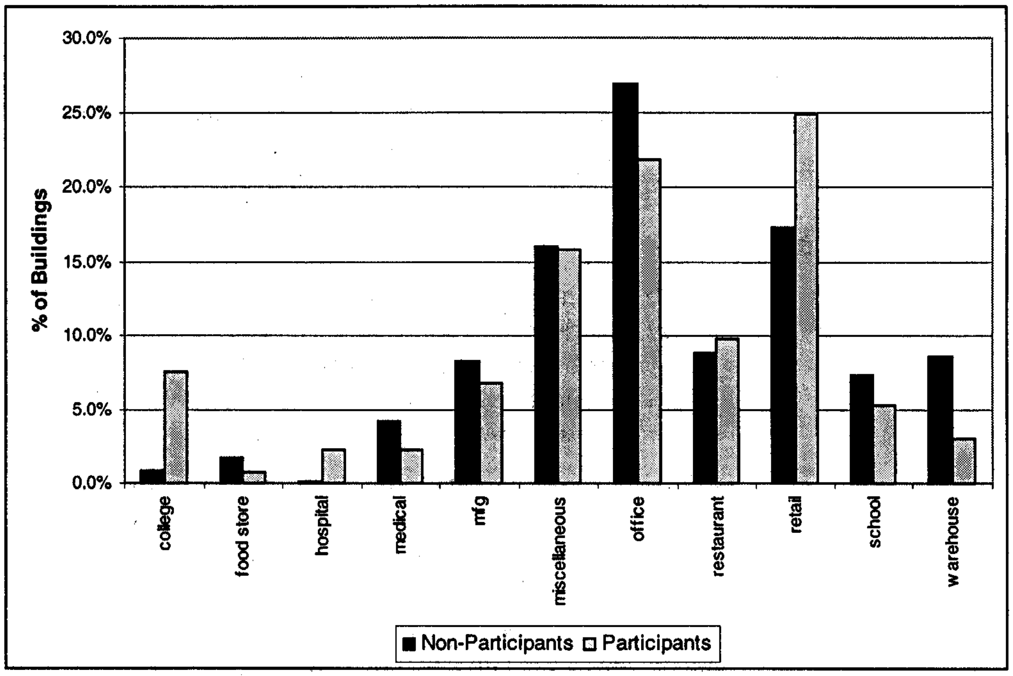


Figure 8: Distribution of Building Types in the Participant and Non-participant Populations

Colleges, hospitals, and retail buildings are over-represented in the participant population compared to new construction in general. Office buildings and warehouses are under-represented.

Data Collection

The data collection effort was one of the largest portions of the project. Twelve on-site surveyors worked with a recruiter for about 10 weeks to collect on-site and telephone survey data on 154 buildings.

Recruiting

A single recruiter was used in the study. Special effort was made to use staff that was experienced in construction and development in order to ensure that the professionals being contacted did not feel that they were speaking with someone who did not understand the basic issues in the field. The approach proved to be a tremendous success.

Table 15 summarizes the recruiting effort. A conversion rate of 55% was achieved. Only 6% refused to participate in the study. This is a reflection of both the effectiveness of the recruiter and the fairly good reputation enjoyed by Edison in this market.

In the table, completed means that the site was successfully recruited and audited. "No contact" means that attempts to contact a decision-maker at the site failed. Dropped indicates that the site was eliminated for one or more of the following reasons:

- The building was not completed and occupied in 1996
- The building could not have qualified for the program
- The building participated in the new construction program in other years.
- The sampling stratum had been filled before the site was recruited

"Terminate in progress" indicates that the site was dropped for one or more of the above reasons after the on-site surveyor learned that the site was outside the scope of the study.

Disposition	Participants	Non-Participants	Total
Completed	73	81	154
Refused	5	11	16
No Contact	1	8	9
Dropped	8	86	94
Terminated in Progress	2	5	7

Table 15: Recruiting Disposition

Decision Maker Surveys

The decision-maker surveys were completed for each audited site. Decision-makers were those individuals involved in the design and construction of the project who could influence the energy efficiency decisions made. This could have been the owners, the developers, the architects, or the engineers. This was done by the recruiter, who made an average of 3.4 calls to 2 different individual decision-makers to complete each survey. Figure 9 shows the distribution of the number of calls necessary to complete each survey.

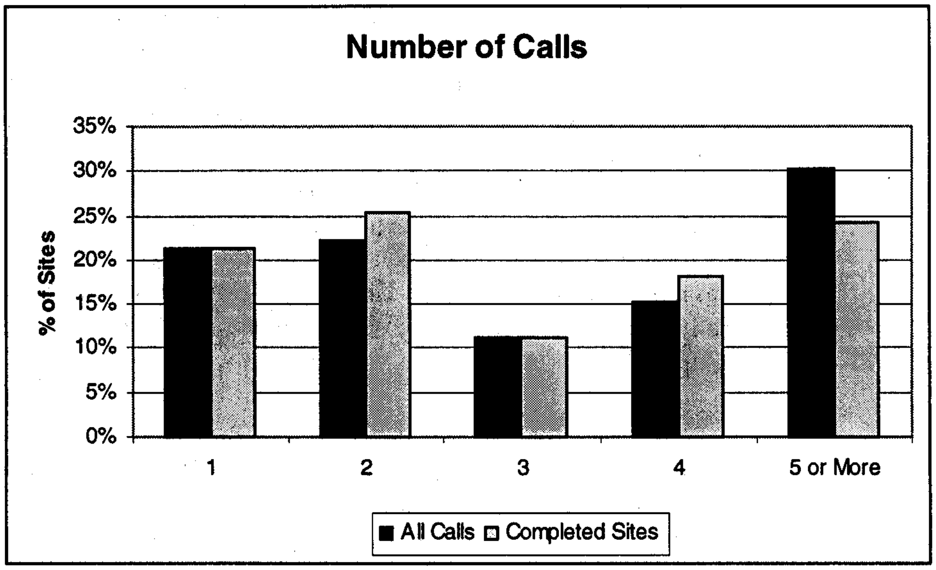


Figure 9: Number of Calls to Complete Each Decision-maker Survey

Figure 10 shows the distribution of the number of individuals that were contacted for the decision-maker surveys. The maximum number of individuals required to complete a survey was 3.

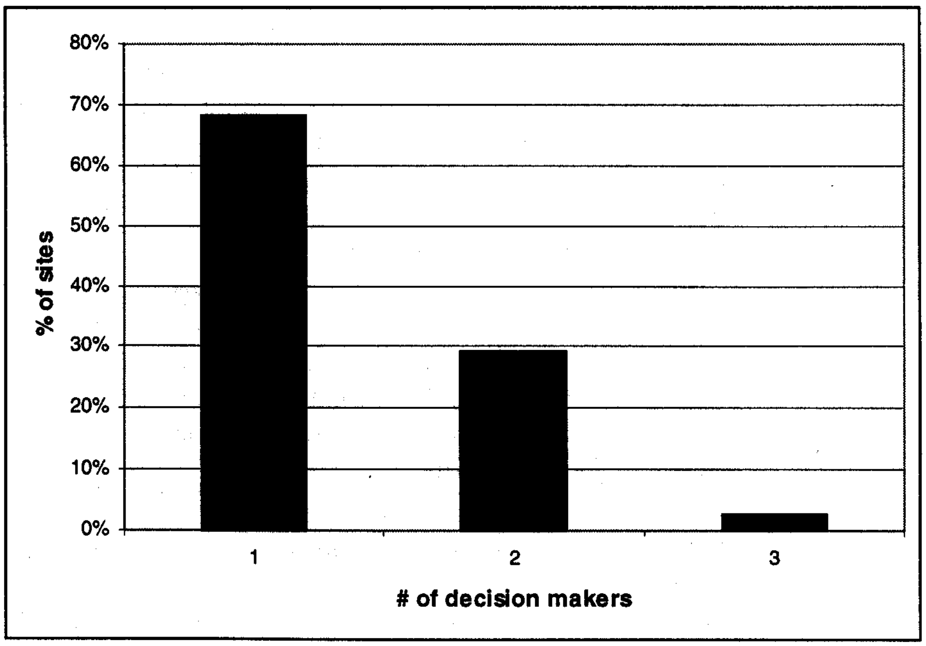


Figure 10: Number of Decision-makers Surveyed

On-Site Surveys

The primary data source for the DOE-2 models was the on-site survey. The survey form was designed so that key modeling decisions on model zoning and equipment/space association were made by the surveyors in the field. The form was designed to follow the logical progression of an on-site survey process. The form started out with a series of interview questions. Conducting the interview first helped orient the surveyor to the building and allowed time for the surveyor to establish a rapport with the customer. Once the interview was completed, an inventory of building equipment was conducted. The survey started with the HVAC systems, and progressed from the roof and/or other mechanical spaces into the conditioned spaces. This progression allowed the surveyor to establish the linkages between the HVAC equipment and the spaces served by the equipment.

Interview Questions.

The interview questions were used to identify building characteristics and operating parameters that were not observable by the surveyor during the course of the on-site survey. The interview questions covered the following topics:

Building functional areas. Functional areas were defined on the basis of operating schedules. Subsequent questions regarding occupancy, lighting, and equipment schedules, were repeated for each functional area.

Occupancy history. The occupancy history questions were used to establish the vacancy rate of the building during 1996. The questions covered occupancy, as a percent of total surveyed floor space, and HVAC operation during the tenant completion and occupancy of the space. Responses to these questions were used to understand building start-up behavior during the model calibration process.

Occupancy schedules. For each functional area in the building, a set of questions were asked to establish the building occupancy schedules. First, each day of the week was assigned to one of three daytypes: full occupancy, partial occupancy, and unoccupied. This was to cover buildings that did not operate on a normal Monday through Friday work week. Holidays and monthly variability in occupancy schedules were identified.

Daily schedules for occupants, interior lighting, and equipment/plug loads. A set of questions was used to establish hourly occupancy, interior lighting, and miscellaneous equipment and plug load schedules for each functional area in the building. Hourly schedules were defined for each daytype. A value, which represents the fraction of the maximum occupancy and/or connected load was entered for each hour of the day. The entry of the schedule onto the form was done graphically.

Daily schedules of kitchen equipment. A set of questions were asked to establish hourly kitchen equipment schedules for each functional area in the building. Hourly schedules were defined for each daytype. A value which represented the equipment-operating mode (off, idle, or low, medium or high volume production) was entered for each hour of the day. The entry of the schedule onto the form was done graphically.

Operation of other miscellaneous systems. General questions on the operation of exterior lighting systems, interior lighting controls, window shading, swimming pools, and spas were covered in this section.

Operation of the HVAC systems. A series of questions were asked to construct operating schedules for the HVAC systems serving each area. Fan operating schedules, and heating and cooling setpoints was entered. Additional questions were used to define the HVAC system controls. The questions were intended to be answered by someone familiar with the operation of the building mechanical systems. The questions covered operation of the outdoor air ventilation system, supply air temperature controls, VAV system terminal box type, chiller and chilled water temperature controls, cooling tower controls, and water-side economizers.

Building-wide water use. A series of questions were used to help calculate the service hot water requirements for the building.

Refrigeration system. The operation of refrigeration systems utilizing remote condensers, which are common in groceries and restaurants, was covered in this section. The systems were divided into three temperature classes, (low, medium and high) depending on the compressor suction temperature. For each system temperature, the refrigerant, and predominant defrost mechanism was identified. Overall system controls strategies were also covered.

Building Characteristics

The next sections of the on-site survey covered observations on building equipment inventories and other physical characteristics. Observable information on HVAC systems, building shell, lighting, plug loads, and other building characteristics were entered, as described below:

Built-up HVAC systems. Make, model number, and other nameplate data were collected on the chillers, cooling towers, heating systems, air handlers, and pumps in the building. Air distribution system type, outdoor air controls, and fan volume controls were also identified.

Packaged HVAC systems. Equipment type, make, model number, and other nameplate data were collected on the packaged HVAC systems in the building.

Zones. Based on an understanding of the building layout and the HVAC equipment inventory, basic zoning decisions were made by the surveyors according to the following criteria:

- ***Unusual internal gain conditions***. Spaces with unusual internal gain conditions, such as computer rooms, kitchens, laboratories were defined as separate zones.
- ***Operating schedules***. Occupant behavior varies within spaces of nominally equivalent use. For example, retail establishments in a strip retail store may have different operating hours. Office tenants may also have different office hours.
- ***HVAC system type and zoning***. When the HVAC systems serving a particular space were different, the spaces were sub-divided according to HVAC system type. If the space was zoned by exposure, the space was

surveyed as a single zone, and a "zone by exposure" option was selected on the survey form.

For each zone defined, the floor area and occupancy type was recorded. Enclosing surfaces were surveyed, in terms of surface area, construction type code, orientation, and observed insulation levels. Window areas were surveyed by orientation, and basic window properties were identified. Interior and exterior shading devices were identified. Lighting fixtures and controls were identified and inventoried. Miscellaneous equipment and plug loads were also inventoried. Zone-level HVAC equipment, such as baseboard heaters, fan coils, and VAV terminals were identified and entered on the form.

Refrigeration systems. Refrigeration equipment was inventoried separately, and associated with a particular zone in the building. Refrigerated cases and stand-alone refrigerators were identified by case type, size, product stored, and manufacturer. Remote compressor systems were inventoried by make, model number, and compressor system type. Each compressor or compressor rack was associated with a refrigerated case temperature loop and heat rejection equipment such as a remote condenser, cooling tower, and/or HVAC system air handler. Remote condensers were inventoried by make, model number, and type. Nameplate data on fan and pump horsepower were recorded. Observations on condenser fan speed controls were also recorded.

Cooking equipment. Cooking equipment was inventoried separately and associated with a particular zone in the building. Major equipment was inventoried by equipment type (broiler, fryer, oven, and so on), size, and fuel type. Kitchen ventilation hoods were inventoried by type and size. Nameplate data on exhaust flowrate and fan horsepower were recorded. Each piece of kitchen equipment was associated with a particular ventilation hood.

Hot water/Pools. Water heating equipment was inventoried by system type, capacity, and fuel type. Observations on delivery temperature, heat recovery, and circulation pump horsepower were recorded. Solar water heating equipment was inventoried by system type, collector area, and collector tilt and storage capacity. Pools and spas were inventoried by surface area and location (indoor or outdoor). Filter pump motor horsepower was recorded. Pool and spa heating systems were inventoried by fuel type. Surface area, collector type, and collector tilt angle data for solar equipment serving pools and/or spas was recorded.

Miscellaneous exterior loads. Connected load, capacity, and other descriptive data on elevators, escalators, interior transformers, exterior lighting, and other miscellaneous equipment were recorded.

Meter Numbers. Additional data were collected in the field to assist in the billing data account matching and model calibration process. This section served as the primary link between the on-site survey and billing data for non-participants. Meter numbers were recorded for each meter serving the surveyed space. If the meter served space in addition to the surveyed space, the surveyor made a judgment on the ratio of the surveyed space to the space served by the meter.

Establishing Component Relationships

In order to create a DOE-2 model of the building from the various information sources contained in the on-site survey, relationships between the information contained in the various parts of the survey needed to be established. In the interview portion of the form, schedule and operations data were cataloged by building functional area. In the equipment inventory section, individual pieces of HVAC equipment: boilers, chillers, air handlers, pumps, packaged equipment and so on were inventoried. In the zone section of the survey, building envelope data, lighting and plug load data, and zone-level HVAC data were collected. The following forms provided the information needed by the software to associate the schedule, equipment, and zone information.

System/Zone Association Checklist. The system/zone association checklist provided a link between each building zone and the HVAC equipment serving that zone. Systems were defined in terms of a collection of packaged equipment, air handlers, chillers, towers, heating systems, and pumps. Each system was assigned to the appropriate thermal zones in accordance with the observed building design.

Interview "Area" / Audit "Zone" Association Checklist. Schedule and operations data gathered during the interview phase of the survey were linked to the appropriate building zone. These data were gathered according to the building functional areas defined previously. Each building functional area could contain multiple zones. The association of the functional areas to the zones, and thereby the assignment of the appropriate schedule to each zone was facilitated by this table.

Short-term Metering

As a part of the overall modeling process, the DOE-2 simulations were calibrated to billing data. In order for a comparison between simulated electricity consumption and billing data to be meaningful, there needs to be a good match between the surveyed space and the space served by Edison meter. At selected sites where the surveyed space and the metered space did not match, short-term metering equipment was installed. An example of such a mismatch is a major tenant improvement or tenant finish in a multi-tenant building, where the Edison revenue meter serves the entire space. Short-term metering equipment was installed on the circuits feeding the surveyed space only, thus serving as a temporary "proxy" meter for the surveyed and modeled space. This data was then used to calibrate the DOE-2 model for the site, instead of billing data.

During the on-site survey, the surveyors collected meter number information, and assessed the match between the space served by the meter(s) and the surveyed space. In situations were a poor match was evident, the surveyors assessed the feasibility of installing short-term metering equipment. The electrical panels serving the surveyed space were identified during the on-site survey. Sites with fairly "clean" circuitry, allowing metering with one or two watt transducers at the whole-panel or switchgear level were identified. If the site appeared to be a reasonable candidate, the surveyor recruited the site contact for short-term metering.

An electrical contractor was dispatched to install the watt transducers on the circuits or panels identified by the surveyor soon after the completion of the on-site survey. The data loggers collected "whole-building" or "whole-space" 5-minute interval demand data were averaged to hourly data and then used to calibrate the DOE-2 models.

A total of 27 sites were initially identified by the surveyors as short-term metering candidates. Twelve sites were successfully recruited for short-term metering. The remaining sites were dropped from short-term metering because of:

- Customer refusal
- Building circuiting not amenable to short-term metering of surveyed space
- Very small load limited value of data (one site)
- Poor system operation limited value of data (one site)

Of the twelve sites scheduled, installation was successful at seven sites. Reasons for dropping sites during installation were:

- Unsafe installation (2 sites)
- Insufficient contractor insurance (1 site)
- Mixed circuits (1 site)
- Sub-metered data available (1 site)

Data gathered for the metered sites were used to calibrate the models. Lighting and occupancy schedules were inferred from the time-series profiles of the metered data. An example of a time-series plot is shown in Figure 11.

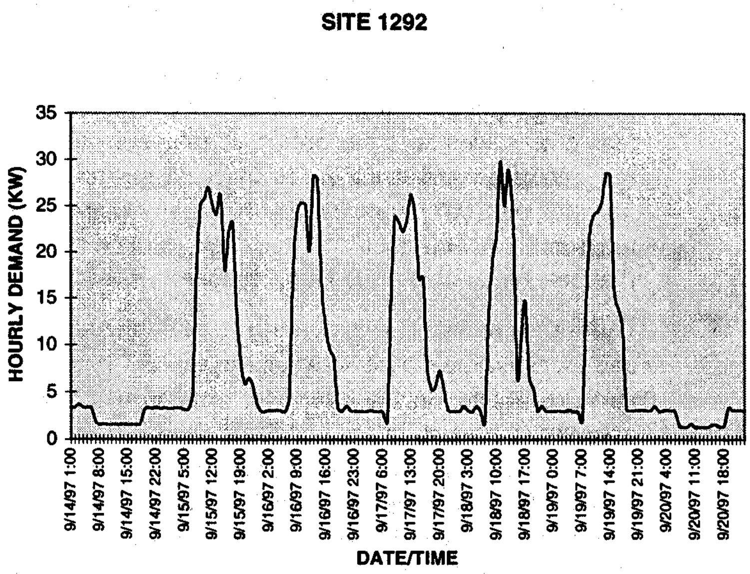


Figure 11: Short Term Monitored Time Series Data

The short-term metered data were also compared to the models on a temperature response basis. Daily average electricity consumption was plotted against daily average temperature for the metered and simulated data. The models were calibrated to match the metered data, as shown in Figure 12:

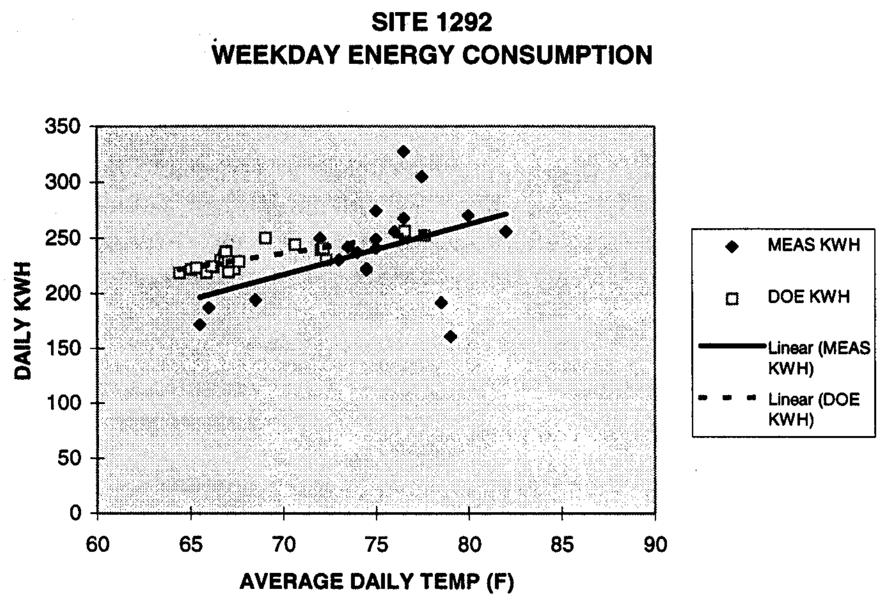


Figure 12: Short Term Temperature Response Comparison

Engineering Models

An automated process was used to develop basic DOE-2 models from data contained in the on-site surveys, Title 24 compliance forms, program information and other engineering data. The modeling software took information from these data sources and created a DOE-2 model. The data elements used, default assumptions, and engineering calculations are described for the Loads, Systems, and Plant portions of the DOE-2 input file as follows.

Loads

Schedules were created for each zone in the model by associating the zones defined in the on-site survey with the appropriate functional area, and assigning the schedule defined for each functional area to the appropriate zone. The software created hourly schedules on a zone-by-zone basis for:

- Occupancy
- Lighting
- Electric equipment
- Gas equipment (primarily kitchen equipment)
- Solar glare
- Window shading
- Infiltration

Occupancy, lighting, and equipment schedules. Each day of the week was assigned to a particular daytype, as reported by the surveyor. Hourly values for each day of the week were extracted from the on-site database according to the appropriate daytype. These values were modified on a monthly basis, according to the monthly building occupancy history.

Solar and shading schedules. The use of blinds by the occupants was simulated by the use of solar and shading schedules. The glass shading coefficient values were modified to account for the use of interior shading devices.

Infiltration schedule. The infiltration schedule was established from the fan system schedule. Infiltration was scheduled "off" during fan system operation, and was scheduled "on" when the fan system was off.

Shell materials. A single-layer, homogeneous material was described which contains the conductance and heat capacity properties of the composite wall used in the building. The thermal conductance and heat capacity of each wall and roof assembly was taken from the Title 24 documents, when available. If the Title 24 documents were not available, default values for the conductance and heat capacity were assigned from the wall and roof types specified in the on-site survey, and the observed R-values. If the R-values were not observed during the on-site survey and the Title 24 documents were not available, an "energy-neutral" approach was taken by assigning the same U-value and heat capacity for the as-built and Title 24 simulation runs.

Windows. Window thermal and optical properties from the building drawings or Title 24 documents (when available) were used to develop the DOE-2 inputs. If these documents were not available, default values for the glass conductance were assigned according to the glass type specified in the on-site survey. If the glass type was not observed during the on-site survey and the Title 24 documents were not available, an "energy-neutral" approach was taken by assigning the same U-value and shading coefficient for the as-built and Title 24 simulation runs.

Lighting kW. Installed lighting power was calculated from the lighting fixture inventory reported on the survey. A standard fixture wattage was assigned to each fixture type identified by the surveyors. Lighting fixtures were identified by lamp type, number of lamps per fixture, and ballast type as appropriate.

Lighting controls. The presence of lighting controls was identified in the on-site survey. For occupancy sensor and lumen maintenance controls, the impact of these controls on lighting consumption was simulated as a reduction in connected load, according to the Title 24 lighting control credits. Daylighting controls were simulated using the "functions" utility in the load portion of DOE-2. Since the interior walls of the zones were not surveyed, it was not possible to use the standard DOE-2 algorithms for simulating the daylighting illuminance in the space. A daylight factor, defined as the ratio of the interior illuminance at the daylighting control point to the global horizontal illuminance was estimated for each zone subject to daylighting control. Typical values for sidelighting applications were used as default values. The daylight factor was entered into the function portion of the DOE-2 input file. Standard DOE-2 inputs for daylighting control specifications were used to simulate the impacts of daylighting controls on lighting schedules. The default daylight factors were adjusted during model calibration.

Equipment kW. Connected loads for equipment located in the conditioned space, including miscellaneous equipment and plug loads, kitchen equipment and refrigeration systems with integral condensers were calculated. Input data were based on the "nameplate" or total connected load. The nameplate data were adjusted using a "rated-load factor," which is the ratio of the average operating load to the nameplate load during the definition of the equipment schedules. This adjusted value represented the hourly running load of all equipment surveyed. Equipment diversity was also accounted for in the schedule definition.

For the miscellaneous equipment and plug loads, equipment counts and connected loads were taken from the on-site survey. When the connected loads were not observed, default values based on equipment type were used.

For the kitchen equipment, equipment counts and connected loads were taken from the on-site survey. Where the connected loads were not observed, default values based on equipment type and "trade size" were used. Unlike the miscellaneous plug load schedules, the kitchen equipment schedules were defined by operating regime. An hourly value corresponding to "off", "idle", or "low," "medium," or "high" production rates were assigned by the surveyor. The hourly schedule was developed from the reported hourly operating status and the ratio of the hourly average running load to the connected load for each of the operating regimes.

For the refrigeration equipment, refrigerator type, count, and size were taken from the on-site survey. Equipment observed to have an "integral" compressor/condenser that is, equipment that rejects heat to the conditioned space, were assigned a connected load per unit size.

Source input energy. Source input energy represented all non-electric equipment in the conditioned space. In the model, the source type was set to natural gas, and a total input energy was specified in terms of Btu/hr. Sources of internal heat gains to the space that were not electrically powered include kitchen equipment, dryers, and other miscellaneous process loads. The input rating of the equipment was entered by the surveyors. As with the electrical equipment, the ratio of the rated input energy to the actual hourly consumption was calculated by the rated load factor assigned by equipment type and operating regime.

