43 B

CUSTOMER ENERGY EFFICIENCY PROGRAM
MEASUREMENT AND EVALUATION PROGRAM

DOUBLE RATIO ANALYSIS
FINAL REPORT

Report Number CIA-93-X01B
September 1993

Measurement and Evaluation
Customer Energy Efficiency Policy & Evaluation Section
Pacific Gas and Electric Company
San Francisco, California

All inquiries should be directed to:

Susan M. Buller
Pacific Gas and Electric Company
CEE Policy & Evaluation, H28E
123 Mission Street
P.O. Box 770000
San Francisco, CA 94177

DOUBLE RATIO ANALYSIS FINAL REPORT

Prepared for

PACIFIC GAS AND ELECTRIC COMPANY
San Francisco, California

Prepared by

XENERGY Inc.
Oakland, California

RLW Analytics, Inc.

September 1993

LIST OF TABLES

LIST OF FIGURES

E

EXECUTIVE SUMMARY

This report presents the results of an evaluation of Pacific Gas and Electric's (PG&E) Commercial, Industrial and Agricultural (CIA) Energy Efficiency Rebate Program. Results from this study were derived by a technique, double ratio analysis, that combines results from short-term end-use metering, calibrated engineering modeling, and PG&E's savings estimates.

It has been recognized that alternative evaluation methods are needed to off set the high cost of metering. This study offers an approach that can be used in future evaluations.

The objectives of the double ratio analysis were to:

1. Evaluate the gross impacts of measures in the commercial building sector for both the Express and the Customized Incentive Programs.

2. Evaluate the effectiveness of the double ratio method for future analysis.

E.1. Approach

This study presents the results of an innovative technique designed to leverage costly metering sites with engineering modeling. Results from the Field Monitoring Study[1] and the HSEM Analysis Study[2] are combined with the PG&E tracking system estimates to derive gross impact savings estimates.

The double ratio estimation approach statistically combines savings estimates from three sources: 1) short-term metering; 2) calibrated engineering modeling; and 3) PG&E tracking system estimates. A summary of the approach, including sample sizes, is depicted in Figure E-1. Two analytical approaches using ratio estimates were

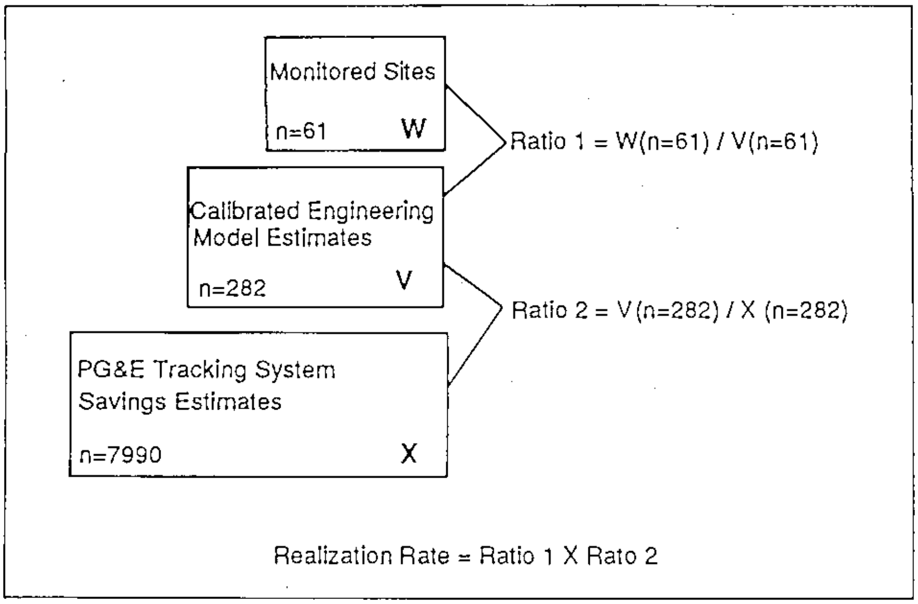
---

[1]PG&E's Commercial, Industrial, and Agricultural (CIA) Retrofit Evaluation Project - Final Short-Term Monitoring Results Report, The Fleming Group, September 1993.

[2]CIA Retrofit Evaluation Project - HSEM Analysis Study Final Report, XENERGY Inc., September 1993

compared: double ratio (i.e., metering, engineering, and tracking system estimates) and single ratio (i.e., metering and tracking system estimates).

Figure E-1
Double Ratio Estimation Process



E.2. Results

The gross impact savings estimates in this study are reported in terms of realization rates. A realization rate is the percentage of energy savings verified in the double ratio impact analysis compared to PG&E's tracking system. A realization rate of 0.90 indicates that 90% of PG&E's savings estimates were verified by measurement.

Table E-1 presents the estimated realization rates for energy savings with the associated 90% confidence intervals.

- Study results show that PG&E's lighting savings are accurate and the savings estimates for HVAC are too high.

- Results from the study show that approximately 82% of the savings can be verified through the double ratio analysis. The realization rate for lighting alone, which represents almost 55% of the program energy savings, is 0.93. The confidence interval indicates that there is a 90% probability that the true realization rate is between 0.67 and 1.19.

- Calibrated engineering models improve the precision of the savings estimates for the Express Program.

Table E-1
CIA Rebate Program - MWh Savings

	Realization Rate	90% Confidence Interval	Evaluation Gross Estimates (GWh)
Lighting			
Both Programs	.93	0.67 - 1.19	473,158
Express	1.07	0.65 - 1.49	278,836
Customized	.75	0.66 - 0.84	186,134
HVAC			
Both Programs	.55	0.20 - 0.90	109,463
Express	.38	0.29 - 0.47	17,292
Customized	.61	0.35 - 0.86	93,645
Refrigeration			
Both Programs[1]	.85	0.44 - 1.26	50,601
Total Lighting and HVAC			
Both Programs	.82	0.60 - 1.04	582,621
Express	.96	0.60 - 1.33	295,450
Customized	.69	0.51 - 0.86	279,779

[1] There were not enough sites to analyze by program type.

E.3. Key Findings

The double ratio approach was more effective for the Express Program than the Customized Program. This is due to the fact that savings estimates for the Customized programs are calculated using formulae customized for each site. Therefore, the engineering model does not add significant information to the calculations.

The Express Program is designed so that savings are the same for each specific measure. The inputs are not specific to the customer. Therefore the engineering model estimates offer a more customer-specific estimate and combined with the metering results, offers a more precise estimate.

The double ratio approach is cost effective for the Lighting Express Program but not the Lighting Customized program.

For the Lighting Express Program, a double ratio analysis with engineering modeling, costs about 60% less than a study using metering alone. However, to obtain the same level of precision for the Customized Program, it would cost almost six times as much as a study which uses metering alone. Other end uses show similar results.

1 INTRODUCTION

1.1. Overview

This report describes the methodology and findings of the "double ratio estimation" analysis. Results from this study represent gross energy and demand savings for 1992 participants. This study is one of several components of the impact evaluation of PG&E's Commercial, Industrial and Agricultural (CIA) Energy Efficiency Rebate program.

1.2. Background

1.2.1. PG&E CIA Rebate Program

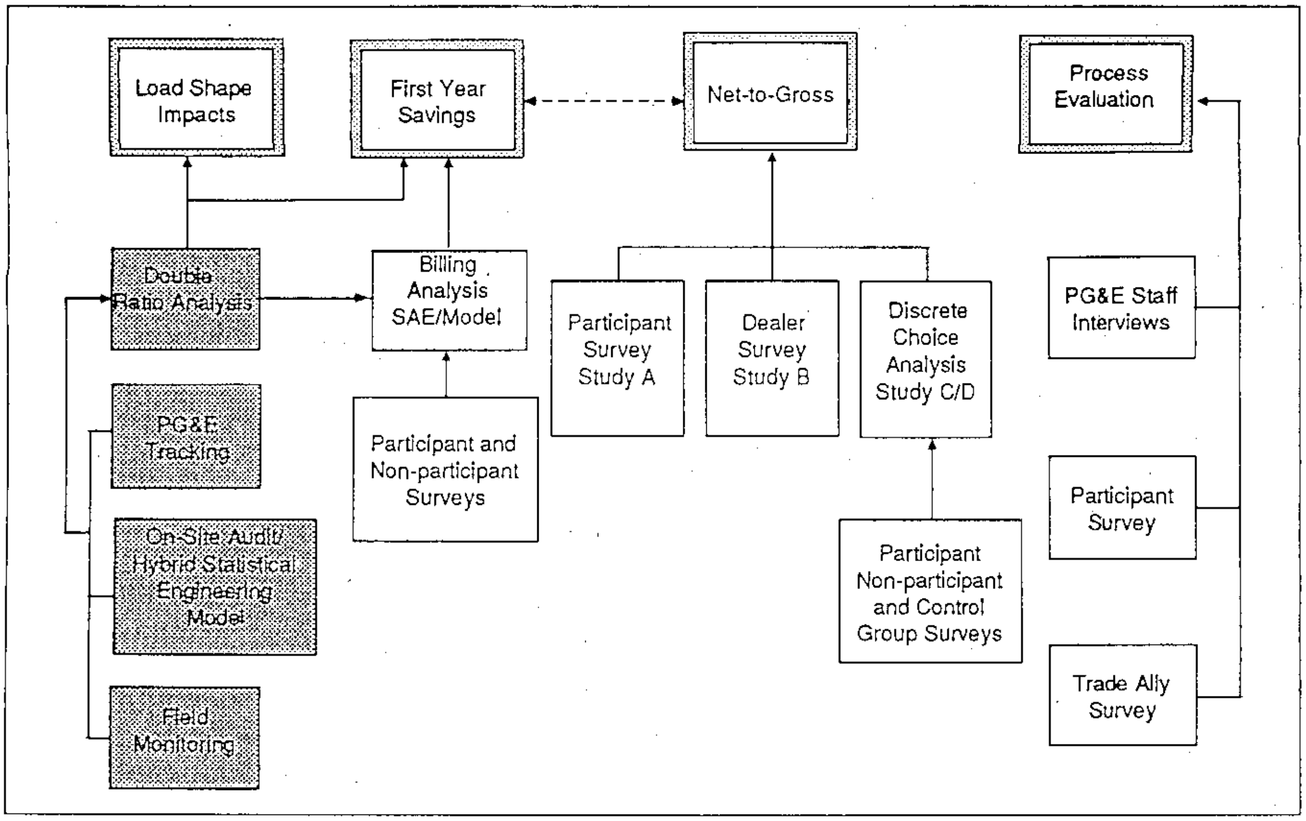
The CIA Rebate program, which began in 1990, provides cash incentives to commercial, industrial and agricultural customers who install a wide range of energy efficiency measures. The program covers end uses such as lighting, HVAC, agricultural measures, motors, refrigeration, and industrial processes. The CIA program offers two types of rebates. One is a customized rebate in which the customer receives a rebate that is directly related to the calculated energy savings for the retrofit. The second incentive is a direct rebate which offers the customer a set dollar amount for specified equipment (now called the "Express" program).

1.2.2. CIA Evaluation Project

The results from the "double ratio" analysis which are presented in this report are a part of a larger, comprehensive program evaluation of the CIA Rebate Program. Figure 1-1 presents a view of the evaluation project.

Results from othe other components of this evaluation, including field measurement results, net-to-gross ratio estimation, billing data analysis, engineering model findings, and process evaluation are found in separate reports. A final, summary report encompassing all components of the evaluation will be issued at the completion of this project. As Figure 1-1 shows, the double ratio analysis is a key component in determining first-year gross savings and load shape impacts.

Figure 1-1
CIA Retrofit Evaluation Project



1.3. Methodology Background

Double ratio estimation, the subject of this report, is an innovative methodology used to derive estimates of gross energy savings for PG&E's CIA program. The research approach statistically combines estimates of participant energy savings from three sources: 1) PG&E program tracking system; 2) calibrated engineering modeling; and 3) short-term monitoring.

The need for this approach has been voiced by DSM professionals for several years. There is general agreement that hourly end-use and appliance metering provide critical information on how DSM measures affect energy consumption. There is also general agreement that engineering analysis and models of energy savings, given the appropriate software and data elements, can provide: 1) estimates of DSM measure savings; and 2) highly useful information on the underlying parameters and assumptions used to implement DSM programs. Lastly, there is general agreement that although program tracking system estimates of participant savings are essential to operating and monitoring DSM programs, their estimates are not, in and of themselves, sufficiently reliable for broad acceptance by policy makers.

The difficulty with employing metering as an exclusive technique for estimating program impacts is that it is relatively costly and time-consuming. In addition to the uncertainty in selecting and recruiting representative sites, metering must be timed in a way that allows for all the major determinants of DSM measure savings to be measured in both the pre- and post installation periods. Further, metering must be in place for a strategic length of time representative of the conditions under which the DSM measures will operate over their lifetimes. Meter equipment failures, data corruption and a host of other factors that can jeopardize data quality, sample size and sample validity, must also be factored into a metering plan.

For these reasons, metering studies tend to be limited to unique situations where it is cost-effective to collect high resolution data on a small number of customers and a narrow range of technologies. For large scale DSM programs such as CIA, with thousands of participants and many qualifying technologies, metering may be too limited and impractical to be used by itself to provide empirical evidence of program achievements.

Engineering models come in many forms. Their common feature is that they combine participant-specific building equipment and end-use information with engineering algorithms to estimate DSM measure energy savings. Engineering model estimates of energy savings tend to be more accurate than program tracking system estimates. This is because simulations calculate end-use consumption in a more complex, and realistic, manner than the streamlined, non-interactive formulas that are used on program applications, especially for programs such as the Express Program. Also, they are usually more data intensive than program tracking system methods of calculating energy savings, and typically require one or more surveys or inspections of participant sites. However, the costs of engineering modeling are significantly lower than metering costs, and thus engineering sample sizes can be substantially larger than metering study samples.

The research design that was employed for this project attempts to simultaneously exploit the strengths, and avoid the weaknesses, of each of these measurement techniques. It was designed to make maximum use of a small sample of costly, high resolution monitoring data by using these data to validate less costly, less precise simulation estimates of savings for a much larger sample of participants. The validated simulation estimates are then used to adjust the low cost, least accurate tracking system estimates of savings available for the entire population of 1992 program participants.

1.4. Methodology Approach

This study reports the results of integrating these three types of savings estimates. The interrelated study sample consists of estimated savings for 61 program participant sites that have been monitored in their pre- and post installation periods, and 282 program participant sites (including the 61 monitored sites) that have received on-site inspection and had their energy use analyzed using a calibrated engineering model. Table 1-1 shows the breakdown of sites by end use and program. For these sites, as well as for the entire 1992 CIA program participant population, program tracking system estimates of energy savings are available.

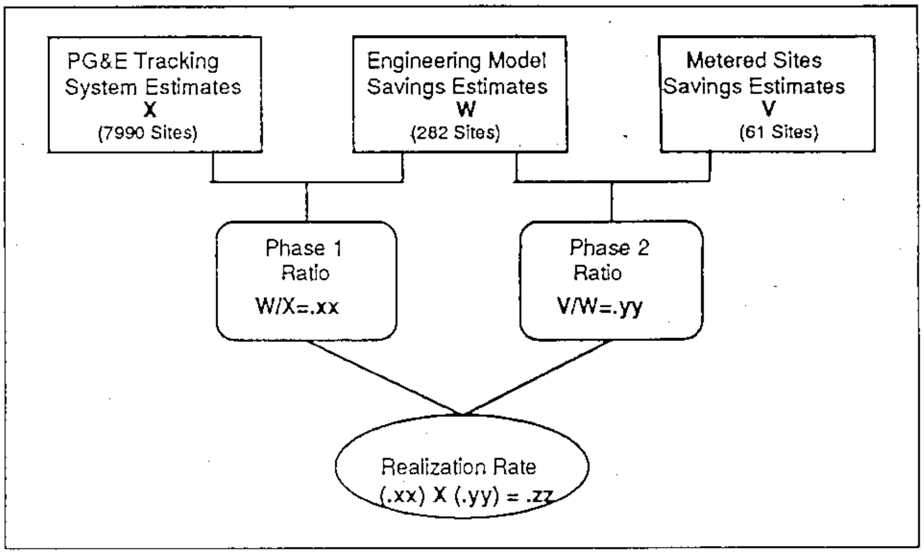
Table 1-1
Double Ratio Estimation Sample Sizes

End Use	Program	Engineering Modeling Sites	Monitored Sites
Lighting	Express	96	16
	Customized	133	36
HVAC	Express	25	0
	Customized	18	9
Refrigeration	Express	4	0
	Customized	6	0
Total		*282*	*61*

Results from this study are reported in terms of realization rates. A realization rate is the ratio of the measured savings to the assumed savings in the tracking system. The double ratio analysis consists of calculating two ratios, 1) the ratio of metering to engineering model savings; and 2) the ratio of engineering model savings to PG&E tracking system estimates. The two ratios are then combined to determine realization rates.

The process to combine the data from various sources is graphically shown in Figure 1-2.

Figure 1-2
Data Integration Process



A two-phase sampling plan was devised for selecting short-term monitoring and engineering model sites from among the population of program participants. The plan specifies an optimal allocation of sampling units, given all the information that was available at the time the sampling plan was designed. The sampling units were allocated among program type and end uses by energy savings strata so that the overall sampling error of the *combined* field monitoring and engineering model estimates of program energy savings was as low as possible.

As shown in the diagram, the first step of the double ratio analysis is the calculation of the ratio of energy savings estimates from engineering model sites to the PG&E tracking system estimates for the same sites.

$$.xx = \frac{Engineering\ Model\ Estimates\ (W)}{Tracking\ System\ Estimates\ (X)}$$

The second phase involves the calculation of the ratio of metering results to the results from the engineering model for each of the metered sites.

$$.yy = \frac{Metering\ Savings\ Estimates\ (V)}{Engineering\ Model\ Estimates\ (W)}$$

In the final statistical analysis, the two ratios which include all three sources of savings estimates are integrated.

$$\text{Realization Rate}(.zz) = (.xx \times .yy)$$

but this is equivalent to metering savings

Tracking.

then why Bother.

It does not make much sense

The integrated results are then expanded to the full 1992 program population. The result of this analysis is a population-weighted realization rate that is interpretable as the percentage of the program tracking system estimate of savings confirmed by empirical measurement.

how

As well as describing the statistical techniques that are used to sample the participants and analyze the energy savings data, this report documents the assumptions used as the basis for the initial sample plan. In this report, these assumptions are examined in light of the findings, and the cost-effectiveness of the research design is analyzed.

The remainder of this report details the findings, assumptions and methodology of the double ratio analysis. Section 2 presents the research findings from this analysis and Section 3 explores the cost effectiveness of the research design and offers recommendations for future studies. Section 4 presents the methodology and data used for the double ratio analysis. Appendix A provides a more detailed description of the methodology and Appendix B contains the final post-stratification and weights as described in Appendix A. Appendix C offers a theoretical discussion of the double ratio analysis. The data used in the analysis is in Appendix D.

2 GROSS IMPACTS FINDINGS

2.1. Introduction

The principal results of this study are presented in the form of realization rates, the ratio between empirically-derived, population-weighted program impacts and the corresponding impacts reported in the CIA program tracking system. The error bound is a 90% confidence upper or lower range from the realization rate. Also accompanying the realization rate is an indicator of the statistical precision of the estimate that is referred to as "relative precision". It quantifies sampling error, or the degree to which an unbiased estimate from a randomly-drawn sample may diverge from a population parameter due to chance in the sample selection process.

