FINAL REPORT
2006–08 RETRO-COMMISSIONING IMPACT EVALUATION

STUDY ID: CPU0020.01

Submitted to

CALIFORNIA PUBLIC UTILITIES COMMISSION
ENERGY DIVISION

Submitted by

SBW CONSULTING, INC.
2820 Northup Way, Suite 230
Bellevue, WA 98004

In association with

ERS, INC.
JAI J. MITCHELL ANALYTICS
KEMA
MODEST SYSTEMS
PWP, INC.
ROBERT THOMAS BROWN & CO.
SAIC

February 8, 2010

ABSTRACT

This report presents the evaluation, measurement, and verification (EM&V) activities for the Commercial Retro-Commissioning (RCx) High Impact Measure (HIM) for the 06–08 IOU/CPUC Program Cycle. The evaluation comprised 225 projects from more than two dozen programs offered by the four IOUs. Project-level sampling was used for both the net and gross impact evaluations, while projects with a large retrofit component were excluded. Samples were developed with the goal of achieving 90/10 precision at the IOU level, using a stratified random sample based on the ex ante estimates of gross MMBTU savings for each completed project, on the assumption that the ex ante estimates would serve as a good predictor of the associated ex post savings. Results were calculated at the IOU level, across all programs that included RCx projects.

There were three analytical components of the overall RCx evaluation:

- A gross impact analysis calculated gross realization rates for individual projects and used the results to calculate a mean realization rate for each utility, by savings parameter (kW, kWh, therms). The impact approach was focused on project-level analysis, with site-specific M&V plans for each project, using a combination of engineering analysis and building simulation methods.
- A net analysis utilized the self-report approach developed and applied to all Large Nonresidential measures and programs to calculate a net-to gross ratio (NTGR) to estimate the degree of program influence in the decision to RCx the facility and the extent of free ridership.
- An investigation of effective useful life (EUL) examined selected measures from prior RCx programs, to determine if each measure was still in place, operational, and yielding savings to estimate EULs for common classes of RCx measures.

Results of the evaluation were based on 50 gross impact sites, 120 net impact sites, and 96 RCx measures implemented in three previous RCx programs.

- The gross impact analysis found an overall realization rate of 0.55 for total energy savings (when measured in MMBTU/year) and gross savings of 335,000 MMBTU/year across the IOUs. We identified 83 significant reasons for differences between the ex post and ex ante savings in the sample, with nearly half of savings reductions attributable to the measures not working any more.
- The net analysis found generally low levels of free ridership across IOUs and projects, with all IOUs averaging NTGRs of 0.80 or greater per project.
- The EUL analysis found that 22 of 96 measures investigated had failed since their installation three to four years ago, leading to an estimated EUL across all RCx measures of eight years based on a linear extrapolation, although there is a wide band of uncertainty around this estimate.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The evaluation, measurement and verification (EM&V) activities for the Commercial Retro-Commissioning (RCx) High Impact Measure (HIM) for the 06–08 Program Cycle are presented in this report. High impact measures are defined as those efficiency measures common across IOU programs that contribute greater than one percent of the entire IOU savings portfolio for reductions in electrical consumption, electrical demand or natural gas consumption. This evaluation reflects the characteristics of RCx as both a technology or measure (the systematic optimization of a building's systems and operation) and a delivery method (the use of an independent retro-commissioning agent to conduct a detailed study to support the optimization process for an entire facility) and comprises 225 projects from more than two dozen programs offered by the four IOUs. The analysis approach to this evaluation was based on a project-level sampling unit that would be used for both the net and gross impact evaluations, while excluding projects that had a large retrofit component. Samples were designed for results to be calculated at the IOU level, across the programs that included RCx projects. They were developed with the goal of achieving 90/10 precision at the IOU level, using a stratified random sample based on the ex ante estimates of gross savings from each completed project, normalized to MMBTU, on the assumption that these ex ante estimates would serve as a good predictor of the associated ex post savings.