Heat gains to space. The heat gains to space were calculated based on the actual running loads and an assessment of the proportion of the input energy that contributed to sensible and latent heat gains. This in turn depended on whether or not the equipment was located under a ventilation hood.

Spaces. Each space in the DOE-2 model corresponded to a zone defined in the on-site survey. In the instance where the "zoned by exposure" option was selected by the surveyor, additional DOE-2 zones were created. The space conditions parameters developed on a zone by zone basis were included in the description of each space. Enclosing surfaces, as defined by the on-site surveyors, were also defined.

Systems

This section describes the methodology used to develop DOE-2 input for the systems simulation. Principal data sources include the on-site survey, Title 24 documents, manufacturers' data, and other engineering references as listed in this section.

Fan schedules. Each day of the week was assigned to a particular daytype, as reported by the surveyor. The fan system on and off times from the on-site survey was assigned to a schedule according to daytype. These values were modified on a monthly basis, according to the monthly HVAC operating hour adjustment. The on and off times were adjusted equally until the required adjustment percentage was achieved. For example, if the original schedule was "on" at 6:00 hours and "off" at 18:00 hours, and the monthly HVAC adjustment indicated that HVAC operated at 50% of normal in June, then the operating hours were reduced by 50% by moving the "on" time up to 9:00 hours and the "off" time back to 15:00 hours.

Setback schedules. Similarly, thermostat setback schedules were created based on the responses to the on-site survey. Each day of the week was assigned to a particular daytype. The thermostat setpoints for heating and cooling, and the setback temperatures and times were defined according to the responses. The return from setback and go to setback time was modified on a monthly basis in the same manner as the fan-operating schedule.

Exterior lighting schedule. The exterior lighting schedule were developed from the responses to the on-site survey. If the exterior lighting was controlled by a time clock, the schedule was used as entered by the surveyor. If the exterior lighting was controlled by a photocell, a schedule, which follows the annual variation in daylength, was used.

System type. The HVAC system type was defined from the system description from the on-site survey. The following DOE-2 system types were employed:

- Packaged single zone (PSZ)
- Packaged VAV (PVAVS)
- Packaged terminal air conditioner (PTAC)
- Water loop heat pump (HP)
- Evaporative cooling system (EVAP-COOL)
- Central constant volume system (RHFS)
- Central VAV system (VAVS)
- Central VAV with fan-powered terminal boxes (PIU)
- Dual duct system (DDS)
- Multi-zone system (MZS)
- Unit heater (UHT)
- Four-pipe fan coil (FPFC)

Packaged HVAC system efficiency. Manufacturers' data were gathered for the equipment surveyed based on the observed make and model number. A database of equipment efficiency and capacity data was developed from an electronic version of the ARI rating catalog. Additional data were obtained directly from manufacturers' catalogs, or the on-line catalog available on the ARI website (www.ari.org). Manufacturers' data on packaged system efficiency is a net efficiency, which considers both fan and compressor energy. DOE-2 requires a specification of packaged system efficiency that considers the compressor and fan power separately. Thus, the manufacturers' data were adjusted to prevent "double-accounting" of fan energy, according to the procedures described in the 1995 Alternate Compliance Method (ACM) manual.

Pumps and fans. Input power for pumps, fans and other motor-driven equipment was calculated from motor nameplate hp data. Motor efficiencies as observed by the surveyors were used to calculate input power. In the absence of motor efficiency observations, standard motor efficiencies were assigned as a function of the motor hp, RPM and frame type. A rated load factor was used to adjust the nameplate input rating to the actual running load. For VAV system fans, custom curves were used to calculate fan power requirements as a function of flow rate in lieu of the standard curves used in DOE-2, as described in the 1995 ACM manual.

Refrigeration systems. Refrigeration display cases and/or walk-ins were grouped into three systems defined by their evaporator temperatures. Ice cream cases were assigned to the lowest temperature circuit, followed by frozen food cases, and all other cases. Case refrigeration loads per lineal foot were taken from manufacturers' catalog data for typical cases. Auxiliary energy requirement data for evaporator fans, anti-sweat heaters, and lighting were also compiled from manufacturers' catalog data. Model inputs were calculated based on the survey responses. For example, if the display lighting was surveyed with T-8 lamps, lighting energy requirements appropriate for T-8 lamps were used to derive the case auxiliary energy input to DOE-2.

Compressor EER data were obtained from manufacturers' catalogs as a function of the suction temperatures corresponding to each of the three systems defined above. These data were used to create default efficiencies for each compressor system. Custom part-load curves were used to simulate the performance of parallel-unequal rack systems.

Service hot water. Service hot water consumption was calculated based on average daily values from the 1995 ACM for various occupancy types. Equipment capacity and efficiency were assigned based on survey responses.

Exterior lighting. Exterior lighting input parameters was developed similarly to those for interior lighting. The exterior lighting connected load was calculated from a fixture count, fixture identification code and the input wattage value associated with each fixture code.

Plant

This section describes the methodology used to develop DOE-2 input for the plant simulation. Principal data sources included the on-site survey, Title 24 documents, manufacturers' data, program data, and other engineering references.

Chillers. The DOE-2 input parameters required to model chiller performance included chiller type, full-load efficiency and capacity at rated conditions, and performance curves to adjust chiller performance for temperature and loading conditions different from the rated conditions. Chiller type was assigned based on the type code selected during the on-site survey. Surveyors also gathered chiller make, model number, and serial number data. These data were used to develop performance data specific to the chiller installed in the building. Program data and/or manufacturers' data were used to develop the input specifications for efficiency.

Cooling towers. Cooling tower fan and pump energy was defined based on the nameplate data gathered during the on-site survey. Condenser water temperature and fan volume control specifications were derived from the on-site survey responses.

Model Calibration

An integral part of DOE-2 model development was the model calibration process. Monthly energy consumption and demand from the DOE-2 models was compared to billing data for the same period to assess the reasonableness of the models. Changes were made to a fixed set of calibration parameters until the models matched the billing data. The goal of the calibration process was to match billing demand and energy data within ± 10 percent on a monthly basis. The overall model calibration process consisted of the following steps:

1. Review and format billing data. Billing data as received from Edison was reformatted as required by the model calibration software.
2. Select relevant accounts. For many of the sites, a number of accounts were provided. Account information such as customer name, address, business type, and meter number was compared to the onsite survey information. The list of accounts that seemed to best match the surveyed space was selected.
3. Assign surveyed to metered space percentage. During the onsite survey, the surveyors were asked to assess the ratio of the space surveyed to the space served by the building meter(s). Billing data records were adjusted to reflect portion of the metered data that applied to the modeled space.
4. Run model. The as-built model was run with actual 1996 and 1997 weather data applicable to the particular site, using the occupancy as reported by the surveyors. Annual simulations for both years were done, and the modeled consumption and demand was aggregated to correspond to the meter read dates from the billing data. The 1997 calibration covered billing data and simulated energy consumption for the first six months of the year. The actual year weather data was provided by SCE.
5. Review kWh and kW comparison. The modeled and metered consumption and demand for each billing period was compared using a graphical data visualization tool. An example output screen from the calibration tool is shown in Figure 13.
6. Reject unreasonable or faulty billing data. Some of the billing data received was incomplete or not well matched to the modeled space. In these cases, the billing data were rejected, and the models were not calibrated.
7. Make adjustments to calibration variables. A fixed set of calibration variables was provided to the modeling calibration team. The calibration parameters, and the range of acceptable adjustments are shown in Table 16. The modelers adjusted the calibration parameters until the modeled results matched the metered results within ± 10 percent for each billing period. This was an iterative process, involving changing the model inputs, repeating the simulation, and reviewing the results. At each iteration, the changes made to the model and the impacts of the change on the model vs. billing data comparison were entered into a calibration log file.

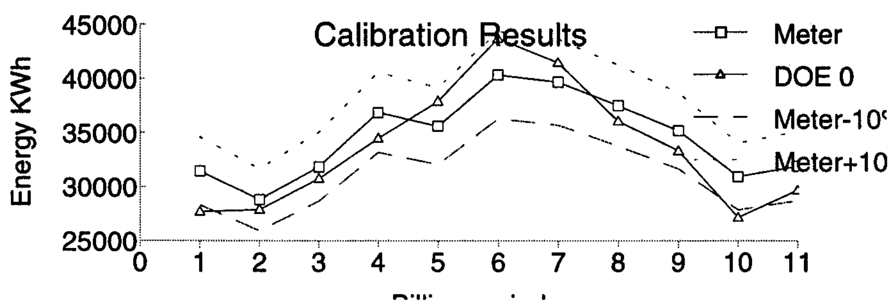


Figure 13: Example Calibration Tool Screen

Calibration Parameter	Adjustment range
Monthly schedule multiplier	.2 – 2
Lighting diversity multiplier	.2 – 2
Plug load diversity multiplier	.2 – 5
Plug load internal heat gains multiplier	.2 – 5
Heating thermostat setpoint	± 5°F
Cooling thermostat setpoint	± 5°F
DHW water use multiplier	.1 – 10
Minimum outside air ratio	.1 - .7, if no additional information
Refrigeration compressor efficiency	± 20%
Heating supply air temp control	discrete choices
Direct evaporative system effectiveness	0.2 - 0.8
Indirect evaporative system effectiveness	0.2 - .07
Heat pump defrost control	discrete choices
Daylight factor	look at hourly reports to verify correct operation
Building azimuth	± 45 degrees

Table 16: Model Calibration Parameters and Acceptable Adjustment Range

In some cases, it was not possible to calibrate the models. When billing or short-term metering data were not available, the modeled results were examined for reasonableness, in terms of annual energy consumption (kWh/SF) by building type and end-use percentage of total consumption. Even when billing data were available, some of the models resisted reasonable attempts to achieve calibration. Rather than making unreasonable adjustment to the models, the models were left uncalibrated or partially calibrated. During calibration, the models were run with actual year weather data provided by SCE from 23 local weather stations located throughout the Edison service territory.

The results of the model calibration process are shown in Figure 14. The modelers were able to successfully calibrate 58% of the models. We were unable obtain useful billing data for 18% of the sites. A total of 24% of the models resisted reasonable attempts at calibration. In other words, for 24% of the sites, billing data was available but the model could not be brought into agreement with the data by making reasonable modifications to the model.

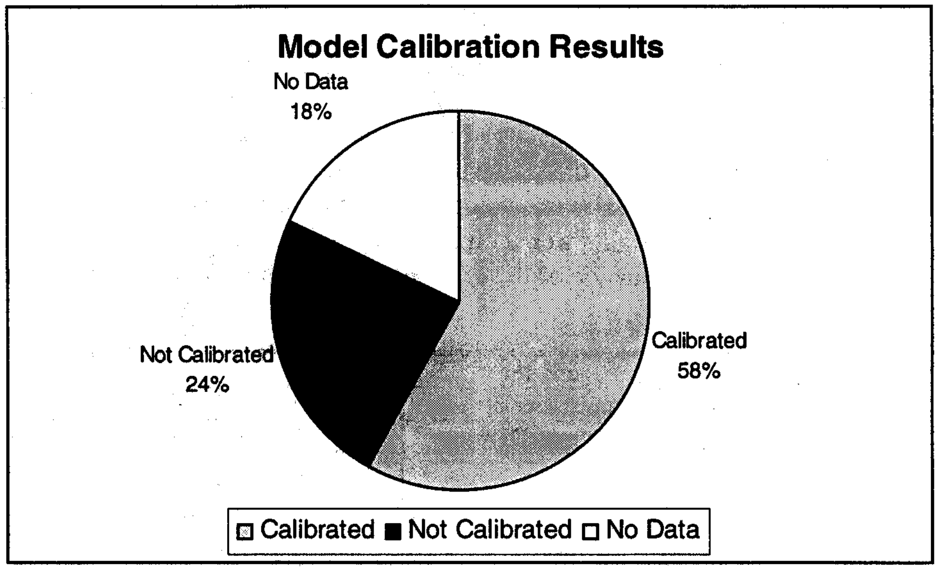


Figure 14: Model Calibration Results

Effects of Model Calibration

To understand the effect of calibrating the models to available billing or short-term metering data, models that were successfully calibrated were projected to the population and compared. That is, only the models that were ultimately calibrated were used in this test. Overall, model calibration had the effect of increasing the measured savings by 6.6%.

The average change in the most popular calibration variables is shown in Table 17.

Calibration Variable	Average Initial Value	Average Final Value
Plug load diversity multiplier	1	2.22
Lighting diversity multiplier	1	1.35
Outdoor air fraction	0.15	0.23
Cooling setpoint	74.2°F	73.9°F
Monthly schedule multiplier	1	1.05

Table 17: Average Change in Calibration Variables

The plug load diversity multiplier was also showed the largest average change of the set of most popular calibration variables. Plug loads were not extensively surveyed, since plug load energy consumption was not addressed by the program or Title 24. The uncertainty in the calculated plug load density and schedule diversity was high, as is the influence of plug loads on total building consumption and demand. However, the impact of plug loads on calculated energy savings is minor.

Model Review and Quality Control

The onsite survey data entry program contained numerous quality control (QC) checks designed to identify invalid building characteristics data during data entry. Once the data were entered, the models were run and the results were reviewed by the surveyor/modeler and senior engineering staff. A building characteristics and model results summary report was created for each site. The model results were compared to a set of QC criteria as shown in Table 18. Data falling outside of the QC range were validated during the QC process.

Building Parameter	Range	Definition
Lighting Power Density	0.9 - 1.9	building wide average
Equipment Power Density	0.1 - 5	building wide average
Cooling Ratio	95 - 200%	capacity from annual run / capacity from sizing run
Cooling EER	8 - 14	capacity weighted cooling efficiency
Wall-U	0.5 - 0.033	area weighted average, includes air film
Roof-U	0.5 - 0.033	area weighted average, includes air film
Win-U	0.3 - 0.88	area weighted average, includes air film
Win-Shading Coefficient	0.35 - 0.88	area weighted average
Win Area	0 - 70%	Percentage of gross wall area associated w/windows, expressed as a true percentage 0 –100
Sky-U	0.3 - 0.9	area weighted average of glazing contained in roof
Sky-Shading Coefficient	0.35 - 0.88	area weighted SC for all horizontal glazing
Sky-Area	0 - 10%	Percentage of gross roof area associated with sky light, expressed as a true percentage 0 –100
LTG Occupancy Sensors	0 - 50%	Percentage of lighting watts controlled by occupancy sensors, expressed as a true percentage 0 –100
LTG Daylighting controls	0 - 50%	Percentage of lighting watts controlled by daylighting sensors, expressed as a true percentage 0 –100
Measures only savings relative to program expectations (participants only)	50% - 150%	measures-only savings / program expectations
Total savings (all sites)	0% - 50%	Savings expressed as a percentage of baseline energy consumption

Table 18: Model Quality Control Criteria

Modeling results were also reviewed by Edison engineering staff. A meeting was held in San Dimas to review results for sites falling outside of the QC range. A number of modeling and data problems were identified during the Edison staff review, adding an additional level of QC to the overall process. These problems were fixed, thus improving the overall accuracy of the modeling process.

Parametrics

Once the models were calibrated and quality checked, a batch process was used to create a series of parametric simulation runs. These runs were used to simulate gross savings for participants and non-participants on a whole-building and measure-class basis. The parametric runs performed for this study are listed below:

As-Built Parametric Run.

Once the models were completed, checked for reasonableness, and/or calibrated, the as-built parametric run was done. Monthly schedule variations resulting from partial occupancy and building startup were eliminated, and the models were run using long-term average weather data from the CEC CTZ long term average weather data files.

Baseline Parametric Run.

Key building performance parameters were reset to a baseline condition to calculate gross energy savings for participants and non-participants. The California Building Energy Efficiency Standard (Title 24) was the primary reference for establishing baseline performance parameters. Title 24 specifies minimum specifications for building attributes such as:

- Opaque shell conductance
- Window conductance
- Window shading coefficient
- HVAC equipment efficiency
- Lighting power density

Title 24 applied to most of the building types covered in the programs covered under this evaluation, with the exception of:

- Hospitals
- Unconditioned space (including warehouses)

Incentives were also offered by the programs for building attributes not addressed by Title 24. In situations where Title 24 does not address building types or equipment covered under the program, baseline parameters equivalent to those used for the program baseline efficiencies were used.

Envelope

Opaque shell U-values were assigned based on Title 24 requirements as a function of climate zone and heat capacity of the observed construction. For windows, Title 24 specifications for maximum relative solar heat gain were used to establish baseline glazing shading coefficients. Fixed overhangs were removed from the baseline building. Glass conductance values as a function of climate zone were applied. For skylights, shading coefficients and overall conductance was also assigned according to climate zone.

Mechanical

Baseline specifications for HVAC equipment efficiency were derived from the Title 24 requirements as a function of equipment type and capacity. Maximum power specifications for fans were established based on Title 24 requirements, which address fan systems larger than 25 hp. Specific fan power was held energy neutral (as built W/CFM = baseline W/CFM) for fan systems under 25 hp. Additionally, all systems larger than 2500 CFM (except for hospitals) were simulated with economizers in the baseline run. All VAV fan systems larger than 50 hp were simulated with inlet vane control. All variable-volume pumps were simulated with throttling valve control.

HVAC system sizing

HVAC system sizing for the as-built case was determined by direct observation of the nameplate capacities of the HVAC equipment. The installed HVAC system capacity was compared to the design loads imposed on the system to determine a sizing ratio for the as-built building. Once established, the sizing ratio was held constant for each subsequent DOE-2 run. A separate sizing run was done prior to the baseline and parametric runs. The peak cooling system size was calculated using the equipment sizing algorithms in DOE-2. The system capacity was reset using the calculated peak cooling capacity, and the as-built sizing ratio. A new system size was calculated for the baseline run and each parametric run.

Lighting

The Title 24 area category method was used to set the baseline lighting power for each zone as a function of the observed occupancy. Task lighting and exit signs were not included in the baseline lighting calculation. A lighting power density appropriate for corridor/restroom/support areas was assigned according to the portion of each space allocated to these areas. All lighting controls were turned off for the baseline simulation.

Additional Parametric Runs

Once the as-built and baseline building models were defined, an additional set of parametric runs were done to estimate the program impact on the lighting, HVAC, and shell / daylighting end-uses. The baseline model was returned to the as-built design in a series of steps outlined as follows:

1. Lighting - measures only. Baseline lighting power densities and controls (except daylighting) for incented measures only were returned to their as-built condition.

2. All Lighting . All baseline lighting power densities and controls (except daylighting) were returned to their as-built condition.

3. Daylighting plus shell - measures only. Run 2 above, plus baseline envelope and daylighting controls for incented measures only returned to their as-built condition.

4. All Daylighting plus shell. Run 2 above, plus all baseline envelope and daylighting controls returned to their as-built condition.

5. HVAC - measures only. Run 4 above, plus HVAC for incented measures only parameters returned to their as-built condition.

6. All HVAC. Run 4 above, plus all HVAC parameters returned to their as-built condition. This run is equivalent to the full as-built run.

Gross Savings

This section presents the gross energy and demand savings estimates of participants. Savings findings for the whole building as well as for lighting, shell/daylighting, and HVAC end-uses are reported.

Some definitions would be helpful to clarify the discussion.

Baseline	A consistent standard of energy efficiency against which all buildings were measured. This was defined as the output of a DOE-2.1E simulation of a building using Title 24 required equipment efficiencies (where applicable) run using the operating schedule found by the on-site surveyor. Where Title 24 did not apply (e.g. hospitals), the baseline that was defined by the program for estimating the program savings was used.
As Built	A DOE-2.1E simulation of a building using all equipment and operating parameters as found by an on-site surveyor.
Savings	The difference between baseline and as built. Positive savings indicate that the building was more efficient – used less energy -- than its base case.
"Better than baseline"	The as built simulation showed less energy consumption than the baseline simulation – more efficient than the base case. Positive savings.
"Worse than baseline"	The as built simulation showed more energy consumption than the baseline simulation – less efficient than the base case. Negative savings.
Time-of-use period	SCE defined time periods for reporting energy usage. See Table 19 for description of each period.

Period	Dates	Days / Times
Summer On-peak	June 1to October 4	Weekdays 12 pm to 6 pm
Summer Mid-peak	June 1 to October 4	Weekdays 8 am to 12 pm and 6 pm to 11 pm
Summer off-peak	June 1 to October 4	Weekdays 11 pm to 8 am. All day weekends and holidays
Winter Mid-peak	October 5 to May 31	Weekdays 8 am to 9 pm
Winter Off-peak	October 5 to May 31	Weekdays 9 pm to 8 am. All day weekends and holidays.

Table 19: Time-of-use periods

Methodology

This project used a statistical methodology called Model-Based Statistical Sampling or MBSS™. MBSS has been used for many evaluation studies to select the sites or projects to be studied and to extrapolate the results to the target population. MBSS has been used for NEES, Northeast Utilities, Consolidated Edison, The New York Power Authority, Wisconsin Electric, Sierra Pacific Power Company, and Washington Power and Light among others. MBSS was used in the end-use metering component of the 1992 evaluation of PG&E's CIA program. A complete description of MBSS methodology is available.[5]

The Sample Design chapter earlier in this report describes the sample designs used in this study. Therefore this section will describe the methods used to extrapolate the results to the target population. Three topics will be described: (a) case weights, (b) balanced stratification to calculate case weights, and (c) stratified ratio estimation using case weights.

Case Weights

We will use the following problem to develop the idea of case weights. Given observations of a variable y in a stratified sample, estimate the population total Y.

Note that the population total of y is the sum across the H strata of the subtotals of y in each stratum. Moreover each subtotal can be written as the number of cases in the stratum times the mean of y in the stratum. This gives the equation:

$$Y = \sum_{h=1}^{H} N_h \mu_h$$

Motivated by the preceding equation, we estimate the population mean in each stratum using the corresponding sample mean. This gives the conventional form of the stratified-sampling estimator, denoted $\hat{Y}$, of the population total Y:

$$\hat{Y} = \sum_{h=1}^{H} N_h \bar{y}_h$$

With a little algebra, the right-hand side of this equation can be rewritten in a different form:

$$\begin{aligned}\hat{Y} &= \sum_{h=1}^{H} N_h \bar{y}_h \\ &= \sum_{h=1}^{H} N_h \left(\frac{1}{n_h} \sum_{k \in s_h} y_k \right) \\ &= \sum_{k=1}^{n} \left(\frac{N_h}{n_h} \right) y_k\end{aligned}$$

[5] *Methods and Tools of Load Research, The MBSS System, Version V.* Roger L. Wright, RLW Analytics, Inc. Sonoma CA, 1996.

Motivated by the last expression, we define the ***case weight*** of each unit in the sample to be $w_k = \frac{N_h}{n_h}$. Then the conventional estimate of the population total can be written as a simple weighted sum of the sample observations:

$$\hat{Y} = \sum_{k=1}^{n} w_k y_k$$

The case weight w_k can be thought of as the number of units in the population represented by unit k in the sample. The conventional sample estimate of the population total can be obtained by calculating the weighted sum of the values observed in the sample.

Table 20 shows an example[6]. In this example, the population of program participants has been stratified into five strata based on the annual savings of each project shown in the tracking system. For example, the first stratum consists of all projects with annual savings less than 101,978 kWh. The maximum kWh in each stratum is called the stratum cut point. There are 339 projects in this stratum and they have a total tracking savings of 8,038,527 kWh. The estimate of gross impact was obtained from the measured savings found in a sample of 85 projects. Column 5 of Table 19 shows that the sample contains 62 projects from the first stratum. Each of these 62 projects can be given a case weight of 339 / 62 = 5.47.

Stratum	Max kWh	Population Size	Total KWh	Sample Size	Case Weight
1	101,978	339	8;038,527	62	5.47
2	278,668	61	10,949,421	9	6.78
3	441,916	35	12,598,315	8	4.38
4	816,615	22	13,654,171	3	7.33
5	4,000,000	12	17,469,244	3	4.00
Total		469	62,709,678	85	

Table 20: Stratification Example

Balanced Stratification

Balanced stratification is another way to calculate case weights. In this approach, the sample sites are sorted by the stratification variable, tracking kWh, and then divided equally among the strata. Then the first stratum cutpoint is determined midway between the values of the stratification variable for the last sample case in the first stratum and the first sample case in the second stratum. The remaining strata cutpoints are determined in a similar fashion. Then the population sizes are tabulated within each stratum. Finally the case weights are calculated in the usual way.

[6] This is an example only. The numbers presented here are not relevant to the study findings.

Table 21 shows an example[7]. In this case the sample of 85 sites has been equally divided among five strata, so there are 17 sites per stratum. Then the stratum cutpoints shown in column two were calculated from the tracking estimates of kWh for the sample sites. Next the population sizes shown in column three were calculated from the stratum cutpoints. The final step was to calculate the case weights shown in the last column. For example, the case weight for the 17 sites in the first stratum is 136 / 17 = 8.

Stratum	Max kWh	Population Size	Total KWh	Sample Size	Case Weight
1	7,948	136	417,368	17	8.00
2	22,361	84	1,211,832	17	4.94
3	63,859	84	3,605,867	17	4.94
4	202,862	73	8,146,886	17	4.29
5	2,883,355	92	49,327,725	17	5.41
Total		469	62,709,678	85	

Table 21: Balanced Stratification

Stratified Ratio Estimation

Ratio estimation is used to estimate the population total Y of the target variable y taking advantage of the known population total X of a suitable explanatory variable x. The ratio estimate of the population total is denoted $\hat{Y}_{ra}$ to distinguish it from the ordinary stratified sampling estimate of the population total, which is denoted as $\hat{Y}$.

Motivated by the identity $Y = BX$, we estimate the population total Y by first estimating the population ratio B using the sample ratio $b = \bar{y}/\bar{x}$, and then estimating the population total as the product of the sample ratio and the known population total X. Here the sample means are calculated using the appropriate case weights. This procedure can be summarized as follows:

$$
\begin{aligned}
\hat{Y}_{ra} &= bX \quad \text{where} \\
b &= \frac{\bar{y}}{\bar{x}} \\
\bar{y} &= \frac{1}{\hat{N}} \sum_{k=1}^{n} w_k y_k \\
\bar{x} &= \frac{1}{\hat{N}} \sum_{k=1}^{n} w_k x_k \\
\hat{N} &= \sum_{k=1}^{n} w_k
\end{aligned}
$$

[7] This is only an example. The numbers presented are not relevant to the study findings.