Relative precision is always expressed as a percentage of the mean estimate for which it is calculated, and always implies that sampling uncertainty can result in a sample estimate that is either higher, or lower, than the true population parameter. For this study, relative precision for each estimate is calculated at the 90 percent level of confidence. This means that there is a 90 percent chance that the true population parameter is within the upper and lower bounds of the sample estimate. The upper and lower error bounds are derived by multiplying the relative precision by the estimate, and subtracting, or adding, this product from the estimate. For example, if the realization rate is 0.90 and the relative precision is 15%, there is a 90 percent chance that the true population parameter is greater than 0.765 and less than 1.035. The error bound in this example is ±.135.

2.2. Findings Summary

A summary of the gross energy realization rates for each end use and program is found in Table 2-1. Demand realization rates are in Table 2-2. These tables present the results of the double ratio analysis which integrates savings estimates from short-term metering, calibrated engineering models, and the PG&E tracking system. Detailed findings for each end use follow in this section.

Table 2-1
Summary Statistics Double Ratio Analysis - Energy Savings

End Use/Program	MWh Savings Realization Rate	90% Confidence Interval	Sample Sizes: Metering	Sample Sizes: Engineering Model
Lighting				
Express	1.07	0.65 - 1.48	16	96
Customized	0.75	0.66 - 0.84	36	133
Both Programs	0.93	0.67 - 1.19	52	229
HVAC				
Express[1]	0.38	0.29 - 0.47	0	25
Customized	0.61	0.35 - 0.86	9	18
Both Programs	0.55	0.20 - 0.90	9	43
Refrigeration				
Both Programs[1,2]	0.85	0.44 - 1.26	0	10
Lighting and HVAC				
Express	0.96	0.36 - 1.33	16	121
Customized	0.68	0.57 - 0.85	45	151
Both Programs	0.81	0.59 - 1.03	61	272

Table 2-2
Summary Statistics Double Ratio Analysis - Demand Savings

End Use/Program	kW Savings Realization Rate	90% Confidence Interval	Sample Sizes: Metering	Sample Sizes: Engineering Model
Lighting				
Express	0.83	0.58 - 1.09	16	96
Customized	0.71	0.60 - 0.82	36	133
Both Programs	0.80	0.57 - 1.03	52	229
HVAC				
Express[1]	0.27	0.19 - 0.36	0	25
Customized	0.56	0.00 - 1.37	9	18
Both Programs	0.32	0.00 - 0.80	9	43
Refrigeration				
Both Programs[1,2]	2.39	1.67 - 3.11	0	10
Lighting and HVAC				
Express	0.69	0.41 - 0.97	16	121
Customized	0.69	0.46 - 0.91	45	151
Both Programs	0.69	0.48 - 0.90	61	272

[1]Single Ratio - No Metering

[2]There were not enough sites to analyze by program type.

2.3. Lighting Measure Program Impacts

Based on the program tracking system, lighting measures constitute 51 percent of the energy savings for the CIA program. Results of the double ratio analysis for lighting measures are presented in Table 2-3. The realization rates indicate that for the Express program, actual MWh savings is 107 percent of that reported in the tracking system. In the Customized program, actual MWh savings is 75 percent of what is reported in the tracking system. For the Express and Customized programs, estimated kW reductions are 83 percent and 71 percent of tracking system estimates, respectively.

Table 2-3
Realization Rates for Lighting Measures

		Express	Customized	Total
Energy (MWh)	Realization Rate	1.07	0.75	0.93
	Error Bound	±.42	±.09	±.26
	Relative Precision	39%	12%	28%
Demand (kW)	Realization Rate	0.83	0.71	0.80
	Error Bound	±.26	±.11	±.23
	Relative Precision	31%	15%	29%

The realization rates for the Express and Customized programs were weighted and combined into a realization rate that represents the empirically-confirmed savings for the CIA program as a whole. Combining the results of both programs, the MWh realization rate for lighting measures for the CIA program as a whole is 0.93 ; for kW reductions the realization rate is 0.80. The relative precision of these estimates at the 90 percent confidence level is 28% and 29%, respectively.

Table 2-4 contains the mean estimates of program energy and demand savings for lighting measures that were installed under the Express and Customized programs. These estimates have all been weighted to their respective 1992 program populations and energy savings strata using the finalized double sampling plan. The relative precision of the ratios of the sample estimates are provided within the table.

Table 2-4
Lighting Estimates - Double Ratio

	1st Phase Modeling to Tracking		2nd Phase Monitoring to Modeling	
	Express	Customized	Express	Customized
Sample Size	96	133	16	36
Short-term Monitoring				
Mean MWh Savings	--	--	53.3	32.2
Mean kW Savings	--	--	10.5	5.2
Engineering Modeling				
Mean MWh Savings	29.1	31.2	42.2	26.2
Mean kW Savings	5.3	5.7	8.1	4.8
Tracking System				
Mean MWh Savings	34.4	51.2	--	--
Mean kW Savings	8.2	8.7	--	--
Results				
Ratio (MWh)	0.85	0.61	1.26	1.23
Relative Precision	18%	8%	22%	21%
Ratio (kW) Savings	0.64	0.66	1.29	1.08
Relative Precision	20%	10%	18%	17%

The findings indicate that engineering model estimates for the Express program are 85 percent of program tracking estimates of energy savings, and 64 percent of program tracking estimates of demand savings. The relative precision of these estimates are 18% and 20%, respectively. The findings also indicate that engineering model estimates of energy and demand savings are systematically smaller than the short-term monitoring estimates. The relative precision of the ratios of engineering modeling to short-term monitoring for MWh and kW are similar to those of the ratios of engineering modeling to program tracking savings.

For the Customized program, the first and second phase ratios indicate that engineering model estimates are lower than program tracking estimates of savings and also lower than short-term monitoring estimates. The relative precision for the MWh and kW estimates for the first-phase sample is somewhat lower than for the second-phase sample. This indicates the Customized program has greater variance in the relationship between engineering model and short-term monitoring estimates than between engineering modeling and the program tracking systems estimates.

Integration of the three estimates of program-related savings is done by multiplying together the first and second phase ratios.[3] The results of these calculations are realization rates which are displayed in Table 2-4.

A comparison of short-term monitoring with program tracking estimates of savings can also be performed. This analysis is displayed in Table 2-5.

Table 2-5
Lighting Estimates - Single Ratio

	Express	Customized
Sample Size	16	36
Short-term Monitoring		
Mean MWh Savings	53.3	32.2
Mean kW Savings	10.5	5.2
Tracking System		
Mean MWh Savings	40.6	49.2
Mean kW Savings	9.6	7.8
Results		
Ratio (MWh)	1.31	0.66
Relative Precision	65%	11%
Ratio (kW)	1.10	0.67
Relative Precision	58%	21%

The findings reveal that the use of double sampling substantially increases the relative precision of the realization rate for the Express program, over the use of the single, short-term monitoring sample.

2.4. HVAC Measure Program Impacts

Based on the program tracking system, HVAC measures account for 22 percent of total CIA program energy savings. Table 2-6 displays the results of integrating the first and second-phase sample findings for the Customized program. As the individual results indicate, there appear to be moderate discrepancies between the actual and program tracking system estimates of MWh savings, and very large discrepancies with respect to kW savings.

[3] See Appendices A and C for methodology details.

Table 2-6
Realization Rates for HVAC

		Customized Only	Customized and Express
Energy (MWh)	Realization Rate	0.61	0.55
	Error Bound	±.25	±.35
	Relative Precision	42%	64%
Demand (kW)	Realization Rate	0.56	0.32
	Error Bound	±.81	±.48
	Relative Precision	146%	151%

Table 2-7 contains the mean estimates of program energy and demand savings for HVAC measures that were installed under the Express programs. As short-term monitoring was not practical for these sites, the ratio estimation results only represent the relationship between engineering model savings and program tracking system savings.

Table 2-7
HVAC (Express) Estimates - Single Ratio

	1st Phase Only Modeling to Tracking
Sample Size	25
Engineering Model	
Mean MWh Savings	3.9
Mean kW Savings	1.7
Tracking System	
Mean MWh Savings	10.4
Mean kW Savings	6.1
Results	
Ratio (MWh)	0.38
Relative Precision	24%
Ratio (kW)	0.27
Relative Precision	31%

The findings indicate that the engineering model can confirm only 38 percent of the MWh savings, and only 27 percent of the kW reductions, for Express program HVAC measures found in the program tracking system. The relative precision of these estimates is 24% and 31%, respectively.

For the Customized program, both first-phase and second-phase samples were available. The findings of the double ratio estimates are displayed in Table 2-8.

Table 2-8
HVAC (Customized) Estimates - Double Ratio

	1st Phase Sample Modeling to Tracking	2nd Phase Sample Monitoring to Modeling
Sample Size	18	9
Short-term Monitoring		
Mean MWh Savings	--	190
Mean kW Savings	--	9.5
Engineering Model		
Mean MWh Savings	206	247
Mean kW Savings	12.7	12.8
Tracking System		
Mean MWh Savings	261	--
Mean kW Savings	456	--
Results		
Ratio (MWh)	0.79	0.77
Relative Precision	35%	33%
Ratio (kW)	0.75	0.74
Relative Precision	115%	64%

This analysis reveals that the program tracking system moderately overstates Customized program HVAC measure MWh savings and kW savings.

As was done for the lighting measures, to test the effectiveness of the double sampling design, a single sample estimate was created using only the validation (second-phase) sample to derive realization rates. This analysis is displayed in Table 2-9.

Table 2-9
HVAC Estimates (Customized) - Single Sample

	Customized
Sample Size	9
Short-term Monitoring	
Mean MWh Savings	190
Mean kW Savings	9.5
Tracking System	
Mean MWh Savings	248
Mean kW Savings	21.4
Results	
Ratio (MWh)	0.77
Relative Precision	22%
Ratio (kW)	0.44
Relative Precision	174%

The findings reveal that the single sample provides a realization rate of 0.77 for energy savings and 0.44 for demand reductions.

As the prior HVAC analyses make clear, program tracking estimates of kW reductions are not in agreement with empirical measurements. Moreover, it appears that the addition of engineering modeled sites to this portion of the study does not contribute positively to the precision of either the MWh or the kW savings results.

2.5. Refrigeration Measure Program Impacts

Based on the program tracking system, refrigeration measures account for almost 4 percent of total CIA program energy savings. Table 2-10 shows the realization rates for refrigeration measures for both the customized and Express programs.

Table 2-10
Realization Rates for Refrigeration

	Both Programs
Realization Rate (MWh)	0.85
Error Bound	±.41
Relative Precision	48%
Realization Rate (kW)	2.39
Error Bound	±.72
Relative Precision	30%

Table 2-11 contains the mean estimates of program energy and demand savings for refrigeration measures that were installed under the Express and Customized programs. As short-term monitoring was not done for these sites, the ratio estimation results only represent the relationship between engineering model savings and program tracking system savings. As before, these estimates have been weighted to their respective 1992 program populations and energy savings strata using the finalized sampling plan.

Table 2-11
Refrigeration Estimates - Single Sample

Sample Size	10
Engineering Model	
Mean MWh Savings	18.9
Mean kW Savings	2.2
Tracking System	
Mean MWh Savings	22.3
Mean kW Savings	0.9

The findings for the refrigeration measures indicate that 85 percent of the reported tracking system estimates of savings are being realized. However, kW reductions appear to be significantly understated in the program tracking system. The relative precision for MWh savings is 48%, and for kW savings it is 30% at the 90 percent confidence level.

2.6. Combined Lighting and HVAC Impacts

To complete the analysis of realization rates, lighting and HVAC measure estimates were combined and weighted to the population of measures based on the finalized sampling plan. Based on program tracking system estimates alone, these two measures make up 73 percent of total CIA program energy savings. The results are displayed in Table 2-12.

Table 2-12
Realization Rates for Lighting and HVAC Measures

		Express	Customized	Total
Energy (MWh)	Realization Rate	096	0.69	0.82
	Error Bound	±.37	±.18	±.22
	Relative Precision	38%	26%	27%
Demand (kW)	Realization Rate	0.69	0.69	0.69
	Error Bound	±.28	±.22	±.21
	Relative Precision	40%	32%	31%

The findings reveal that 96 percent of the lighting and HVAC energy savings for the Express program that are reported on the program tracking system can be confirmed through empirical measurement. Sixty-nine percent of the Customized program savings can be confirmed. The relative precision of these estimates are 38% and 26%, respectively. In total, 82 percent of the program tracking system estimates of energy savings for the CIA program can be confirmed; the relative precision of this estimate is 27%.

Regarding demand reductions, about 69 percent of the lighting and HVAC demand reduction reported on the program tracking system for both the Express and Customized Programs can be confirmed. The relative precisions are of 40% and 32% respectively, at the 90 percent confidence level. Combining the two programs also leads to a realization rate for demand reduction of 0.69; the relative precision of this estimate is 31%at the 90 percent confidence level.

3 OTHER KEY FINDINGS

3.1. Introduction

This section presents a discussion of the analysis technique with recommendations for future studies using the double ratio analysis. The error ratios achieved in this study, which have implications for future studies, are presented. A discussion of sampling and its associated cost-effectiveness follows. And finally, recommendations for future studies are discussed.

3.2. Comparison of Ratios

Various analytical options are available for estimating gross impact savings. This study combined two ratios that derived data from 1) comparing metering results to engineering model results and 2) comparing engineering results to tracking system results. The two ratios are then combined into one realization rate. However, because this is an innovative technique, we need to understand the value of the information gained by combining these three sources of savings estimates. This section looks at the each end-use analysis by comparing the results of just using a single ratio to estimate savings.

Ratio 1 is the ratio of short-term end-use metering to tracking system estimates and Ratio 2 is the double ratio estimate developed for this study. In this study, there were sufficient sample sizes to compute single ratios for the lighting Express and Customized programs, and also the HVAC Customized program.

Table 3-1
Comparison of Double and Single Ratios:
Realization Rates and (Relative Precision)

		Ratio 1 Metering to Tracking System	Ratio 2 Double Ratio
Lighting Express			
	MWh	1.31	1.07
		65%	39%
	kW	1.10	0.83
		58%	31%
Lighting Customized			
	MWh	0.66	0.75
		11%	12%
	kW	0.67	0.71
		21%	15%
HVAC Customized			
	MWh	0.77	0.61
		22%	42%
	kW	0.44	0.56
		174%	146%

As seen in the table, the addition of the estimates from the engineering model in the double ratio estimates improves the savings estimates in the Express program. However, it does not improve the estimates for the Customized programs. This is best explained by the fact that savings estimates for the Customized programs are calculated by formula customized for each customer. Therefore, the engineering model does not add significant information to the calculations. However, the Express program is designed so that savings use the same end-use formula for each customer. The inputs are not specific to the customer. Therefore, the engineering model estimates offer a more customer-specific estimate and combined with the metering results, offers a more precise estimate.

As noted in the previous section, the discrepancy on HVAC peak demand savings are due to the fact that PG&E does not claim peak demand benefits from control measures such as energy management systems.

3.3. Recommendations for Future Studies

The various results of this data integration study, the realization rates of the Express and Customized programs, the error ratios, or the cost-effectiveness analysis, indicate the technique for data integration that was used for this project. It consisted of double sampling and double ratio estimation which can be highly effective for minimizing impact evaluation cost and maximizing the relative precision of the statistical estimates of savings. However, the results also indicate that such an approach must be well-planned and properly adapted to the program in question.

The results suggest that double sampling was effective for the Express program, but not for the Customized program. There are several plausible explanations for this finding. For example, it is possible that the Customized program measures, or the end uses in the buildings themselves, are too complex to be handled well by a calibrated engineering model. Alternatively, it is possible that the site-specific data that is collected by the engineering model survey is simply not sufficient to provide the data that is needed for accurately estimating impacts for these sites. In addition, site specific information is used to determine PG&E savings and calculate rebates for the Customized Program. Therefore the site specific information is already included in the tracking system savings estimate and makes the first phase analysis redundant, not providing much added value.

Another explanation for finding that the double ratio approach works well in one program context but not in another, may have little to do with the inherent limitations in any given measurement technique. Rather, it may have to do with *coordination* of the methods used among the three estimation techniques used to arrive at energy and demand savings.

Many examples of divergent assumptions and savings definitions emerged in the course of this study. For example, the program tracking system may define savings for a given technology to be the difference in energy use between a measure that meets the most recent government code or industry standard. This may be different from the new measure sanctioned by the utility program. However, empirical techniques can, by definition, only measure the change in energy use from the pre-existing base technology to the new technology. Thus, to match the program tracking system definition of savings, the empirical techniques must extrapolate the results. Clearly, there is a great potential for adding additional, unexplained variance to the savings estimates when confronted with such a situation. In this study, saving estimates were calculated from the same "pre-condition" as defined in the tracking system. Nonetheless, there was still an extrapolation of the actual pre-condition. Unless the definitions and protocols for estimating savings are coordinated among the techniques, the savings estimates from each method of measurement can vary widely.

Based on the experience of this study of PG&E's CIA program, it appears that to ensure the success of an impact evaluation study that employs double sampling and double ratio estimation, several steps should be followed. These are:

a) Assess the quality of the program tracking system estimates of savings. If the estimates are based on customer-specific data and detailed algorithms that are flexible enough to accurately capture customer-to-customer variations in savings, it is likely that there will be little benefit to adding engineering models to the research design. However, if the tracking system estimates of savings are based on non-customer specific default values and general-purpose algorithms, it is likely that adding a larger sample of less-costly engineering models to the research design will provide increased, cost-effective precision to the metering findings;

b) Determine whether the utility will benefit by undertaking some mode of field monitoring or load metering to estimate DSM impacts for a given program;

c) Calculate the cost of implementing a monitoring/metering study -- in particular, calculate the marginal (unit) cost of installing metering at a typical site;

d) Calculate the number of sites that it will be necessary to meter for the evaluation to achieve a desired level of statistical precision;

e) Coordinate the energy savings definitions between the measurement techniques to assure that everything accounted for in the tracking system is also accounted for, in a like manner, in the engineering model and metering studies;

f) Select an engineering model, and whatever data collection protocol is used to gather data for the model, based on whether it is adequate for calculating accurate savings for the DSM measures that are being studied; and,

g) Use prior studies to develop planning assumptions for determining sample size and sampling strategies.

h) Before/after monitoring will be especially appropriate in programs that include an audit of pre-retrofit conditions and calculate the tracking system savings using this information. If savings are based on the efficiency of standard replacement equipment, then only post retrofit monitoring may be sufficient.

3.4. Estimated Error Ratios

Given the importance of error ratios in developing an efficient sampling plan, it is useful to compare the error ratios that were assumed at the beginning of this study with the error ratios that were actually achieved by the study. The *assumed* ratios used in the sample design are in Table 4-1. The *achieved* error ratios are displayed in Table 3-2.

Table 3-2
Achieved Error Ratios

Engineering Model-to-Tracking System		
End Use	MWh	kW
Lighting (Express)	84%	75%
Lighting (Customized)	52%	62%
HVAC (Customized)	98%	N/A
Refrigeration	57%	N/A
Short-term Monitoring-to-Engineering Model		
End Use	MWh	kW
Lighting (Express)	46%	44%
Lighting (Customized)	78%	61%
Short-term Monitoring-to-Tracking System		
End Use	MWh	kW
Lighting (Express)	110%	96%
Lighting (Customized)	37%	60%
HVAC (Customized)	40%	N/A

In general, these findings indicate that the achieved error ratios are not, on average, very far off from the assumed error ratios of .60 for lighting measures, 1.0 for HVAC measures and .60 for refrigeration measures. However, the error ratios for specific samples can vary substantially from the assumed values. Higher values than were initially assumed indicate that the data fit more poorly than expected, while lower values indicate that the data fit better than expected.