There were three analytical components of the overall RCx evaluation:

- A gross impact analysis calculated gross impacts for individual projects, and used the results to calculate a mean realization rate and unit energy savings (UES) for each utility, by savings parameter (kW, kWh, therms). The impact approach was focused on project-level analysis, with site-specific M&V plans for each project, using a combination of engineering analysis and building simulation methods.
- A net analysis utilized the self-report approach developed and applied to all Large Nonresidential measures and programs to estimate the degree of program influence in the decision to retro-commission the facility. Net to gross ratios (NTGRs) were calculated for individual projects, and were used to calculate fuel-weighted NTGRs for each IOU.
- An EUL investigation examined selected measures from prior RCx programs, to determine if each measure was still in place, operational, and yielding savings to estimate EULs for common classes of RCx measures.

Results of the evaluation were based on 50 gross impact sites, 120 net impact sites, and 96 RCx measures implemented in 3 previous RCx programs.

Gross Impact Results

The gross impact analysis found an overall realization rate of 0.55 for total energy savings (MMBTU/year) and gross savings of 335,000 MMBTU/year across the IOUs. Gross savings realization rates and ex post UES values by kW, kWh and therms for each IOU are shown in Table 1 and **Error! Not a valid bookmark self-reference.**, while MMBTU results are summarized in Figure 1.

Table 1: Gross Savings Realization Rates

			Gross Savings Realization Rate			90% Relative Precision		
Utility	Population	Sample	kW	kWh	Therms	kW	kWh	Therms
PG&E	135	24	0.31	0.45	0.53	0.51	0.23	0.24
SCE	58	13	2.07	0.94	N/A	0.66	0.17	N/A
SCG	28	10	N/A	N/A	0.93	N/A	N/A	0.23
SDG&E	4	3	2.60	1.23	0.21	0.04	0.06	0.01

Table 2: Gross Ex Post Unit Energy Savings

			Gross Ex Post Unit Energy Savings			90% Relative Precision		
Utility	Population	Sample	kW	kWh	Therms	kW	kWh	Therms
PG&E	135	24	13	178,355	7,334	0.51	0.23	0.24
SCE	58	13	30	383,712	462	0.66	0.17	1.56
SCG	28	10	6	28,781	23,735	0.00	0.00	0.23
SDG&E	4	3	129	606,849	11,454	0.04	0.06	0.01