The conventional 90 percent confidence interval for the ratio estimate of the population total is usually written as

$$
\begin{aligned}
\hat{Y}_{ra} &\pm 1.645\sqrt{V\left(\hat{Y}_{ra}\right)} \quad \text{where} \\
V\left(\hat{Y}_{ra}\right) &= \sum_{h=1}^{H} N_h^2 \left(1-\frac{n_h}{N_h}\right)\frac{s_h^2(e)}{n_h} \\
s_h^2(e) &= \frac{1}{n_h-1}\sum_{k\in s_h}\left(e_k-\bar{e}_h\right)^2 \\
e_k &= y_k - b x_k
\end{aligned}
$$

We can calculate the relative precision of the estimate $\hat{Y}_{ra}$ using the equation

$$
rp = \frac{1.645\sqrt{V\left(\hat{Y}_{ra}\right)}}{\hat{Y}_{ra}}
$$

MBSS theory has led to an alternative procedure to calculate confidence intervals for ratio estimation, called model-based domains estimation. This method yields the same estimate as the conventional approach described above, but gives slightly different error bounds. This approach has many advantages, especially for small samples, and has been used throughout this study.

Under model-based domains estimation, the ratio estimator of the population total is calculated as usual. However, the variance of the ratio estimator is estimated from the case weights using the equation

$$
V\left(\hat{Y}_{ra}\right) = \sum_{k=1}^{n} w_k\left(w_k-1\right)e_k^2
$$

Here w_k is the case weight discussed in Section 6.5.1 and e_k is the sample residual $e_k = y_k - b x_k$. Then, as usual, the confidence interval is calculated as

$$
\hat{Y}_{ra} \pm 1.645\sqrt{V\left(\hat{Y}_{ra}\right)}
$$

and the achieved relative precision is calculated as

$$
rp = \frac{1.645\sqrt{V\left(\hat{Y}_{ra}\right)}}{\hat{Y}_{ra}}
$$

The model-based domains estimation approach is often much easier to calculate than the conventional approach since it is not necessary to group the sample into strata. In large samples, there is generally not much difference between the case-weight approach and the conventional approach. In small samples the case-weight approach seems to perform better. For consistency, we have come to use model-based domains estimation in most work.

This methodology generally gives error bounds similar to the conventional approach. Equally, the model-based domains estimation approach can be derived from the conventional approach by making the substitutions:

$$\begin{aligned} \overline{e}_h &\approx 0 \\ s_h^2(e) &\approx \frac{1}{n_h}\sum_{k \in s_h} e_k^2 \end{aligned}$$

In the first of these substitutions, we are assuming that the within-stratum mean of the residuals is close to zero in each stratum. In the second substitution, we have replaced the within-stratum variance of the sample residual e, calculated with $n_h - 1$ degrees of freedom, with the mean of the squared residuals, calculated with n_h degrees of freedom.

Model-based domains estimation is appropriate as long as the expected value of the residuals can be assumed to be close to zero. This assumption is checked by examining the scatter plot of y versus x. It is important to note that the assumption affects only the error bound, not the estimate itself. $\hat{Y}_{ra}$ will be essentially unbiased as long as the case weights are accurate.

Gross Savings Expansions

Each building in the sample was modeled as described in the Engineering Models section. A baseline, as built, and savings estimate was developed for every building in the sample. The sample of baseline, as built, and savings estimates was projected to the participant population using model-based statistical methods described above.

Energy Impact Findings

Whole Building

SCE's whole building gross energy savings was 42,730 MWh. The relative precision of the estimate was ±8.8%. This represents a gross realization rate of 116.4% of verified annual savings. Table 22 shows the estimated savings by time-of-use period.

Period	Energy Savings (MWh)	Energy Rel. precision
Annual	**42,730**	**± 8.8%**
Summer On-Peak	4,196	± 9.6%
Summer Mid-Peak	4,679	± 8.7%
Summer Off-Peak	6,838	± 9.5%
Winter Mid-Peak	14,220	± 9.4%
Winter Off-Peak	12,800	± 11.2%

Table 22: Whole Building Energy Savings by Time-of-use period

The participant group was more energy efficient than the non-participant comparison group. Figure 15 shows the savings of participants and non-participants expressed as a percentage of each group's baseline usage.

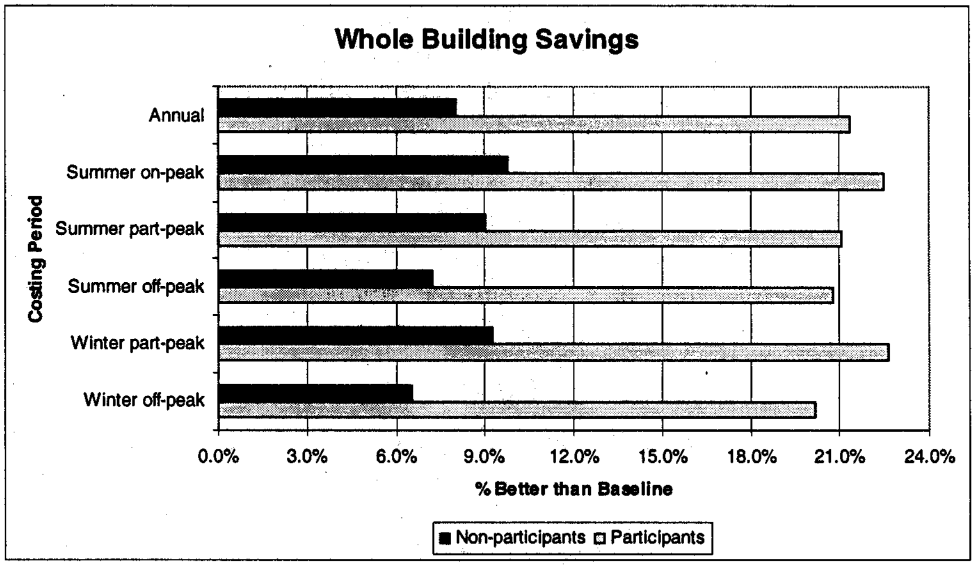


Figure 15: Participant and Non-participant Energy Savings Relative to Baseline

As Figure 15 shows, the participant group was 21.4% better than baseline on average. The non-participant comparison group was 8.1% better than baseline. The level of efficiency relative to the baseline remains fairly constant throughout the year.

End-Use Savings

Three end-uses were examined as part of this study, lighting, HVAC, and shell / daylighting. Those sites that had savings were projected to the population to arrive at the total savings estimate. Note that the sum of the end-use savings may not add exactly to 1 due to rounding. In each of the figures describing end-use savings, the percentages are of the whole building baseline. The percentage scale in the figures is an indicator of the contribution to overall savings of each end-use. Figure 16 shows the breakdown of annual energy savings by end-use.

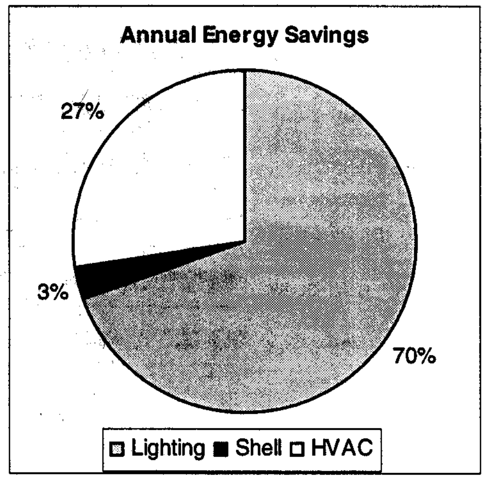


Figure 16: Composition of Energy Savings

Lighting

The lighting end-use accounted for 69% of the annual energy savings of the participant group, or 29,580,000 kWh. Table 23 shows the savings and relative precision by time-of-use period.

Period	Savings (MWh)	Relative Precision
Annual	29,580	± 12.6%
Summer on-peak	2,746	± 13.2%
Summer part-peak	3,136	± 12.8%
Summer off-peak	4,539	± 12.8%
Winter part-peak	10,290	± 13.1%
Winter off-peak	8,865	± 12.4%

Table 23: Lighting Energy Savings by Time-of-use period

Figure 17 shows the participant and non-participant lighting savings relative to baseline consumption by time-of-use period. The lighting savings of participants was 64% greater than the non-participants. This was the smallest difference among the three end-uses studied.

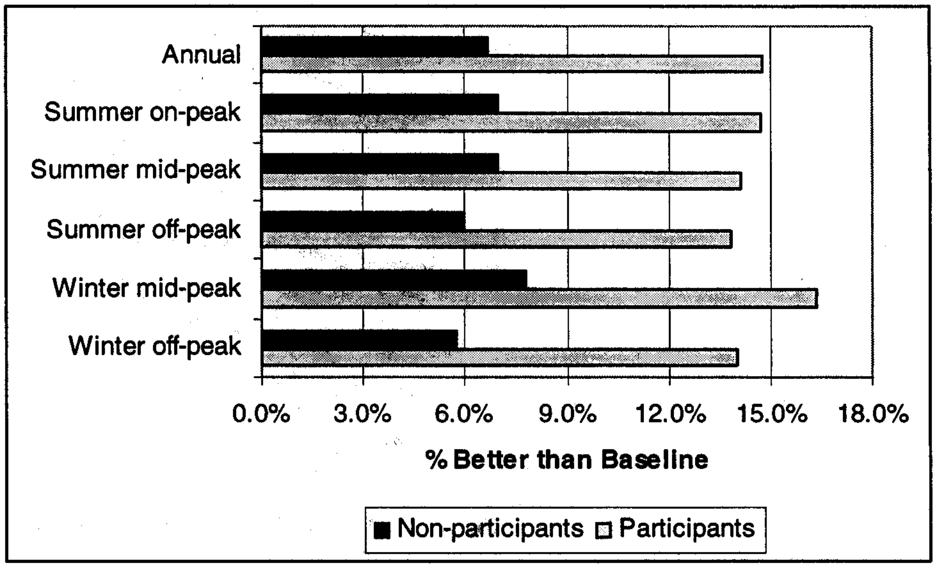


Figure 17: Lighting Energy Savings Relative to Baseline

HVAC

The HVAC end-use accounted for 27.6% of the participant group savings, or 11,720 MWh. Table 24 shows the savings and relative precision by time-of-use period.

Period	Savings (MWh)	Relative Precision
Annual	11,720	± 30.3%
Summer on-peak	1,191	± 20.2%
Summer part-peak	1,282	± 22.2%
Summer off-peak	2,007	± 31.9%
Winter part-peak	3,635	± 31.3%
Winter off-peak	3,602	± 41.5%

Table 24: HVAC Energy Savings by Time-of-use period

Figure 18 shows the participant and non-participant HVAC savings relative to baseline consumption by time-of-use period. The HVAC end-use savings for participants was 81% greater than that for non-participants. This was the end-use where the program affected the greatest change in efficiency.

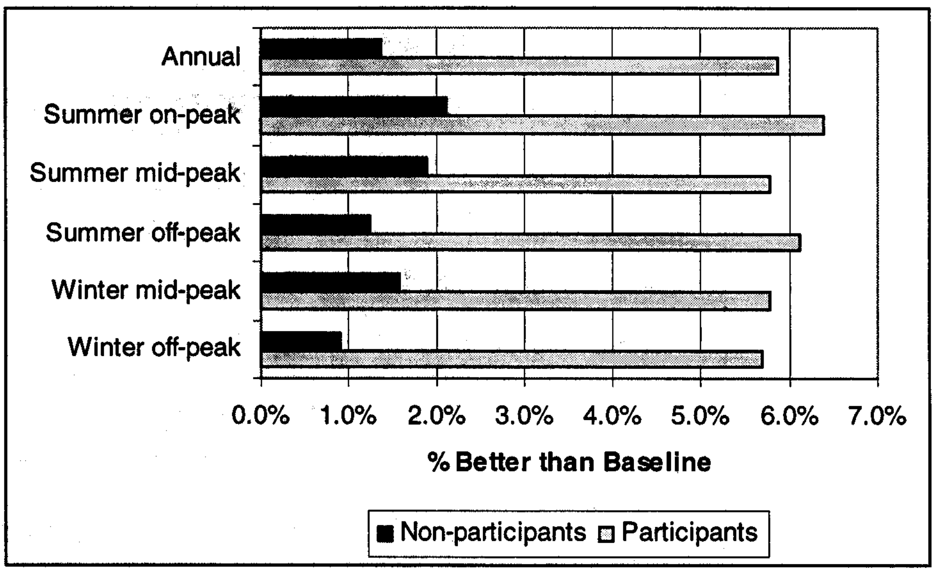


Figure 18: HVAC Energy Savings Relative to Baseline

Shell & Daylighting

The shell / daylighting control end-use accounted for 3.3% of the participant group savings, or 1,378 MWh. Table 25 shows the savings and relative precision by time-of-use period.

Period	Savings (MWh)	Relative Precision
Annual	1,378	± 36.8%
Summer on-peak	252	± 27.1%
Summer mid-peak	252	± 42.5%
Summer off-peak	284	± 36.1%
Winter mid-peak	271	± 61.4%
Winter off-peak	319	± 41.7%

Table 25: Shell & Daylighting Energy Savings by Time-of-use period

Figure 19 shows the participant and non-participant shell & daylighting savings relative to baseline consumption by time-of-use period. The participants' shell & daylighting savings was 71% greater than the non-participants' shell & daylighting savings.

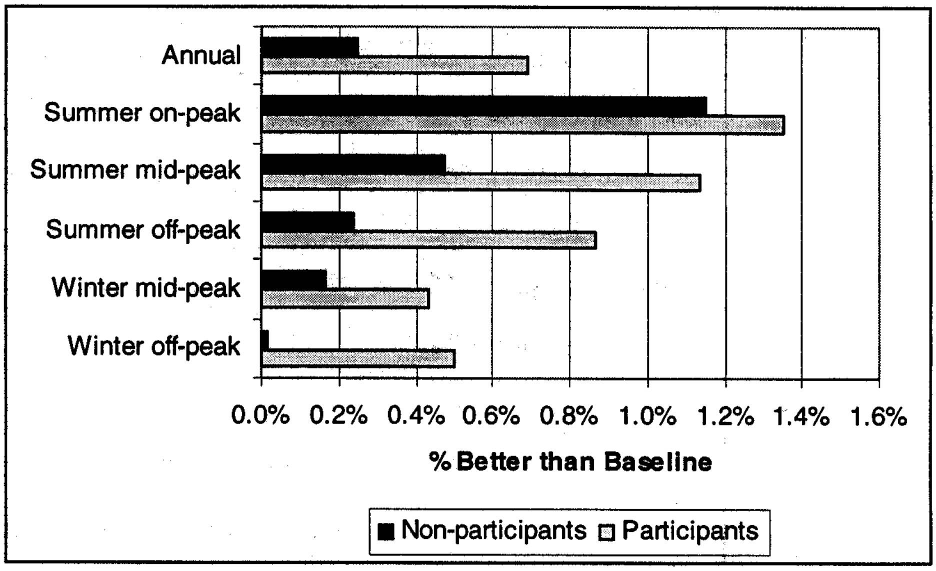


Figure 19: Shell & Daylighting Energy Savings by Time-of-use period

Demand Impact Findings

Whole Building

SCE's whole building gross demand savings was 10.130 MW. The relative precision of the estimate was ±11.0%. This represents a gross realization rate of 115.0% of verified summer on-peak demand savings. Table 26 shows the estimated savings by time-of-use period.

Period	Demand Savings (MW)	Demand Rel. Precision
Summer On-Peak	10.13	± 11.0%
Summer Mid-Peak	9.73	± 11.1%
Summer Off-Peak	9.53	± 11.1%
Winter Mid-Peak	9.59	± 10.4%
Winter Off-Peak	9.25	± 10.7%

Table 26: Whole Building Demand Savings by Time-of-use period

The participant group was more energy efficient than the non-participant comparison group. Figure 20 shows the savings of participants and non-participants expressed as a percentage of each group's baseline demand.

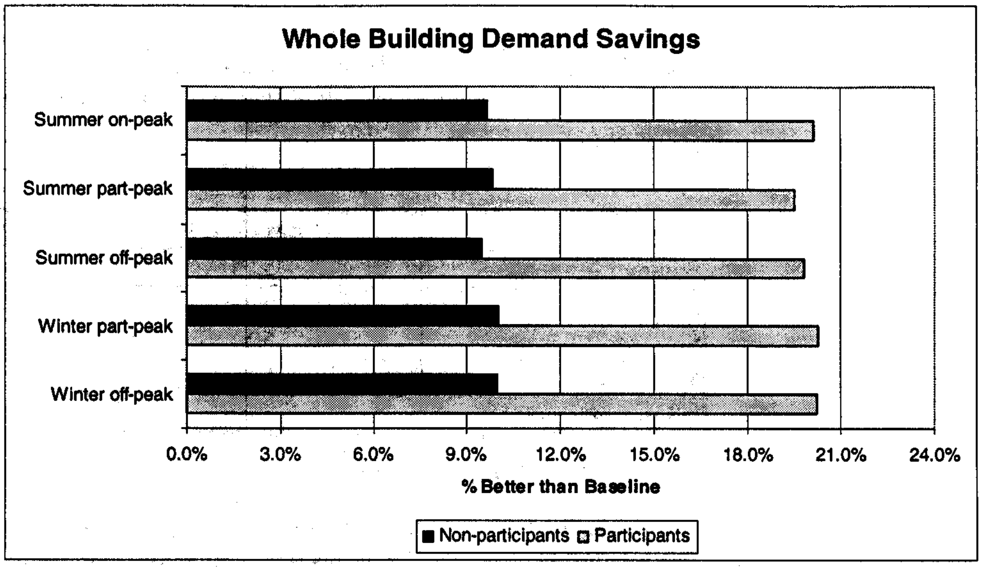


Figure 20: Participant and Non-participant Savings Relative to Baseline

As Figure 20 shows, the participant group was 20.2% better than baseline on average. The non-participant comparison group was 9.7% better than baseline. The level of efficiency relative to the baseline remains fairly constant throughout the year.

End-Use Demand Savings

Three end-uses were examined as part of this study, lighting, HVAC, and shell / daylighting. Those sites that had savings were projected to the population to arrive at the total savings estimate. Note that the sum of the end-use savings may not add exactly to 1 due to rounding. In each of the figures describing end-use savings, the percentages are of the whole building baseline. The percentage scale in the figures is an indicator of the contribution to overall savings of each end-use. Figure 21 shows the breakdown of summer peak demand savings by end-use.

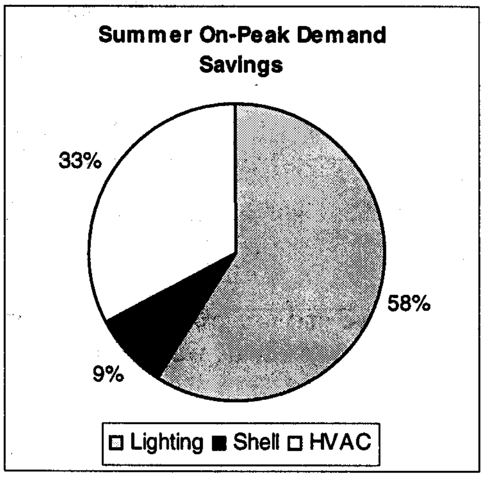


Figure 21: Composition of Summer Peak Demand Savings

Lighting

SCE's lighting end-use gross demand savings was 5.94 MW. The relative precision of the estimate was ±13.4%. Table 27 shows the estimated savings by time-of-use period.

Period	Savings (MW)	Relative Precision
Summer on-peak	5.94	±13.4%
Summer mid-peak	5.48	± 13.8%
Summer off-peak	5.53	±14.1%
Winter mid-peak	5.51	± 13.7%
Winter off-peak	5.32	± 13.9%

Table 27: Lighting Summer On-Peak Demand Savings by Time-of-use period

The participant group was more energy efficient than the non-participant comparison group. Figure 22 shows the savings of participants and non-participants expressed as a percentage of each group's baseline demand. The lighting end-use participants saved 57% more than the non-participants relative to baseline.

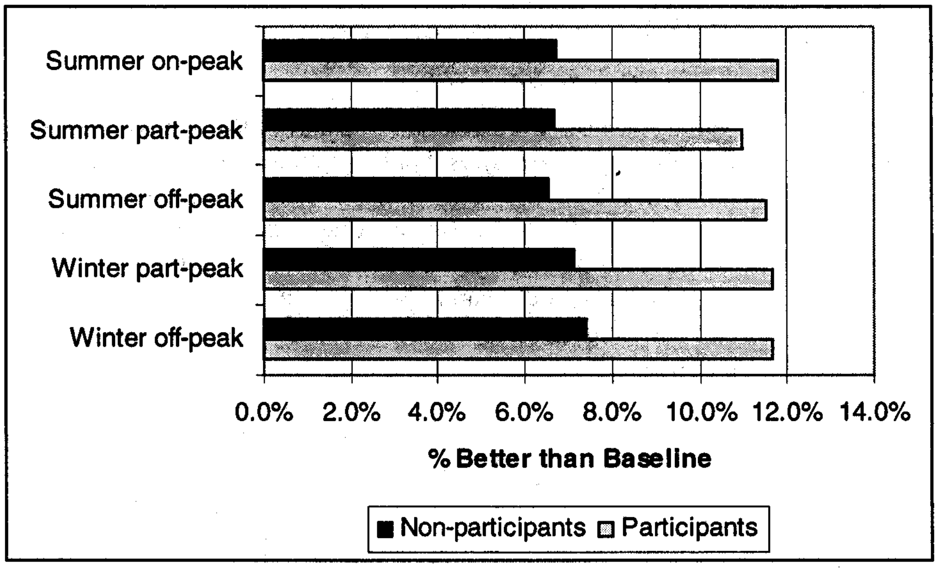


Figure 22: Lighting Summer On-Peak Demand Savings Relative to Baseline

HVAC

SCE's HVAC end-use gross demand savings was 3.30 MW. The relative precision of the estimate was ±21.6%. Table 28 shows the estimated savings by time-of-use period.

Period	Savings (MW)	Relative Precision
Summer on-peak	3.30	±21.6%
Summer mid-peak	3.33	± 21.4%
Summer off-peak	3.21	±21.5%
Winter mid-peak	3.19	± 24.5%
Winter off-peak	3.23	± 23.2%

Table 28: HVAC Summer On-Peak Demand Savings by Time-of-use period

The participant group was more energy efficient than the non-participant comparison group. Figure 23 shows the savings of participants and non-participants expressed as a percentage of each group's baseline demand. The HVAC end use is where the most dramatic effect of the program is seen. The participant group savings was 75% larger than the non-participant group savings for this end use.

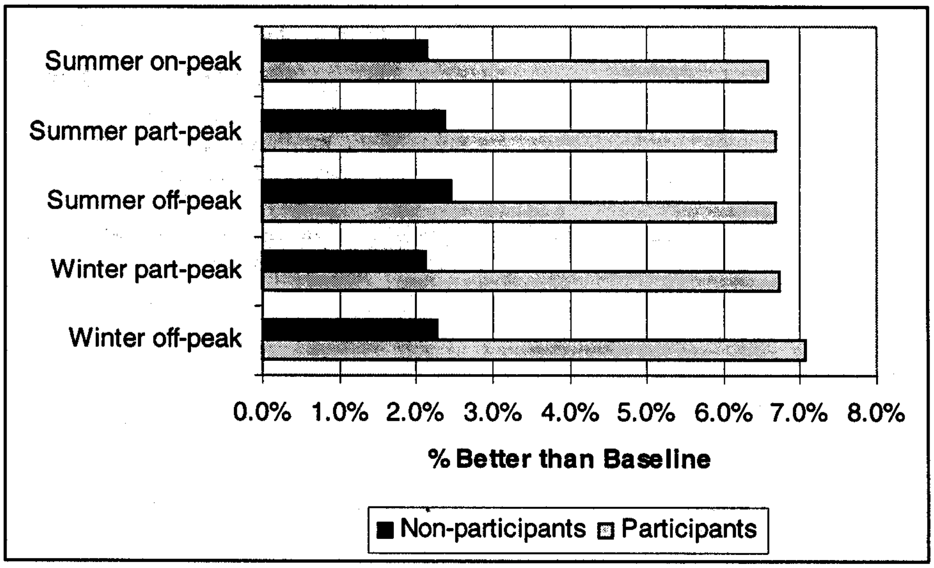


Figure 23: HVAC Summer On-Peak Demand Savings Relative to Baseline

Shell & Daylighting

SCE's shell and daylighting control end-use gross demand savings was 0.86 MW. The relative precision of the estimate was ±24.4%. Table 29 shows the estimated savings by time-of-use period.

Period	Savings (MW)	Relative Precision
Summer on-peak	0.86	±24.4%
Summer mid-peak	0.90	± 31.5%
Summer off-peak	0.77	±31.1%
Winter mid-peak	0.87	± 31.8%
Winter off-peak	0.63	± 39.8%

Table 29: Shell & Daylighting Summer On-Peak Demand Savings by Time-of-use Period

The participant group was more energy efficient than the non-participant comparison group. Figure 24 shows the savings of participants and non-participants expressed as a percentage of each group's baseline demand. The participant group saved 44% more relative to baseline than the non-participant group for this end-use. This was the smallest difference of any of the end uses.

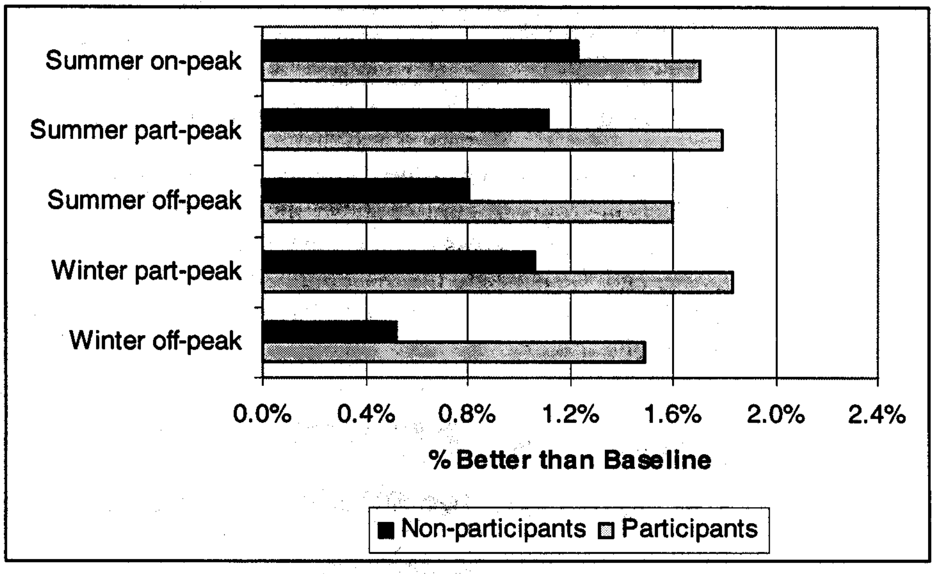


Figure 24: Shell & Daylighting Summer On-Peak Demand Savings Relative to Baseline

Net Impact Findings

Two different methodologies were followed in the net-to-gross analysis: a relatively simple difference of differences approach and a more complex econometric approach. In the difference of differences methodology, the net-to-gross ratio was calculated by comparing the gross savings relative to baseline of the program participants to the gross savings relative to baseline of the non-participants. In the econometric approach, the net-to-gross ratio was calculated by using regression modeling techniques to estimate the net savings due to the program for each of the program participants.