For lighting measures in the Express program, it appears that the achieved error ratios for the impact sample are higher than assumed originally. However, the reverse is true for the validation sample. For HVAC measures, the error ratio is very close to the planning assumption. For directly comparing short-term monitoring to tracking system estimates for lighting measures, it appears that the Express error ratio is higher than expected and the Customized error ratio is lower than expected.

The results for the lighting measures in the Express program suggest that double sampling is an improvement over single sampling. The error ratios for both energy and demand savings are substantially smaller in the second-phase validation sample than in the first-phase impact sample. Moreover, the error ratios for short-term monitoring-only are even larger. This indicates that engineering model estimates are good predictors of short-term monitoring estimates, and are much better predictors than program tracking system estimates.

On the other hand, the results for the lighting measures in the Customized program reveal that the first-phase error ratios are smaller than the second-phase ratios. This indicates that the engineering model is not particularly good at predicting short-term monitoring estimates. This is reinforced by the short-term monitoring relationship to the program tracking system estimates. In short, for the Customized program the program tracking system estimates of savings are more closely aligned with the short-term monitoring estimates than are the engineering model estimates. However, this does not mean that the program tracking estimates are more *unbiased* than the engineering model estimates. It simply means that program tracking estimates contain less *variability* with respect to short-term monitoring than do the engineering model estimates.

3.5. Double Sampling Cost-effectiveness

To complete this study, the cost-effectiveness of the double sampling approach is compared to the cost-effectiveness of a single sampling approach that employs short-term monitoring, only. This analysis hinges on the relative cost of collecting engineering model data versus short-term monitoring field data. For the purposes of this analysis it is assumed that engineering model data costs $1,000 per site, and short-term monitoring data costs either $5,000, $10,000 or $15,000 per site. Table 3-3 displays the results of the cost-effectiveness analysis.

Table 3-3
Double Sampling Cost-Effectiveness Analysis

RETROFIT PROGRAM				
	Error Ratio	Unit Cost	Sample Size	Total Cost
Impact	84%	$1,000	66	$66,000
Validation	46%	$5,000	16	$80,000
Total				$146,000
One-Sample	%110	$5,000	53	$265,000
Case 2	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	84%	$1,000	81	$81,000
Validation	46%	$10,000	14	$140,000
Total				$221,000
One-Sample	110%	$10,000	53	$530,000
Case 3	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	84%	$1,000	92	$92,000
Validation	46%	$15,000	13	$195,000
Total				$287,000
One-Sample	110%	$15,000	53	$795,000
CUSTOMIZED PROGRAM				
Case 1	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	52%	$1,000	51	$51,000
Validation	78%	$5,000	34	$170,000
Total				$221,000
One-Sample	37%	$5,000	6	$30,000
Case 2	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	52%	$1,000	67	$67,000
Validation	78%	$10,000	32	$320,000
Total				$387,000
One-Sample	37%	$10,000	6	$60,000
Case 3	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	52%	$1,000	80	$80,000
Validation	78%	$15,000	31	$465,000
Total				$545,000
One-Sample	37%	$15,000	6	$90,000

Using the field data cost assumptions and the achieved error ratios for lighting measures, the findings indicate that to attain a 25 percent level of relative precision, the double sampling plan is cost-effective for the Express program, but not cost-effective for the Customized program. These results closely match what the analysis of the achieved error ratios suggested, namely that the engineering model data fit the short-term monitoring data for the Express program, but not for the Customized program.

4 METHODOLOGY

4.1. Introduction

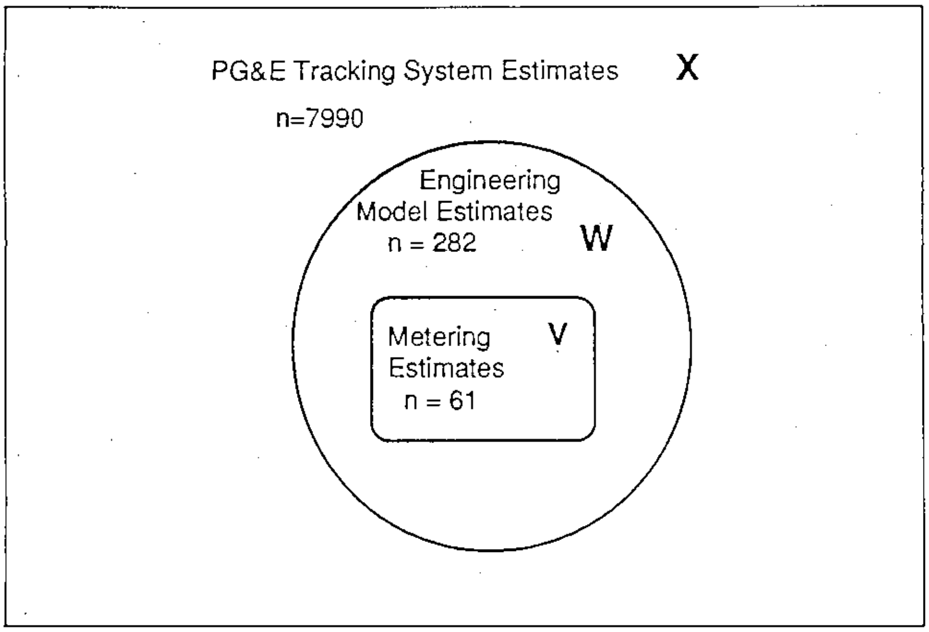
This section is a discussion of the method and data used in the Double Ratio Analysis. A brief discription of the steps used to determine energy and demand savings for this methodology is presented. Discussion of the data used in the analysis is also discussed in this section. In addition, summary information concerning the tracking system, monitoring and calibrated engineering modeling sites are presented. Detailed information and technical discussions concerning the double ratio methodology are found in Appendices A and C.

4.2. Double Ratio Estimation

The methodology used in this study is referred to as Double Ratio Estimation. It is a technique which utilitzed a small sample of expensive end-use metering sites, a larger sample of calibrated engineering modeling sites and the PG&E tracking system savings estimates. The estimates from these three sources are combined and expected to produce results with better precision than just using metering and tracking system estimates.

In this report, the short-term monitoring sites are referred to as "validation" sites. This denotes that, by having monitoring equipment installed to measure DSM savings, these sites provide savings estimates that are of the highest level of reliability and accuracy. They function to "validate" the two alternative estimates of savings, i.e. calibrated engineering model estimates and program tracking estimates. The validation sites are designated symbolically by the term *V*. The calibrated engineering model sites are referred to as "impact" sites, denoting their central role in this study as the major source of empirical estimates of program impacts. These sites are designated symbolically by the term *W*. Figure 4-1 shows how the three data sources relate to each other in terms of the samples.

Figure 4-1
Double Ratio Estimation - Embedded Sample



Two hundred and eighty-two engineering model sites make up the phase-one "impact" sample, W., and the 61 short-term monitored sites make up the phase-two "validation" sample, V. The phase-one ratio for the engineering modeled sites is calculated by dividing the program-weighted savings for the engineering modeled sites, W, by the savings estimated from the program tracking system for the same sites, X:

$$\text{1st-Phase Ratio} = (W/X) = R_1.$$

Similarly, the phase-two ratio is calculated by dividing the program-weighted savings for the short-term monitoring sites, V, by the engineering modeled program-weighted estimates of savings for these sites, W:

$$\text{2nd-Phase Ratio} = (V/W) = R_2.$$

These ratios are calculated separately, by end use and program type for both demand and energy savings. For each end use/program type analysis, the two ratios are combined by multiplying them together:

$$\text{Combined Ratio} = (R_1)^*(R_2) = \text{Realization Rate}$$

A realization rate for energy savings between zero and one implies that the empirical estimate of actual savings is that fraction of the estimate of the tracking system; a realization rate greater than one means that the program tracking system estimate of energy savings is understated. Lastly, a realization rate with a value of 1 means that all the savings that are reported on the program tracking system, no more and no less, have been confirmed via empirical measurement.

The standard error of each individual ratio is calculated in a similar manner to the calculation of the standard error of the weighted ratio estimate of a stratified sample. However, the standard error of the product of the weighted ratios, the realization rate, requires a more complex calculation. Technical details of this procedure are found in Appendix A.

4.3. Double Ratio Estimation Data

4.3.1. PG&E Tracking System Data

The PG&E program tracking system database is the central element in this evaluation, because it is the savings estimates of this database that are to be statistically adjusted by the study findings. Being the focus of this study, it is also the key element in the development of the sampling plan.

Detailed information about each CIA project undertaken by PG&E in 1992 is contained in the program tracking system database. The information for each application includes the name, address and account number for each program participant, as well as various processing dates for the application. For each type of measure on each application there exists specific information such as the end-use code, number of measures, rebate amount, and engineering estimates of the first-year kW and kWh savings.

Table 4-1 displays program tracking system summary information, recorded as of December 31, 1992, for program activity in 1992. These data are displayed by major end uses. Table 4-2 shows the project activity by type of program.

Table 4-1
1992 CIA Projects by End Use[1]

	# of Projects	% of Total Projects	GWh Savings	% of Total GWh	MW Savings	% of Total MW
Lighting	5963	58.4%	249.2	50.5%	61.6	59.9%
Space Cond.	1644	16.1%	106.3	21.5%	18.2	17.7%
Agriculture	1682	16.5%	82.9	16.8%	17.0	16.5%
Refrigeration	383	3.8%	21.4	4.3%	1.7	1.7%
Motors	231	2.3%	1.3	0.3%	0.1	0.1%
Process	138	1.4%	31.2	6.3%	3.9	3.8%
Control	132	1.3%	0.3	0.1%	0.0	0.0%
Hot Water	15	0.1%	0.6	0.1%	0.2	0.2%
Food Service	8	0.1%	0.2	0.0%	0.1	0.1%
Boiler	6	0.1%	0.2	0.0%	0.1	0.1%
Total	*10,202*	*100.0%*	*493.6*	*100.0%*	*102.9*	*100.0%*

Table 4-2
1992 CIA Projects by Program[1]

	# of Projects	% of Total Projects	GWh Savings	% of Total GWh	MW Savings	% of Total MW
Express Program						
Lighting	4,454	56.7%	144.0	59.0%	44.0	60.4%
Agriculture	1,484	18.9%	63.7	26.1%	12.8	17.6%
Space Cond.	1,317	16.8%	25.5	10.4%	15.1	20.7%
Motors	231	2.9%	1.3	0.5%	0.1	0.1%
Refrigeration	228	2.9%	8.3	3.4%	0.7	1.0%
Control	132	1.7%	0.3	0.1%	0	0.0%
Process	12	0.2%	1.2	0.5%	0.2	0.3%
Total	*7,858*	*100.0%*	*244.2*	*100.0%*	*72.9*	*100.0%*
Customized Program						
Lighting	1,509	64.4%	105.2	42.2%	17.6	58.7%
Space Cond.	327	14.0%	80.8	32.4%	3.1	10.3%
Agriculture	198	8.4%	19.2	7.7%	4.2	14.0%
Refrigeration	155	6.6%	13.1	5.3%	1	3.3%
Process	126	5.4%	30	12.0%	3.7	12.3%
Hot Water	15	0.6%	0.6	0.2%	0.2	0.7%
Food Service	8	0.3%	0.2	0.1%	0.1	0.3%
Boiler	6	0.3%	0.2	0.1%	0.1	0.3%
Total	*2,344*	*100.0%*	*249.4*	*100.0%*	*30*	*100.0%*

[1] The data used in this analysis is from the PG&E tracking system as of December 31, 1992 and does not reflect the entire 1992 energy savings by PG&E. The final 1992 savings estimates can be found in the *"Annual Summary Report on Demand Side Management Programs in 1992 and 1993"*, PG&E, March 1993.

As the table indicates, lighting projects made up the majority of the activities in each program types; average project savings for lighting is 32.2 MWh for the Express program, and 69.6 MWh for the Customized Program. In all, lighting measures account for 51 percent of all program energy savings. The next most important end use other than agriculture is space conditioning, consisting primarily of ventilation and air conditioning measures (HVAC). Average energy savings for these measures is 19.4 MWh for the Express program and 247 MWh for the Customized program. This end use accounts for approximately 22 percent of total program energy savings.

For these tabulations, agricultural measures that involved saving energy through "water conservation" were removed from the data set. The energy savings associated with these measures, which include water metering and upgrading of irrigation systems, cannot be adequately quantified with field monitoring and engineering studies. As a result, these measures were excluded from the data integration element of the CIA evaluation.

To begin developing the sampling plan, all CIA projects were categorized by end use and by custom versus retrofit programs using the most current program tracking system data available. As complete information for 1992 program activities was not available until after the end of the year, the sampling plan for this study was developed using tentative estimates of the number of program projects and program savings, by end use and program type.

In an attempt to accurately assess the value of each project's impact to PG&E, project and measure impacts were characterized both by energy and demand savings, and by lifecycle avoided costs. Avoided costs were calculated with a formula that incorporates the life of the measures and the assumed energy (kWh) and peak demand (kW) savings. These costs can provide a more realistic estimate of the value of the resource than annual energy savings. This is because they weight the resource savings by the number of years the measure is expected to last, as well as its contribution to easing capacity requirements.

The revised program tracking system estimates were used throughout 1992 to develop tentative data integration sampling plans, and each revised sampling plan was used as a general guide for selecting study sites. Coupled with the difficulty in recruiting targeted sites, the tentative sampling plans resulted in actual study samples that did not conform to the optimal sample design. Hence, to retroactively make the study samples representative of the 1992 program population, the sample was post-stratified and each observation was assigned a case weight. Post-stratification and the accompanying case weights were derived from the final program sampling plan that was created from the program tracking system database completed as of December 31, 1992. Details of the post-stratification weighting scheme are found in Appendix A. The post-stratification weights used in the analysis are found in Appendix B.

4.3.2. Short-term Monitoring Data

The two methods of estimating gross program savings for this research project, i.e. engineering model and short-term monitoring, require different kinds of field data. The process of obtaining short-term monitoring estimates of savings began with installing data loggers to collect hourly data from the original end-use measure for a minimum of one or two weeks prior to the installation of the program-funded DSM measure. If the total area to be retrofitted could not be directly monitored, then a representative area was selected and monitored.

After the retrofit, data were collected from the new measure for a minimum of two weeks. The pre and post-installation data were then analyzed by calculating electric load reduction, equivalent full load hours, and other parameters that determine energy use. Adjustments were made to account for burnt out lamps, annual use and other factors that can influence savings estimates. Once completed, the annualized results were extrapolated from the monitored area to the entire building area that was affected by the retrofit. To make the estimates conform to the same baseline that was used by the tracking system to estimate energy savings, adjustments were made to specific technologies to take into account current practice. More detailed information can be found in the *CIA Field Monitoring Results Report*, September 1993, The Fleming Group.

4.3.3. Calibrated Engineering Model Data

On-site inspection data was collected by arranging for building auditors to visit participant sites. Auditors used a customized engineering model survey instrument to inspect the retrofit-affected area, noting the number of pieces of equipment and their energy-related characteristics and efficiency ratings. In addition, the survey is used to collect detailed information on other end uses at the site. Through in-person interviews with building managers, building operations data is collected and key data are verified.

For this evaluation, on-site inspections were conducted prior to, and following the installation of DSM measures. The second inspection was meant to document the retrofit and any other changes that might be relevant to the calculation of energy savings. When completed, the engineering model was used to disaggregate end-use energy shares and to derive estimates of that portion of the change in energy use that was attributable to the retrofit measures. The model's module that is used for estimating building heating and cooling loads is based on ASHRAE's Cooling Load Temperature Difference (CLTD) Method. As with the short-term monitoring estimates, the engineering model estimates of savings were adjusted for specific technologies, depending on whether the tracking system savings were based on pre-installation conditions or current practice. A complete description of the model calibration results and engineering model savings estimates are found in the *CIA Retrofit Evaluation HSEM Analysis Report*, September 1993, XENERGY, Inc.

Having used prior load research for PG&E's service territory to calibrate the engineering model, none of the short-term monitoring data were used to adjust or modify the engineering model estimates, nor were the tracking system estimates used as feedback for the estimates of either of the two techniques. This point is important to bear in mind in considering the present research design. It strengthens the validity of the findings because it means that the three techniques for estimating energy savings were implemented independently. Hence, comparisons of their findings provide unambiguous evidence of how well they approximate each other.

A

DOUBLE RATIO METHODOLOGY

A.1. Introduction

The purpose of this appendix is to present the methodology of the double ratio analysis. The theoretical foundation of the methodology is found in Appendix C.

A.2. Sampling and Estimation Strategies

The sampling strategy for this study is known as "two-phase" or "double sampling". Although a description of this strategy can be found in many textbooks, it is not widely used because its application to many research studies is limited. Double sampling is motivated by the main goal of the research design for this project, which is to integrate high cost, high resolution data with lower cost data in a way that exploits the strengths, and avoids the weaknesses, of each mode of data. It is of greatest benefit when there is a substantial difference in the cost of collecting data, and when these different data will all be used for generating estimates of the same variable.

For present purposes, this variable is energy savings, and the different data from two evaluation methodologies, i.e. calibrated engineering modeling and short-term monitoring. All things being equal, if the variances of the estimates that are derived from the two techniques have a constant relationship to each other, the larger the differences in the costs of the data, the greater the benefits of double sampling. Benefits can be expressed in one of two ways; either as cost savings that arise from needing fewer sample points to attain a given level of statistical precision; or, as gains in statistical precision from using a given number of sample points.

As well as using double sampling to achieve a high level of precision for a given number of sample points, an estimation strategy known as "ratio expansion" has also been incorporated into the data integration methodology to help achieve the highest possible levels of statistical precision. Unlike the more common method, mean-per-unit expansion, for extrapolating sample estimates to a population, ratio expansion involves calculating the ratio of two estimates of the same (or highly representative) variables, and applying this ratio to the population estimate. It can be shown mathematically that ratio estimation can, under certain statistical conditions, greatly increase the efficiency of an estimate.

To tie together the sampling technique of double sampling, and the estimation technique of ratio expansion, two ratios, and their product, are calculated. This procedure quantifies the relationship between the three estimates of energy savings and results in a parameter that is interpretable as *the percentage of the program tracking system estimate of energy savings that can be confirmed based on the integration of two independent, empirical measurements of energy savings (i.e. engineering model and short-term monitoring).* This parameter, which represents empirically-confirmed savings, is referred to as the "realization rate" of program tracking estimates of savings.

A.3. Stratification and Sample Allocation Plan

To implement the double sampling and ratio estimation strategies, this study employs a sample stratification plan formulated using model-based statistical sampling (MBSS). MBSS resembles conventional stratified sampling in that the outcome of the sampling plan is a prescription for how many observations of certain types should be drawn for a study to achieve a targeted level of statistical precision. Like conventional stratified sampling, MBSS is used to determine the best and most practical number of strata, the boundaries of the strata and the number of sample points that should be allocated to each strata. For this study, the subject under investigation is participant energy savings, and thus,this is the primary variable used for stratification.

Conceptually, MBSS differs from conventional sampling in one major respect -- it doesn't create a stratification plan based on the outcome variable alone. MBSS creates a stratification plan based on the relationship of the outcome variables to one or more related variables. As in double sampling and ratio expansion, it can be demonstrated mathematically that if certain statistical assumption are met, an MBSS sampling plan will require fewer observations than other sampling techniques to achieve a targeted precision level, or obversely, higher levels of precision can be achieved for a given sample size.