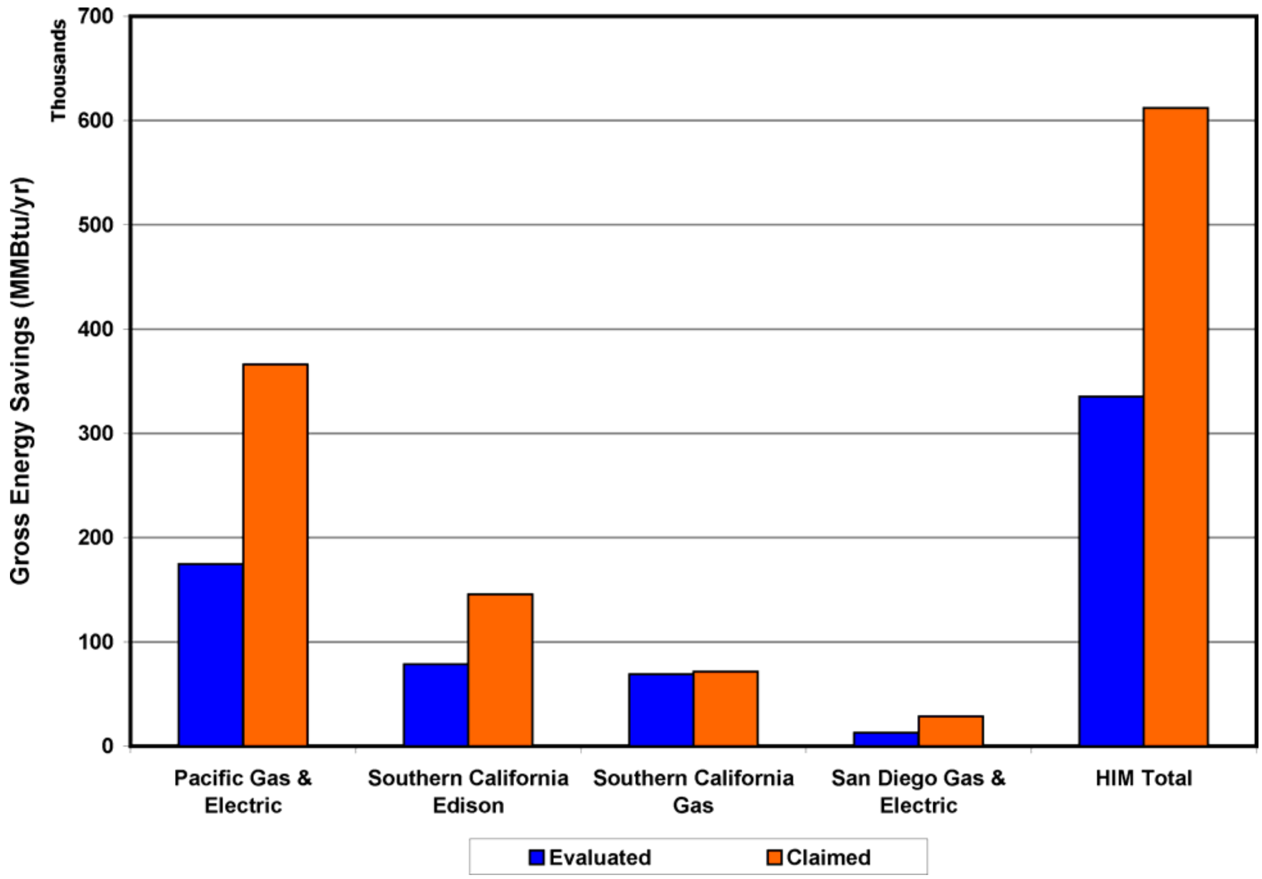


Figure 1: Comparison of evaluated and claimed gross energy savings

Key findings from these tables and graphs include the following:

- The number of RCx projects varied widely between IOUs. PG&E claimed 135 projects, or 60% of the RCx population of 225 projects. In contrast, SDG&E claimed four projects, all of which occurred at the same university campus. Figure 3 illustrates how, on an energy basis, the PG&E RCx projects accounted for about half of the claimed and evaluated savings.
- PG&E projects as a whole had relatively low realization rates for both electric and gas energy savings (0.45 and 0.53), compared to SCE and SDG&E, which had realization rates that varied between gas and electric savings. SCG's single realization rate for therms was 0.93.

To investigate major reasons why the claimed and evaluated gross savings were different, we reviewed the site M&V reports and calculations for all gross impact sites and identified 83 significant reasons for differences between the ex post and ex ante savings in the sample, with nearly half of savings reductions attributable to the measures not working any more.

We also developed a scheme to standardize RCx measures by the building system and general strategy addressed in an effort to better understand the classes of measures and systems that yield RCx savings. Generally, total savings appeared to be distributed fairly evenly among the classes, so that no one class dominated. We had hoped through this analysis to identify particular measure or system classes that might be particularly ripe targets for achieving RCx savings, but the results revealed no obvious targets for future programs.

Net Impact Results

Overall, NTGR scores for the RCx HIM are relatively high, reflecting the continued influence of a variety of programs on the motivation and ability of organizations to pursue RCx projects. With only a single exception, NTGR scores averaged more than 0.50 for every fuel type in every stratum, and the overall mean was significantly higher for all IOUs, as shown in Table 3.

Table 3: NTGR by IOU

			Net to Gross Ratios				90% Relative Precision			
Utility	Population	Sample	Projects	kW	kWh	Therms	Projects	kW	kWh	Therms
PG&E	135	73	0.81	0.76	0.80	0.86	0.04	0.04	0.03	0.02
SCE	58	29	0.86	0.78	0.86	0.91	0.03	N/A	0.02	0.01
SCG	28	15	0.92	N/A	N/A	0.92	0.01	N/A	N/A	0.01
SDG&E	4	3	0.80	0.75	0.75	0.68	0.00	0.00	0.00	0.00

Reasons for the NTGR scores include the following.

- Programs that cover all or part of the cost of the RCx study reduce the risk associated with an RCx project significantly and lead many organizations to proceed with the project. Incentives that cover the cost of the study received the highest mean rating for all program influences cited by respondents—even higher than incentives for implementing recommended measures.
- RCx programs also make projects possible by helping offset funding cutbacks, staffing shortages, and reductions in maintenance budgets, particularly in public institutions, but also in hard-hit private sectors such as office buildings and the hospitality industry.
- The most significant non-program influences in the decision to pursue RCx projects appear to be government or corporate policies that require or encourage implementation of energy efficiency or other "green" measures.