Difference of Differences Net-to-Gross

The following table summarizes the findings from the difference of differences analysis. The table shows the savings and the estimated net-to-gross ratio for both annual energy and summer peak demand savings.

	Net Savings	Net-to-gross Ratio	Relative Precision
Annual Energy	26,620 MWh	62.3%	±22.0%
Summer Peak Demand	5.27 MW	52.0%	±24.9%

Table 30: Difference of Differences Net-to-gross Ratios

The table also shows the relative precision of each estimate.[8] For example, in the case of annual energy, the net-to-gross ratio was estimated to be 62.3% with a relative precision of ±22.0%. The error bound for the 90% confidence interval for the true net-to-gross ratio is equal to 22.0% of the estimate, i.e. to ±13.7%. The 90% confidence interval for the true net-to-gross ratio can be calculated using the equation:

$$0.623 \pm 0.623 \times 0.220 = (0.486, 0.760)$$

We can be quite confident that this interval contains the true net-to-gross ratio that would have been obtained by developing onsite surveys and building engineering simulation models for all program participants and a very large sample of non-participants using the methodology of this study. The confidence interval reflects sampling variability and random measurement error but does not reflect any possible systematic measurement error that might be repeated throughout the data collection and engineering simulation.

[8] The standard error reflects the standard deviation of the estimate in repeated sampling. The error bound at the 90% level of confidence is 1.645 times the standard error. The confidence interval is the estimate plus or minus the error bound. The relative precision is the error bound divided by the estimate itself.

Econometric Net-to-Gross

The following table summarizes the net-savings findings from the econometric analysis. The top portion of the table shows the estimated net-to-gross ratio for both annual energy and summer peak demand savings. These are the net savings found for the participants, excluding any non-participant spillover savings. The middle portion of the table shows the results including the estimated spillover savings for non-participants. Finally, the lower portion shows a conservative estimate of net savings. Each of these is described below.

Direct Net Savings	**Net Savings**	**Net-to-gross Ratio**	**Relative Precision**
Annual Energy	26,383 MWh	61.7%	±10.2%
Summer Peak Demand	5.578 MW	55.1%	±11.1%
Net Savings with Spillover & Double Mills Ratio	**Net Savings**	**Net-to-gross Ratio**	**Relative Precision**
Annual Energy	41,005 MWh	96.0%	±30.5%
Summer Peak Demand	10,818 MW	106.8%	±21.5%
Conservative Estimate & Double Mills Ratio	**Net Savings**	**Net-to-gross Ratio**	**Relative Precision**
Annual Energy	31,273 MWh	73.2%	na
Summer Peak Demand	9.008 MW	88.9%	na

Table 31: Econometric Net-to-gross Ratios

Table 31 shows the relative precision of each estimate. For example, the net annual energy savings was estimated to be 26,383 and the net-to-gross ratio was estimated to be 61.7%, both with a relative precision of ±10.2%. The error bound for the 90% confidence interval for true annual energy savings is the estimate plus or minus10.2% of the estimate, i.e.,

$$26{,}383 \pm 26{,}383 \times 0.102 = (23{,}692, 29{,}074)$$

Similarly, the 90% confidence interval for the true net-to-gross ratio is

$$0.617 \pm 0.617 \times 0.102 = (0.554, 0.680)$$

Table 31 also shows estimates of net savings that include the spillover savings for non-participants. Both program participants and non-participants were asked about the level of their interaction with SCE and the impact of SCE on the design and equipment choices for the project. The econometric analysis showed a positive, statistically significant relationship between the impact reported by non-participants and the energy and demand efficiency of their building. The resulting econometric model was used to estimate the impact of the program on the efficiency of each non-participant building in the sample. Then we used standard statistical methods to expand the non-participant savings from the sample to the population of new construction in order to estimate the total spillover savings and to evaluate the statistical precision of the estimate.

With spillover included, the estimated net-to-gross ratio was 96.0% with a relative precision of ±30.5%. The relatively poor standard error arose from the spillover component of savings due to unavoidable limitations in the Dodge database. The statistical expansion techniques employed in this study give much more accurate estimates of the direct net savings of program participants than the spillover savings of non-participants. This is because the estimate of participant savings takes advantage of the relatively strong association between the program tracking estimates of savings for each site and the net savings predicted from the engineering analysis and the econometric model. Unfortunately, the variables in the Dodge data base, e.g., the square footage of the planned project, are only weakly associated with the net savings predicted for each non-participant site from the engineering analysis and the econometric model. Therefore, the estimates of spillover savings have a wider margin of statistical error.

To be conservative, the total net savings including spillover can be written down based on the relative precision of the estimate. The conservative estimate was obtained by reducing the point estimate by 1.28 times the standard error of the estimate. This strategy factors in the statistical allowance for error and is very conservative. For example, in the case of annual energy savings, the calculation was

$$41{,}005\left(1-\frac{1.28}{1.645}\times .305\right)=31{,}273$$

The ratio 1.28 / 1.645 was used to convert the ordinary two-sided error bound to the error bound of a one-sided confidence interval calculated at the 90% level of confidence. The probability is high, about 0.9, that this procedure underestimates the true value of savings.

This approach gave a net to gross ratio of 73.2% for energy and 88.9% for peak demand. No relative precision is associated with these results since the relative precision is already reflected in the estimate itself.

Difference of Differences Methodology

This section describes the difference of differences methodology. For simplicity we will discuss the methodology used to analyze annual energy savings. An analogous approach was used to analyze summer peak demand savings.

The following table summarizes the derivation of the net-to-gross ratio for annual energy. The analysis starts with the baseline and as-built energy consumption of the participants and non-participants. All of these results are reported in million kWh and were obtained by statistically expanding the sample data to the population of 1996 program participants. For example, the table shows that we would estimate that all program participants would have an aggregate annual consumption of 157.2 million kWh, based on the as-built simulation runs developed for the sites in the participant sample. By contrast, if we expand the as-built simulation runs of the non-participants to the same participant population, we would expect an aggregate annual consumption of 147.8 million kWh.

	Participants	Nonparticipants	Net Savings
Baseline (1 million kWh)	199.9	160.7	
As Built (1 million kWh)	157.2	147.8	
Savings (1 million kWh)	42.7	12.9	26.6
Savings (% of baseline)	21.4%	8.1%	13.3%
Net to Gross Ratio			62.3%

Table 32: Summary of Difference of Differences Calculation

Considering only the as-built results, the participants would appear to be less energy efficient than the non-participants. However, this fails to control for differences between the two samples. The preceding table shows that the baseline results were 199.9 million kWh using the participant sample and only 160.7 million kWh using the non-participant sample. Both samples were designed to be representative of the population of 1996 program participants. However we would expect differences in the baseline results from the two samples due to sampling variability. Moreover, difficulty in obtaining large non-participant sample sites to match the large participants in the program may have led to some systematic difference between the participant and non-participant samples. In fact, the observed difference corresponds to a t-statistic of about 2,[9] implying that the difference is just barely significant at the 5% level of significance. This makes it difficult to conclude whether the difference in the baseline results are purely sampling error or whether they reflect some sampling bias.

Regardless, for a more meaningful comparison, the as-built energy use should be considered relative to the baseline. The table shows the gross savings, calculated as the difference between the baseline and the as-build energy use. Calculated this way, the gross savings relative to baseline were 42.7 million kWh using the participant sample and 12.9 million kWh using the non-participant sample. In proportion to the respective baseline energy use of each sample, the gross savings were 21.4% for the participant sample and 8.1% for the non-participant sample.

In the difference of differences approach, the net savings can be estimated as the difference between the percentage savings of the participants and non-participants. In this case the net savings is 13.3% of baseline use. Multiplying 199.9 million kWh by 13.3%, the net savings of the population of 1996 program participants can be estimated to be 26.6 million kWh.

The net savings of the program participants can also be calculated using the following equation.

[9] The standard errors were about 12 million kWh for the baseline energy use from the participant sample and about 14 million kWh for the baseline energy use from the non-participant sample.

$$\left(\frac{147.8}{160.7}\right)(199.9) - 157.2 = 26.6$$

Here the first factor is the as-built energy use relative to the baseline energy use of the non-participants. This is used to adjust the baseline energy use of the participants. Then the net savings is calculated by subtracting the as-built energy use of the participants. Finally, the net savings is found to be 13.3% of the baseline energy use of the participants. The two approaches for calculating net savings are mathematically equivalent.

Finally the net-to-gross ratio can be calculated by dividing the net savings (13.3%) by the participants' gross savings (21.4%). This gives the difference of differences estimate of 62.3% for the net-to-gross ratio for annual energy.

Error Bound Methodology for the Difference of Differences Estimate

In the preceding section, it was shown that the difference of differences estimate of net savings can be expressed as an adjustment of the participant sample results based on the ratio between the as-built and baseline results observed from the non-participant sample. This is an extension of the technique of ratio estimation that is common in survey sampling.[10] The error bound and relative precision of the difference of differences estimate of net savings can be estimated using techniques similar to the methods of standard ratio estimation. In this section, we will describe the approach.

First some notation. Let $r_{np} = \left(\frac{147.8}{160.7}\right)$ denote the ratio between the as-built and baseline energy use obtained in the non-participant sample. Let $r_b = \left(\frac{199.9}{160.7}\right)$ denote the ratio between the baseline energy use of the participants relative to the non-participants. For any sample site, participant or non-participant, let e denote the difference between the as-built energy use of the site and the product of r_{np} and the baseline energy use of the site.

Now the error bound of the difference of differences estimate of net savings can be estimated in three steps:

1. Calculate e for each site in the participant sample and use standard techniques to expand the results to the target population. Let A denote the error bound of the result, calculated the usual way.
2. Calculate e for each site in the non-participant sample and use standard techniques to expand the results to the target population. Let B denote the error bound of the result, calculated the usual way.
3. Estimate the error bound of the difference of differences estimate of net savings using the equation $\sqrt{A^2 + r_b^2 B^2}$.

[10] See, for example, Chapter 6 of *Sampling Techniques*, by W. A. Cochran, Wiley and Sons, third edition, 1977.

The preceding methodology can be derived from a standard Taylor's series approximation to the sampling distribution of the difference of differences estimator.

Rationale for the Econometric Net-to-Gross Methodology

The econometric methodology can be regarded as an extension of a simple comparison of the efficiency choice of non-participant and participants through the difference of difference methodology. A coefficient of the participation indicator variable reflects the difference in efficiency choice between a participant and a non-participant. Other variables are included in the model to control for other factors that are associated with efficiency choice.

The inclusion of these variables can improve the statistical model in two ways:

1. Reduce potential bias, and
2. Provide improved statistical precision.

The potential bias arises if the model omits an explanatory variable that (a) is related to efficiency choice, and (b) is correlated with participation. For example, suppose a particular type of builder or designer tends to build a more efficient building and also tends to participate in the program. Then the difference of difference approach would tend to overestimate the actual impact of the program. This is sometimes called self-selection bias.

As another example, suppose that some of the non-participants have incorporated efficiency measures into the current building that they learned from participating in the program in prior years. In this case the difference of difference approach would underestimate the actual impact of the program. This can be called bias due to spillover.

Therefore, under most circumstances the difference of difference approach provides a biased estimate of the actual program impact. The size of the bias depends on the balance between any positive bias due to self-selection and related factors versus any negative bias due to spillover and similar factors. The only circumstances under which the difference of difference approach would give an unbiased estimate are either (a) if both self-selection and spillover are negligible, or (b) if they are exactly equal. Both of these assumptions seem unlikely, especially for a program deliberately designed to influence general practice in new construction, so a more powerful methodology is needed to obtain an unbiased estimate of net savings.

The econometric methodology seeks to obtain an unbiased estimate of net savings by including both program variables and other explanatory variables in a multivariate regression model. If the model is accurately specified and if the program variables and other explanatory variables are not multicollinear, then the model will provide an unbiased estimate of the net program savings among the participants as well as the spillover impact among the non-participants. This is the primary motivation for a multivariate regression analysis.

The econometric approach can also improve statistical precision by including explanatory variables that significantly affect efficiency choice. If an explanatory variable has a significant relationship with efficiency, then its inclusion in the model may significantly decrease the residual variance, or unexplained variance, of the model, and in turn, provide more statistically reliable estimates of net savings and spillover impacts.

Conversely, there are reasons for excluding all variables that do not have a significant relationship with efficiency. The inclusion of such variables needlessly tends to reduce the statistical precision of the results and makes the models unnecessarily complex and difficult to interpret. Therefore, we seek to include all truly relevant variables but drop the irrelevant variables. Necessarily, this is an iterative process, but a well-defined and objective procedure can be followed to obtain the final model and resulting estimates of net savings and spillover impacts.

Explanatory Variables

The following table summarizes the data elements used to develop the potential explanatory variables for the econometric analysis. The table shows the source of each data element and gives a brief description of the relevance of each data element to the econometric analysis. The data elements collected by phone were all contained in the decision-maker survey.

Data Element	*Collection*	*Rationale*
Building Type	On-site	Different types of buildings may be built to different efficiency standards. This was seen in the 1994 study.
Project Type	Phone	New construction may be built more efficiently than additions or renovations
Building ownership	Phone	Owner occupants may be more concerned with efficiency than developers / landlords.
Construction circumstances	Phone	Same as above
Owner input	Phone	More owner input makes owner attitudes more important with respect to efficiency choices.
Pre-existing plans	Phone	Standard designs reduce the likelihood of efficiency measures in response to the program.
Investment Criteria	Phone	Investment criteria may affect willingness to install efficiency measures
Signif. Of energy costs	Phone	Significance of energy costs may influence efficiency choice.
Signif. Of energy eff	Phone	Significance of energy efficiency may influence decision to install higher eff. equipment
Awareness of program	Phone	Awareness may lead to spillover.
Interaction with utility on this project	Phone	Interaction with SCE may lead to spillover
Influence of utility on this project	Phone	Influence of SCE may lead to spillover.
Interaction with utility on previous projects	Phone	Interaction with SCE may lead to spillover
Influence of utility on previous projects	Phone	Influence of SCE may lead to spillover

Table 33: Variables Considered for Econometric Analysis

Decision-maker findings

The primary purpose of the decision-maker survey was to provide data for the econometric analysis of net savings. This section summarizes the data that was collected. The purpose of this task was to perform reality checks on the data, as well as to present the survey data. The reality checks included verifying if:

- The data was collected from the right people
- The data leads to sensible conclusions

Data Quality Checks

Figure 25 shows the responses to the 7-point scale question asking about the decision-maker's involvement in project decisions. In Figure 25, a response of 1 indicates little input on design decisions and a 7 indicates significant input on design decisions. As the figure shows, people who were significantly involved in project decisions were the people who responded to the survey.

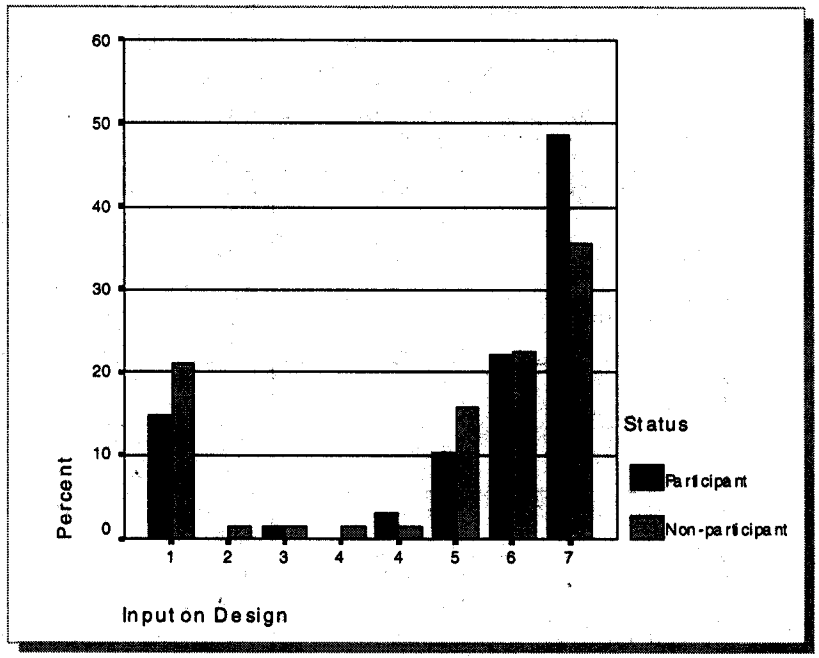


Figure 25: Reported Input on the Design Process

A comparison of participants and non-participants on the type of construction provides an additional check of the comparability of the participant and non-participant samples. There were no significant differences in the distribution of types of construction between the two groups. Figure 26 shows the type of construction by participant and non-participant status.

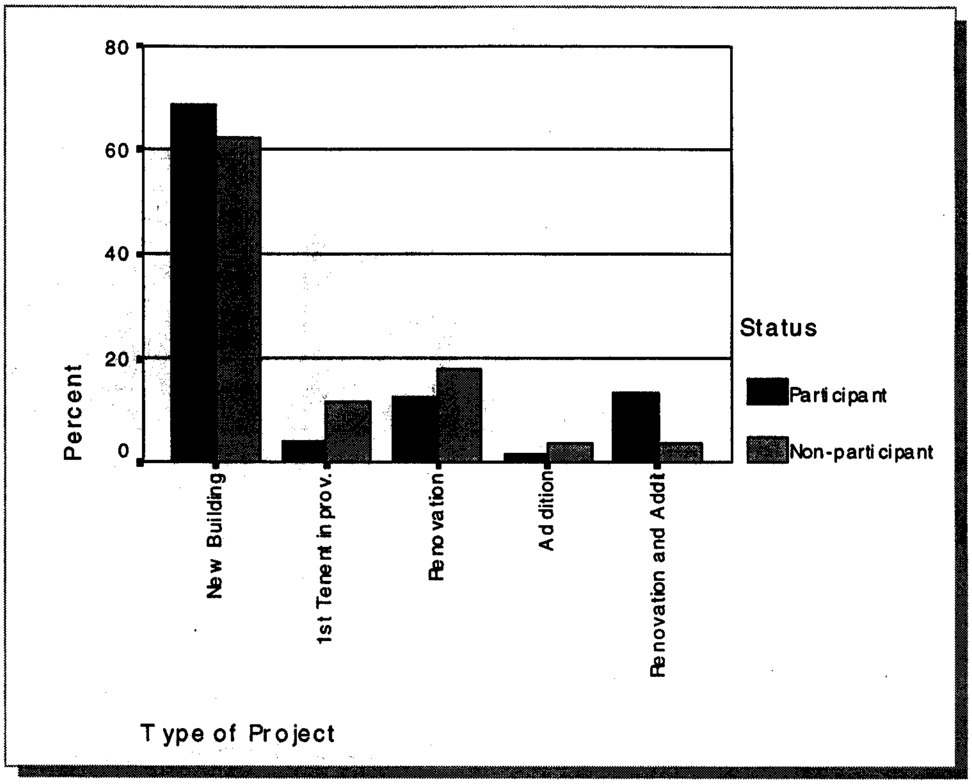


Figure 26: Type of Construction Project

Findings

Respondents were asked about their financial decision-making criteria. While the most common response for both groups was that a combination of criteria were used to make decisions. Participants were found to take a longer view of energy investments. Based on a chi-square test, program participants were significantly more likely to base decisions on a payback, or lifecycle cost, calculation. Non-participants were significantly more likely to base their decisions on lowest first cost. Figure 27 shows the responses to this question.

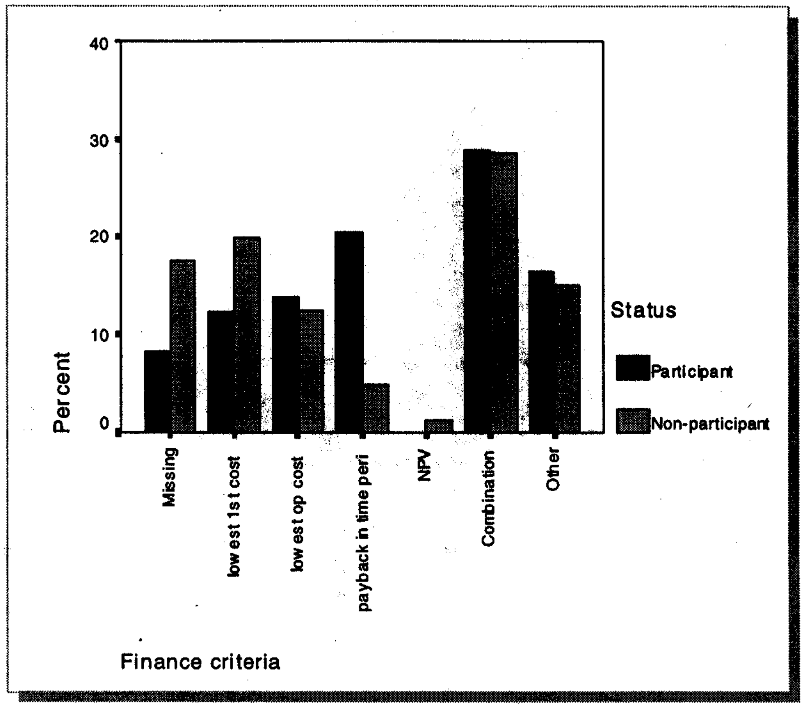


Figure 27: Financial Decision-making Criteria

Not surprisingly, participants report a greater awareness of the program. The participants report a mean awareness of 3.9 (median = 4) on a 7-point scale. A "1" indicates little familiarity with SCE's program and a "7" indicates complete familiarity with the program. Non-participants report a mean awareness of 2.55 (median = 2). The distribution of responses is shown in Figure 28.

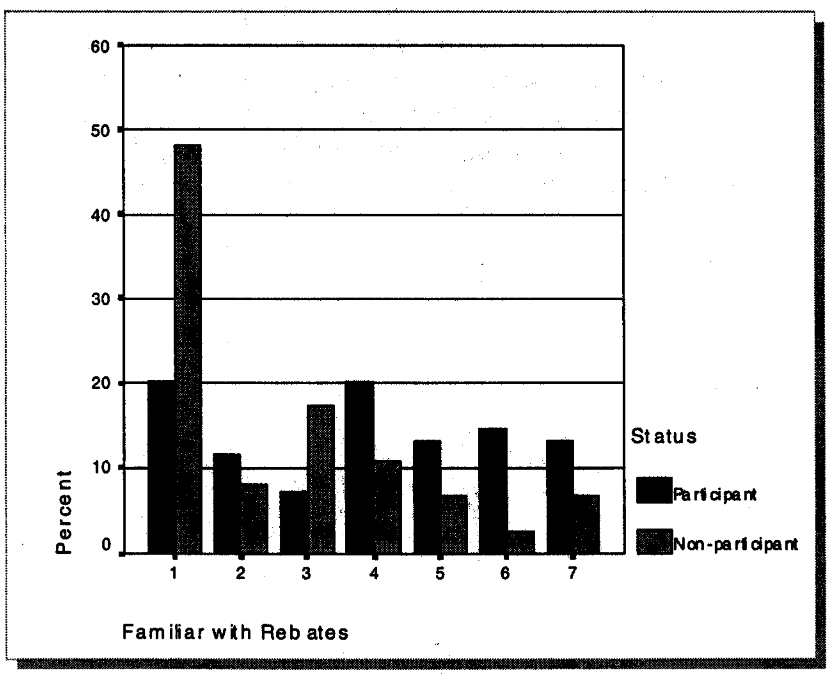


Figure 28: Familiar with Rebates

Participants report greater interaction with and influence of Edison on project decisions. Table 34 shows the mean and median scores on a 7-point scale. A "1" on the 7-point scale indicates no interaction or influence and a "7" indicates significant influence or interaction. Figure 29 and Figure 30 show the distribution of influence and interaction scores, respectively.