Each of the two, interrelated sampling plans that were designed using MBSS were formulated around a theoretical model that described the relationship between the three types of estimates of participant energy savings. The critical parameter in these models is referred to as the "error ratio". An index of dispersion, it provides information about how well the predicted values of the outcome variable, i.e. energy savings, are correlated with the actual values of the outcome variable.

The MBSS technique employs the error ratio to develop an optimal sampling plan. If the association between different estimates of savings is strong, the error ratio is assumed to be small and the sample sizes needed to attain a given error bound are likely to be correspondingly small. If the association is weak, the error ratio is assumed to be large and thus the required sample size to attain a given error bound or level of precision will be relatively large.

The MBSS technique assumes that error ratios may vary across energy savings estimation techniques, as well as across end uses. The error ratios for the short-term monitoring sampling plan are calculated as the population sum of the standard deviations of the residuals in the relationship between short-term monitoring and engineering model estimates of savings. It is divided by the population sum of the expected values of dependent variable, that is, engineering model estimates of savings. The error ratios for the engineering model sampling plan are calculated similarly, only with respect to the tracking system estimates.

The nature of sampling plans is such that the data that must be used to develop the sampling plans does not exist until after the data collection is completed. Thus, sampling plans must rely on prior information or expert judgment, the accuracy of which can only be measured following the execution of the sampling plan. Based on engineering judgment and prior evaluation studies, error ratios for this evaluation were assumed for each end use and each type of site study. Table A-1 presents the assumed error ratios for the first-phase sample, where engineering model estimates of energy savings are compared to program tracking estimates of savings.

Table A-1
Engineering Model-to-Tracking Error Ratios

End Use	Error Ratio
Lighting	0.6
HVAC	1.0
Refrigeration	0.6

These values indicate that tracking system estimates of savings are expected to be more accurate for lighting and refrigeration measures, which are generally not seasonally sensitive, than for HVAC (space conditioning) measures, whose loads may vary greatly from day to day and season to season. Both the engineering model and the short-term monitoring estimates of savings are expected to provide relatively similar estimates of actual energy savings. Therefore, the error ratios for the second-phase sample, where short- term monitored energy savings data are compared to engineering model estimates of savings, were assumed to be the same as those of the first phase sample.

After finalizing the double sampling plan using the complete 1992 program tracking system database, the 61 validation sites and the 282 impact sites were reclassified from their initial strata into the final study strata. This process, known as sample "post-stratification", partially compensates for the fact that the achieved study samples are not as efficient as the one that was dictated by the final, optimal sampling plan.

A.4. Sample Case Weights

The statistical methodology for this study centers around calculating the percentage of the program-related energy savings that can be confirmed by empirical measurement. This percentage is referred to as the realization rate of the program tracking system estimate of savings. Before the realization rate can be calculated, the population-weighted energy savings must be calculated for the engineering model, the short-term monitoring, and the program tracking system samples of program projects. Other than the three project estimates of savings themselves, the major element of these calculations is the project "case weight".

Once reclassified, a case weight is assigned to each observation; it is used to expand the sample result to the population. The case weight associated with each sample project indicates the number of projects in the tracking system represented by that project. It is mathematically equivalent to the procedure used in conventional stratified sampling, where the sample mean of each strata is weighted by the strata's representation in the population. The case weight, like the strata weight, ensures that the sample average provides an unbiased estimate of the corresponding average in the entire program population.

The case weight is calculated by dividing the number of program projects for an end use and program type in a given strata , by the number of sample points for that end use and program type and strata. For example, if the largest strata for lighting projects in the direct retrofit program contained 249 projects, and the sample consisted of 16 of these projects, the weight assigned to each of these 16 projects would be 15.6. If *w* denotes the case weight of each sample project and *x* denotes any particular project savings, then the weighted average of savings for the population, X, is calculated using the equation:

$$X = \sum wx \,/\, \sum w$$

The equation indicates that the weighted average is calculated by multiplying the savings of each sample project by the case weight of the project, summing the results across all sample projects, and dividing by the sum of the case weights across all sample projects.

A.5. Sampling Background

Double sampling, sometimes called two-phase sampling, is a technique for extending the quality of information provided by limited amounts of data, thereby economizing on study resources. Double sampling entails drawing two interpenetrating samples. The smaller second-phase subsample provides the most defensible measurement, while the larger first-phase sample strengthens the sampling precision of the results.

The double-sampling strategy was applied to the CIA evaluation. The data-collection plan called for the measurement of savings using relatively expensive field monitoring in a limited number of sites. To strengthen the statistical precision of the field monitoring results, the evaluation plan called for much less expensive calibrated engineering model simulations in added sites.

Ratio estimation was used to develop realization rates from this information. Ratio estimation has been used in several prior projects to link monitoring results with tracking information. However, the present study is believed to be its first application to an impact-evaluation study that uses double-sampling.

Table A-2
Example - Lighting Express Program

Database	Number	Avg. Savings(MWh)		Ratio
Tracking	4,454	32.3		
Impact	96	34.4	29.1	84.5%
Validation	16	42.2	53.3	126.4%
Combined		32.3	34.5	106.8%

The lighting measures in the Express Program will be used to illustrate the approach. See Table A-2. Our tracking system contains 4,454 Express Program applications for lighting measures, with an average first-year savings of 32.3 MWh per project. The impact sample includes 96 of these projects and 16 of these 96 projects are also included in the validation sample.

Table A-2 summarizes the information from the building simulations and field monitoring. The first row of the table shows the average savings of 32.3 MWh for the 4,454 applications in the tracking system. The second row summarizes the information for the impact sites. Two different estimates are given. The first number, 34.4 MWh, is the average value of the tracking-system savings for the 96 impact sites; while the second number, 29.1 MWh, is the average savings determined from the calibrated engineering model analysis of the 96 impact sites.

The ratio between the engineering model and tracking estimates can be regarded as a realization rate. In our example, the ratio of 84.5% implies that the savings found in the engineering model analysis were 84.5% of the corresponding tracking estimates. In other words, the calibrated engineering model results indicate that the tracking estimates should be adjusted by multiplying them by 0.845. This ratio might be called the engineering model realization rate.

The next step is to adjust the engineering model results using the monitoring results found in the validation sample. The third row of Table A-2 summarizes the information from the validation sample. Again two different estimates of savings are given for the 16 validation sites. This time the two estimates are from the engineering model and field monitoring, respectively. The ratio of 126.4% indicates that field monitoring savings were 126.4% of the corresponding engineering model estimates. In other words, the monitoring results indicate that the engineering model estimates should be adjusted by multiplying them by 1.264. This is called the monitoring adjustment factor. The engineering model realization rate can be adjusted to reflect the monitoring results by multiplying it by the monitoring adjustment factor. The combined result is

$$84.55 \times 126.4\% = 106.8\%$$

It indicates that, based on the engineering model and monitoring results, the tracking estimates of savings should be adjusted by multiplying them by 1.068. This can be called the verified realization rate.

The final step is to apply the verified realization rate back to the tracking system. This step is shown in the bottom row of Table A-2. After adjustment by the verified realization rate, the average savings will be 106.8% of 32.3 MWh, which is equal to 34.5 MWh per project.

These techniques can also be used when only a single sample is available. In the present study, for example, Express HVAC and refrigeration measures were included in the impact sample but not in the validation sample. In this situation, a realization rate can be developed from the impact information, but of course there is no validation phase to true up any bias that might be present in the impact results.

A.6. Sampling and Weighting

At the beginning of this project sampling plans were developed for both the impact and validation samples. The sampling plans were based on 1991 tracking data and were stratified by end-use category and the size of the tracking-estimate of savings. The sampling plans were periodically revised to reflect the most current tracking information. It was not possible to follow the sample designs exactly because of difficulties in identifying upcoming projects with sufficient lead time. So the sampling plans were used as a guide in selecting sites for the impact and validation samples.

At the analysis stage, the final impact and validation samples were post-stratified to reflect the size of the tracking-estimate of savings within each end-use category of interest. Continuing the Lighting Express example of the previous section, Table A-3 illustrates the approach. In this example, the 4,454 projects in the tracking system have been grouped nto five strata according to the MWh savings of each project as shown in the tracking system. Columns 2-3 show the range of savings used to define each stratum, and column 4 shows the number of projects in the tracking system within this range. For example, stratum 1 contains 2,015 projects with savings from 0 MWh up to and including 8.3 MWh.

The range of savings is also used to stratify the sample of 96 impact projects. To provide comparability with the tracking system, the stratification is based on the tracking estimate of savings for each sample project. As column 5 shows, 23 sample projects were found to fall in stratum 1.

Table A-3
Example - Stratification and Weighting

Stratum	Estimated Saving		Number of Projects		Case Weight
	From	To	Tracking	Impact	
1	0.0	8.3	2,015	23	87.6
2	8.3	20.3	1,059	20	53.0
3	20.3	48.4	694	20	34.7
4	48.4	135.6	437	17	25.7
5	135.6	1,200.0	249	16	15.6
Total			*4,454*	*96*	

This information is used to calculate the case weight shown in the last column of Table B-3. The case weight within each stratum is simply the number of projects in the tracking system divided by the number of projects in the sample. In stratum 1, for example , the case weight is calculated as:

$$\frac{2,015}{24} = 87.6$$

The same procedure is used to post-stratify the validation sample. If, for example, the first stratum was found to contain 4 validation sites, the corresponding case weight would be:

$$\frac{2,015}{23} = 503.8$$

The case weight associated with each sample project indicates the number of projects in the tracking system represented by that project. The case weights are used to calculate the average savings for the sample. In Table A-2, for example, case weights were used to calculate the results in columns 3-4 for the impact and validation samples. The purpose of the case weights is to help ensure that the sample average provides an unbiased estimate of the corresponding average in the entire tracking system.

The case weights are actually used as follows. If w denotes the case weight of each sample project and x denotes any particular measure of savings for each project, then the weighted average of saving is calculated using the equation:

$$\bar{x} = \frac{\sum wx}{\sum w}$$

Here the symbol sum denotes the summation over all projects in the sample, so the equation indicates that the weighted average is calculated by multiplying the savings of each sample project by the case weight of the project, summing the results across all sample projects, and dividing by the sum of the case weights across all sample projects.

The use of case weights is equivalent to the usual procedures used in stratified sampling. In fact, the above equation is algebraically identical to the more familiar but somewhat more complicated equation:

$$\bar{x} = \frac{\sum N_h \bar{x}_h}{\sum N_h}$$

In this equation N_h denotes the number of projects in each stratum, $\bar{x}_h$ denotes the mean savings of the sample projects in each stratum and $\sum$ denotes the summation over all strata.

A.7. Statistical Precision with a Single Sample

The last methodological issue concerns the statistical precision of the reported realization rates. This issue will be discussed in this section and the following section. This section will discuss statistical precision in the case of a single sample, e.g., an impact sample. The next section will extend the one-sample methodology to double sampling, e.g., both an impact and validation sample.

In the single-sample situation, an appropriate modeling or monitoring technique is used to measure the savings in a stratified sample of projects. The realization rate is estimated by calculating a single ratio relating the measured savings to the average value of the tracking-estimates of savings for the sample. The approach is similar to the calculation of the calibrated engineering model realization rate described earlier.

In addition to the realization rate itself, a measure of the statistical precision is also provided. As an example, a realization rate might be reported as 85% ± 15%. Loosely speaking, this implies that the true realization rate is expected to be between 70% and 90%. The quantity ± 15% is called an error bound. The error bound reflects uncertainty regarding the extent to which the realization rate estimated from the sample provides an accurate representation of the true realization rate in the population of all projects in the tracking system.

The error bound itself is related to a second measure of statistical accuracy called the relative precision. The relative precision is simply the error bound divided by the estimated realization rate. In the preceding example the relative precision would be

$$\frac{\pm 15\%}{85\%} = \pm 18\%$$

This would imply that, after the realization rate is applied to the tracking information, the results are expected to be within ± 18% of the actual savings. Note that it is easy to confuse the error bound with the relative precision, since they are both measured as percentages.

The relative precision and error bound are derived following the principles of statistical sampling. The methodology has been adopted to ratio estimation with stratified sampling. This following discussion will give a nontechnical overview, with an emphasis on the underlying factors affecting the expected relative precision. The discussion will assume the 90% level of confidence and will assume that the sample is efficiently allocated among the strata. The focus will be on the relative precision, which in turn can be used to calculate the error bound.

In the single-sample case, the expected relative precision is determined by the size of the sample and the variability in the population. For ratio estimation with a single sample of size, the equation is:

$$rp = 1.654\left(\frac{er}{\sqrt{n}}\right)$$

Here *er* denotes a measure of population variability called the error ratio. For example, consider the impact sample for lighting measures in the Express Program. In this case, the error ratio is 84% and the sample size is equal to 96. Given these parameters, the expected relative precision would be approximately:

$$rp = 1.654\left(\frac{er}{\sqrt{96}}\right)$$

Then the expected error bound can be calculated as the product of the expected realization rate and the expected relative precision. For example, if the expected realization rate is 85%, then the expected error bound would be:

$$85\% \times 14 = \pm 12\%$$

Two additional comments: First, the preceding calculations are most appropriate in planning a project. Once the sample data are collected, more elaborate procedures are used to calculate the relative precision and error bound than is actually achieved. In the case of the Lighting Express Program, the final error bound was ± 15%. In this case the difference between the achieved error bound of + 15% and the expected error bound of ± 12% is primarily due to the difference between the final sample and the optimal allocation among strata. This difference arose from problems in identifying the required projects prior to their implementation.

The second comment concerns the fact that in the Lighting Express example the realization rate and error ratio found in the impact sample were almost equal, 85% and 84% respectively. This is purely coincidental. The realization rate and error ratio are independent parameters and generally are not equal.

Returning to the underlying principles, it has been seen that the expected error bound is determined from the error ratio and the sample size. The error bound will be small if the error ratio is sufficiently small or the sample is sufficiently large. The error ratio has the stronger effect since the sample size operates through the square root.

The error ratio reflects the strength of the association between the two measures of savings that make up the realization rate. If the association is strong, the error ratio is small and the error bound is correspondingly small, depending on the sample size. The association is strong if, after adjusting for the estimated realization rate, the tracking estimates generally give accurate project-by-project estimates of the measured savings observed in the sample.

Figure A-1
Example of an Error Ratio of 84%, Lighting Express

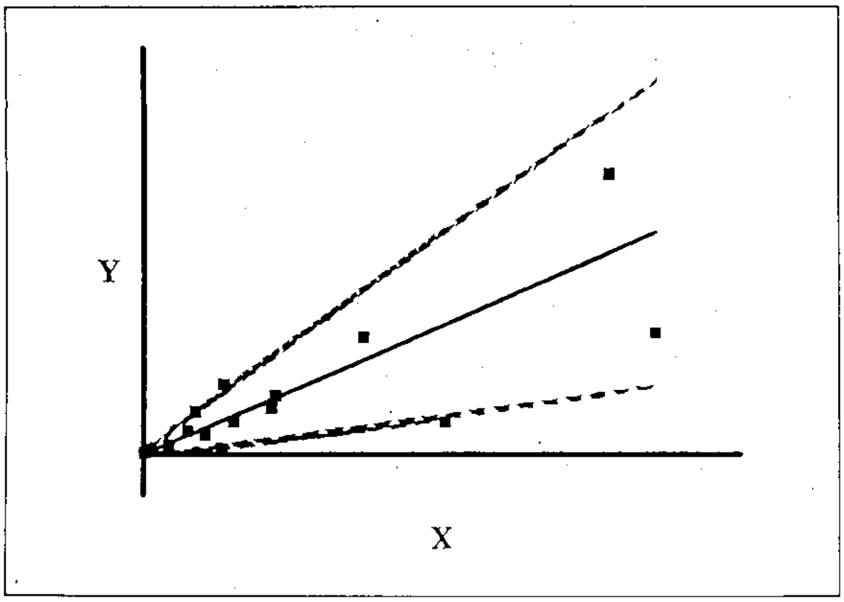


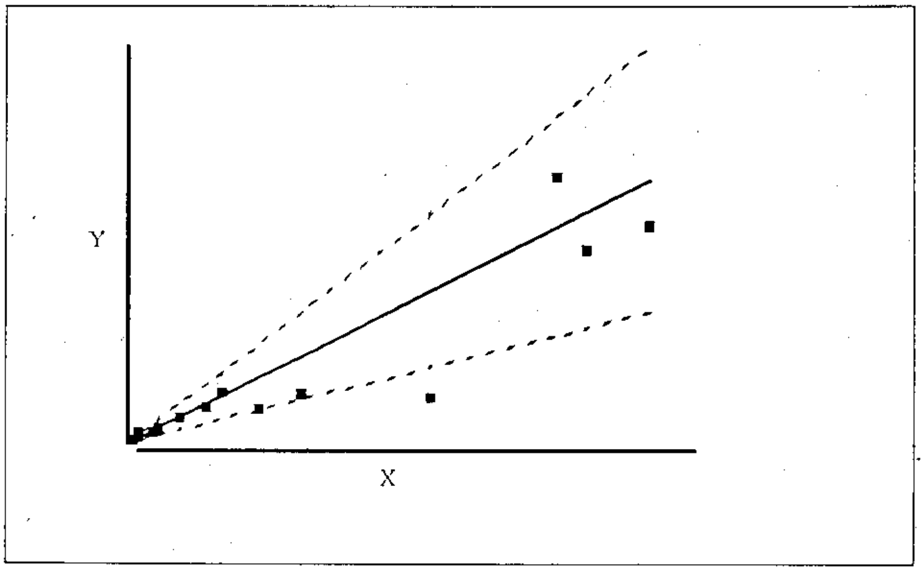
Figure A-1 illustrates the error ratio of 84% that was found in the Lighting Express impact sample. The variable plotted on the horizontal axis is the engineering estimate of savings from the tracking system. The vertical variable is the corresponding savings measured in the calibrated engineering model analysis. Each point represents a particular project in the impact sample.

The solid line is drawn from the origin through the point that represents the average value of the tracking estimates and the average value of the measured savings. The slope of this line is equal to the realization rate. This line represents the expected value of the measured savings for each project.

The measured savings of each project also has a standard deviation which represents its variability around the expected value. In most cases, the standard deviation increases with the expected value of the measured savings. In the simplest case, the standard deviation is directly proportional to the expected value. In this case, the error ratio is the standard deviation divided by the expected value. For example, if the standard deviation is equal to 84% of the expected value, then the error ratio is equal to 84%. Figure A-1 illustrates this example. The dashed lines are drawn around the solid line at plus and minus one standard deviation, i.e., at ± 84% of the height of the solid line.

Now it should be clear why the error ratio is so important. If the error ratio is small, then the measured points will generally fall very close to the solid line. Figure A-2 provides an example of an error ratio of 46%. This example is taken from the validation sample for the Lighting Express measures. When the error ratio is smaller, a smaller sample will provide a good estimate of the realization rate. In the extreme case that the error ratio is zero, even one measured point would be adequate.

Figure A-2
Example of an Error Ratio of 46%, Lighting Express



A.8. Statistical Precision with Double Sampling

The single-sampling principles described in the preceding section also provide the basis for the statistical precision of ratio estimation with double sampling. In the present application of double sampling, the information collected in the impact sample is used to estimate the engineering model realization rate, while the information in the validation sample determined the monitoring adjustment factor. The product of these two factors gives the verified realization rate. The issue to be discussed in this section is the statistical precision of the verified realization rate.

As in the case of a single sample, it is important to distinguish between two measures of statistical precision, the error bound, and the relative precision. The relative precision is usually calculated first and then used to find the error bound.

In the double-sampling case, the expected relative precision is related to:

(1) the error ratio and sample size in the impact sample, denoted er_1 and n_1
(2) the error ratio and sample size in the validation sample, denoted er_2 and n_2 respectively.

Here er_1 reflects the association between the tracking-estimates of savings and the calibrated engineering model estimates of savings, while er_2 reflects the association between the calibrated engineering model estimates of savings and the field monitoring estimates of savings. Then the expected relative precision is:

$$rp = 1.645\sqrt{\left(\frac{er_1^2}{n_1} + \frac{er_2^2}{n_2}\right)}$$

As an example, consider the Lighting Express case again. In this example, the impact error ratio is 84% and the sample size is 96, while the validation error ratio is 46% and the sample size is 16. Thus the relative precision is:

$$rp = 1.645\sqrt{\left(\frac{(84\%)^2}{96} + \frac{(46\%)^2}{16}\right)} = \pm 23\%$$

A contradiction seems to arise when this example is compared to the example in the prior section. The prior discussion showed that the impact sample alone would give an expected relative precision of about ± 14%. Here, when the validation sample is added, the precision drops to ± 23%. How can the precision be poorer when the added information is considered?

The answer is that the two results are not actually comparable. When the impact sample is used alone, it is necessary to assume that the impact results are unbiased. The expected precision of ± 14% is conditional on this crucial assumption. In the double-sampling situation, it is no longer assumed that the impact results are unbiased, although it is assumed that the validation results are unbiased. In fact, the validation sample is used to correct any bias that might occur. Thus the precision of ± 23% seems to be poorer, but the assumption that the field monitoring is unbiased may be more realistic than the assumption that the calibrated engineering model analysis is unbiased.

A more meaningful comparison is between the expected precision of the double-sample design with the precision that would have be expected using field monitoring alone, without the support from the calibrated engineering model impact sample. In this case the relevant error ratio reflects the association between the tracking estimates and the field monitoring results. In the Lighting Express example, this error ratio is estimated to be 110%. Using this parameter, the expected precision from a field monitoring sample of 16 projects would be:

$$1.645\left(\frac{110\%}{\sqrt{16}}\right) = \pm 44\%$$

This example illustrates why double sampling is often cost effective. In the impact sample, the error ratio is rather large (84%) but it is offset by a fairly large sample. The large sample can be afforded since the unit cost of the calibrated engineering model analysis is relatively small. In the validation sample, on the other hand, the error ratio is smaller (46%) because it reflects the engineering model results so that the sample can be kept small. This is appropriate since the monitoring used in the validation sample generally has a rather high unit cost.

The engineering model results improve the overall precision as long as the error ratio between the calibrated engineering model estimates and the field-monitoring results is substantially smaller than the error ratio between the tracking estimates and the field-monitoring results, e.g., 46% versus 110%.

Results like these are especially useful in planning new evaluation studies. Somewhat more complex calculations have been used to evaluate the statistical precision that has been achieved in our samples. These more complex calculations are usually favored because they require fewer simplifying assumptions than the methodology discussed in Section A.7.

A.9. Estimated Error Ratios

In Section A.3 it was shown that the statistical precision expected from a single sample depends on the size of the sample and the variability in the population. For stratified ratio estimation, the population variability can be measured by the error ratio which reflects the strength of the association between the two estimates of savings used to calculate the ratio, e.g., the tracking estimates of savings and the calibrated engineering model estimates. These error ratios are not directly related to the error bounds achieved by a specific sample since they assume a near-optimal sample design. However, like the error bounds, the error ratios can be estimated from the available sample data.

Table A-4 summarizes the error ratios that have been estimated from the data available for the validation sample. Results are presented for each end-use category for both energy and demand impacts. These error ratios measure the strength of the association between the tracking estimates of savings and the estimates developed from the engineering model analysis. Rather high error ratios, representing a poor association, were expected since the tracking estimates are thought to give a relatively poor prediction of actual savings. The values reported in Table A-4 are in the range of 57% to 99%, which is substantially higher than was expected. This information can be used to plan the sample sizes for new evaluation studies using calibrated engineering model techniques of data collection.

The error ratios shown in Table A-4 seem to follow a systematic pattern. The error ratios appear to be lower for lighting measures than for HVAC measures. This may be because simple engineering estimates are more accurate for lighting than for HVAC. The results for lighting indicate that the error ratios may be lower in the Customized Program than in the Express Program. This is consistent with the fact that the tracking estimates in the Customized Program usually reflect more site-specific information about operating hours. In the refrigeration category the results are based on data pooled across both programs. The error ratio appears to be high, suggesting the refrigeration tracking estimates are poor predictors of the calibrated engineering model estimates.

Table A-4
Example - Error Ratios

End-use	MWh	kW
Lighting Express	84%	75%
HVAC Express	56%	76%
Lighting Customized	61%	70%
HVAC Customized	114%	NA
Refrigeration - Both Programs	157%	NA

In the case of double sampling, the expected statistical precision depends on the sample sizes and error ratios in each of the two phases, as discussed in Section A.4. The first-phase error ratios reflect the strength of the association between the tracking estimates and calibrated engineering model results in the impact sample. The estimated first-phase error ratios have been shown in Table A-4.

The second-phase error ratios, in turn, describe the strength of the association between the calibrated engineering model estimates of savings and the field-monitoring estimates. The second phase error ratios can be estimated from the validation sample. Results are available for the energy and demand savings of lighting measures in the Express and Customized Programs. The third column of Table A-5 summarizes the results. To facilitate comparisons, the second column of Table A-5 repeats the first-phase impact error ratios from Table A-4.

The validation sample can also be used to estimate error ratios between the tracking estimates of savings and the field-monitoring estimates. These error ratios are used to predict the statistical precision in a one-sample experimental design using monitoring alone, without the engineering model estimates. The last column of Table A-6 reports these results.

The results shown for the Express Program are very promising for double sampling. The error ratios for both energy and demand are substantially smaller in the second-phase validation sample than in the first-phase impact sample. Moreover the error ratios for monitoring-only are even larger. These results indicate that the calibrated engineering model estimates are good predictors of the field-monitoring estimates and are much better than the tracking estimates.

The results are quite different for the Customized Program. In this case the validation error ratios are larger than the impact error ratios. This indicates that the calibrated engineering model information is not providing particularly good estimates of the field-monitoring results. This conclusion is reinforced by the monitoring-only results. The monitoring-only error ratios are smaller than the validation error ratios. This indicates that the tracking estimates of impact provide more information than the engineering model results.

It should be emphasized that these results do not imply that the unadjusted tracking estimates are as accurate as the calibrated engineering model estimates. To compare the accuracy of the unadjusted tracking and engineering model estimates it is necessary to consider both bias and variability. The error ratios reflect variability, while bias is indicated by the realization rates discussed in in the body of the report. The realization rates indicate that there is more bias in the tracking estimates of savings than in the calibrated engineering model results.

Table A-5
Example - Estimated Error Ratios for Lighting Express

End use	Impact	Validation	Monitoring Only
Express MWh	84%	46%	110%
Express kW	75%	49%	96%
Customized MWh	52%	78%	37%
Customized kW	62%	61%	60%

Table A-6
Example - Gammas Found in the Impact Sample

End use	MWh	kW
Lighting Express	0.9	0.9
VAC Express	0.4	0.3
Customized Lighting	0.5	0.9
Customized HVAC	0.2	1.5
Refrigeration	0.3	NA

Table A-7
Example - Estimated Gammas for Lighting Measures

End use	Impact	Validation	Monitoring Only
Express MWh	0.9	1.2	0.7
Express kW	0.9	0.8	0.6
Customized MWh	0.5	1.2	1.0
Customized kW	0.9	0.9	0.7

Values close to 1 indicate that stratification by estimated savings will be highly useful, with higher sampling fractions for larger projects. Values close to 0 indicate that a representatively allocated sample will be most efficient. Generally the estimated values are close to expectation.

A.10. Optimal Experimental Design

A statistical study is said to be optimally designed if it is expected to provide a specified level of precision at the least cost. In the context of this report, we seek the best allocation of resources between the first-phase calibrated engineering model work and the second-phase field monitoring.

Section A-4 described an equation for estimating the expected relative precision from estimates of the first- and second-phase error ratios, together with the sample sizes selected for the two phases. With some algebraic manipulation, together with suitable assumptions about the cost of the impact and validation samples, this equation can be used to design the optimal double-sampling study. The cost of the best double-sampling study can be compared to the cost of a single-sample approach that achieves the same expected precision using monitoring alone.

The results of this analysis will depend on the relative cost per unit for the impact and validation samples. The analysis presented in this section assumes that the engineering model work can be carried out for $1,000 per impact project, while the field monitoring may cost $5,000 or $10,000 or even $15,000 per validation project. Of course the actual costs will depend on the characteristics of each particular program, as well as the approach to the calibrated engineering model and field monitoring data collection and analysis.

To provide specific examples, we have combined these cost assumptions with the specific error ratios given in Table A-4. The findings are summarized in Tables A-8 through A-11. All of these applications assume that the expected relative precision is set at ± 25%.

Table A-8 is based on the error ratios for the energy savings (MWh) of lighting measures in the Express Program. The table shows three cases, with unit cost in the validation sample equal to $5,000 $10,000 and $15,000 per sample project, respectively. In each case, column four reports the sample sizes for the impact and validation samples under the optimal double-sampling approach. Column four also shows the sample size required to provide the same relative precision using a one-sample approach using monitoring alone. For example, in case 1, the best double-sampling experimental design requires 66 impact projects and 16 validation projects, while a one-sample approach giving the same expected precision requires 53 monitoring projects.

Column five of Table A-8 shows the total costs in the various situations. The effectiveness of the double-sampling strategy can be seen by comparing the total cost of the impact and validation samples with the total cost of the equivalent one-sample approach.

Table A-8
Sample Sizes for Varying Cost - Lighting Express MWh

Case 1	Error Ratio	Unit Cost	Sample Size	Total Cost
Impact	84%	$1,000	66	$66,000
Validation	46%	$5,000	16	$80,000
Total				***$146,000***
One-Sample	110%	$5,000	53	$265,000
Case 2	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	84%	$1,000	81	$81,000
Validation	46%	$10,000	14	$140,000
Total				***$221,000***
One-Sample	110%	$10,000	53	$530,000
Case 3	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	84%	$1,000	92	$92,000
Validation	46%	$15,000	13	$195,000
Total				***$287,000***
One-Sample	110%	$15,000	53	$795,000

In case 1, for example, the optimal double-sampling approach would cost $146,000 whereas the equivalent one-sample design would cost $265,000. So, in this case, the double-sampling approach gives the same precision at a substantially lower cost. The savings are even larger in cases 2 and 3 which assume relatively higher monitoring costs per unit.

Table A-9 is based on the error ratios for the demand (kW) savings of lighting measures in the Express Program. Since the validation and one-sample error ratios are larger than in Table A-7, the required sample sizes and total costs are corresponding larger. And the savings from the double-sampling approach are also larger.

The results shown in Table A-10 are quite different. This analysis is based on the error ratios estimated for MWh savings in the Customized lighting program. In this situation, the best double-sampling approach is much more expensive that the one-sample approach using monitoring alone. This is because the error ratio in the validation sample is larger than in the one-sample approach. In other words, the calibrated engineering model information is not beneficial because it fails to reduce the error ratio associated with monitoring.

A less extreme situation is shown in Table A-11, Lighting Customized kW savings. In this case, the validation error ratio is smaller than the one-sample error ratio so the calibrated engineering model information has some benefit. But in each of the three cases, the benefit is not large enough to offset the cost.

Table A-9
Sample Sizes for Varying Cost - Lighting Express kW

Case 1	Error Ratio	Unit Cost	Sample Size	Total Cost
Impact	75%	$1,000	61	$61,000
Validation	49%	$5,000	18	$90,000
Total				***$151,000***
One-Sample	96%	$5,000	40	$200,000
Case 2	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	75%	$1,000	77	$77,000
Validation	44%	$10,000	16	$160,000
Total				***$237,000***
One-Sample	96%	$10,000	40	$400,000
Case 3	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	75%	$1,000	89	$89,000
Validation	49%	$15,000	15	$225,000
Total				***$314,000***
One-Sample	96%	$15,000	40	$600,000

Table A-10
Sample Sizes for Varying Cost - Lighting Customized MWh

Case 1	Error Ratio	Unit Cost	Sample Size	Total Cost
Impact	61%	$1,000	70	$70,000
Validation	93%	$5,000	48	$240,000
Total				***$310,000***
One-Sample	53%	$5,000	12	$60,000
Case 2	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	61%	$1,000	93	$93,000
Validation	93%	$10,000	45	$450,000
Total				***$543,000***
One-Sample	53%	$10,000	12	$120,000
Case 3	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	61%	$1,000	109	$109,000
Validation	93%	$15,000	43	$645,000
Total				***$754,000***
One-Sample	53%	$15,000	12	$180,000

Table A-11
Sample Sizes for Varying Cost - Lighting Customized kW

Case 1	Error Ratio	Unit Cost	Sample Size	Total Cost
Impact	70%	$1,000	71	$71,000
Validation	73%	$5,000	33	$165,000
Total				***$236,000***
One-Sample	53%	$5,000	12	$60,000
Case 2	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	70%	$1,000	91	$91,000
Validation	73%	$10,000	30	$300,000
Total				***$391,000***
One-Sample	53%	$10,000	12	$120,000
Case 3	**Error Ratio**	**Unit Cost**	**Sample Size**	**Total Cost**
Impact	70%	$1,000	108	$108,000
Validation	73%	$15,000	29	$435,000
Total				***$543,000***
One-Sample	53%	$15,000	12	$180,000

A.11. Analysis Data

This section will provide additional information about the data used in this report. The tracking data will first be described, followed by a brief description of the calibrated engineering model and field-monitoring data.

Table A-12 summarizes our tracking information for the Express Program. The table shows the number of applications and savings for each of seven end-use categories. In the Express Program, there were 4,454 applications with lighting measures. The total savings recorded in the tracking system for these measures was 144 GWh energy and 44 MW demand. The lighting measures represent about 60% of the total energy and demand savings of the Express Program.

Table A-13 shows similar information for the Customized Program. In this case, lighting and space conditioning save 186 GWh of energy and 20.7 MW of demand, or 75% of total energy and 70% of total demand.

Table A-12
PG&E 1992 Tracking System Summary - Express Program
(as of December 31, 1992)

End-use Category	Number of Applications	Savings	
		GWh	MW
Agriculture	1,484	63.7	12.8
Control	132	.3	0.0
Lighting	4,454	144.0	44.0
Motors	231	1.3	0.1
Process	12	1.2	0.2
Refrigeration	228	8.3	0.7
HVAC	1,317	25.5	15.1
Total		*244.2*	*72.9*

Table A-13
PG&E 1992 Tracking System Summary - Express Program
(as of December 31, 1992)

End-use Category	Number of Applications	Savings	
		GWh	MW
Agriculture	198	19.2	4.2
Boiler	6	0.2	0.1
Food Service	8	0.2	0.1
Hot Water	15	0.6	0.2
Lighting	1,509	105.2	17.6
Process	126	30.0	3.7
Refrigeration	155	13.1	1.0
HVAC	327	80.8	3.1
Total		*249.4*	*30.0*

A.12. References

"Measuring DSM Impacts: End-Use Metering and the Engineering Calibration Approach," Mike Townsley and Roger Wright, "End Use Load Information and Its Role in DSM", The Fleming Group, Syracuse, 1990.

"End-use load information in the commercial sector, The Energy Information Project," Dick Richards, Roger L. Wright and Curt D. Puckett, "End Use Load Information and Its Role in DSM", The Fleming Group, Syracuse, 1990.

"Effective Sample Designs in DSM Evaluation, A Look at Data from Prior Studies," prepared by RLW Analytics for Northeast Utilities, February 3, 1992.

B

CASE WEIGHTS

Tables B-1 through B-8 in this appendix contain the post-stratification and case weights for end uses by program which are used in the double ratio analysis. Details of the methodology approach are in Appendix A.

Table B-1
Express Program Lighting Impact Sites
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Impact	
1	0.0	8.3	2,015	23	87.6
2	8.3	20.3	1,059	20	53.0
3	20.3	48.4	694	20	34.7
4	48.4	135.6	437	17	25.7
5	135.6	1,200.0	249	16	15.6
Total			*4,454*	*96*	

Table B-2
Express Program Lighting Validation Sites
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Validation	
1	0.0	12.9	2,572	3	857.3
2	12.9	34.5	966	0	N/A
3	34.5	86.6	505	2	252.5
4	86.6	185.1	269	7	38.4
5	185.1	2,000.0	142	4	35.5
Total			*4,454*	*16*	

Table B-3
Customized Program Lighting Impact Sites
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Impact	
1	0.0	7.5	382	30	12.7
2	7.5	21.0	387	38	10.2
3	21.0	51.1	292	26	11.2
4	51.1	124.0	256	19	13.5
5	124.0	6,000.0	192	20	9.6
Total			*1,509*	*133*	

Table B-4
Customized Program Lighting Validation Sites
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Validation	
1	0.0	7.5	382	6	63.6
2	7.5	21.0	387	8	48.4
3	21.0	51.1	292	7	41.7
4	51.1	124.0	256	7	36.6
5	124.0	6,000.0	192	8	24.0
Total			*1,509*	*36*	

Table B-5
HVAC Impact Sites (Both Programs)
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Impact	
1	0.0	1.4	247	7	35.3
2	1.4	4.1	270	7	38.6
3	4.1	8.7	269	4	67.3
4	8.7	24.0	266	4	66.5
5	24.0	700.0	265	3	88.3
Total			*1,317*	*25*	

Table B-6
Customized Program HVAC Impact Sites
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Impact	
1	0	17.6	76	4	19.0
2	17.6	69.9	72	3	24.0
3	69.9	165.6	65	3	21.7
4	165.6	388.8	61	6	10.2
5	388.8	9,000.0	53	2	26.5
Total			*327*	*18*	

Table B-7
Customized Program HVAC Validation Sites
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Validation	
1	0.0	17.6	76	2	38.0
2	17.6	69.9	72	1	72.0
3	69.9	165.6	65	1	65.0
4	165.6	388.8	61	3	20.3
5	388.8	9,000.0	53	2	26.5
Total			*327*	*9*	

Table B-8
Refrigeration Validation Sites (Both Programs)
Post-Stratification and Case Weights

Stratum	Estimated Savings		Number of Projects		Case Weight
	From	To	Tracking	Validation	
1	0.0	10.6	178	1	178.0
2	10.6	30.2	94	6	15.7
3	30.2	68.1	59	1	59.0
4	68.1	284.7	36	1	36.0
5	284.7	3,000.0	16	1	16.0
Total			*383*	*10*	

C FOUNDATIONS OF DOUBLE SAMPLING

1 Introduction

Double sampling, also called two-phase sampling, is an important extension of ordinary ratio estimation.[1] In ordinary ratio estimation, additional statistical precision is achieved by taking advantage of the association between the target variable of interest, say y, and an explanatory variable x that is known thoughout the population.

We say that x provides supporting information for y. Two conditions are required for ratio estimation to be effective:

1. x must be a good predictor of y in the sense that x explains much of the variation in y from one sampling unit to another, and

2. The average or total value of x must be known in the population of interest.

→ citation ?

For example, ratio estimation is being used in some impact evaluation studies. In these applications y usually is the impact of a retrofit project measured through end-use metering, and x is the engineering estimate of the impact that is available in the program tracking system. Ratio estimation will work well if the engineering estimates are highly correlated with the measured impact. Double sampling, on the other hand, holds promise even when the tracking estimates are relatively poor.