	Participants		Non-Participants	
	Mean	Median	Mean	Median
Influence of Edison on project	3.84	3	2.48	1
Interaction with Edison on project	4.03	4	2.38	1

Table 34: Influence and Interaction on Project

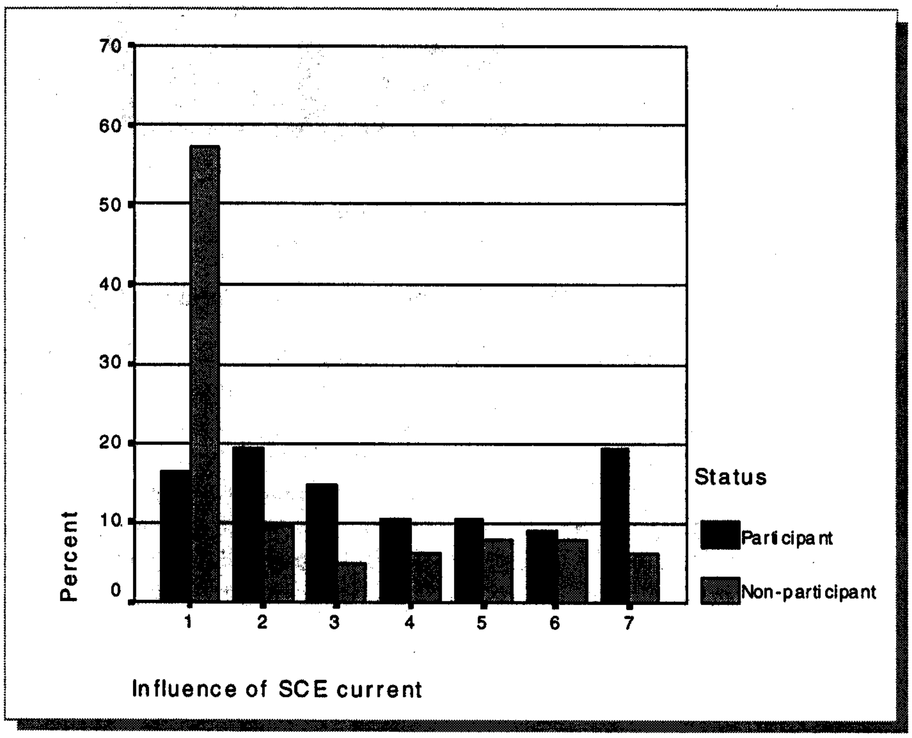


Figure 29: Influence of Edison on Project Decisions

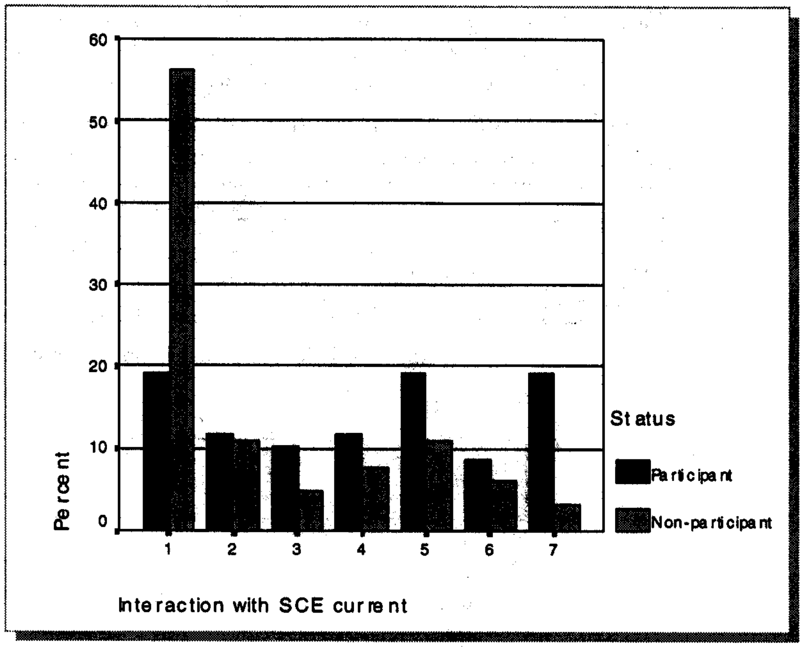


Figure 30: Interaction with Edison on Project Decisions

Respondents were asked about their interaction with and the influence of Edison on past projects. These responses were compared to the responses for the project being studied. Figure 31 and Figure 32 show that influence and interaction remain more or less unchanged over time. In all cases, the median response is 0, i.e. no change in influence or interaction. Participants have a slightly positive mean response, meaning an increase in influence and interaction, while non-participants have a slightly negative mean response, meaning less interaction and influence than in the past.

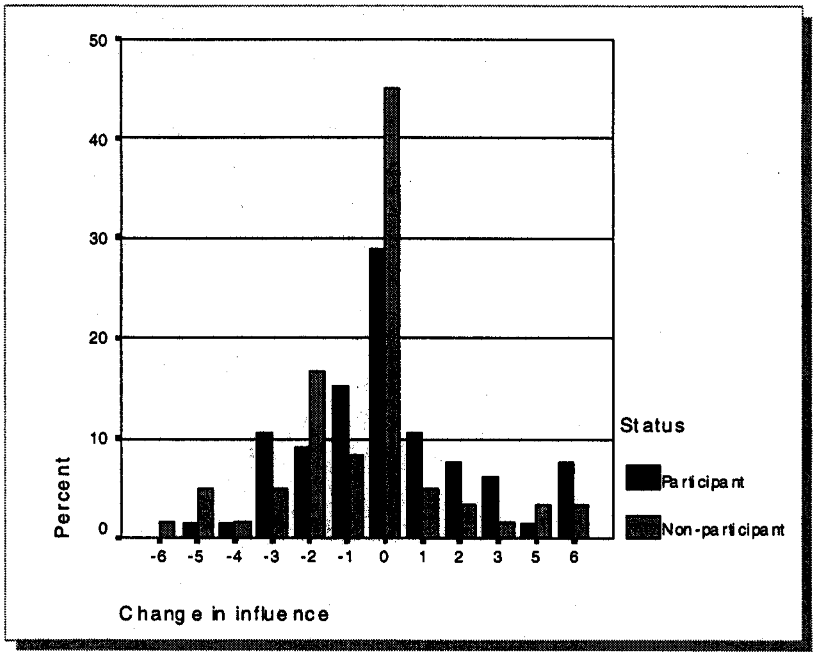


Figure 31: Change in Influence over Time

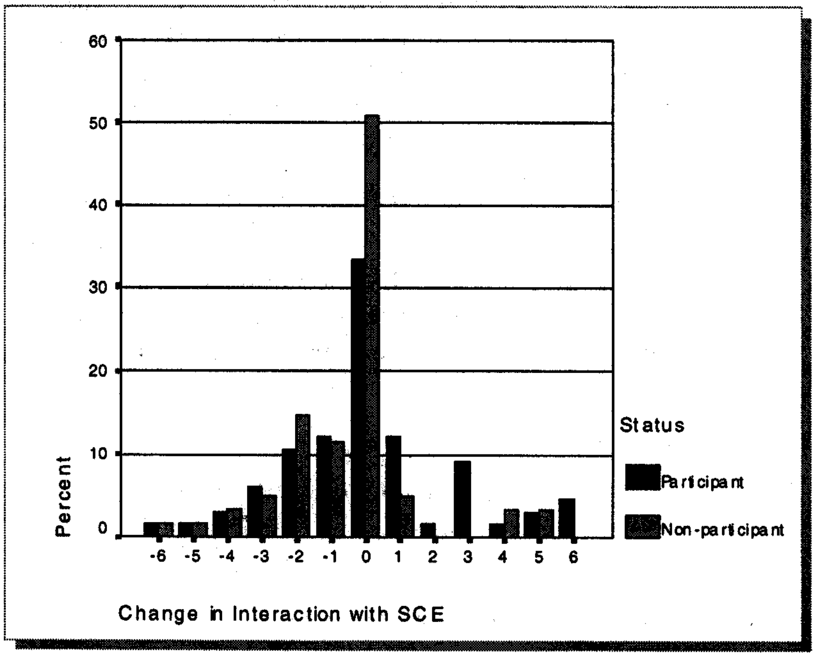


Figure 32: Change in Interaction Over Time

General Methodology for Data Screening and Analysis

A systematic process was followed to specify the final logistic and efficiency choice models. The present section summarizes how each of the following issues were addressed. Additional details will be found in other sections of the report, especially the following sections of this chapter.

- Weather adjustment
- Background variables such as economic activity
- Missing data points
- Missing or unusable billing data
- Missing responses to questions
- Outliers and data screens
- Model specification
- Cross sectional variation
- Time series variation
- Participant self selection
- Omitted factors
- Estimation of net impacts
- Errors in measuring variables
- Autocorrelation
- Heteroscedasticity
- Collinearity
- Influential data points
- Statistical Precision

Weather adjustment

This was handled in the engineering modeling. The model calibration used actual weather concurrent with the available billing data. Then all models were run using typical meteorological weather data. In this way the gross savings determined by the engineering models reflected normal weather conditions expected in each climate zone.

Background variables such as economic activity

This was also handled in the engineering modeling. The schedules used in the models were based on the levels of building use observed in the onsite survey. The schedules were held fixed in calculating the gross savings. Therefore the savings can be regarded as representing the actual savings obtained under the economic activity found at the time of the onsite surveys.

Missing data points

Sites that refused to participate in the study were replaced using a randomly drawn sample of backup sites. The level of refusal was rather low, as discussed earlier in this report.

Missing or unusable billing data

Whenever possible, the engineering models were calibrated to the available billing data. However, many of the projects studied in this evaluation were actually renovations or additions to existing buildings. In many of these cases, the available billing data described the whole building rather than the actual space that was renovated or added. In these cases, when it was practical we installed special metering equipment to collect load data for use in calibration. When this was not practical, the models were used without calibration.

Missing responses to questions

When a decision-maker did not know or refused to answer a particular question, we tried to identify a more appropriate respondent. If this failed, we recorded the response as 'don't know' or 'refused'. In the case of questions with categorical answers, we treated all such answers as a distinct category of response and created a corresponding indicator variable. In the case of the questions that were answered on a seven-point scale, we coded the response as 0 and created a corresponding indicator variable.

Outliers and data screens

The full sample was retained throughout the analysis. Studentized residuals were used to identify outliers. A site was considered to be an outlier if its studentized residual was greater than three in absolute value. A separate indicator variable was used to represent each such outlier in the model. The coefficient of this indicator variable indicated how much the dependent variable deviated from its expected value for the particular outlier. The statistical significance of these indicator variables were used to identify outliers that were statistically significant.

Model specification

A systematic approach was followed so that each model would be properly specified. The primary concern was to avoid bias arising from specification error – omitted variables, outliers, omitted statistical interactions, etc. We also sought to obtain a parsimonious final model that included only statistically significant variables. The following sections trace the approach, indicate some of the tests and graphical displays that were used to examine intermediate models, and compare the models that were examined. The entire process of refining the models is documented in SPSS command files.

Cross sectional variation

Cross-sectional variation was addressed throughout the sample design and experimental approach as well as in the modeling. The sample design was based on the experience of the 1994 evaluation study and sought to represent the full diversity of participants in the program, and a matched sample of non-participants. The sample size and stratification were chosen to yield statistically reliable estimates of the overall savings of the program. The experimental approach was built around engineering surveying and modeling techniques that were designed to capture the full range of actual building equipment types and schedules found in the population. The gross analysis was designed to determine the actual gross savings of each site, controlled for the actual equipment and use of the site. The net-to-gross analysis was designed to control for additional factors affecting the decision making process.

Time series variation

In the gross analysis, time series variation was controlled by the simulation methodology. The gross savings were calculated by simulating the building with and without the energy efficiency measures but holding other equipment and schedules fixed as observed. Time-series variation was not an issue in the net-to-gross regression analysis since all observations reflected the same time period.

In other words, the regression modeling addressed variation from one same site to another, but not from one time point to another.

Participant self selection

Self selection was addressed in the net-to-gross analysis by developing a logistics model for the probability of participating, and then using the resulting double inverse Mills ratios as added explanatory variables in the efficiency choice models. The statistical significance and effect of the inverse Mills ratios were estimated and reported.

Omitted factors

Two factors might be discussed: the use of Title 24 documentation and billing data. The study sought to use both Title 24 documentation and billing data to the extent practical. When either Title 24 documentation or billing data was available, it was used to improve the accuracy of the engineering models. This approach allowed us to maintain the full sample even when these data were unavailable.

The evaluation of the 1994 program clearly demonstrated the difficulty of obtaining Title 24 documentation, especially for the non-participants. In order to avoid high refusal rates and the concomitant risk of nonresponse bias, we only insisted on Title 24 documentation for sites that used the tailored lighting approach or the performance-based approach to Title 24 compliance.

Billing data was used to calibrate each individual engineering model whenever possible. However, as described elsewhere, the available billing data did not always reflect the space affected by the new construction. In some of these cases, we sought to supplement the billing data with our own metering. Nevertheless, some of the sites did not have actual usage data. In such cases we trusted that the engineering models were accurate without calibration. To confirm this assumption, we compared the gross savings determined before and after calibration for the sites with billing data or our metering. This analysis confirmed that the pre-calibration models were very accurate.

Estimation of net impacts

The combination of statistical sampling, onsite surveys, site-specific engineering models, econometric analysis, and statistical expansion was carefully designed to provide an unbiased and statistically reliable estimate of net program savings. In particular, the decision-maker survey was designed to isolate self-selection bias and the long-run impact of the program on design practice. The model was specified to include any observable and statistically significant effects of the program on the energy efficiency of both participants and non-participants.

Errors in measuring variables

In the onsite surveys and engineering modeling we sought to obtain an accurate representation of each individual sample site. Past experience suggested that serious errors could arise from failing to model the space in the building actually affected by the new construction, or by failing to accurately describe some of the equipment and schedules of use. The present study addressed these problems by improved training and communication with the surveyors, earlier retrieval and review of program files, having the surveyors themselves responsible for the data entry and modeling, and having the surveyors develop the model for a site soon after completing its survey. The engineering team met with SCE's program managers and reviewed the site-specific models in detail. We also redesigned the decision-maker survey, streamlined the process used to recruit each site and complete the decision maker survey, and assigned the responsibility for the whole process to a single, very competent person. All of these measures resulted in much more accurate data going into the econometric analysis than in the prior study.

Autocorrelation

Autocorrelation was not an issue since, as explained above, the analysis was cross sectional.

Heteroscedasticity

Heteroscedasticity – the tendency of larger projects to have greater variation – was addressed in both the sample design and efficiency-choice regression models.

The MBSS methodology used in the sample design addressed heteroscedasticity by modeling the variation in savings as a function of the tracking estimate of savings or the square footage of each site and then using an efficiently stratified sampling plan to increase the probability of selecting large sites. This ensures that the sample is effectively focused where the savings are greatest, while retaining an unbiased representation of small and large projects alike.

The efficiency-choice regression models were specified to minimize the danger of heteroscedastisity by defining the dependent variable as the gross savings as a fraction of the baseline energy use. This specification is closely related to the weighted-least-square methodology resulting from the assumption that the residual variation in gross savings is proportional to the baseline energy use of each site. Graphical scatter plots of the studentized residuals were examined to confirm the absence of heteroscedasticity. In addition, a statistical test of homogeneity of variance was carried out to measure the statistical significance of differences in the variance of the residuals grouped by building type and by the level of efficiency predicted by the model.

Collinearity

Multicollinearity is generally a less serious problem in a cross sectional analysis than in a time series analysis. Our methodology was designed to protect against the type of problem that might arise in a cross sectional analysis. Extreme multicollinearity can cause computational problems. Several of the indicator variables used in the regression models were perfectly collinear. This occurred, for example, if a respondent who failed to answer a given question also failed to answer a second question. In this case the missing-response indicators would be perfectly collinear. The SPSS software used in the analysis identifies and reports these instances and automatically drops one of the variables from the analysis. The software also provides a warning if the multicollinearity is strong enough to affect the numerical accuracy of the estimated coefficients. In practice there was no indication of a serious problem with numerical accuracy.

When explanatory variables have strong but not extreme multicollinearity, it is important to guard against obtaining biased results. Omitted-variable bias can arise if one of the correlated variables is dropped from the model. We guarded against this possibility by systematically comparing the estimated coefficients of our various models and looking for other indicators such as large shifts in statistical significance.

Influential data points

We followed diagnostic procedures recommended by Belsley, Kuh and Welsh.[11] Our key indicator of an influential observation was the studentized residual which can be related to the t-distribution. We also examined normal probability plots, partial-regression leverage plots for each explanatory variable, and other case-specific measures of influence. When an influential observation was identified, we included an indicator variable in the analysis that was 1 for the influential observation and 0 for all other cases in the sample. We retained this variable if it was statistically significant in the final model.

Statistical Precision

In each regression model, we used standard logistics or least-squares techniques to calculate the standard error and statistical precision of each coefficient. We used the standard MBSS statistical techniques described in the Gross Savings chapter to expand to the econometric estimates for each sample site to the population and to measure the statistical precision of the results.

Overview of the Econometric Net-to-Gross Methodology

Under the econometric approach, the net-to-gross ratio was calculated in the following seven steps. For simplicity we will discuss the methodology used to analyze annual energy savings. An analogous approach was used to analyze summer peak demand savings.

1) **Dependent Variable:** For each site in the combined participant-non-participant sample, calculate the efficiency choice of each site; this is the difference between the baseline and as-built energy use as a fraction of the baseline energy use. The efficiency choice was the dependent variable, i.e., the *y*-variable, in the regression analysis.

2) **Analysis Data Base:** For each site in the combined participant-non-participant sample, create an indicator variable for program participation, and indicator variables reflecting the responses to the categorical questions in the decision-maker survey. Create indicator variables to identify missing data to each of the decision-maker questions. Create indicator variables to identify the building-type categories. Include the scale response variables from the decision-maker survey as additional potential explanatory variables.

3) **Logistic Regression Model:** Develop a logistic regression model to estimate the probability that each sample site is a participant. Use the preceding indicator variables as well as the scale response variables as

[11] D. A. Belsley, E. Kuh and R. E. Welsch, *Regression Diagnostics*, Wiley, 1980.

possible explanatory variables in the model. Examine the model for outliers and other violations of the assumptions of logistics regression. Drop explanatory variables that are not statistically significant. Use the simplified logistics model to calculate the predicted probability that each site in the combined sample is a participant. Then use the predicted probabilities to calculate inverse Mills ratios in order to correct for possible self-selection bias.

4) **Efficiency choice Regression Model:** Formulate a regression model explaining the variation in efficiency choice as a function of various variables describing the participants and non-participants. The explanatory variables included the following:

    (a) The indicator variable for program participation,

    (b) Indicators describing the type of building,

    (c) Indicators for the decision makers planning process and priorities, concern about energy, etc.

    (d) Scale variables measuring the degree of interaction with SCE and the amount of influence SCE had on the design of this project and of past projects, and

    (e) The inverse Mills and double inverse Mills ratios, and

    (f) Indicators for potential outliers.

5) **Model Diagnostics and Simplification:** Examine suitable graphs and statistics to determine the adequacy of the regression model. Simplify the regression model by dropping statistically insignificant variables. Add statistically significant interaction variables.

6) **Net Savings:** Use the simplified regression model to estimate the net savings attributable to the program for each sample participant, after statistically controlling for the efficiency choice of non-participants, any significant differences between participants and non-participants in the other explanatory variables, and self selection via the inverse Mills and double inverse Mills ratios. Then use the statistical sampling methods to expand the net savings attributable to the program for each sample participant to the population of 1996 program participants, as described in the Gross Savings chapter. Finally, calculate the error bound and relative precision of the results using the usual statistical sampling methods.

7) **Spillover:** Use the simplified regression model to estimate the spillover effect of the program for each sample non-participant. Then use standard statistical sampling methods to expand the net savings attributable to the program for each sample non-participant to the population of 1996 non-participants, using Dodge new construction data. Finally, calculate the error bound and relative precision of the results using the usual statistical sampling methods.

The Data Base for the Econometric Analysis

The analysis database consisted of 153 sample observations with twenty variables. Fifty-one additional indicator variables were created to reflect the building types, categorical survey information and missing responses to specific questions. Several additional indicator variables were created to represent individual sample sites that appeared to be outliers in the preliminary residual analysis. Additional variables were created within the analysis for statistical interactions, for the Mills ratios, and for various diagnostic tests.

The Logistic Regression Model

As previously indicated, the objective of this task was to develop a logistic regression model to estimate the probability that each sample site is a program participant.

Table 35 summarizes the final logistic model. The column labeled B is the regression coefficient for each explanatory variable. A positive value indicates a higher probability of being a program participant whereas a negative value indicates a lower probability. For example, a restaurant was more likely to be a program participant, whereas a site owned by the Federal government was relatively unlikely to be a program participant.

Large coefficients were obtained for several indicator variables for missing data – ownership missing, input missing, and current interaction missing - as well as for the indicator variable for site 2769. These variables had a relatively large standard error (S.E.) indicating that the coefficient was not measured very reliably. Moreover, their statistical significance (Sig) was generally poor, e.g., greater than 0.10. So these variables are probably not very important in the model. [12]

[12] These variables were retained in the final model because the output of the regression analysis indicated deleting each of these variables would cause a statistically significant drop in the log likelihood ratio.

Explanatory Variable	B	S.E.	Sig
restaurant	2.56	1.25	0.04
retail	1.55	0.71	0.03
warehouse	3.90	1.97	0.05
Owned by Fed Govt	-4.16	1.65	0.01
Ownership missing	-15.57	99.66	0.88
Built by Owner for tenent	-1.99	0.86	0.02
Other	2.34	1.40	0.10
No Preexisting Plans	1.30	0.70	0.06
payback in time period	4.11	1.18	0.00
Fin Criteria missing	-3.32	1.34	0.01
Commisioning Yes	1.50	0.84	0.07
Commisioning missing	4.00	1.59	0.01
Familiar with Rebates	0.53	0.15	0.00
Input on Design	0.31	0.17	0.07
Input missing	17.96	22.90	0.43
Energy Costs	-0.89	0.38	0.02
Energy Efficiency	1.17	0.39	0.00
Interaction with SCE current	0.71	0.24	0.00
Curr Inter missing	-9.52	22.77	0.68
Influence of SCE current	-0.37	0.22	0.10
Interaction with SCE past	0.43	0.15	0.01
Indicator for c2769	-20.62	99.66	0.84
Constant	-8.97	1.83	0.00

Table 35: Logistic Model Coefficients

The preceding model was developed in the following steps.

1. Estimate a logistic regression model relating the dependent variable – the indicator of program participation – to all of the potential explanatory variables. Measure the fit, save the diagnostic statistics, and examine the diagnostic graphs. This analysis suggested that site 2769 was an outliner.

2. Estimate a second regression model relating the dependent variable to all of the potential explanatory variables, plus an indicator variable for site 2769. Measure the fit, save diagnostic statistics, and examine diagnostic graphs.

3. Use backward stepwise regression to eliminate the statistically insignificant variables from the preceding model. Use a p-value of 0.05 for adding variables and 0.10 for deleting variables.

4. Estimate the simplified model shown above, measure its fit, save its diagnostic statistics, and examine its diagnostic graphs.

The following figure is a sequence plot of the studentized residuals. The first 73 cases are the participants and the remaining cases are the non-participants. The residuals for the participants tend to be centered around 1 whereas the residuals for the non-participants tend to be centered around –1. This is due to the binary character of the response variable. Case 131 in the database - Site 2665 – has a studentized residual less than –3 and might be considered to be a second outlier. Otherwise, the residuals appear to be random.

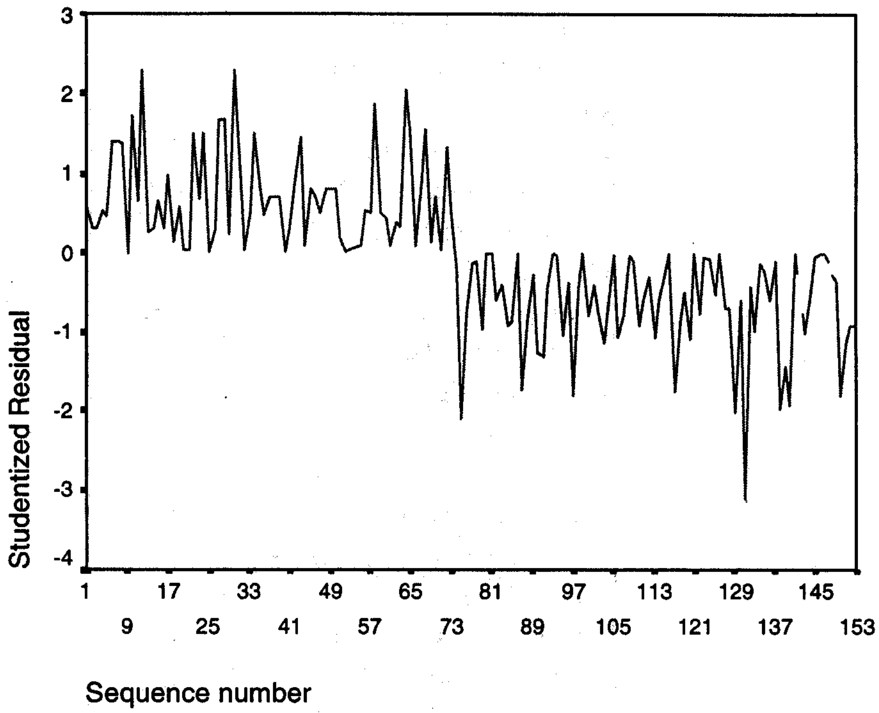


Figure 33: Logistic Model Studentized Residuals

Figure 34 shows a normal probability plot for the studentized residuals of the model. This is a tool to assess the hypothesis of a normal probability distribution that is the basis of the logistics analysis. If the hypothesis of a normal probability distribution is valid, then the plotted points should lie along the straight line. A failure of the residuals to be normally distributed may be indicated if the plotted points deviate substantially from the line. The figure shows that the observed residual below –3 would have been unexpected under the normal distribution. Otherwise the figure supports the hypothesis of a normal probability distribution.[13]

[13] An indicator variable for Site 2665 was not used in the logistic model since when it was added, the Q-Q plot appeared to violate the normal distribution.

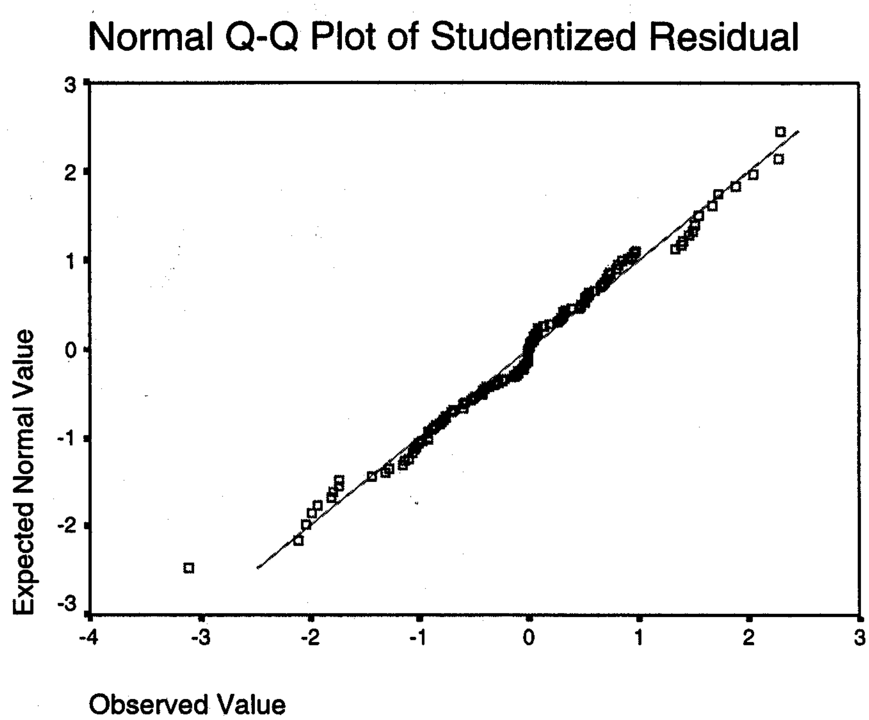


Figure 34: Normal Q-Q Plot of Studentized Residuals

The following table provides a common measure of the goodness of fit of the final model called the classification table. Of the 80 non-participants in the sample, the program correctly predicted that 71 were non-participants, for a score of 89% correct. Of the 73 participants, the program correctly predicted that 58 were participants, for a score of 79% correct. The overall score was 84%.