In the context of impact evaluation, double sampling is appropriate when

1. The available tracking information is only weakly correlated with measured savings, but

2. It is possible to collect good site-specific engineering information in a first-phase sample that is substantially larger than the second-phase sample in which y is observed.

[1] The methods discussed throughout this note apply equally well to generalized regression estimation.

In the CIA evaluation study being conducted for PG&E, site-specific engineering engineering information is being collected and analyzed using building simulation and related techniques in a relatively large sample called the evaluation sample. Due to cost considerations, end-use metering is restricted to a substantially smaller sample called the validation sample which is a subset of the evaluation sample. Double sampling statistical theory was used in the scoping study to plan the CIA evaluation, and double sampling methodology will play an important role in the analysis.

The purpose of this technical note is to discuss the double-sampling statistical methodology both in sample design and in analysis.

2 Optimal Sample Design for Double Sampling

Double sampling is discussed in Chapter 12 of Cochran [1] and in Chapter 9 of Sarndal et. al. [2]. But the CIA application has required the development of additional techniques. Our approach to double sampling will be grounded on the Model-Based Statistical Sampling approach, e.g., Chapter 12 of Sarndal and [3].

2.1 The Two-Phase Ratio Estimator

Let y be the target variable which is observed only in the second-phase sample, let x be a predictor of y based on information that is collected in a larger first-phase sample that includes the second phase sample, and let w be a predictor of y based on the information that is available throughout the population. We want to estimate the population total value of y, denoted Y, from the sample information and the known population total of w, denoted W. Alternatively, we want to estimate the *realization rate* which is defined to be the fraction $B = Y/W$. Since W is known, the problem of estimating Y is equivalent to the problem of estimating the realization rate B.

We will estimate Y or B using the two-phase ratio estimators:

$$\begin{aligned} \hat{Y} &= W\left(\frac{\bar{x}_1}{\bar{w}_1}\right)\left(\frac{\bar{y}_2}{\bar{x}_2}\right) \qquad (1) \\ &= W b_1\, b_{2.1} \end{aligned}$$

$$\begin{aligned}\hat{B} &= \hat{Y}/W \\ &= b_1\, b_{2.1}\end{aligned}$$

Here $b_1 = \bar{x}_1/\bar{w}_1$ is the ratio between the mean of x and the mean of w observed in the relatively large first-phase sample. The product Wb_1 corrects W using b_1, thereby giving an approximately unbiased estimate of the population total X.

Similarly, $b_{2.1} = \bar{y}_2/\bar{x}_2$ is the ratio between the mean of y and the mean of x observed in the smaller second-phase sample. $b_{2.1}$ is used to further correct the product Wb_1, thereby giving an approximately unbiased estimate of the population total Y.

The double sampling ratio estimator is a function of four sample means: $\bar{w}_1$, $\bar{x}_1$, $\bar{x}_2$, and $\bar{y}_2$. If simple random sampling or stratification with proportional allocation is used in each of the two phases, then these are ordinary unweighted sample means. If, however, either phase uses stratification with optimal or near optimal allocation, then each of these statistics must be properly weighted to reflect the sample design.

Let $\pi_{1k} = Pr(k \in s_1)$ be the probability that case k is included in the first-phase sample s_1, and let $\pi_{2k} = Pr(k \in s_2)$ be the probability that case k is included in the second-phase sample s_2. [2] Using these inclusion probabilities, we define the weighted means to be:

$$\begin{aligned}\bar{w}_1 &= \frac{\sum_{s_1} \pi_{1k}^{-1} w_k}{\sum_{s_1} \pi_{1k}^{-1}} \\ \bar{x}_1 &= \frac{\sum_{s_1} \pi_{1k}^{-1} x_k}{\sum_{s_1} \pi_{1k}^{-1}} \\ \bar{x}_2 &= \frac{\sum_{s_2} \pi_{2k}^{-1} x_k}{\sum_{s_2} \pi_{2k}^{-1}} \\ \bar{y}_2 &= \frac{\sum_{s_2} \pi_{2k}^{-1} y_k}{\sum_{s_2} \pi_{2k}^{-1}}\end{aligned} \tag{2}$$

2.2 The Double-Sampling Model

In ordinary ratio estimation, the MBSS approach to sample design is to

[2] π_{2k} is the unconditional inclusion probability, i.e., $\pi_{2k} = \pi_{1k} \times Pr(k \in s_2 | k \in s_1)$. Sometimes it is necessary to work with a somewhat different quantity $\pi_{2k}^{*} = Pr(k \in s_1) \times Pr(k \in s_2|s_1)$. See Section 9.2 of Sarndal.

formulate a model for the relationship between y and x. The assumed model is used to select a suitable sample design. Once the sample design is specified, the data analysis closely follows traditional survey sampling practice. This approach provides good statistical efficiency while retaining good protection against bias from possible model misspecification.

In ordinary ratio estimation, the MBSS model usually consists of two equations, called the primary and secondary equations. Double sampling requires a more complex model, involving four equations. For any sampling unit k, let

$$\begin{aligned} y_k &= \beta_{2.1}\, x_k + \epsilon_{2.1k} \\ &= E(y_k) + \epsilon_{2.1k} \end{aligned} \tag{3}$$

Here $E(y_k)$ is the expected value of y_k given the value of the stronger predictor x_k.

We assume that $E(y_k)$ is related to the weaker predictor w_k following another equation

$$E(y_k) = \beta_2\, w_k + \epsilon_{1k} \tag{4}$$

In the preceding equations, ϵ_{1k} and $\epsilon_{2.1k}$ are assumed to be independent random variables with zero expected value. Their standard deviations are allowed to depend on the sampling unit. We write

$$\begin{aligned} sd(\epsilon_{1k}) &= \sigma_{1k} = \sigma_0 x_k^{\gamma_1} \\ sd(\epsilon_{2.1k}) &= \sigma_{2.1k} = \sigma_0 x_k^{\gamma_2} \end{aligned} \tag{5}$$

This model is used to define the first-phase error ratio to be

$$\begin{aligned} er_1 &= \frac{\sum_U sd(\epsilon_{1k})}{Y} \\ &= \frac{\sum_U \sigma_{1k}}{Y} \end{aligned} \tag{6}$$

Note that the error ratio is sometimes defined in terms of $\sum_U E(y_k)$ instead of Y. The use of Y simplifies analysis for the realization rate B.

The second-phase error ratio is defined similarly:

$$er_2 = \frac{\sum_U \sigma_{2.1k}}{Y}. \tag{7}$$

2.3 Optimal Model-Based Sample Design

In our context, the basic results for two-phase sample design are the following:

Result 1:

As previously defined, let $\pi_{1k} = Pr(s_1)$ be the probability that case k is included in the first-phase sample s_1, and let $\pi_{2k} = Pr(s_2)$ be the probability that case k is included in the second-phase sample s_2. Let ANV denote the approximate anticipated variance, i.e., the expected value of the approximate sampling variance of $\hat{Y}$ or $\hat{B}$ considering the model. Then the approximate anticipated variance is the sum of two terms corresponding to the first and second phases of the sample design:

$$\begin{aligned} ANV(\hat{Y}) &= \sum_U \left(\frac{1}{\pi_{1k}} - 1\right) \sigma_{1k}^2 + \sum_U \left(\frac{1}{\pi_{2k}} - 1\right) \sigma_{2.1k}^2 \\ ANV(\hat{B}) &= \frac{ANV(\hat{Y})}{W^2} \end{aligned} \tag{8}$$

Proof:

The proof is similar to Sarndal's proof of Equation 12.2.12 in Section 12.2.

Result 2:

The approximate anticipated variance is minimized by a sample design for which

$$\begin{aligned} \pi_{1k} &= \frac{n_1\, \sigma_{1k}}{\sum_U \sigma_{1k}} \\ \pi_{2k} &= \frac{n_2\, \sigma_{2.1k}}{\sum_U \sigma_{2.1k}} \end{aligned} \tag{9}$$

Here n_1 and n_2 are the sample sizes in the first and second phases of the sample design. In words, under an optimal design, the inclusion probabilities in each phase are proportional to the corresponding standard deviations. In practice, near-optimal sample designs can be developed through model-based stratification as described in Section 12.4 of Sarndal.

Result 3:

Let the minimum value of the approximate anticipated variance be denoted ANV_0. Also let

$$\begin{aligned} \delta_1 &= \frac{N \sum_U \sigma_{1k}^2}{\left(\sum_U \sigma_{1k}\right)^2} \\ &= 1 + cv_{\sigma_1}^2 \end{aligned} \tag{10}$$

Let δ_2 be defined in a similar way using $\sigma_{2.1k}$. Then:

$$\begin{aligned} ANV_0(\hat{Y}) &= \frac{1}{n_1}\left(1 - \frac{n_1\,\delta_1}{N}\right)\left(\sum_U \sigma_{1k}\right)^2 \\ &+ \frac{1}{n_2}\left(1 - \frac{n_2\,\delta_2}{N}\right)\left(\sum_U \sigma_{2k}\right)^2 \\ ANV_0(\hat{B}) &= \frac{ANV_0(\hat{Y})}{W^2} \end{aligned} \tag{11}$$

Result 4:

The preceding result can be rephrased in terms of relative precision and error ratios. Let the approximate anticipated relative precision rp be defined to be

$$rp = \frac{z_{\alpha/2}\sqrt{ANV_0(\hat{Y})}}{Y} \tag{12}$$

$$= \frac{z_{\alpha/2}\sqrt{ANV_0(\hat{B})}}{B} \tag{13}$$

Here $z_{\alpha/2}$ denotes the z-value corresponding to a specified level of confidence, e.g., $z_{\alpha/2} = 1.645$ for the 90% level of confidence. Then, under an optimal sample design,

$$rp = z_{\alpha/2}\sqrt{\frac{1}{n_1}\left(1 - \frac{n_1\delta_1}{N}\right) er_1^2 + \frac{1}{n_2}\left(1 - \frac{n_2\delta_2}{N}\right) er_2^2} \tag{14}$$

If n_1 and n_2 are small relative to N, then rp is approximately

$$rp = z_{\alpha/2}\sqrt{\frac{er_1^2}{n_1} + \frac{er_2^2}{n_2}} \tag{15}$$

It is useful to note that both (13) and (14) can be written in the simple form

$$rp = \sqrt{rp_1^2 + rp_2^2} \tag{16}$$

with rp_1 and rp_2 defined to be the anticipated relative precision associated with phase 1 and phase 2 respectively.

Result 5:

Assume that we want to choose n_1 and n_2 to achieve an optimal allocation of resources between the first and second phases of the sample. Let c_1 be the variable cost per sample unit in the first phase, and let c_2 be the additional cost per sample unit in the second phase, so that the total variable cost is $C = n_1\,c_1 + n_2\,c_2$. Let

$$d = \left(\frac{er_1}{er_2}\right)\sqrt{\frac{c_2}{c_1}} \tag{17}$$

provided that this is greater than 1; otherwise let $d = 1$. Also assume the use of an optimal sample design within each phase following Result 2.

To achieve an optimal allocation for the anticipated variance or relative precision, the sample sizes should satisfy the equation:

$$n_1 \doteq d\,n_2 \tag{18}$$

Moreover, if n_1 and n_2 are small relative to n and if the required relative precision rp is specified, the sample size should also satisfy

$$n_2 = z_{\alpha/2}^2\left(\frac{er_1^2/d + er_2^2}{rp^2}\right) \tag{19}$$

3 Analysis of Two-Phase Sample Data

This section gives several results applicable to the analysis of data collected following a two-phase sample design. We assume that $\hat{Y}$ is given by the two-phase ratio estimator defined in Section 2.1. We also assume that either simple random sampling or stratified sampling is used in each phase. However, we do not assume that the sample design is model-based or near optimal.

Result 6:

Assume that a simple random sample design is used in each phase. The double-sampling ratio estimator $\hat{Y}$ is approximately unbiased for Y, with the approximate variance

$$\begin{aligned} AV(\hat{Y}) &= N^2\left(1-\frac{n_1}{N}\right)\frac{S_1^2}{n_1} \\ &+ N^2\left(1-\frac{n_2}{n_1}\right)\frac{S_2^2}{n_2} \end{aligned} \tag{20}$$

Here

$$\begin{aligned} B_2 &= Y/W = \sum_U y_k / \sum_U w_k \\ B_{2.1} &= Y/X = \sum_U y_k / \sum_U x_k \\ S_1^2 &= \sum_U \frac{(y_k - B_2 w_k)^2}{N} \\ S_2^2 &= \sum_U \frac{(y_k - B_{2.1} x_k)^2}{N} \end{aligned} \tag{21}$$

The variance can be estimated as

$$\begin{aligned} \hat{V}(\hat{Y}) &= N^2\left(1-\frac{n_1}{N}\right)\frac{s_1^2}{n_1} \\ &+ N^2\left(1-\frac{n_2}{n_1}\right)\frac{s_2^2}{n_2} \end{aligned} \tag{22}$$

where

$$
\begin{aligned}
b_2 &= \sum_{s_2} y_k \Big/ \sum_{s_2} w_k \\
b_{2.1} &= \sum_{s_2} y_k \Big/ \sum_{s_2} x_k \\
s_1^2 &= \sum_{s_2} \frac{(y_k - b_2 w_k)^2}{n_2} \\
s_2^2 &= \sum_{s_2} \frac{(y_k - b_{2.1} x_k)^2}{n_2}
\end{aligned} \tag{23}
$$

Result 7:

Assume that a stratified sample design has been followed in phase 1, and a simple random sample has been selected from each stratum of the phase 1 sample in phase 2. Then Result 6 can be used to evaluate the variance in each stratum, and the results can be totaled across all strata.

Alternatively, the estimated variance can be rewritten in terms of inclusion probabilities and sample residuals as follows:

$$
\begin{aligned}
\hat{V}(\hat{Y}) &= \sum_{s_2} \frac{1}{\pi_{2k}} \left(\frac{1}{\pi_{1k}} - 1 \right) e_{2k}^2 \\
&+ \sum_{s_2} \frac{1}{\pi_{2k}} \left(\frac{1}{\pi_{2k}} - \frac{1}{\pi_{1k}} \right) e_{2.1k}^2
\end{aligned} \tag{24}
$$

Here

$$
\begin{aligned}
b_2 &= \sum_{s_2} \pi_2^{-1} y_k \Big/ \sum_{s_2} \pi_2^{-1} w_k \\
b_{2.1} &= \sum_{s_2} \pi_2^{-1} y_k \Big/ \sum_{s_2} \pi_2^{-1} x_k \\
e_{2k} &= y_k - b_2 w_k \\
e_{2.1k} &= y_k - b_{2.1} x_k
\end{aligned} \tag{25}
$$

$$ \tag{26} $$

and π_{1k} and π_{2k} are defined as in Section 2.1.

Result 8:

Assuming the model of Section 2.2, the preceding result can be reconciled with Result 1 of Section 2.3. The first step is to combine equations (3) and (4) to obtain

$$\begin{aligned} y_k &= \beta_2 w_k + \epsilon_{1k} + \epsilon_{2.1k} \\ &= \beta_2 w_k + \epsilon_{2k} \end{aligned} \tag{27}$$

where we define $\epsilon_{2k} = \epsilon_{1k} + \epsilon_{2.1k}$. If, furthermore, we denote $sd(\epsilon_{2k}) = \sigma_{2k}$, then the assumption that ϵ_{1k} and $\epsilon_{2.1k}$ are independent implies that

$$\sigma_{2k}^2 = \sigma_{1k}^2 + \sigma_{2.1k}^2 \tag{28}$$

Using an approximation similar to the approximation used to develop Result 1, the sample residuals e_{2k} and $e_{2.1k}$ in (23) can be replaced by the random components ϵ_{2k} and $\epsilon_{2.1k}$ of equations (3) and (26). Taking the expected value under the sample design and under the model,

$$\begin{aligned} E\hat{V}(\hat{Y}) &= \sum_U \left(\frac{1}{\pi_{1k}} - 1\right) \sigma_{2k}^2 + \sum_U \left(\frac{1}{\pi_{2k}} - \frac{1}{\pi_{1k}}\right) \sigma_{2.1k}^2 \\ &= \sum_U \left(\frac{1}{\pi_{1k}} - 1\right) (\sigma_{1k}^2 + \sigma_{2.1k}^2) + \sum_U \left(\frac{1}{\pi_{2k}} - \frac{1}{\pi_{1k}}\right) \sigma_{2.1k}^2 \\ &= \sum_U \left(\frac{1}{\pi_{1k}} - 1\right) \sigma_{1k}^2 + \sum_U \left(\frac{1}{\pi_{2k}} - 1\right) \sigma_{2.1k}^2 \end{aligned} \tag{29}$$

which is equal to (8)

References

[1] Cochran, W.G., *Sampling Techniques,* Third Edition, John Wiley & Sons, 1977.

[2] Sarndal, C.E., Swensson, B., and Wretman, J., *Model Assisted Survey Sampling,* Springer-Verlag, 1992.

[3] Wright, R.L., "Estimating required sample sizes for model-assisted survey sampling," *Uses of Auxiliary Information in Surveys,* Statistics Sweden, October 5, 1992.

D

SITE DATA

Analysis Data by Site

TRACKING DATA BASE								
Case	Prog	End Use	Sq Feet	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
1	D	Light	21000	474	7000	29.85	51.00	208978
2	D	Light	25542	516	4700	37.56	43.10	176513
3	D	Light	330000	560	4000	11.20	10.90	44800
4	D	Light	52800	1500	3400	41.03	34.00	139500
5	D	Light	125434	1271	1800	65.07	16.30	117130
6	D	Light	6041	157	3400	3.00	2.49	10205
7	D	Light	100000	15	3400	0.26	0.21	870
8	D	Light	166000	1986	3400	33.43	27.70	113648
9	D	Light	0	176	1800	4.65	2.10	8376
10	D	Light	0	1143	4100	222.75	222.70	913257
11	D	Light	40000	50	4000	35.00	34.10	140000
12	D	Light	200000	4987	4000	81.24	79.30	324946
13	D	Light	46666	993	3400	28.85	23.90	98097
14	D	Light	12000	292	6742	11.33	16.00	76418
15	D	Light	0	996	3400	26.95	22.30	91632
16	D	Light	28000	941	7000	36.87	63.00	258092
17	D	VAC	1000	1	1200	0.95	0.70	1136
18	C	Light	34500	1	5610	24.01	24.0	134704
19	C	Light	144000	1963	3263	107.94	108.00	352209
20	C	Light	5000	83	6278	5.07	4.80	31823
21	C	Light	10271	62	1634	10.70	0.00	17485
22	C	Light	4000	35	5096	3.20	3.20	16313
23	C	Light	0	64	2860	3.39	3.40	9701
24	C	Light	20000	37	4017	3.58	2.80	14394
25	C	Light	2400	27	8760	2.05	2.1	17975
26	C	Light	315000	447	5595	27.61	18.90	154505
27	C	Light	23200	283	6504	23.73	23.70	154345
28	C	Light	12594	1	3614	15.38	15.40	55588
29	C	Light	3800	1	5460	2.96	2.90	16148
30	C	Light	8000	18	2340	2.52	2.50	5897
31	C	Light	10000	341	5075	17.10	15.50	87534
32	C	Light	320292	1	804	54.00	0.00	43440
33	C	Light	122000	30	6390	4.07	4.10	25995
34	C	Light	50000	35	4380	2.09	2.10	9173
35	C	Light	1500	13	3276	1.20	1.20	3935
36	C	Light	2500	14	2756	1.15	1.2	3159
37	C	Light	0	58	8760	4.85	4.90	42447
38	C	Light	0	126	8760	7.01	7.10	61364

	TRACKING DATA BASE							
		End	Sq	Number	Oper.	Del_Kw	System	System
Case	Prog	Use	Feet	Measures	Hours	Connected	Peak_Kw	Kwh
39	C	Light	0	1	4380	5.74	0.00	25132
40	C	Light	62500	692	4680	65.39	54.00	306044
41	C	Light	8000	12	8760	0.47	0.50	4100
42	C	Light	167415	120	4790	5.00	5.00	23948
43	C	Light	0	140	6440	10.17	4.80	65494
44	C	Light	1000	9	3307	0.55	0.60	1806
45	C	Light	16000	122	6037	12.50	12.50	75482
46	C	Light	3000	32	3127	2.75	2.80	8608
47	C	Light	0	241	5026	15.20	15.20	76395
48	C	Light	60000	1	5840	23.48	23.50	137123
49	C	Light	40415	105	2496	33.03	0.00	82447
50	C	Light	55080	668	8760	28.10	28.10	246156
51	C	Light	22000	1	5248	32.50	32.10	170560
52	C	Light	2000	1	3016	1.50	1.5	4518
53	C	Light	5200	75	5420	4.16	3.50	22529
54	C	VAC	275000	1	1200	145.29	49.00	174345
55	C	VAC	35000	4	1203	83.21	0.00	100098
56	C	VAC	100000		300	25.35	47.40	80184
57	C	VAC	91000	10	644	445.82	0.00	287107
58	C	VAC	400000	1	2282	179.00	0.00	408429
59	C	VAC	219540	11	1396	236.00	0.00	329461
60	C	VAC	44000	1	2340	16.48	0.00	163263
61	C	VAC	1500000	1	8760	189.87	155.00	1663293
62	C	VAC	7952	180	572	19.15	6.50	10951
63	D	Light	8000	78	4700	6.37	7.30	30030
64	D	Light	1200	21	1800	6.49	2.80	11676
65	D	Light	5000	19	4000	1.32	1.20	5262
66	D	Light	61315	21	4000	1.83	1.80	7308
67	D	Light	10000	24	3400	1.94	1.60	6600
68	D	Light	2600	24	4800	2.69	1.90	12912
69	D	Light	100000	4340	4700	56.22	64.50	264245
70	D	Light	0	94	4700	5.38	6.10	25274
71	D	Light	66000	10	4700	5.96	6.80	28000
72	D	Light	30000	90	1800	10.80	7.74	19440
73	D	Light	2000	6	4000	0.41	0.40	1650
74	D	Light	0	1	3400	15.19	12.60	51640
75	D	Light	30000	79	4700	47.06	54.00	221200
76	D	Light	1500	12	7000	1.78	3.00	12425

TRACKING DATA BASE

Case	Prog	End Use	Sq Feet	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
77	D	Light	3600	8	4700	3.37	3.90	15817
78	D	Light	2000	46	4700	2.66	2.90	12521
79	D	Light	75000	296	4000	5.38	5.10	21522
80	D	Light	20045	454	1800	7.29	3.20	13120
81	D	Light	0	34	4700	7.75	8.90	36448
82	D	Light	205708	3106	3400	215.80	178.80	733722
83	D	Light	70000	2840	3400	18.38	15.20	62480
84	D	Light	150000	1992	3400	136.23	112.90	463172
85	D	Light	100000	71	4000	49.70	48.60	198800
86	D	Light	18455	26	4000	0.83	0.80	3328
87	D	Light	33950	240	4000	54.72	53.40	218880
88	D	Light	2000	38	4700	1.83	2.10	8580
89	D	Light	30000	179	4100	12.22	12.20	50120
90	D	Light	4000	10	4000	2.61	2.50	10452
91	D	Light	1500	19	3400	0.70	0.40	2383
92	D	Light	3399	30	4000	6.05	6.00	24202
93	D	Light	38000	32	4700	4.62	5.30	21706
94	D	Light	20000	5	1800	5.84	2.50	10513
95	D	Light	1700	52	3400	0.64	0.90	2160
96	D	Light	5000	79	4700	2.53	2.90	11909
97	D	Light	0	175	4000	16.93	16.50	67732
98	D	Light	0	293	4700	7.01	8.10	32968
99	D	Light	3000	16	4700	0.48	0.50	2240
100	D	Light	6427	90	4700	4.52	5.20	21231
101	D	Light	5800	138	3400	2.21	1.80	7498
102	D	Light	8000	202	4700	7.05	8.10	33128
103	D	Light	3000	30	4700	1.75	2.00	8220
104	D	Light	30000	34	4000	2.96	2.90	11832
105	D	Light	6515	330	3400	11.52	9.50	39160
106	D	Light	1500	14	4700	9.80	11.20	46060
107	D	Light	64473	51	3400	1.46	1.20	4965
108	D	Light	0	245	3400	16.69	10.50	43146
109	D	Light	6600	361	4700	11.51	13.20	54109
110	D	Light	30000	48	4000	33.60	32.80	134400
111	D	Light	0	234	4000	26.24	12.00	104956
112	D	Light	55000	82	4000	3.69	3.60	14760
113	D	Light	24157	1765	4700	12.29	14.10	57749
114	D	Light	8600	23	3400	1.04	0.90	3519

	TRACKING DATA BASE							
		End	Sq	Number	Oper.	Del_Kw	System	System
Case	Prog	Use	Feet	Measures	Hours	Connected	Peak_Kw	Kwh
115	D	Light	28000	432	3400	9.66	8.00	32832
116	D	Light	375000	100	4000	6.88	6.70	27500
117	D	Light	1500	18	3400	0.45	0.40	1527
118	D	Light	0	189	5767	3.84	4.40	22171
119	D	Light	2800	68	4800	3.03	2.90	14554
120	D	Light	3500	24	4000	2.09	2.00	8352
121	D	Light	0	16	4000	1.15	1.10	4608
122	D	Light	9000	29	3400	0.55	0.50	1885
123	D	Light	0	27	4700	0.88	1.00	4131
124	D	Light	22526	10	4700	0.33	0.40	1530
125	D	Light	2600	10	4000	8.23	8.00	32900
126	D	Light	100000	1000	4000	45.00	43.90	180000
127	D	Light	0	39	3400	32.73	27.10	111274
128	D	Light	250000	40	4000	23.59	23.00	94360
129	D	Light	25956	16	4000	1.18	1.20	4736
130	D	Light	15365	558	4700	32.65	37.40	153450
131	D	Light	40000	30	4700	21.00	24.10	98700
132	D	Light	69700	35	3400	3.51	3.00	11938
133	D	Light	9000	53	3400	2.08	10.90	44977
134	D	Light	3000	12	3068	1.38	1.40	4243
135	D	Light	750	10	4000	2.28	2.20	9120
136	D	Light	0	2	280	2.64	0.20	740
137	D	Light	0	25	4700	15.27	17.50	71750
138	D	Light	162226	1086	1880	81.22	37.10	152699
139	D	Light	4100	285	4800	5.43	6.20	26075
140	D	Light	6000	1	1248	0.35	0.10	436
141	D	Light	6000	12	3400	1.23	1.00	4194
142	D	Light	2850	32	4700	1.87	2.10	8768
143	D	Refrig	0	4	5000	0.72	0.30	3576
144	D	Refrig	15000	7	4380	2.69	2.00	11803
145	D	Refrig	1200	20	5000	1.38	0.60	6900
146	D	Refrig	43700	1434	8660	8.04	5.60	69637
147	D	VAC	0	1	1200	11.69	8.30	14121
148	D	VAC	24000	8	1200	2.54	1.80	3011
149	D	VAC	6000	2	800	32.91	15.60	26332
150	D	VAC	1560	13	1200	1.62	1.20	1947
151	D	VAC	34000	6	180	136.43	14.50	24558
152	D	VAC	3000	3	180	68.22	7.30	12279

	TRACKING DATA BASE							
Case	Prog	End Use	Sq Feet	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
153	D	VAC	1200	1	1200	0.75	0.50	900
154	D	VAC	20000	17	1200	2.18	1.50	2611
155	D	VAC	5000	60	1200	30.00	21.30	36000
156	D	VAC	0	4	900	15.68	8.20	14115
157	D	VAC	2000	2	1200	4.12	2.90	4943
158	D	VAC	0	1	1200	0.81	0.60	973
159	D	VAC	1800	2	1200	5.30	5.30	8943
160	D	VAC	1500	1	1200	0.40	0.30	480
161	D	VAC	2700	1	1200	3.78	2.70	4534
162	D	VAC	0	1	1200	6.64	4.70	7963
163	D	VAC	4500	5	1200	0.41	0.40	695
164	D	VAC	2000	11	1200	1.26	0.90	1511
165	D	VAC	0	4	1200	0.35	0.20	420
166	D	VAC	1200	5	1200	0.68	0.50	811
167	D	VAC	863	32	4800	0.41	1.20	1947
168	D	VAC	1479	15	1200	1.89	1.30	2271
169	D	VAC	4100	2	1200	6.82	4.80	8186
170	D	VAC	2850	3	1200	10.23	7.30	12279
171	C	Light	18000	93	2920	5.58	0.00	16249
172	C	Light	46378	367	5808	38.17	38.90	221680
173	C	Light	8500	70	4344	6.16	6.20	26759
174	C	Light	2500	27	5460	1.84	1.80	10057
175	C	Light	25000	300	4751	15.00	15.00	71264
176	C	Light	7000	7	8760	1.61	1.50	14112
177	C	Light	2575	28	3712	2.40	2.40	8908
178	C	Light	20000	369	3874	51.13	42.10	198096
179	C	Light	2400	31	8736	1.78	1.70	15576
180	C	Light	6000	240	5475	10.04	6.90	54984
181	C	Light	1200	14	8264	0.75	0.80	6192
182	C	Light	16000	42	5824	4.14	4.20	24088
183	C	Light	8000	188	3638	12.40	12.40	45105
184	C	Light	57609	616	3050	48.04	50.50	146516
185	C	Light	18000	82	8760	4.63	4.70	40577
186	C	Light	34969	265	5419	81.76	66.10	443084
187	C	Light	2000	25	4693	1.80	1.80	8447
188	C	Light	4500	69	4190	5.95	6.00	24930
189	C	Light	20000	196	3120	13.02	9.50	40621
190	C	Light	44700	89	7313	6.90	6.90	50459

Analysis Data by Site

	TRACKING DATA BASE								
		End	Sq	Number	Oper.	Del_Kw	System	System	
Case	Prog	Use	Feet	Measures	Hours	Connected	Peak_Kw	Kwh	
191	C	Light	0	1515	5296	82.00	82.00	434288	
192	C	Light	3000	100	6205	6.68	6.00	41472	
193	C	Light	5000	47	4368	4.04	4.00	17625	
194	C	Light	20000	133	4679	10.80	10.80	50537	
195	C	Light	9500	21	5840	0.55	0.00	3211	
196	C	Light	12000	102	8200	9.85	9.90	80803	
197	C	Light	900	7	2805	0.60	0.60	1689	
198	C	Light	30000	356	7288	33.20	33.20	241962	
199	C	Light	85000	214	6226	23.50	23.50	146312	
200	C	Light	0	100	4680	10.30	10.80	48204	
201	C	Light	0	5	4380	3.00	0.00	13140	
202	C	Light	7500	28	6032	3.78	3.80	22777	
203	C	Light	107000	43	4015	8.87	0.00	35613	
204	C	Light	0	136	3120	14.53	14.60	45342	
205	C	Light	3000	15	3311	1.30	1.30	4305	
206	C	Light	1850	22	2826	1.89	1.90	5347	
207	C	Light	0	21	5460	2.40	1.90	13082	
208	C	Light	0	1	4380	1.28	0.00	5606	
209	C	Light	1500	23	2600	1.02	0.20	2649	
210	C	Light	3500	31	3432	2.45	2.40	8402	
211	C	Light	0	6	5096	0.59	0.60	3018	
212	C	Light	1100	26	4576	0.86	1.00	3935	
213	C	Light	7128	54	6188	2.24	0.00	13845	
214	C	Light	1000	9	4700	0.77	0.80	3603	
215	C	Light	25000	250	3400	38.08	36.10	129460	
216	C	Light	2200	66	3484	3.05	2.90	10628	
217	C	Light	12000	523	5400	6.88	6.70	37125	
218	C	Light	18000	346	3744	30.43	22.80	113914	
219	C	Light	0	144	4866	1.95	0.00	9500	
220	C	Light	1000	50	5460	3.59	0.90	19597	
221	C	Light	2000	6	3380	1.13	1.20	3834	
222	C	Light	0	36	4576	3.86	3.60	17666	
223	C	Light	23435	223	4264	25.32	21.80	107984	
224	C	Light	5200	62	4420	4.88	5.40	21584	
225	C	Light	28500	733	8760	15.58	22.20	136488	
226	C	Light	2500	46	3120	2.12	2.10	6618	
227	C	Light	1300	20	3120	1.10	0.90	3430	
228	C	Light	1550	18	5096	0.73	0.80	3709	

TRACKING DATA BASE								
Case	Prog	End Use	Sq Feet	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
229	C	Light	30000	113	8760	5.98	6.00	52420
230	C	Light	44404	369	8760	9.96	5.20	87220
231	C	Light	863	11	6650	1.59	1.60	10595
232	C	Light	1400	42	5200	1.81	1.80	9391
233	C	Light	2700	72	2744	4.48	4.40	12304
234	C	Light	600	22	3276	0.68	0.80	2228
235	C	Light	45000	53	5460	4.85	3.90	26481
236	C	Light	77000	214	6120	2.73	1.90	16705
237	C	Light	24169	38	3400	1.18	0.40	4003
238	C	Light	2800	30	4800	1.59	1.70	7630
239	C	Light	0	50	8760	1.59	1.60	13972
240	C	Light	3000	87	7000	9.52	11.00	66623
241	C	Light	0	48	3640	1.97	2.00	7158
242	C	Light	250000	324	8760	14.91	14.60	130594
243	C	Light	4000	60	3400	4.59	5.20	15604
244	C	Light	5200	181	4000	7.67	7.50	30698
245	C	Light	2000	30	3640	1.29	1.20	4696
246	C	Light	15000	91	3120	4.84	5.00	15102
247	C	Light	2000	24	4368	3.04	0.00	13284
248	C	Light	1400	14	2468	1.43	1.20	3518
249	C	Light	20263	533	5759	11.60	11.6	66805
250	C	Light	2000	36	6188	1.54	1.50	9554
251	C	Light	0	41	3400	3.27	0.00	11126
252	C	Light	1630	15	2340	1.67	1.30	3906
253	C	Light	1100	13	5252	0.74	0.70	3866
254	C	Light	10000	206	4000	29.38	22.40	117526
255	C	Light	3600	24	8760	1.08	1.10	9460
256	C	Light	5310	52	3640	4.25	4.30	15477
257	C	Light	6500	126	6205	7.74	7.7	48007
258	C	Light	0	22	3588	2.31	2.00	8271
259	C	Light	47120	510	4965	27.30	19.80	135537
260	C	Light	2500	22	3952	2.20	2.30	8689
261	C	Light	0	10	2808	0.86	0.90	2415
262	C	Light	12000	6	6534	0.36	0.40	2352
263	C	Light	150773	1	2280	37.50	0.00	85500
264	C	Light	10000	4	2614	1.23	0.00	2240
265	C	Light	10000	85	6190	5.40	5.40	33402
266	C	Light	72080	719	3400	41.16	34.00	139927

	TRACKING DATA BASE							
		End	Sq	Number	Oper.	Del_Kw	System	System
Case	Prog	Use	Feet	Measures	Hours	Connected	Peak_Kw	Kwh
267	C	Light	47120	676	4000	28.95	23.00	115815
268	C	Refrig	6000	2	2700	4.42	0.00	11925
269	C	Refrig	6000	64	4732	2.78	0.00	13155
270	C	Refrig	2000	2	4004	7.29	0.00	29184
271	C	Refrig	35125	3	3650	14.11	0.00	51497
272	C	Refrig	12000	13	8760	2.99	0.00	26180
273	C	Refrig	6000	1	2920	5.96	0.0	17395
274	C	VAC	174110	1	336	1140.00	0.00	383040
275	C	VAC	1116719	1	3400	33.82	0.00	114975
276	C	VAC	40000	1	3636	0.00	0.00	1
277	C	VAC	150203	1	5000	61.39	0.00	306935
278	C	VAC	83000	1	1200	46.76	0.00	56109
279	C	VAC	21300	1	1200	123.67	0.00	148409
280	C	VAC	3900	3	6205	5.63	0.00	34937
281	C	VAC	125000	1	1200	320.00	0.00	383998
282	C	VAC	3000	1	5975	1.81	0.0	10819

	HSEM DATA BASE								
	PRE TO POST					STD TO POST			
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
1	460	4581	13.07	10.32	59855	4581	9.28	7.33	42499
2	516	4949	26.94	21.28	128351	4107	22.35	17.66	106501
3	560	8760	16.23	16.23	142201	8760	12.55	12.55	109947
4	1500	4598	64.00	50.56	294260	4598	55.00	43.45	252880
5	1271	1275	86.62	0.00	110465	945	71.55	0.00	67646
6	156	2311	14.79	11.68	34183	2311	16.16	12.77	37349
7	15	6664	0.34	0.27	2279	6664	0.20	0.16	1333
8	1952	3388	71.03	56.11	240614	3388	54.51	43.07	184668
9	176	2637	2.76	0.67	7268	2637	2.76	0.67	7268
10	1116	8350	64.43	64.43	538016	7169	50.43	50.43	429485
11	50	2849	6.93	5.47	19732	2849	7.34	5.80	20903
12	4987	3926	108.64	85.82	426487	3926	99.53	78.63	390731
13	937	2957	25.21	19.91	74535	2957	25.21	19.91	74535
14	292	4513	11.30	8.93	51011	4509	11.42	9.03	51509
15	996	3570	45.82	36.19	163545	3570	26.39	20.85	94216
16	941	6794	22.91	17.53	155627	6794	22.91	17.53	155627
17	1	1200	0.00	0.00	0	1200	0.70	0.70	840
18	1	4581	14.58	11.52	66781	4581	11.72	9.26	53698
19	1778	3388	16.38	12.94	55496	3388	16.38	12.94	55496
20	93	3645	5.36	4.23	19518	3645	5.36	4.23	19518
21	62	4451	9.27	9.11	41261	4451	9.27	9.11	41261
22	35	4507	3.12	2.46	14054	4507	3.12	2.46	14054
23	64	3388	4.61	3.64	15610	3388	4.61	3.64	15610
24	37	3753	2.95	2.33	11074	3753	2.69	2.12	10076
25	27	6920	1.67	1.32	11585	6920	1.32	1.05	9156
26	466	2649	25.68	20.29	68042	2649	25.68	20.29	68042
27	275	5047	35.99	28.43	181636	5047	35.99	28.43	181636
28	1	2960	11.71	9.25	34644	2960	6.31	4.99	18679
29	1	4654	2.74	2.16	12733	4654	2.32	1.83	10797
30	19	2491	1.81	1.43	4496	2491	1.81	1.43	4496
31	279	4616	10.91	8.62	50352	4616	10.91	8.62	50352
32	1	787	43.80	6.54	34490	787	43.80	6.54	34490
33	30	2311	4.01	3.17	9268	2311	4.01	3.17	9268
34	35	4015	2.13	0.00	8532	4015	2.13	0.00	8532
35	13	2706	1.20	0.95	3236	2706	0.96	0.76	2603
36	15	3634	1.41	1.12	5139	3634	1.41	1.12	5139
37	57	4654	4.37	1.97	20328	4654	4.37	1.97	20328
38	119	6920	6.31	4.98	43633	6920	6.31	4.98	43633