	Predicted		
Observed	**Non-participant**	**Participant**	**Percent Correct**
Non-participant	71	9	88.75%
Participant	15	58	79.45%
Overall			84.31%

Table 36: Logistic Model Participation Prediction

Two other measures were calculated reflecting the goodness of fit of the logistics model. The Nagelkerke R-squared statistic was 68% - indicating that the model explained 68% of the total variance. The statistical significance of the model was .0000 – indicating that the model was statistically very significant.

Another way to assess the simplified model is to compare its goodness of fit to the full model developed in step 2 of the analysis. This analysis indicated that the variables that were deleted from the full model were not statistically significant as a group. This suggests that the simplified model is an adequate summary of the relationship between program participation and the variables developed from the decision-maker survey. From all of the preceding analysis, we can conclude that the simplified model is a good predictive model for program participation.

The simplified logistic model was then used to estimate the probability that each site in the sample might have been a participant as a function of the characteristics of the site and the information about the decision-making process. For each site, let Z represent the numerical result of substituting the values of the explanatory variables into the logistic equation. Then the estimated probability is calculated using the equation

$$\hat{p} = \frac{e^Z}{1+e^Z}$$

The inverse Mills ratio was calculated as

$$C = \left[\frac{(1-\hat{p}) \times \ln(1-\hat{p})}{\hat{p}} + \ln(\hat{p})\right] \qquad \text{for participants, and}$$

$$C = -1 \times \left[\frac{\hat{p} \times \ln(\hat{p})}{1-\hat{p}} + \ln(1-\hat{p})\right] \qquad \text{for non-participants.}$$

The double inverse Mills ratio was calculated by multiplying C by the indicator variable for program participation.[14] These variables were labeled *Mills ratio* and *Double Mills ratio*, respectively.

Efficiency Choice Regression Model for Annual Energy

The objective of this task was to develop a regression model to estimate the efficiency choice of each sample site, participant and non-participant. The efficiency choice of each sample site was measured as the difference between as built and baseline use as a fraction of baseline use.

Table 37 summarizes the final efficiency choice model. The column labeled B is the regression coefficient for each explanatory variable. A positive value indicates a higher efficiency choice whereas a negative value indicates a lower efficiency choice. For example, the model indicates that a program participant tended to have a 0.132 higher efficiency choice than a non-participant. The econometric standard error of this estimate was 0.030 indicating that the error bound at the 90% level of confidence was 1.645 * 0.030 = 0.049. The 90% confidence interval for the true value is 0.132 ± 1.645 * 0.030 = (0.083, 0.181). The program participant coefficient was statistically significant at the .000 level of significance.

The next explanatory variable is based on the seven-point scale variable reflecting SCE's influence on the current project as reported by the decision-maker. The actual response was coded 1 to 7, denoting very weak to very strong,

[14] Net Savings Estimation: *An Analysis of Regression and Discrete Choice Approaches*, Prepared for the CADMAC Subcommittee on Base Efficiency, Prepared by Xenergy, Inc. Madison WI, by M. Goldberg and K. Train, Revised March 1996.

respectively. The variable, labeled Curr Influ_nonpart, measured the influence reported by a non-participant. The value of the variable was equal to the actual 1 to 7 response for a non-participant and zero for a participant. We will discuss the role of the participant and influence variables in detail in a later section.

The remaining variables represent other factors that were found to have a statistically significant effect on efficiency choice. The coefficients of these variables are generally reasonable. The model indicates that:

- Restaurants and retail buildings are less efficient
- Buildings owned by the Federal government or state or local governments are more efficient
- Buildings without preexisting plans are more efficient
- Building built by the owner for a tenant are more efficient
- Buildings using the lowest operating cost financial criterion or a financial criteria that is a combination of several factors are more efficient

The model indicates higher efficiency if the decision maker was familiar with the rebates, and had an input on the design. The coefficients of all of these variables are plausible.

Four outliers were identified in the analysis. Site 1253 had significantly lower efficiency choice that expected, whereas sites 807, 497 and 1276 all had higher efficiency choice than expected given the other factors included in the model. Both inverse Mills ratios were statistically significant, indicating that the variation explained by the corrections for self selection was greater than expected by chance.

Coefficients

		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.058	.037		-1.580	.117
	Program Participant	.132	.030	.397	4.354	.000
	Curr Influ_nonpart	.017	.008	.184	2.251	.026
	Restaurant	-.090	.041	-.134	-2.183	.031
	Retail	-.055	.023	-.145	-2.346	.020
	School	.074	.030	.152	2.495	.014
	Owned by Fed Govt	.143	.064	.136	2.226	.028
	Owned by State or Local Govt	.058	.034	.110	1.717	.088
	Other Ownership	.214	.085	.146	2.519	.013
	Built by Owner for tenent	.068	.031	.132	2.178	.031
	Lowest op cost	.077	.030	.156	2.550	.012
	Combi. Crit.	.069	.023	.188	3.007	.003
	Familiar with Rebates	.012	.006	.159	2.238	.027
	Input on Design	.012	.005	.182	2.508	.013
	Input missing	.171	.053	.241	3.211	.002
	indicator for c1253	-.857	.126	-.414	-6.820	.000
	indicator for c807	.565	.118	.273	4.786	.000
	Indicator for c497	.373	.119	.180	3.137	.002
	Indicator for c1276	.307	.119	.148	2.573	.011
	Mills ratio	.037	.017	.235	2.114	.036
	Double Mill	-.090	.025	-.337	-3.577	.000

Table 37: Energy Model Coefficients

The following table provides several measures of the goodness of fit of the final model. The adjusted R square was .589 indicating that the model explains almost 60% of the total variation in efficiency choice. The F-statistic was 9.478, corresponding to a statistical significance of 0.000, indicating that the model as a whole was highly significant.

Model Summary

				Change Statistics			
	R	R Square	Adjusted R Square	F Change	df1	df2	Sig. F Change
	.768	.589	.527	9.478	20	132	.000

Table 38: Energy Model Summary

The figure below shows a normal probability plot for the deviancies of the final model. This is a tool to assess the hypothesis of a normal probability distribution that is the basis of the efficiency-choice regression analysis. If the hypothesis of a normal probability distribution is valid, then the plotted points should lie along the straight line. The figure suggests that this assumption is valid.

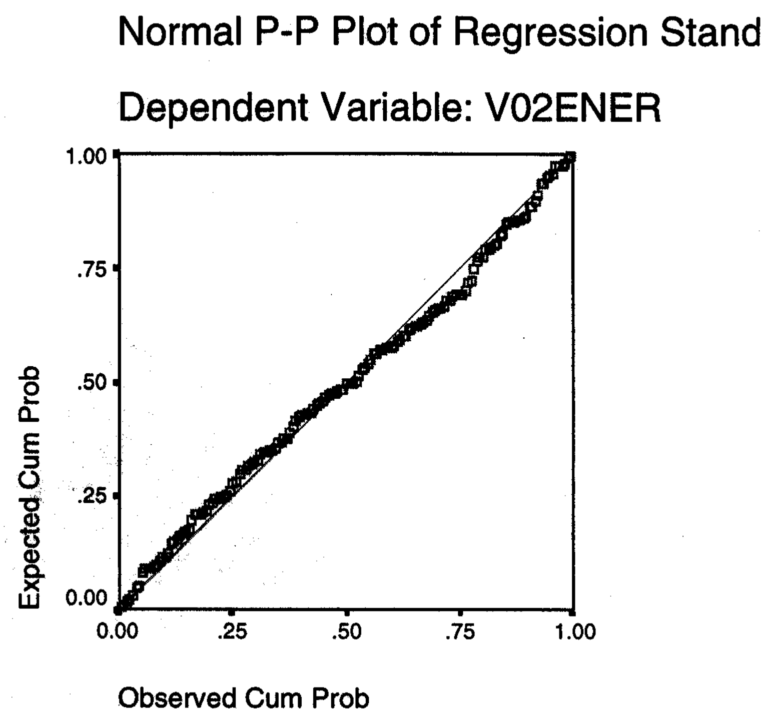


Figure 35: Energy Model Diagnostic Plot

The following figure shows a more conventional histogram of the standardized residuals of the model. Again the assumption of a normal distribution appears to be generally satisfactory. This evidence, together with the relatively large size of the sample, indicates that standard measures of statistical significance should be valid.

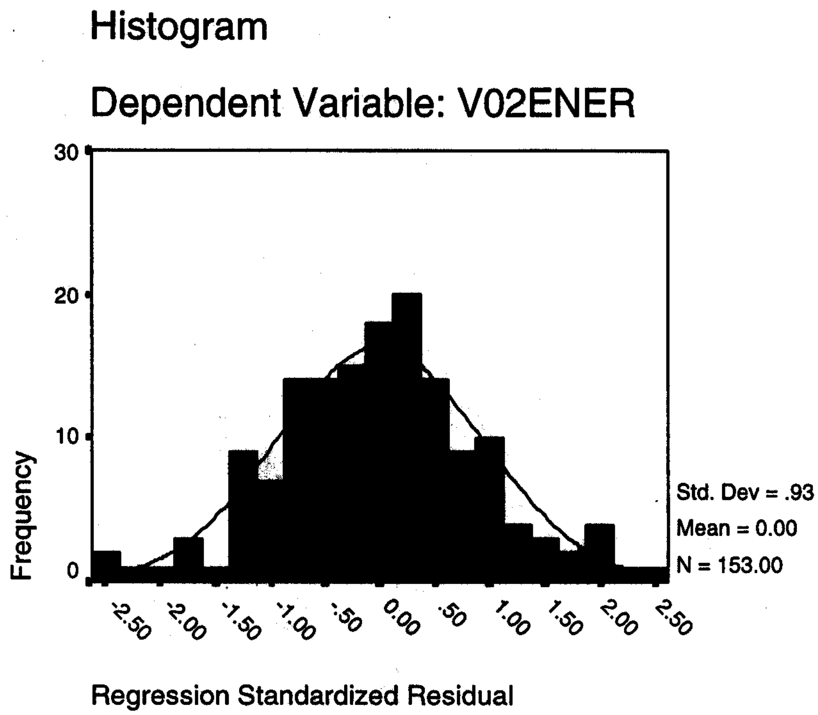


Figure 36: Histogram of Standardized Residuals from Energy Model

Figure 37 is a sequence plot of the studentized residuals of the final model. The residuals appear to be random and to indicate no remaining outliers.

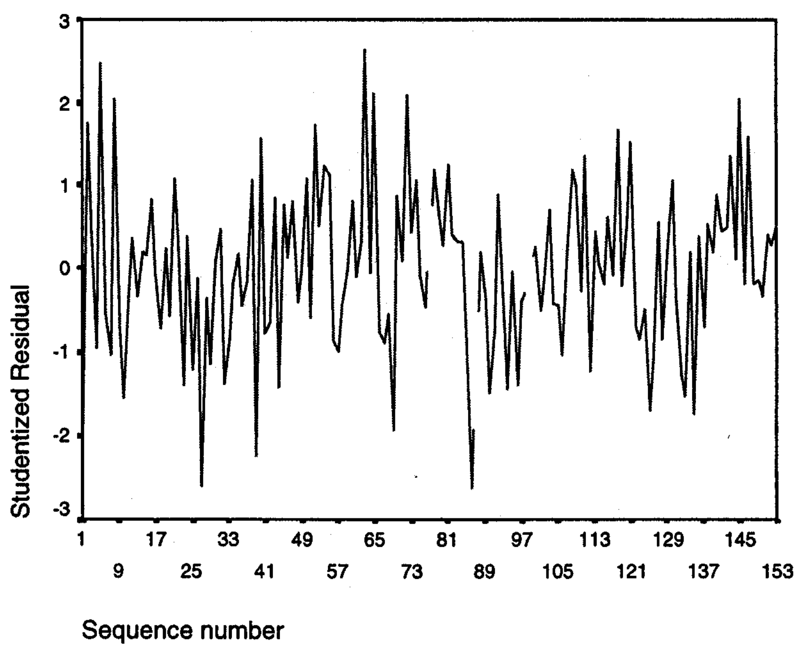


Figure 37: Energy Model Diagnostic Plot of Studentized Residuals

The remaining graph shows a scatter plot of the residuals compared to predicted values. The important issue is not the range of predicted values on the horizontal axis, but rather the range of the residuals on the vertical axis. Again this graph shows that the residuals are randomly distributed. Moreover, it shows that the residuals are homoscedastic. In other words, the variance of the residuals seems to be independent of the predicted values. To confirm this finding, we measured the homogeneity of the variance of the residuals grouped by building type and by the energy efficiency predicted by the model. The significance level of these results was .613 and .244 respectively, indicating no statistically significant heteroscedasticity in the residuals.

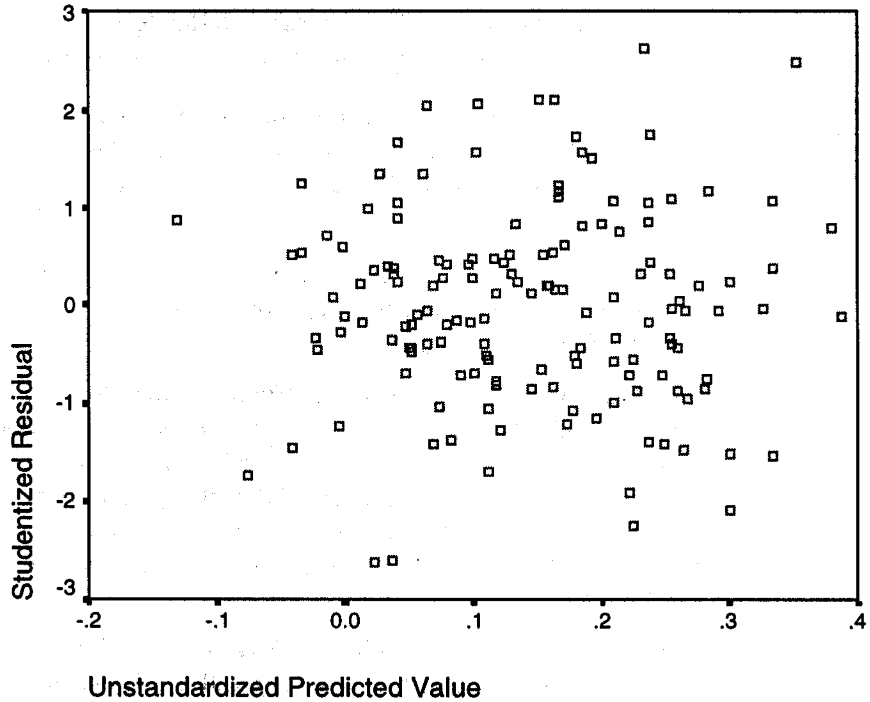


Figure 38: Scatter Plot of Studentized Residuals and Predicted Values

Another way to assess the simplified model is to compare its goodness of fit to the full model developed in the first step of the analysis. This analysis indicated that the variables that were deleted from the full model were not statistically significant as a group. This suggests that the simplified model is an adequate summary of the relationship between efficiency choice and the variables developed from the decision-maker survey. From all of the preceding analysis, we can conclude that the simplified model is a good predictive model for efficiency choice.

Comparison of Models

In seeking the most complete and parsimonious model for the energy efficiency choice, a sequence of regression models were examined. The following general steps were followed to obtain the final model.

1. Estimate a linear regression model relating the dependent variable – the efficiency choice of each site – to all of the potential explanatory variables. Measure the fit, save the diagnostic statistics, and examine the diagnostic graphs. This analysis suggested that sites 1253, 807, and 497 might be outliers. Appropriate indicator variables were added.

2. Use backward stepwise regression to eliminate the statistically insignificant variables from the preceding model. Use a p-value of 0.05 for adding variables and 0.10 for deleting variables.

3. Estimate the simplified model shown above, measure its fit, save its diagnostic statistics, and examine its diagnostic graphs. This analysis suggested that site 1276 is a high leverage case. An appropriate indicator variable was added and indicated that this site was also a significant outlier.

4. Estimate the final regression model relating the dependent to the reduced set of explanatory variables plus indicator variables for the four possible outliers. Measure the fit, save diagnostic statistics, and examine diagnostic graphs.

5. Test whether any omitted variables were statistically significant, including other statistical interaction variables between program participation and program interaction or influence.

The following table shows the coefficient of the program participation variable and its standard error for each of the energy efficiency-choice models that were estimated.[15] The table traces how the value of the coefficient changed as various variables were added or dropped. All of the models were based on the same underlying data. Models 1 through 9 trace the steps that were taken to obtain the final model. Models 10-13 reflect additional work to confirm the validity of the final model.

In models 1-3 we were seeking to identify and deal with outliers that might bias the results. The approach was to start with a full model reflecting all candidate explanatory variables, look at the various diagnostic statistics and graphs to check the validity of the model, and introduce corrections to any problems that are indicated. Our objective was to get a good model that passes the diagnostic statistics before working to simplify the model.

The first model included all of the candidate explanatory variables, including two outlier indicator variables, for cases 2769 and 2665. Case 2769 was the outlier identified in the logistic regression model. The studentized residuals of this model indicated that 1253 and 807 were also outliers, using 3.0 as the critical value. Model 2 was similar to model 1 but included these two added outliers. In Model 3, an indicator was added for c497, another outlier. Through these three models the coefficient was volatile, rising from .09 to .22, and the standard error was consistently high, about .4 or .5. With the addition of the four

[15] Many of the models included statistical interaction variables between program participation and other variables. These variables would affect the estimate of program impact.

outliers, the residual of the model passed the diagnostic tests very well, but the model contained many statistically insignificant variables.

In step 4 of the analysis, backward stepwise regression was used to eliminate the statistically insignificant variables. This analysis went through 40 automated iterations. In each iteration, the most statistically insignificant variable was dropped and the model was re-estimated.

Model 5 was the reduced model obtained from the stepwise regression. The elimination of many statistically insignificant variables from Models 1-3 improved the standard error of B.

Model 5 included indicators for three outliers, c1253, c807, and c497. The studentized residuals of this model passed the diagnostic tests. However, c1276 was a high leverage case. Model 6 was obtained by adding the corresponding indicator variable. Further analysis revealed that c1276 was the only site built on speculation, so V09.3 was dropped from the model, giving the final model, Model 7, reported in the preceding section. The steps from Model 5 to Model 7 had no effect on the participant coefficient or associated standard error but eliminated the spurious speculation effect from the model.

	B	S.E.	Description
1	0.088	0.483	First full model, c2769 and c2665
2	0.222	0.415	c1253, c807 added
3	0.203	0.395	c497 added
4	na	na	Stepwise regression, 40 iterations
5	0.132	0.030	Reduced model from stepwise
6	0.132	0.030	Add c1276
7	0.132	0.030	Drop V09.3, Final Model
8	0.191	0.049	Test statistical interactions
9	0.189	0.052	Test missing response indicators

Table 39: Energy Model Development Summary

Models 1-3 included several measures of the interaction of the decision maker with the utility and the influence of the utility on building design, both for participants and non-participants. Only one of these variables was retained in the final model, Model 7. Models 8-9 were run to confirm that the omitted variables did not have a statistically significant effect on efficiency choice. The omitted variables were added as a block to Model 7, and then the indicators for missing responses were added as a second block. The analysis confirmed that these variables were not statistically significant.

Analysis of Program Impact and Spillover

The final energy efficiency model was described in an earlier section. The first two explanatory variables reflected program participation and the influence of SCE on the current project with a non-participant. The Double Mills ratio was also statistically significant and included in the final model. The remaining variables all reflected factors other than the program. So the analysis of the program impact and possible spillover was based on the first two variables in the model, together with the Double Mills variable. By using the multivariate regression model, these results are adjusted for non-program factors that appear to influence the efficiency choice.

Neglecting the Double Mills variable, the efficiency choice regression model can be written as follows:

$$\text{Expected efficiency} = .132 * \text{Participant} + .017 * \text{influence on non-participant} + \text{other factors}$$

Here, the participant variable was 1 for a participant and 0 for a non-participant. The influence on a non-participant was measured on a seven point scale, with 1 indicating very weak influence and 7 indicating very strong influence. This variable was equal to 0 for a participant.

The energy efficiency model can be used to estimate the impact of the program on any particular sample site. This is done by calculating the difference between the expected energy efficiency predicted by the model and the energy efficiency that would be expected for the site in the absence of the program. In the absence of the program, the participant variable would be equal to 0. So the impact of the program on the expected energy efficiency can be calculated for a program participant as

$$\text{Added Efficiency} = 0.132$$

In other words, for a participant, the program increased the expected building efficiency by 0.132 regardless of the reported influence.

The model can also be used to estimate the impact of the program on the expected energy efficiency for a program non-participant. In the absence of the program we can set the rated influence to the lowest value of response, i.e., to 1.

This implies that the program increased the expected energy efficiency for a non-participant by:

$$\text{Added Efficiency} = 0.017 * (\text{influence on non-participant} - 1)$$

In other words, for a non-participant, the program appeared to increase the expected building efficiency by 0.017 times the level of influence of SCE less the level of influence expected in the absence of the program.

Table 40 shows the added efficiency due to the program implied by the model. The variable across the top of the table represents level of influence, if the site was a non-participant. The values in the table show the increase in expected efficiency due to the program, for both participants and non-participants, evaluated using the preceding two equations.

Curr Infl	1	2	3	4	5	6	7
Part	0.132	0.132	0.132	0.132	0.132	0.132	0.132
Non-part	0.000	0.017	0.034	0.051	0.068	0.085	0.102

Table 40: Added Efficiency Due to Program

This shows that for a participant, the program increased the expected efficiency by .132. In other words, the percent efficiency of the site relative to baseline was .132 higher than in the absence of the program, independent of the strength of the current influence. For participants, the percent efficiency did not vary with the current influence since this variable was not statistically significant for participants, perhaps due to lack of variation in the level of response. For a non-participant that was very weakly influenced by SCE, there was no increased efficiency due to the program, but for a non-participant that was very strongly influenced by SCE, the program increased the expected efficiency by .102. The expected efficiency choice of a participant is always higher than the expected efficiency choice of a non-participant. However, for those non-participants that reported a very strong influence, the efficiency choice was almost as high as for the participants.

This suggests that the program has two impacts. First the program has a direct net impact on the participants. Second, the program appears to have an indirect or spillover impact on the non-participants.

Under certain assumptions, the double Mills variable should also be considered to obtain an unbiased estimate of participant savings. The adjustment was obtained by multiplying the value of the double Mills variable by the corresponding coefficient in the efficiency choice model. Since the coefficient was -.09, the double Mills adjustment decreased the expected efficiency choice of participants. For the participants, the double Mills had a mean of .7 and a range from 0 to 3.4. So the average reduction was 0.062 but the maximum reduction was .306. The magnitude of adjustment was greatest for program participants that were most unlikely to participate under the logistic regression model previously discussed.

The inclusion of this adjustment is debatable. The double Mills adjustment can be shown to eliminate bias that might arise if the impact of the program on participants is randomly distributed. Unfortunately, simulations have shown that the validity of the adjustment is highly dependent on the form of the random distribution. In some simulated experiments, the double Mills adjustment actually increases the bias rather than decreasing the bias.[16] Because of the problems associated with the Double Mills ratio, the net savings estimates were calculated both with and without the Double Mills ratio.

The next step in the analysis process was to use the energy-efficiency regression model to estimate the net direct impact of the program. For each participant we calculated the net annual kWh savings due to the program by multiplying the base annual energy use of the site by the estimated increase in efficiency due to the program, calculated from the preceding equation. Results were calculated both with and without the Double Mills adjustment. Then these results were expanded to the population of program participants.

---

[16] Net Savings Estimation: *An Analysis of Regression and Discrete Choice Approaches*, Prepared for the CADMAC Subcommittee on Base Efficiency, Prepared by Xenergy, Inc. Madison WI, by M. Goldberg and K. Train, Revised March 1996.

The final step was to use the energy-efficiency regression model to estimate the direct impact and spillover impact of the program. In this analysis, we worked with the non-participants in the sample. For each non-participant, we calculated the net annual kWh savings due to the program by multiplying the base annual energy use of the site by the estimated increase in efficiency due to the program, calculated from the preceding non-participant equation. Then we used the Dodge database to expand the sample non-participants to the population of new construction. To ensure a conservative estimate of spillover, we made an adjustment to factor out any participant sites that may have been present in the Dodge database. To accomplish this, the sample spillover was projected to both the participant and new construction population and the participant population estimate was subtracted from the new construction population estimate.

Table 41 shows the net savings estimate and the estimate of spillover, together with the relative precision of each estimate. The difference-of-differences net savings was 26,621 MWh. The econometric approach yielded a direct net savings of 26,383 MWh. These results were calculated without the Double Mills adjustment. The difference between the two estimates can be thought of as the effect of self-selection bias not accounted for in the difference-of-differences approach.

Table 41 also shows the calculation of total net savings, counting both participant savings and spillover to non-participants. These results were adjusted for the Double Mills factor. The table first shows the net participant savings after the Double Mills adjustment. Then the table shows the spillover savings in the full Dodge database. Finally the table shows the adjustment to factor out any participant sites included in the Dodge database. This analysis yielded a net savings of 41,005 MWh and a net to gross ratio of 96%, with a relative precision of ± 31%.

To be conservative, the total net savings including spillover can be written down based on the relative precision to obtain a conservative estimate of total net savings, of 31,273 MWh. The conservative estimate was obtained by reducing the point estimate by 1.28 times the standard error of the estimate. This strategy factors in the statistical allowance for error and is very conservative. The probability is high, about 0.9, that this procedure underestimates the true value of savings. No relative precision is associated with these results since the relative precision is already reflected in the estimate itself.

	Estimate	Net-to-Gross Ratio	Relative Precision
Net Savings of Participants	26,383 MWh	61.7%	±10.2%
Adjusted Net Participant Savings	13,760 MWh		±17.0%
Spillover in NP population	29,146 MWh		±42.2%
Spillover in P population	1,901 MWh		±32.5%
Spillover Estimate	27,245 MWh		±45.2%
Total Savings	41,005 MWh	96.0%	±30.5%
Conservative Savings	31,273 MWh	73.2%	--

Table 41: Net Energy Savings and Spillover Estimates

Summer Demand Regression Model

The objective of this task was to develop a regression model to estimate the summer demand efficiency choice of each sample site. The analysis followed the same steps as the efficiency choice for annual energy, reported in the preceding sections. The dependent variable is the summer peak demand savings divided by the summer peak baseline demand of each model.