	HSEM DATA BASE								
	PRE TO POST					STD TO POST			
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
39	1	4015	4.10	0.00	16478	4015	4.10	0.00	16478
40	692	3029	25.56	20.20	77428	3029	25.56	20.20	77428
41	12	6576	0.34	0.25	2210	6576	0.34	0.25	2210
42	108	4321	4.78	3.78	20673	4321	4.78	3.78	20673
43	147	4654	11.15	5.02	51880	4654	11.15	5.02	51880
44	3	3191	0.47	0.37	1484	3191	0.47	0.37	1484
45	122	5074	11.00	8.69	55817	5074	11.00	8.69	55817
46	32	2301	3.40	2.41	7822	2301	3.40	2.41	7822
47	278	2849	12.51	9.88	35646	2849	12.51	9.88	35646
48	1	4451	29.79	23.53	132573	4451	26.06	20.59	115989
49	105	942	31.05	4.91	29238	942	31.05	4.91	29238
50	611	6773	6.14	4.75	41605	6773	6.14	4.75	41605
51	1	4832	30.52	24.11	126647	4074	25.73	20.33	106777
52	15	3185	1.11	0.88	3536	3185	1.11	0.88	3536
53	75	4174	2.18	1.72	9079	4174	2.18	1.72	9079
54	1	1200	16.00	16.00	19200	1200	16.00	16.00	19200
55	4	180	216.00	0.00	38880	180	216.00	0.00	38880
56	1	240	240.00	0.00	57600	240	240.00	0.00	57600
57	10	544	283.00	0.00	153864	544	283.00	0.00	153864
58	1	1363	513.39	0.00	699856	1030	513.39	0.00	528644
59	12	1057	1226.04	86.63	1295498	1057	1226.04	86.63	1295498
60	1	360	124.30	0.00	44748	360	124.30	0.00	44748
61	1	2359	454.00	78.80	1070990	2359	454.00	78.80	1070990
62	180	562	19.50	0.00	10951	562	19.50	0.00	10951
63	75	3316	6.9	5.4	22881	3329	5.9	4.7	19643
64	21	2515	4.2	3.3	10565	2474	3.7	2.9	9153
65	14	3781	0.9	0.2	3403	3732	0.9	0.2	3359
66	21	6976	2.0	1.6	13952	6976	2.0	1.6	13952
67	19	4182	0.6	0.4	2509	4182	0.6	0.4	2509
68	24	3111	1.6	0.2	4978	3111	1.6	0.2	4978
69	.4106	4254	109.3	86.3	464947	4254	109.3	86.3	464947
70	97	2186	11.8	6.1	25800	2190	10.1	5.2	22117
71	10	3018	-10.8	-8.5	-32599	2976	5.4	4.2	16073
72	92	1630	12.2	0.1	19883	1632	10.4	0.1	16974
73	12	1190	0.1	0.1	119	1190	0.1	0.1	119
74	1	3247	21.0	16.5	68183	3268	15.1	11.9	49344
75	79	4360	1.1	0.9	4796	4360	1.1	0.9	4796
76	12	2688	2.3	0.8	6183	2688	2.3	0.8	6183

	HSEM DATA BASE								
	PRE TO POST					STD TO POST			
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
77	7	2682	4.0	3.2	10728	2682	4.0	3.2	10728
78	46	3371	4.0	3.2	13484	3371	4.0	3.2	13484
79	292	4492	13.1	7.8	58847	4420	12.5	7.3	55250
80	469	1950	16.1	1.6	31392	1941	11.8	1.1	22905
81	36	4434	3.5	2.7	15519	4434	3.5	2.7	15519
82	3106	2670	344.5	272.1	919702	2670	292.9	231.3	781987
83	2204	2842	17.3	13.7	49161	2848	12.3	9.8	35031
84	1992	3022	244.0	192.8	737433	3044	211.3	167.0	643189
85	71		0.0	0.0	0		0.0	0.0	0
86	20	2980	0.2	0.1	596	2980	0.2	0.1	596
87	240	4613	4.6	2.1	21221	4613	4.6	2.1	21221
88	38	3035	-0.2	-0.2	-607	3181	2.7	2.1	8588
89	179	2761	16.1	6.8	44453	2761	16.1	6.8	44453
90	10	4380	-0.2	0.0	-876	4140	-0.2	0.0	-828
91	19	3010	0.8	0.6	2408	3010	0.8	0.6	2408
92	29	1754	1.0	0.2	1754	1754	1.0	0.2	1754
93	64	113	14.4	0.0	1633	113	14.4	0.0	1633
94	5	3680	0.3	0.0	1104	3680	0.3	0.0	1104
95	16	2720	1.7	1.3	4624	2769	1.6	1.3	4431
96	79	2352	3.2	2.5	7527	2380	2.7	2.1	6426
97	175	4781	9.1	2.9	43506	4781	9.1	2.9	43506
98	293	3022	12.2	8.0	36865	3027	10.3	6.8	31180
99	16	2371	1.4	0.8	3320	2371	1.4	0.8	3320
100	98	3850	5.2	4.2	20019	3835	4.4	3.5	16874
101	96	2646	3.1	2.4	8203	2711	2.6	2.1	7049
102	202	4153	12.3	9.7	51085	4142	10.5	8.2	43486
103	36	4075	3.8	3.0	15485	4075	3.8	3.0	15485
104	26	3625	2.6	2.0	9424	3625	2.6	2.0	9424
105	330	2948	14.1	11.1	41566	2952	12.1	9.5	35721
106	14	3445	-1.5	-1.2	-5168	3341	-1.8	-1.4	-6014
107	55	3435	0.4	0.4	1374	3435	0.4	0.4	1374
108	254	2967	14.0	11.1	41535	2967	14.0	11.1	41535
109	372	3902	12.5	9.8	48773	3898	10.9	8.6	42492
110	48	2368	5.2	0.5	12315	2368	5.2	0.5	12315
111	280	5811	18.2	5.8	105765	5811	18.2	5.8	105765
112	74	5361	5.6	4.4	30022	5361	5.6	4.4	30022
113	1660	3991	17.7	13.9	70633	3991	17.7	13.9	70633
114	14	5122	2.1	0.0	10756	5122	2.1	0.0	10756

	HSEM DATA BASE								
	PRE TO POST					STD TO POST			
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
115	448	6683	11.3	8.9	75520	6683	11.3	8.9	75520
116	86	1280	2.6	0.7	3329	1280	2.6	0.7	3329
117	16	2268	0.5	0.3	1134	2268	0.5	0.3	1134
118	191	5475	4.3	1.9	23544	5475	4.3	1.9	23544
119	65	4893	2.3	1.8	11253	4893	2.3	1.8	11253
120	24	4336	1.0	0.0	4336	4336	1.0	0.0	4336
121	16	3084	0.5	0.4	1542	3087	0.3	0.2	926
122	29	2746	0.5	0.5	1373	2410	0.3	0.3	723
123	28	4005	1.6	0.0	6408	4005	1.6	0.0	6408
124	11	5012	0.6	0.3	3007	5012	0.6	0.3	3007
125	10	2914	-0.8	-0.7	-2331	2914	-0.8	-0.7	-2331
126	1000	3387	51.0	20.0	172755	3387	51.0	20.0	172755
127	38	4019	9.3	0.0	37379	4019	9.3	0.0	37379
128	40	861	306.9	37.7	264310	861	306.9	37.7	264310
129	16	4068	4.5	0.0	18308	4068	4.5	0.0	18308
130	168	3149	3.3	2.4	10391	3149	3.3	2.4	10391
131	28	3525	38.1	35.4	134309	3525	38.1	35.4	134309
132	35	668	56.4	1.1	37673	668	56.4	1.1	37673
133	53	1010	25.6	2.9	25850	1010	25.6	2.9	25850
134	12	6796	1.3	1.0	8835	6796	1.3	1.0	8835
135	12	3488	0.5	0.4	1744	3488	0.5	0.4	1744
136	2	0	0.0	0.0	0	0	0.0	0.0	0
137	25	4330	-21.5	0.0	-93128	4330	-21.5	0.0	-93128
138	1086	1785	82.0	1.9	146351	1785	82.0	1.9	146351
139	205	5201	3.8	2.9	19764	5062	3.6	2.7	18222
140	1	0	0.0	0.0	0	0	0.0	0.0	0
141	12	5805	0.4	0.0	2322	5805	0.4	0.0	2322
142	27	4930	1.2	1.0	5916	4930	1.2	1.0	5916
143	4	5363	0.4	0.3	2145	5363	0.4	0.3	2145
144	7	167	127.3	0.0	21271	167	127.3	0.0	21271
145	20	6028	5.2	4.6	31225	6028	5.2	4.6	31225
146	1434	82	380.4	0.0	31240	0	0.0	0.0	0
147	1	62	62.0	6.1	3865	62	62.0	6.1	3865
148	8	758	137.9	57.6	104503	159	31.0	7.2	4939
149	2	202	74.5	0.0	15064	202	74.5	0.0	15064
150	26	454	5.3	2.3	2406	198	3.1	0.7	614
151	6	139	82.7	0.0	11531	139	82.7	0.0	11531
152	3	-32	26.1	0.8	-834	-32	26.1	0.8	-834

	HSEM DATA BASE								
	PRE TO POST					STD TO POST			
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
153	1	1475	2.5	3.8	3688	1300	0.6	0.8	780
154	17	1339	-9.2	-9.0	-12319	1348	8.3	8.4	11191
155	40	952	3.3	3.3	3140	952	3.3	3.3	3140
156	2	-760	-19.3	1.8	14661	7146	2.2	1.8	15722
157	2	231	-7.6	1.7	-1755	-817	0.9	0.4	-735
158	1	-182	-7.8	0.3	1418	1200	0.8	0.3	900
159	2	-402	-3.4	0.3	1368	1284	0.8	1.2	1027
160	1	4905	0.4	3.1	1962	621	1.4	1.4	870
161	1	3648	0.5	2.7	1824	3888	0.5	2.9	1944
162	1	1545	5.9	6.0	9116	1180	2.2	1.6	2597
163	2	-661	-5.4	3.0	3572	627	0.6	0.3	376
164	11	1965	1.1	2.5	2162	821	3.4	3.2	2792
165	4	-1101	-5.0	3.5	5503	920	2.8	2.1	2577
166	5	770	1.3	1.0	1001	715	0.2	0.1	143
167	32	84	6.3	0.0	530	84	6.3	0.0	530
168	15	-760	-1.7	1.3	1292	-130	-0.3	0.1	39
169	2	1	37.5	0.0	47	1	37.5	0.0	47
170	2	4	35.5	0.0	132	4	35.5	0.0	132
171	80	2555	4.8	0.0	12264	2555	4.8	0.0	12264
172	367	4415	48.1	38.0	212372	4415	48.1	38.0	212372
173	64	3233	5.5	4.3	17783	3233	5.5	4.3	17783
174	16	4460	1.3	1.0	5798	4460	1.3	1.0	5798
175	300	5472	11.8	9.2	64573	5460	10.8	8.5	58973
176	7	1667	4.7	1.8	7837	1667	4.7	1.8	7837
177	28	4537	2.3	1.2	10435	4537	2.3	1.2	10435
178	337	1848	43.0	17.7	79455	6009	7.0	12.2	42062
179	25	6889	1.8	1.4	12401	7167	1.4	1.1	10034
180	142	3852	9.3	3.5	35824	3852	9.3	3.5	35824
181	13	6931	0.8	0.7	5545	6931	0.8	0.7	5545
182	55	5153	5.2	4.2	26795	5153	5.2	4.2	26795
183	97	3277	7.6	6.0	24904	3298	6.0	4.8	19785
184	411	3032	26.1	20.6	79128	3032	26.1	20.6	79128
185	88	2978	6.9	5.5	20545	2978	6.9	5.5	20545
186	265	4543	64.3	50.8	292115	4543	64.3	50.8	292115
187	24	4770	2.1	1.6	10017	4782	1.8	1.4	8608
188	69	5131	6.5	4.9	33351	5131	6.5	4.9	33351
189	196	3385	10.0	7.8	33849	3385	10.0	7.8	33849
190	80	3545	5.4	4.2	19145	3533	4.1	3.2	14484

	HSEM DATA BASE								
	PRE TO POST					STD TO POST			
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
191	1626	3459	106.2	83.9	367367	3475	84.5	66.7	293619
192	100	1129	6.0	1.4	6774	1129	6.0	1.4	6774
193	47	3606	4.0	3.1	14425	3606	4.0	3.1	14425
194	120	4050	9.4	7.4	38068	4050	9.4	7.4	38068
195	21	4033	2.2	0.0	8873	4033	2.2	0.0	8873
196	105	5647	9.7	7.7	54778	5642	8.4	6.7	47392
197	7	2672	0.6	0.6	1603	2580	0.5	0.5	1290
198	701	5772	22.5	17.8	129861	5763	17.7	14.0	102005
199	214	6037	9.9	7.3	59762	6066	9.7	7.2	58845
200	80	5371	5.9	5.5	31691	5482	4.6	4.4	25218
201	5	3950	3.1	0.0	12246	3950	3.1	0.0	12246
202	28	3270	3.7	3.0	12099	3270	3.7	3.0	12099
203	43	5989	5.8	0.0	34615	5989	5.8	0.0	34615
204	163	2323	13.9	11.0	32296	2315	11.9	9.4	27545
205	16	2448	1.4	1.1	3427	2448	1.4	1.1	3427
206	26	2716	1.7	1.4	4618	2414	1.4	1.1	3379
207	21	4593	2.1	1.7	9645	4654	1.8	1.4	8378
208	69	4756	17.9	0.0	85131	4756	17.9	0.0	85131
209	20	2510	1.0	0.8	2510	2510	1.0	0.8	2510
210	31	2801	2.1	1.7	5883	2801	2.1	1.7	5883
211	6	5010	0.5	0.4	2505	4408	0.5	0.4	2204
212	13	4513	1.0	0.8	4513	4513	1.0	0.8	4513
213	54	3886	3.0	2.4	11659	3886	3.0	2.4	11659
214	8	3378	0.8	0.7	2702	3157	0.7	0.6	2210
215	208	1818	12.8	10.1	23269	1818	12.8	10.1	23269
216	66	3337	2.9	2.3	9676	3337	2.9	2.3	9676
217	523	3321	25.3	20.0	84013	3314	23.9	18.8	79213
218	346	2996	36.3	25.8	108747	3025	32.1	22.5	97092
219	144	4742	7.1	0.0	33666	4742	7.1	0.0	33666
220	50	4593	2.5	2.0	11482	4750	2.1	1.7	9974
221	6	4050	1.0	0.8	4050	4050	1.0	0.8	4050
222	36	4576	2.9	2.6	13059	4576	2.9	2.6	13059
223	223	3020	19.3	15.3	58294	3018	15.6	12.4	47084
224	61	2668	5.8	4.6	15473	2668	5.8	4.6	15473
225	658	5124	25.4	13.3	130159	5137	25.5	13.4	131005
226	52	2588	2.5	1.9	6469	2588	2.5	1.9	6469
227	18	5090	0.1	0.1	509	5090	0.1	0.1	509
228	24	3666	0.9	0.7	3299	3666	0.9	0.7	3299

	HSEM DATA BASE									
	PRE TO POST					STD TO POST				
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
229	113	6364	5.8	4.5	36912	6438	4.3	3.4	27684
230	57	3933	6.0	1.6	23598	3933	6.0	1.6	23598
231	11	6978	1.2	1.0	8374	6978	1.2	1.0	8374
232	44	4915	1.9	1.5	9338	4915	1.9	1.5	9338
233	72	2816	4.1	3.2	11547	2816	4.1	3.2	11547
234	22	3071	0.9	0.7	2764	3071	0.9	0.7	2764
235	53	3220	4.1	2.2	13204	3220	4.1	2.2	13204
236	214	6059	3.9	0.9	23629	6059	3.9	0.9	23629
237	36	7754	1.4	1.3	10856	7754	1.4	1.3	10856
238	30	3933	1.6	1.2	6292	3933	1.6	1.2	6292
239	47	3917	1.5	1.1	5875	3917	1.5	1.1	5875
240	87	4726	8.8	7.0	41585	4726	8.8	7.0	41585
241	48	3503	2.0	1.6	7006	3503	2.0	1.6	7006
242	324	6803	11.2	8.9	76190	6803	11.2	8.9	76190
243	52	2156	1.9	1.2	4097	2156	1.9	1.2	4097
244	181	3349	4.3	3.5	14400	3349	4.3	3.5	14400
245	30	7284	1.1	0.9	8012	7284	1.1	0.9	8012
246	88	4136	8.2	8.6	33867	4694	5.3	6.4	24932
247	24	3942	3.3	0.0	13009	3942	3.3	0.0	13009
248	14	2238	1.6	1.3	3581	2167	1.4	1.1	3034
249	533	4311	14.2	11.2	61221	4311	14.2	11.2	61221
250	36	5215	1.9	1.5	9908	5215	1.9	1.5	9908
251	41	3400	5.1	2.1	17317	3400	5.1	2.1	17317
252	15	2998	1.5	1.2	4497	2998	1.5	1.2	4497
253	15	4653	1.2	1.1	5584	4653	1.2	1.1	5584
254	208	3032	22.7	17.9	68817	3018	19.6	15.4	59153
255	24	4734	2.3	1.1	10889	4734	2.3	1.1	10889
256	52	2961	3.7	2.9	10954	2961	3.7	2.9	10954
257	78	4894	8.6	6.8	42090	4894	8.6	6.8	42090
258	22	2998	1.6	1.3	4797	2998	1.6	1.3	4797
259	510	3185	23.4	18.6	74532	3185	23.4	18.6	74532
260	29	3444	1.7	1.4	5854	3444	1.7	1.4	5854
261	10	2390	1.0	0.8	2390	2403	0.8	0.6	1922
262	6	2803	0.4	0.3	1121	2803	0.4	0.3	1121
263	1	0	0.0	0.0	0	0	0.0	0.0	0
264	4	146	9.7	0.3	1414	146	9.7	0.3	1414
265	74	3270	5.0	3.8	16348	3270	5.0	3.8	16348
266	719	3205	33.0	26.2	105764	3205	33.0	26.2	105764

Analysis Data by Site

	HSEM DATA BASE								
	PRE TO POST					STD TO POST			
Case	Number Measures	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh	Oper. Hours	Del_Kw Connected	System Peak_Kw	System Kwh
267	676	3197	31.1	24.6	99439	3197	31.1	24.6	99439
268	2	73	80.5	0.0	5894	73	80.5	0.0	5894
269	64	36	69.5	0.0	2498	36	69.5	0.0	2498
270	2	86	90.5	0.0	7807	86	90.5	0.0	7807
271	3	49	330.1	0.0	16053	49	330.1	0.0	16053
272	14	88	26.0	0.0	2287	88	26.0	0.0	2287
273	1	79	67.7	0.0	5357	79	67.7	0.0	5357
274	1	51	477.6	31.4	24352	51	477.6	31.4	24352
275	1	364	238.0	-4.4	86555	364	238.0	-4.4	86555
276	1	156	133.0	0.0	20718	156	133.0	0.0	20718
277	1	306	334.7	68.2	102439	306	334.7	68.2	102439
278	1	51	228.3	1.5	11575	51	228.3	1.5	11575
279	1	556	183.6	3.3	102056	556	183.6	3.3	102056
280	3	77	29.6	0.0	2265	77	29.6	0.0	2265
281	1	86	657.8	0.0	56833	86	657.8	0.0	56833
282	1	185	42.2	0.0	7786	185	42.2	0.0	7786