Table 42 summarizes the final efficiency choice model for summer peak demand. Many of the variables and coefficients in this model are similar to the efficiency choice model for energy. The model was developed in the same fashion as the energy efficiency choice model. The model includes seven indicators for outliers. The model for demand does not include the Mills or Double Mills variables since these variables were not statistically significant.[17]

[17] The two Mills variables had a joint statistical significance of 0.212. The significance of the Mills and Double Mills variables individually was .214 and .080 respectively.

Coefficients

		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.054	.025		-2.159	.033
	Program Participant	.111	.022	.374	5.152	.000
	Curr Influ_nonpart	.013	.006	.158	2.164	.032
	exhibit	.093	.048	.112	1.958	.052
	other	.066	.034	.125	1.953	.053
	Restaurant	-.086	.035	-.143	-2.418	.017
	School	.072	.026	.166	2.772	.006
	Private Owned and rented	.051	.021	.154	2.372	.019
	Owned by Corp for Franchise or sub	.134	.058	.145	2.315	.022
	Owned by Fed Govt	.161	.055	.173	2.948	.004
	Owned by State or Local Govt	.080	.029	.169	2.748	.007
	Lowest op cost	.074	.026	.168	2.830	.005
	Combi. Crit.	.065	.020	.199	3.216	.002
	Commisioning missing	-.101	.050	-.143	-2.025	.045
	Familiar with Rebates	.016	.005	.229	3.484	.001
	Input missing	.108	.039	.171	2.752	.007
	indicator for c1253	-.628	.106	-.341	-5.912	.000
	indicator for c807	.588	.102	.320	5.740	.000
	Indicator for c497	.306	.102	.166	2.989	.003
	Indicator for c1276	.295	.104	.160	2.848	.005
	indicator for c2572	-.386	.103	-.210	-3.746	.000
	Indicator for c2382	.347	.103	.188	3.370	.001
	Indicator for c2688	.318	.115	.173	2.761	.007

Table 42: Demand Model Coefficients

Table 43 gives several measures of the goodness of fit of the model. The adjusted R square was .54 indicating that the model explains over 50% of the total variation in efficiency choice. The F-statistic was 9.294, corresponding to a statistical significance of 0.000, indicating that the model as a whole was highly significant.

Model Summary

	R	R Square	Adjusted R Square	Change Statistics			
				F Change	df1	df2	Sig. F Change
	.782	.611	.546	9.294	22	130	.000

Table 43: Demand Model Summary

Figure 39 that shows a normal probability plot for the deviancies of the final model. This is a tool to assess the hypothesis of a normal probability distribution that is the basis of the regression analysis. If the hypothesis of a normal probability distribution is valid, then the plotted points should lie along the straight line. The figure suggests that this assumption is valid.

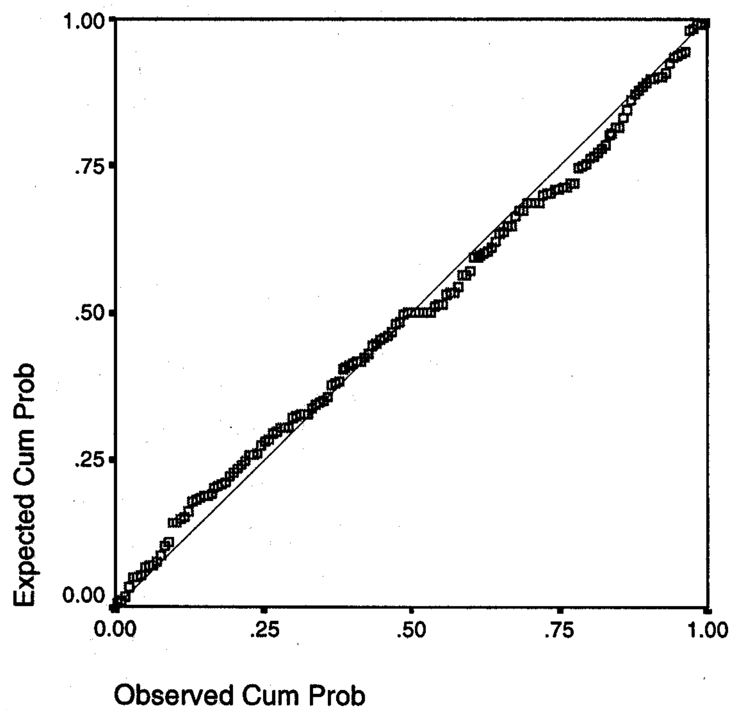


Figure 39: Demand Model Diagnostic Plot

Figure 40 shows a more conventional histogram of the standardized residuals of the model. Again the assumption of a normal distribution appears to be generally satisfactory. This evidence, together with the relatively large size of the sample, indicates that standard measures of statistical significance should be valid.

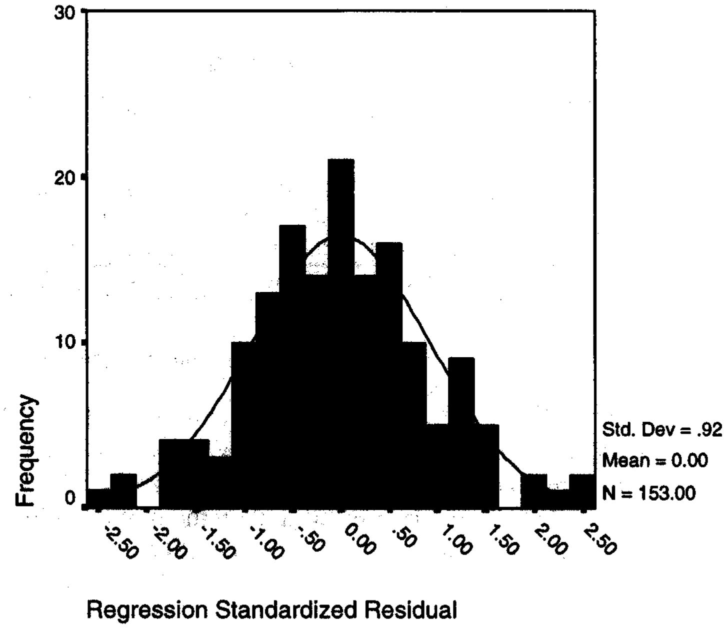


Figure 40: Demand Model Histogram of Residuals

The following figure is a sequence plot of the studentized residuals of the final model. The residuals appear to be random and to indicate no remaining outliers.

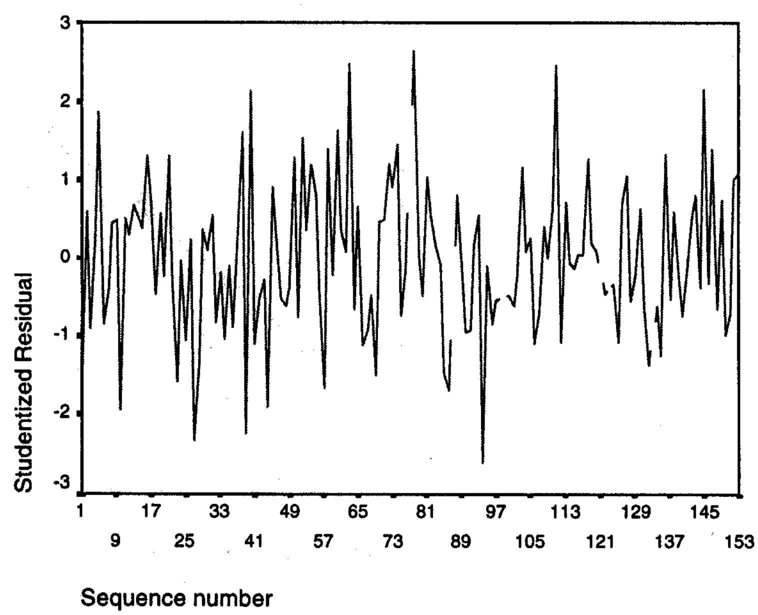


Figure 41: Demand Model Residual Plot

The remaining graph shows a scatter plot of the residuals compared to predicted values. The important issue is not the range of predicted values on the horizontal axis, but rather the range of the residuals on the vertical axis. Again this graph shows that the residuals are randomly distributed. Moreover, it shows that the residuals are homoscedastic. In other words, the variance of the residuals seems to be independent of the predicted values. To confirm this finding, we measured the homogeneity of the variance of the residuals grouped by building type and by the energy efficiency predicted by the model. The significance level of these results was .546 and .406 respectively, indicating no statistically significant heteroscedasticity in the residuals.

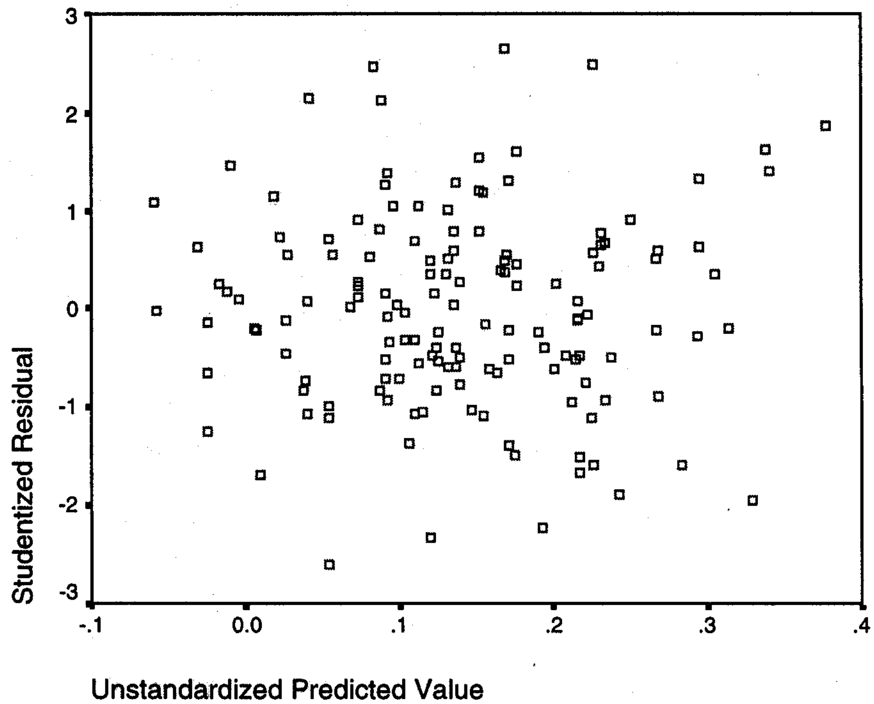


Figure 42: Demand Model Residual Scatter Plot

Another way to assess the simplified model is to compare its goodness of fit to the full model developed in the first step of the analysis. This analysis indicated that the variables that were deleted from the full model were not statistically significant as a group. This suggests that the simplified model is an adequate summary of the relationship between efficiency choice and the variables developed from the decision-maker survey. From all of the preceding analysis, we can conclude that the simplified model is a good predictive model for efficiency choice.

Analysis of Program Impact and Spillover

The final demand efficiency model was discussed in the preceding section. The first two explanatory variables reflected program participation and the influence of SCE on the current project with a non-participant. The remaining variables all reflected factors other than the program. So the analysis of the program impact and possible spillover was based on the first two variables in the. By using the multivariate regression model, these results are adjusted for non-program factors that appear to influence the efficiency choice.

The efficiency choice regression model can be written as follows:

Expected efficiency = .111 * Participant

+ .013 * influence on non-participant

+ other factors

Here, the participant variable was 1 for a participant and 0 for a non-participant. The influence on a non-participant was measured on a seven point scale, with 1 indicating very weak influence and 7 indicating very strong influence. This variable was equal to 0 for a participant.

The energy efficiency model can be used to estimate the impact of the program on any particular sample site. This is done by calculating the difference between the expected energy efficiency predicted by the model and the energy efficiency that would be expected for the site in the absence of the program. In the absence of the program, the participant variable would be equal to 0. So the impact of the program on the expected energy efficiency can be calculated for a program participant as

$$\text{Added Efficiency} = 0.111$$

In other words, for a participant, the program increased the expected building efficiency by 0.011.

The model can also be used to estimate the impact of the program on the expected energy efficiency for a program non-participant. In the absence of the program we can set the rated influence to the lowest value of response, i.e., to 1. This implies that the program increased the expected energy efficiency for a non-participant by:

$$\text{Added Efficiency} = 0.013 * (\text{influence on non-participant } -1)$$

In other words, for a non-participant, the program appeared to increase the expected building efficiency by 0.013 times the level of influence of SCE less the level of influence expected in the absence of the program.

Table 44 shows the added efficiency due to the program implied by the model. The variable across the top of the table represents level of influence, if the site was a non-participant. The values in the table show the increase in expected efficiency due to the program, for both participants and non-participants, evaluated using the preceding two equations.

Curr Infl	1	2	3	4	5	6	7
Part	0.111	0.111	0.111	0.111	0.111	0.111	0.111
Non-part	0.000	0.013	0.026	0.039	0.052	0.065	0.078

Table 44: Added Efficiency Due to Program

This shows that for a participant, the program increased the expected efficiency by .111. In other words, the percent efficiency of the site relative to baseline was .111 higher than in the absence of the program, independent of the strength of the current influence. For a non-participant that was very weakly influenced by SCE, there was no increased efficiency due to the program, but for a non-participant that was very strongly influenced by SCE, the program increased the expected efficiency by .078. The expected efficiency choice of a participant is always higher than the expected efficiency choice of a non-participant. However, for those non-participants that reported a very strong influence, the efficiency choice was almost as high as for the participants.

This suggests that the program has two impacts. First the program has a direct net impact on the participants. Second, the program appears to have an indirect or spillover impact on the non-participants.

The next step was to use the demand-efficiency regression model to estimate the net direct impact of the program. For each participant we calculated the net annual kW savings due to the program by multiplying the base annual demand of the site by the estimated increase in efficiency due to the program, calculated from the preceding equation. Then these results were expanded to the population of program participants.

The final step was to use the demand -efficiency regression model to estimate the direct impact and spillover impact of the program. In this analysis, we worked with the non-participants in the sample. For each non-participant, we calculated the net annual kW savings due to the program by multiplying the base annual demand of the site by the estimated increase in efficiency due to the program, calculated from the preceding non-participant equation. Then we used the Dodge database to expand the sample non-participants to the population of new construction. To ensure a conservative estimate of spillover, we made an adjustment to factor out any participant sites that may have been present in the Dodge database. To accomplish this, the sample spillover was projected to both the participant and new construction populations and the participant population estimate was subtracted from the new construction population estimate.

Table 45 shows the net savings estimate and the estimate of spillover, together with the relative precision of each estimate. The difference-of-differences net savings was 5,266 kW. The econometric approach yielded a direct net savings of 5,578 kW. The double Mills ratio was not a significant predictor of demand and was not in the demand model.

Table 45 also shows the calculation of total net savings, counting both participant savings and spillover to non-participants. The table shows the spillover savings in the full Dodge database. Finally the table shows the adjustment to factor out any participant sites included in the Dodge database. This analysis yielded a net savings of 10.818 MW.

To be conservative, the net estimate including spillover can be written down by 1.28 times the standard error of the estimate, yielding a conservative estimate of net savings of 9.008 MW. This strategy factors in the statistical allowance for error and is very conservative. The probability is high, about 0.9, that this procedure underestimates the true value of savings. No relative precision is associated with these results since the relative precision is already reflected in the estimate itself.

	Estimate	Net-to-Gross Ratio	Relative Precision
Net Savings of Participants	5.578 MW	55.1%	±11.1%
Adjusted Net Participant Savings	5.578 MW		±11.1%
Spillover in NP population	5.621 MW		±39.8%
Spillover in P population	0.381 MW		±31.7%
Spillover Estimate	5.24 MW		±42.8%
Total Savings	10.818 MW	106.8%	±21.5%
Conservative Net Savings	9.008 MW	88.9%	--

Table 45: Net Demand Savings and Spillover Estimates

Recommendations for Future Studies

The methodology used for this study has proven to be very successful. RLW Analytics and AEC were able to collect and analyze large amounts of detailed data quickly using this methodology. To be sure, this was not an inexpensive endeavor, but it has produced characteristic and energy use information that is also very valuable for studies of market transformation, building characteristics, and other market research.

The key improvements made here from the 1994 PG&E/SCE evaluation include:

- **The use of the same staff to survey buildings and build engineering models.** This approach allowed RLW Analytics and AEC to build much better models because the data was collected with a full understanding of the needs of the models. Also, because the person who developed the model was on-site, a much better "reality check" could be done using the judgement of the engineer.
- **Building the engineering model shortly after the site visit.** In the 1994 study, several months passed before the modeling staff could review the field data, greatly increasing the chance that errors could not be adequately corrected. In this study, the initial models were built within days or weeks of the site visit. This, combined with the point above, greatly improved the quality of the models because the building was much fresher in the mind of the modeler.
- **The use of scale variables in the econometric models.** In the 1994 study, a binary variable was used to indicate "partial participation" (a non-participant with spillover). This crude approach to a subtle issue contributed to the econometric model's inability to identify non-participant spillover. In this study, a series of scale variables were used to isolate spillover. This more sensitive approach was successful in measuring "partial participation."
- **A single, experienced construction professional was used to recruit and survey design professionals and building owners.** The use of someone who understood the industry was the primary reason that such a high participation rate was observed. This also helped with survey completion and data quality because the respondents felt as though the surveyor understood the subject matter and could speak on their level.
- **More active involvement by the study sponsors.** This study was truly a collaborative process between the SCE team and the RLW Analytics / AEC team. The active involvement of many talented people at SCE, the Heschong-Mahone Group, and the involvement of members of the CADMAC New Construction subcommittee greatly contributed to the smooth flow of the project and to the quality of the final results.

Most of the cost and effort in this study involved the data collection and engineering model building tasks. There are several steps that could be taken in those areas to improve the cost effectiveness of the study:

- **Improvements in the model building software.** Further work to integrate the data entry, model building, and calibration modules of the software would increase the throughput and reduce the human intervention needed to turn survey data into DOE models. Because this system was developed independently of this project, work on these issues is ongoing. Future studies using the RLW Analytics / AEC team would benefit from these improvements.

- **Electronic data entry**. Related to the above point, the use of handheld computers to record survey data would streamline data entry and move quality control checks to the survey site, where the errors could most easily and accurately be corrected.
- **"Codify" engineering judgement**. A major factor in the data collection cost was the use of experienced engineers to collect the data. To the extent that some of the engineering judgement could be captured in the software, lower cost staff could be used in the data collection. This is a fine line to walk, as reductions in surveyor experience and skill could contribute to degradation in the quality of data.
- **Capture decision-maker data as the program runs.** One of the biggest challenges in this type of study is to ask a decision-maker about events that occurred as long as two years prior. The data collected for the econometric analysis could be significantly improved by collecting this data at the time the project is done. This would require a standard survey to be developed by CADMAC and administered by the utility sponsoring the program.
- **Revision of the CADMAC protocols on sampling**. To the extent that CADMAC sponsored regulatory studies like this one continue after January 1, 1998, a revision of the sampling protocols would benefit future studies. The great wisdom of the CADMAC committee was evident in their approval of the waiver to allow this study's variance from the protocols. The results of the study show that this sampling approach is effective in capturing the required information at a significantly lower cost than would be required by a sample complying with the current protocol.

Table of Contents

CADMAC PROTOCOLS TABLE 6

Southern California Edison
Study ID # 543

	Energy		Demand	
	Participant Group	Comparison Group	Participant Group	Comparison Group
	(per sqft in kwh/sqft/year)		*(per sqft in w/sqft)*	
Energy Usage				
Base Usage	199,900,000	160,700,000	50,260	38,430
Base usage per square foot	24.92	14.55	6.27	3.48
Impact Year Usage	157,100,000	147,700,000	40,130	34,710
Impact Year Usage per sqft	19.58	13.38	5.00	3.14
Gross Load Impact	42,730,000	12,940,000	10,130	3,724
Gross Load Impact per sqft	5.33	1.17	1.26	0.34
Net Load Impact	31,273,000	na	9,008	na
Net Load Impact per sqft	3.90	na	1.12	na
% Load Impact	21.4%	8.1%	20.2%	9.7%
% Load Impact per sqft	21.4%	8.1%	20.2%	9.7%
Gross Realization Rate	116.0%	na	115.0%	na
Net Realization Rate	84.9%	na	102.3%	na
Net-to-Gross Ratios				
Load Impacts	73.1%	na	88.9%	na
Load Impact per sqft	73.1%	na	88.9%	na
Square Footage				
Pre-Installation	8,021,983	11,041,805	8,021,983	11,041,805
Post-Installation	8,021,983	11,041,805	8,021,983	11,041,805
90% Precision				
Base Usage	10.2%	14.1%	11.1%	10.7%
Base usage per sqft	10.2%	14.1%	11.1%	10.7%
Impact Year Usage	11.7%	14.9%	11.8%	11.3%
Impact Year Usage per sqft	11.7%	14.9%	11.8%	11.3%
Gross Load Impact	8.8%	30.6%	11.0%	22.3%
Gross Load Impact per sqft	8.8%	30.6%	11.0%	22.3%
Net Load Impact	30.5%	na	21.5%	na
Net Load Impact per sqft	30.5%	na	21.5%	na
80% Precision				
Base Usage	7.9%	11.0%	8.7%	8.3%
Base usage per sqft	7.9%	11.0%	8.7%	8.3%
Impact Year Usage	9.1%	11.6%	9.2%	8.8%
Impact Year Usage per sqft	9.1%	11.6%	9.2%	8.8%
Gross Load Impact	6.9%	23.8%	8.6%	17.4%
Gross Load Impact per sqft	6.9%	23.8%	8.6%	17.4%
Net Load Impact	23.8%	na	16.8%	na
Net Load Impact per sqft	23.8%	na	16.8%	na

Table 6, continued

Measure Counts	Measure counts are not applicable to the design of this program

Population by Building Type	
Building Type	Percent of Population
college	8%
food store	1%
hospital	2%
medical	2%
mfg	7%
miscellaneous	16%
office	22%
restaurant	10%
retail	25%
school	5%
warehouse	3%
Total	**100%**

CADMAC Protocols Table 6

Southern California Edison
Study ID # 543

	Energy		Demand	
	Participant Group	Comparison Group	Participant Group	Comparison Group
	(per bldg in kwh/bldg/year)		*(per bldg in kw/bldg)*	
Energy Usage				
Base Usage	199,900,000	160,700,000	50,260	38,430
Base usage per bldg	1,525,954.20	65,914.68	383.66	15.76
Impact Year Usage	157,100,000	147,700,000	40,130	34,710
Impact Year Usage per bldg	1,199,236.64	60,582.44	306.34	14.24
Gross Load Impact	42,730,000	12,940,000	10,130	3,724
Gross Load Impact per bldg	326,183.21	5,307.63	77.33	1.53
Net Load Impact	31,273,000	na	9,008	na
Net Load Impact per bldg	238,725.19	na	68.76	na
% Load Impact	21.4%	8.1%	20.2%	9.7%
% Load Impact per bldg	21.4%	8.1%	20.2%	9.7%
Gross Realization Rate	116.0%	na	115.0%	na
Net Realization Rate	84.9%	na	102.3%	na
Net-to-Gross Ratios				
Load Impacts	73.1%	na	89%	na
Load Impact per bldg	73.1%	na	89%	na
Buildings				
Pre-Installation	131	2,438	131	2,438
Post-Installation	131	2,438	131	2,438
90% Precision				
Base Usage	10.2%	14.1%	11.1%	10.7%
Base usage per bldg	10.2%	14.1%	11.1%	10.7%
Impact Year Usage	11.7%	14.9%	11.8%	11.3%
Impact Year Usage per bldg	11.7%	14.9%	11.8%	11.3%
Gross Load Impact	8.8%	30.6%	11.0%	22.3%
Gross Load Impact per bldg	8.8%	30.6%	11.0%	22.3%
Net Load Impact	30.5%	na	21.5%	na
Net Load Impact per bldg	30.5%	na	21.5%	na
80% Precision				
Base Usage	7.9%	11.0%	8.7%	8.3%
Base usage per bldg	7.9%	11.0%	8.7%	8.3%
Impact Year Usage	9.1%	11.6%	9.2%	8.8%
Impact Year Usage per bldg	9.1%	11.6%	9.2%	8.8%
Gross Load Impact	6.9%	23.8%	8.6%	17.4%
Gross Load Impact per bldg	6.9%	23.8%	8.6%	17.4%
Net Load Impact	23.8%	na	16.8%	na
Net Load Impact per bldg	23.8%	na	16.8%	na

CADMAC PROTOCOLS TABLE 7
Southern California Edison
Study ID 543

A. OVERVIEW INFORMATION

1. Study title and study ID number
Impact Evaluation of Southern California Edison 1996 Non-residential New Construction Programs. Study ID number 543.

2. Program and year
SCE 1996 Non-residential New Construction Programs.

3. End uses measures
The study was directed primarily to the total load of the affected space. Shell measures, lighting, and HVAC were also examined.

4. Methods and models used
This study used an integrated combination of model-based statistical sample design, onsite audits, site-specific DOE-2 engineering models calibrated to billing data, short-term end use metering, econometric analysis and statistical expansion. See report body for methodological discussion.

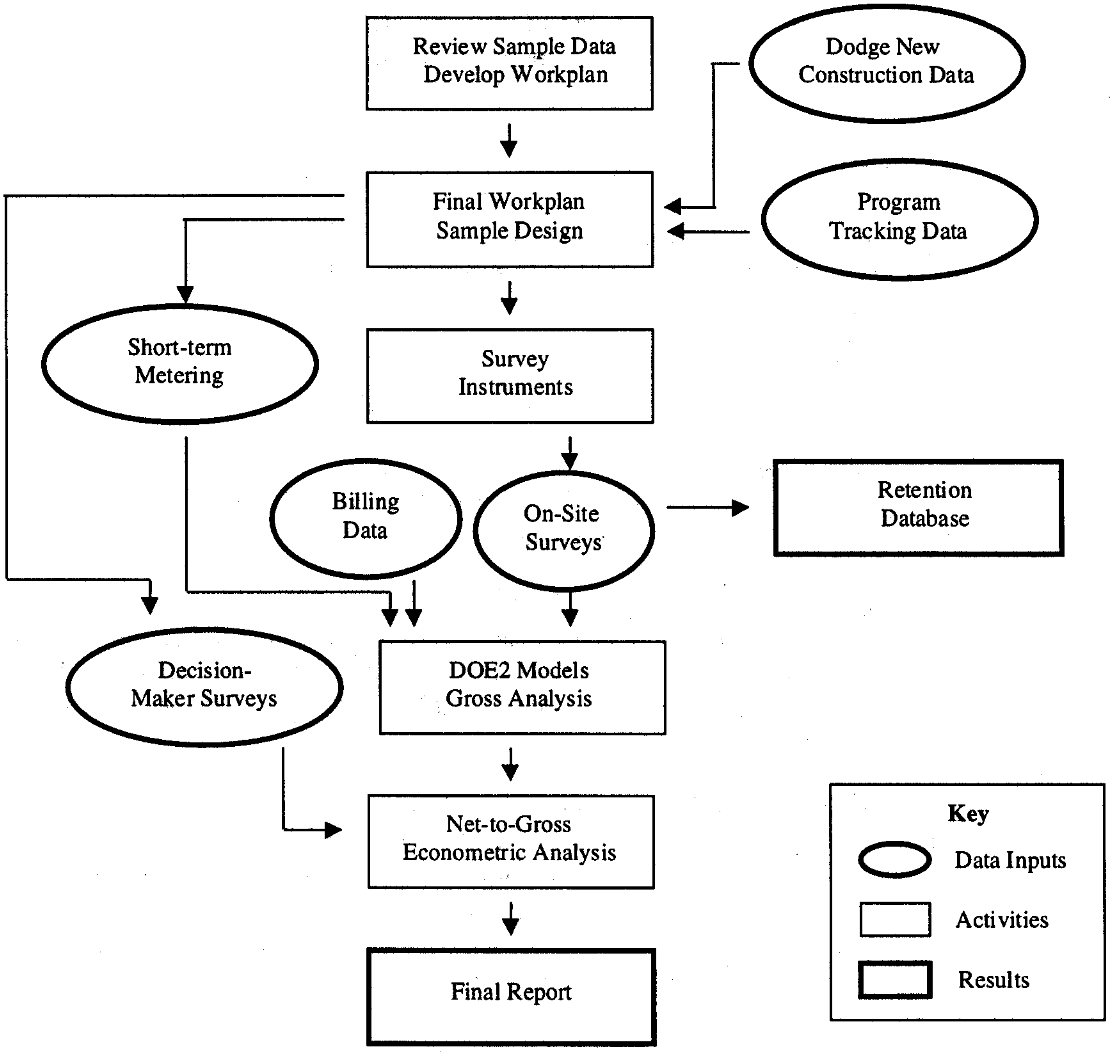
5. Participant and comparison group definitions
Participants were sites that received a rebate during the 1996 program year. Non-participants were completed new construction in 1996 that did not receive a rebate.

6. Analysis sample sizes
Gross analysis: 154 buildings. Net to gross analysis: 153 buildings.

B. DATABASE MANAGEMENT

1. Data elements
The following figure shows the relationship between the data elements and major tasks used in the study. The principle data elements were the Decision Maker Survey, the on-site audits, the short-term monitoring the billing data used to calibrate the DOE-2 models used in the gross analysis, the program tracking systems, and the Dodge data base describing new construction. Additional instruments were used for recruiting.

The short-term metering, onsite audits and billing data support the DOE-2 modeling which supports the gross analysis. The primary purpose of the Decision-Maker Survey was to support the net to gross analysis. The Dodge database and the program tracking systems were used in the sample design and in the statistical expansion of the sample findings. The relationship between these elements is described in the report.



2. Specific data sources

The Decision-Maker Survey collected data regarding:

- The degree of program participation
- The specific nature of influences on key design decisions
- Whether their design decisions would have been taken in the absence of the program.

The on-site survey was used to obtain an independent, realistic, observation of the ECM conditions and performance. The on-site survey instrument was designed to provide the information needed to simulate energy use and demand for each building by a minimum of five different scenarios. For maximum validity, the field data collection was aimed at directly observable data. Special attention was paid to Title 24 specifications and program measures throughout the building. The on-site visits also helped to assess the suitability of each site for potential short-term metering.

For details, see the report body.

3. Data attrition process

Please see D3 below.

See report section on net impact findings.

4. Data quality checks
Strict quality control measures were carried out throughout the data collection phase of the project. They consisted of a number of range, consistency, and sanity checks on the collected data, as well as random spot-checks on auditors in the field. These procedures are discussed in detail in the report section on engineering models and data collection.

5. Data collected but not used
None.

C. SAMPLING

1. Sampling procedures and protocols
The primary sampling frame was the Dodge database of new construction in California in 1995. This sample was screened for service area and building type and matched to the participant tracking database. The participant sample was stratified by the program estimate of savings, the non-participant sample was stratified by estimated square footage and building type. Model based statistical sampling (MBSS™) methods were used to construct the strata and choose the sample sizes. See the report section on sample design.

2. Survey information
See report text and answer D 3 below.

3. Statistical descriptions
Standard descriptive statistics are misleading for a stratified ratio estimation since weighting is necessary to obtain meaningful results and the methods described in the report are needed to evaluate statistical precision. The report provides statistical results for all key variables that are properly expanded to the population, together with suitable error bounds at the 90% level of confidence.

D. DATA SCREENING AND ANALYSIS

1. Outliers, missing data, and weather adjustment

The full sample was retained throughout the analysis. Studentized residuals were used to identify outliers. A site was considered to be an outlier if its studentized residual was greater than three in absolute value. A separate indicator variable was used to represent each such outlier in the model. The coefficient of this indicator variable indicated how much the dependent variable deviated from its expected value for the particular outlier. The statistical significance of these indicator variables were used to identify outliers that were statistically significant.

Sites that refused to participate in the study were replaced using a randomly drawn sample of backup sites. The level of refusal was rather low, as discussed earlier in this report.

Weather adjustment was handled in the engineering modeling. The model calibration used actual weather concurrent with the available billing data. Then all models were run using typical meteorological weather data. In this way the gross savings determined by the engineering models reflected normal weather conditions expected in each climate zone.

2. Control for background variables

The experimental design provided two types of control: (a) engineering models which provided 'same-building' comparisons, and (b) the net-to-gross analysis which compared the results of the engineering models for the participant and non-participant subsamples. The engineering models provided the first 'line of defense' against biased findings. The engineering models were used to compare the 'as-built' building to the 'baseline' building. Here the baseline referred to a building that just complied with Title 24 code. The engineering models were normalized for weather. The occupancy schedules were based on the onsite information describing the normal occupancy of the building on a daily and monthly basis.

This led to our estimates of weather-normalized gross savings. The net-to-gross analysis, in turn, compared the gross savings found from the engineering models for the participant and non-participant subsamples. The net to gross analysis used econometric techniques to estimate the naturally occurring level of efficiency that would have been built in the absence of the program. The econometric analysis included additional explanatory variables to control for self-selection bias and other differences between participants and non-participants.

All of these procedures were designed to get a reliable, unbiased estimate of the net impact of the programs. In particular, the experimental approach was designed to control for the effect of changes in economic or political activity. Increased operating hours would increase the gross savings for both the participants and non-participants but be controlled for in the net savings.

3. Screening procedures

The tables below summarize the screening procedures used to arrive at the final analysis datasets. In the case of the onsite audits, 154 buildings were recruited for the audit. Of these, 7 were dropped for a variety of reasons, e.g., the owner decided not to allow the audit, or the site did not meet the criteria (new commercial construction eligible for the program and occupied in 1996). This left 154 buildings in the final gross analysis. See the report section on the gross impact findings.

OnSite Audits		
Recruited	161	buildings
Audited	154	buildings
Final Data	154	buildings
Models	154	buildings
Used	154	buildings

Decision Maker Surveys		
Recruited	161	buildings
Completed	153	buildings
Used	153	buildings

The above table also shows the disposition of the Decision-Maker surveys. The objective of the Decision-Maker survey was to interview one or more key decision-makers for each building. Seven of the surveys was dropped from the analysis because of missing survey data or because the building was dropped from the gross analysis. This left 153 buildings in the final net-to-gross analysis. See the report section on the net impact findings.

4. Regression statistics

The following table shows the participation decision model.

Explanatory Variable	B	S.E.	Sig
restaurant	2.56	1.25	0.04
retail	1.55	0.71	0.03
warehouse	3.90	1.97	0.05
Owned by Fed Govt	-4.16	1.65	0.01
Ownership missing	-15.57	99.66	0.88
Built by Owner for tenent	-1.99	0.86	0.02
Other	2.34	1.40	0.10
No Preexisting Plans	1.30	0.70	0.06
payback in time period	4.11	1.18	0.00
Fin Criteria missing	-3.32	1.34	0.01
Commisioning Yes	1.50	0.84	0.07
Commisioning missing	4.00	1.59	0.01
Familiar with Rebates	0.53	0.15	0.00
Input on Design	0.31	0.17	0.07
Input missing	17.96	22.90	0.43
Energy Costs	-0.89	0.38	0.02
Energy Efficiency	1.17	0.39	0.00
Interaction with SCE current	0.71	0.24	0.00
Curr Inter missing	-9.52	22.77	0.68
Influence of SCE current	-0.37	0.22	0.10
Interaction with SCE past	0.43	0.15	0.01
Indicator for c2769	-20.62	99.66	0.84
Constant	-8.97	1.83	0.00

The following table shows the annual energy efficiency choice model.

Coefficients

		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.058	.037		-1.580	.117
	Program Participant	.132	.030	.397	4.354	.000
	Curr Influ_nonpart	.017	.008	.184	2.251	.026
	Restaurant	-.090	.041	-.134	-2.183	.031
	Retail	-.055	.023	-.145	-2.346	.020
	School	.074	.030	.152	2.495	.014
	Owned by Fed Govt	.143	.064	.136	2.226	.028
	Owned by State or Local Govt	.058	.034	.110	1.717	.088
	Other Ownership	.214	.085	.146	2.519	.013
	Built by Owner for tenent	.068	.031	.132	2.178	.031
	Lowest op cost	.077	.030	.156	2.550	.012
	Combi. Crit.	.069	.023	.188	3.007	.003
	Familiar with Rebates	.012	.006	.159	2.238	.027
	Input on Design	.012	.005	.182	2.508	.013
	Input missing	.171	.053	.241	3.211	.002
	indicator for c1253	-.857	.126	-.414	-6.820	.000
	indicator for c807	.565	.118	.273	4.786	.000
	Indicator for c497	.373	.119	.180	3.137	.002
	Indicator for c1276	.307	.119	.148	2.573	.011
	Mills ratio	.037	.017	.235	2.114	.036
	Double Mill	-.090	.025	-.337	-3.577	.000

The following table shows the summer peak demand efficiency choice model.

Coefficients

		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.054	.025		-2.159	.033
	Program Participant	.111	.022	.374	5.152	.000
	Curr Influ_nonpart	.013	.006	.158	2.164	.032
	exhibit	.093	.048	.112	1.958	.052
	other	.066	.034	.125	1.953	.053
	Restaurant	-.086	.035	-.143	-2.418	.017
	School	.072	.026	.166	2.772	.006
	Private Owned and rented	.051	.021	.154	2.372	.019
	Owned by Corp for Franchise or sub	.134	.058	.145	2.315	.022
	Owned by Fed Govt	.161	.055	.173	2.948	.004
	Owned by State or Local Govt	.080	.029	.169	2.748	.007
	Lowest op cost	.074	.026	.168	2.830	.005
	Combi. Crit.	.065	.020	.199	3.216	.002
	Commisioning missing	-.101	.050	-.143	-2.025	.045
	Familiar with Rebates	.016	.005	.229	3.484	.001
	Input missing	.108	.039	.171	2.752	.007
	indicator for c1253	-.628	.106	-.341	-5.912	.000
	indicator for c807	.588	.102	.320	5.740	.000
	Indicator for c497	.306	.102	.166	2.989	.003
	Indicator for c1276	.295	.104	.160	2.848	.005
	indicator for c2572	-.386	.103	-.210	-3.746	.000
	Indicator for c2382	.347	.103	.188	3.370	.001
	Indicator for c2688	.318	.115	.173	2.761	.007

See the report section on net impact findings for details.

5. Specification of Models

The "Engineering Models" section of the report describes the DOE-2 engineering models used to estimate the gross savings. The "Net Impact Findings" section of the report describes the econometric models that were used in the net to gross analysis.

Heterogeneity: The DOE-2 engineering models were designed to represent the heterogeneity of sites in the program. The models were designed to represent all building types, functional zones and equipment types encountered in the sample sites. The econometric models were designed to explain the variation in efficiency choice from one site to another.

Time series variation: In the gross analysis, time series variation was controlled by the simulation methodology. The gross savings were calculated by simulating the building with and without the energy efficiency measures but holding other equipment and schedules fixed as observed. Time-series variation was not an issue in the net-to-gross regression analysis since all observations reflected the same time period. In other words, the regression modeling addressed variation from one same site to another, but not from one time point to another.

Self selection: Self selection was addressed in the net-to-gross analysis by developing a logistics model for the probability of participating, and then using the resulting double inverse Mills ratios as added explanatory variables in the efficiency choice models. The statistical significance and effect of the inverse Mills ratios were estimated and reported.

Omitted factors: Two factors might be discussed: the use of Title 24 documentation and billing data. The study sought to use both Title 24 documentation and billing data to the extent practical. When either Title 24 documentation or billing data was available, it was used to improve the accuracy of the engineering models. This approach allowed us to maintain the full sample even when these data were unavailable.

The evaluation of the 1994 program clearly demonstrated the difficulty of obtaining Title 24 documentation, especially for the non-participants. In order to avoid high refusal rates and the concomitant risk of nonresponse bias, we only insisted on Title 24 documentation for sites that used the tailored lighting approach or the performance-based approach to Title 24 compliance.

Billing data was used to calibrate each individual engineering model whenever possible. However, as described elsewhere, the available billing data did not always reflect the space affected by the new construction. In some of these cases, we sought to supplement the billing data with our own metering. Nevertheless, some of the sites did not have actual usage data. In such cases we trusted that the engineering models were accurate without calibration. To confirm this assumption, we compared the gross savings determined before and after calibration for the sites with billing data or our metering. This analysis confirmed that the pre-calibration models were very accurate.

Net impacts The combination of statistical sampling, onsite surveys, site-specific engineering models, econometric analysis, and statistical expansion was carefully designed to provide an unbiased and statistically reliable estimate of net program savings. In particular, the decision-maker survey was designed to isolate self-selection bias and the long-run impact of the program on design practice. The model was specified to include any observable and statistically significant effects of the program on the energy efficiency of both participants and non-participants.

6. Errors in measuring variables

In the onsite surveys and engineering modeling we sought to obtain an accurate representation of each individual sample site. Past experience suggested that serious errors could arise from failing to model the space in the building actually affected by the new construction, or by failing to accurately describe some of the equipment and schedules of use. The present study addressed these problems by improved training and communication with the auditors, earlier retrieval and review of program files, having the auditors themselves responsible for the data entry and modeling, and having the auditors develop the model for a site soon after completing its survey. The engineering team met with SCE's program managers and reviewed the site-specific models in detail. We also redesigned the decision-maker survey, streamlined the process used to recruit each site and complete the decision maker survey, and assigned the responsibility for the whole process to a single, very competent person. All of these measures resulted in much more accurate data going into the econometric analysis than in the prior study.

7. Autocorrelation

Does Not Apply. All regression analysis was cross-sectional.

8. Heteroscedasticity

Heteroscedasticity – the tendency of larger projects to have greater variation – was addressed in both the sample design and efficiency-choice regression models.

The MBSS methodology used in the sample design addressed heteroscedasticity by modeling the variation in savings as a function of the tracking estimate of savings or the square footage of each site and then using an efficiently stratified sampling plan to increase the probability of selecting large sites. This ensures that the sample is effectively focused where the savings are greatest, while retaining an unbiased representation of small and large projects alike.

The efficiency-choice regression models were specified to minimize the danger of heteroscedastisity by defining the dependent variable as the gross savings as a fraction of the baseline energy use. This specification is closely related to the weighted-least-square methodology resulting from the assumption that the residual variation in gross savings is proportional to the baseline energy use of each site. Graphical scatter plots of the studentized residuals were examined to confirm the absence of Heteroscedasticity. In addition, a statistical test of homogeneity of variance was carried out to measure the statistical significance of differences in the variance of the residuals grouped by building type and by the level of efficiency predicted by the model..

9. Collinearity

Multicollinearity is generally a less serious problem in a cross sectional analysis than in a time series analysis. Our methodology was designed to protect against the type of problem that might arise in a cross sectional analysis. Extreme multicollinearity can cause computational problems. Several of the indicator variables used in the regression models were perfectly collinear. This occurred, for example, if a respondent who failed to answer a given question also failed to answer a second question. In this case the missing-response indicators would be perfectly collinear. The SPSS software used in the analysis identifies and reports these instances and automatically drops one of the variables from the analysis. The software also provides a warning if the multicollinearity is strong enough to affect the numerical accuracy of the estimated coefficients. In practice there was no indication of a serious problem with numerical accuracy.

When explanatory variables have strong but not extreme multicollinearity, it is important to guard against obtaining biased results. Omitted-variable bias can arise if one of the correlated variables is dropped from the model. We guarded against this possibility by systematically comparing the estimated coefficients of our various models and looking for other indicators such as large shifts in statistical significance.

10 Influential data points

We followed diagnostic procedures recommended by Belsley, Kuh and Welsh.[1] Our key indicator of an influential observation was the studentized residual, which can be related to the t-distribution. We also examined normal probability plots, partial-regression leverage plots for each explanatory variable, and other case-specific measures of influence. When an influential observation was identified, we included an indicator variable in the analysis that was 1 for the influential observation and 0 for all other cases in the sample. We retained this variable if it was statistically significant in the final model.

11 Missing data

See answer D.1. above.

12 Precision

In each regression model, we used standard logistics or least-squares techniques to calculate the standard error and statistical precision of each coefficient. We used the standard MBSS statistical techniques described in the Gross Savings chapter to expand to the econometric estimates for each sample site to the population and to measure the statistical precision of the results.

E. DATA INTERPRETATION AND APPLICATION

1 Method of net to gross analysis

The net impact was calculated as the participant gross impact less the naturally occurring impact predicted by the econometric model. The econometric model in turn was estimated by comparing the efficiency choice of the participants to the control group. Thus the approach was essentially equivalent to comparing the participants to the control group and adjusting for any uncontrolled differences between the two groups. We also estimated spillover impacts, which are discussed in the "Net Impact Findings" section of the report.

2. Process and rational used in net to gross analysis

The econometric analysis was designed to isolate the naturally occurring efficiency choice by comparing the efficiency choice found in the participant and non-participant samples, and adjusting the results for uncontrolled differences between the participants and non-participants, as well as for self selection.

[1] D. A. Belsley, E. Kuh and R. E. Welsch, *Regression Diagnostics*, Wiley, 1980.

CADMAC PROTOCOLS TABLE 11

LOAD IMPACT RESULTS FOR USE IN PLANNING AND FORECASTING

For Non-Residential New Construction Incentives Program
First Year Load Impact Evaluation – Whole Building Savings
Southern California Edison
SCE Study No. 543

1. **Base Energy Usage:** The primary purpose of both the engineering and statistical models was to produce estimates of energy savings in kWh and kW due to the non-residential new construction programs. Base energy usage was arbitrarily defined by the researchers for purposes of this study. Therefore, no estimates of base energy use are provided for forecasting.

2. **Determination of Net Program Impacts:** The applicability of net-to-gross estimates derived in this study to forecasts of future program impacts depends on several factors, including: 1) the differences in characteristics between the general population and the study sample; 2) the generalizability of the net-to-gross statistical models; 3) market changes that affect net-to-gross ratio determinants. Net-to-gross estimates were developed at the whole building level and reported in Tables 6 & 7. The estimates were produced using weights that were specific to the population of 1996 non-residential new construction. To the extent that any of the characteristics of the new construction population changes from year to year or the new construction population differs from the general building population, the results are not transferable. Changes in program design and general construction practice can influence the types of customers who participate and the types of technologies that are covered by the programs. The estimates were developed for a population with a given program structure and state of building practice. To the extent that either of these things change, the results are not transferable. Long-term market changes were beyond the scope of this study. Due to the probable changes in market conditions over time, specific net impact results developed for the 1996 Non-residential new Construction programs are not transferable for use in long-term forecasting.

3. **Load Impacts:** Gross kWh per ft2 per year is 5.33 kWh; gross kW per ft2 is 0.0012 kW. The study found an energy net-to-gross ratio of 0.731 and a demand net-to-gross ratio of 0.889. The load impacts were calculated from a mix of prescriptive and custom incentive packages. These savings estimates cannot be applied to other program forecasts where the mix of custom and prescriptive incentive packages is different from the 1996 Non-Residential New Construction sample or where the prescriptive requirements are different from the 1996 program.

SOUTHERN CALIFORNIA EDISON COMPANY RETROACTIVE WAIVER FOR 1996 NON-RESIDENTIAL NEW CONSTRUCTION PROGRAM (Study ID Number 543)

Date Approved: August 20, 1997

Summary of Edison Request

This waiver requests deviations from the Protocols by Southern California Edison (Edison) for its 1996 Nonresidential New Construction Impact Study. Edison seeks approval to:

1. Achieve requisite precision and confidence levels with a reduced sample size
2. Permit the use of short-term whole premise metering in addition to or instead of billing data for calibration of building simulation models (DOE-2) and eliminate the requirement for a minimum of 9 months of billing data
3. Use two different methodologies to estimate program net savings impacts and specify selection criteria to determine which of the two estimates will be used to calculate earnings for this program.

In the remainder of this waiver, items (1) to (3) above are referenced by their item number.

PROGRAM SUMMARY Nonresidential New Construction Program

Number of Participants (coupons)	130
Administrative Costs	$919K
Incentive Costs	$2,834K
Total Program Costs	$3,753K
Net Resource Benefits	$12,081K
Earnings	$1,297K

Proposed Waiver

Edison seeks CADMAC approval to: (see Table A for summary)

(1) Achieve requisite precision and confidence levels with a reduced sample size

Parameter

Table C-8, Item #1 Sample design, which refers to Table 5, Section C Sample Design for First Load Impact Year, which specifies minimum sample sizes for nonresidential impact evaluations. (Similar requirements for Participant and Comparison Groups)

Protocol Requirement

The Protocols specify that if there are less than 350 program participants, sample size will attempt a census. If there are more than 350 program participants, sample size for participants will be sufficiently large to achieve a minimum precision of plus/minus 10% at 90% confidence level, based on total annual energy use. In any case, samples must have at least 150.

SOUTHERN CALIFORNIA
EDISON
An EDISON INTERNATIONAL℠ Company

May 15, 1998

Dr. Donald K. Schultz
Office of Ratepayer Advocates
1227 "O" Street, 4th Floor
Sacramento, CA 95814-5840

Re: Notice of Erratum—Study 543, Table 6
(Impact Evaluation of the 1996 Nonresidential New Construction Program)

Dear Don:

We have discovered an error in the Table 6 reporting of the net realization rate for energy and demand savings for this program. The correction is described below and is marked and documented on a copy of Table 6 attached to this memo.

The calculation was incorrectly made using the verified gross energy and demand savings as the denominator, rather than the verified net energy and demand savings. Thus, the two percentages reported in Table 6 for the net realization rate for the participant group for Energy Usage (the twelfth row of data in the table) should be changed as follows:

Energy / Participant Group / Energy Usage / Net Realization Rate
was reported as 84.9% should be changed to 92.5%

Demand/ Participant Group/ Energy Usage / Net Realization Rate
was reported as 102.3% should be changed to 109.4%

You may wish to mark this change in your copy of the study or enclose this notification within your copy of the study.

Sincerely,

Marian V. Brown

Marian V. Brown
Manager
Measurement & Evaluation

Attachment

cc: All parties who received copies of Study 543 from Edison

300 N. Lone Hill Ave.
San Dimas, CA 91773

CADMAC PROTOCOLS TABLE 6

Southern California Edison
Study ID # 543

	Energy		Demand	
	Participant Group	Comparison Group	Participant Group	Comparison Group
	(per sqft in kwh/sqft/year)		*(per sqft in w/sqft)*	
Energy Usage				
Base Usage	199,900,000	160,700,000	50,260	38,430
Base usage per square foot	24.92	14.55	6.27	3.48
Impact Year Usage	157,100,000	147,700,000	40,130	34,710
Impact Year Usage per sqft	19.58	13.38	5.00	3.14
Gross Load Impact	42,730,000	12,940,000	10,130	3,724
Gross Load Impact per sqft	5.33	1.17	1.26	0.34
Net Load Impact	31,273,000	na	9,008	na
Net Load Impact per sqft	3.90	na	1.12	na
% Load Impact	21.4%	8.1%	20.2%	9.7%
% Load Impact per sqft	21.4%	8.1%	20.2%	9.7%
Gross Realization Rate	116.0%	na	115.0%	na
Errors → Net Realization Rate	92.5% ~~81.9%~~	na	109.4% ~~103.3%~~	na
Net-to-Gross Ratios				
Load Impacts	73.1%	na	88.9%	na
Load Impact per sqft	73.1%	na	88.9%	na
Square Footage				
Pre-Installation	8,021,983	11,041,805	8,021,983	11,041,805
Post-Installation	8,021,983	11,041,805	8,021,983	11,041,805
90% Precision				
Base Usage	10.2%	14.1%	11.1%	10.7%
Base usage per sqft	10.2%	14.1%	11.1%	10.7%
Impact Year Usage	11.7%	14.9%	11.8%	11.3%
Impact Year Usage per sqft	11.7%	14.9%	11.8%	11.3%
Gross Load Impact	8.8%	30.6%	11.0%	22.3%
Gross Load Impact per sqft	8.8%	30.6%	11.0%	22.3%
Net Load Impact	30.5%	na	21.5%	na
Net Load Impact per sqft	30.5%	na	21.5%	na
80% Precision				
Base Usage	7.9%	11.0%	8.7%	8.3%
Base usage per sqft	7.9%	11.0%	8.7%	8.3%
Impact Year Usage	9.1%	11.6%	9.2%	8.8%
Impact Year Usage per sqft	9.1%	11.6%	9.2%	8.8%
Gross Load Impact	6.9%	23.8%	8.6%	17.4%
Gross Load Impact per sqft	6.9%	23.8%	8.6%	17.4%
Net Load Impact	23.8%	na	16.8%	na
Net Load Impact per sqft	23.8%	na	16.8%	na

$$\frac{\text{Ex Post Net Energy Savings}}{\text{Verified Net Energy Savings}} = \frac{31{,}273{,}000}{33{,}826{,}490} = 92.5\%$$

$$\frac{\text{Ex Post Net Demand Savings}}{\text{Verified Net Demand Savings}} = \frac{9{,}008}{8{,}232} = 109.4\%$$

5/13/98