EUL Results

Comprising a heterogeneous mix of measures and actions, the measures addressed by the EUL study totaled 96 of 100 measures originally implemented at these sites through the three programs. The EUL analysis found that 22 of these measures had failed since their installation three to four years ago, leading to an estimated EUL across all RCx measures of eight years based on a linear extrapolation. Note, however, that there is a large uncertainty band around this estimate.

1. INTRODUCTION AND PURPOSE OF STUDY

The evaluation, measurement and verification (EM&V) activities for the Commercial Retro-Commissioning (RCx) High Impact Measure (HIM) described in this report comprise three research designs, each of which was described either in the original RCx Evaluation Plan, the RCx HIM Evaluation Plan or the RCx Effective Useful Life Study Plan.

The RCx evaluation evolved somewhat differently than other contract group evaluations, primarily because of the unique characteristics of RCx, which is a combination of a technology or measure (the systematic optimization of a building's systems and operation) and a delivery method (the use of an independent retro-commissioning agent to conduct a detailed study to support the optimization process) Because of this, the RCx evaluation, unlike other 2006–08 evaluations, was always measure-based. The initial assignment of programs to the RCx contract group was based on the fact that these programs shared a focus on the RCx technology/measure and the associated delivery method. For these programs, we, as the RCx HIM evaluation contractor, developed an approach to sampling and analysis based on a project-level sampling unit that would be used for both the net and gross impact evaluations.

It was recognized fairly early in the 2006–08 evaluation cycle that there were other programs with a substantial RCx component—notably the University of California/California State University/Investor Owned Utility (UC/CSU/IOU) and various Local Government partnerships—which were formally made part of the RCx evaluation after the shift to a HIM-based approach. But it was not until a later review of program tracking data that we discovered the extent to which other programs had included RCx projects. Specifically, there were seven Pacific Gas and Electric (PG&E) programs that were found to include more than 50 RCx projects, which subsequently became part of the RCx evaluation population.

At the same time, many of the programs initially included in the RCx evaluation ended up falling far short of their initial participation goals, which would have rendered program-level results of less value. As a result, the HIM based approach, with a project-level analysis of RCx as a single measure, was designed for all RCx projects, with the goal of achieving reliable estimates of gross and net savings for each of the four IOU RCx portfolios. Although early attempts were made to assemble information on RCx projects likely to be completed directly from the program implementers, the final sampling from for the RCx evaluation was assembled from the program tracking records submitted by the IOUs in March 2009. These data were used to develop the HIM sample frame in support of the calculation of gross realization rates and net-to-gross ratios (NTGRs) at the IOU level.

While this evaluation treats all RCx actions as a single measure, RCx is, in reality, an overarching strategy that consolidates many disparate actions into individual projects. RCx projects generally consist of a custom-engineered, site-specific combination of recommended actions designed to optimize all or a portion of a facility's energy systems. Depending on the program design, some programs supported projects that encompassed low-cost RCx actions, such as changing control setpoints, as well as conventional retrofit measures, such as installing efficient light fixtures. The RCx sample excluded projects that were exclusively retrofit. HIM sample frame information was sufficient to solidly define the population of RCx projects, but was insufficient for us to classify individual measures. As shown in the results section, we were able to disaggregate savings for some individual recommendations in the sample projects. This disaggregation was fairly rough, but it was sufficient to draw some general conclusions about which types of recommendations yielded the most savings, to inform future program planning efforts.

While Effective Useful Life (EUL) analysis was generally outside the scope of the 2006–08 impact evaluations, the California Public Utilities Commission (CPUC) Energy Division (ED) determined such an analysis was warranted in the case of RCx due to the high level of uncertainty surrounding program EUL claims. Since little or no prior study had been done of RCx EULs, we were directed to analyze the EUL of measures installed through earlier RCx programs.

Table 4 below summarizes the three analytical components of the overall RCx Evaluation: the gross impact analysis (report section 6), the net impact analysis (section 7) and the EUL study (section 8). Each analytical component is discussed briefly below.

Table 4: Retro-Commissioning Evaluations

	Gross Savings	Net Savings	Effective Useful Life
Evaluation Method→	Field Measurements	Self Report	Field Measurements
Report Section			
3.	Saving Realization Rate		
4..		Net-To-Gross Ratio (NTGR)	
5..			Effective Useful Life (EUL)

1.1. Summary of Evaluation Activities

1.1.1. Gross Savings Evaluation

The approach described in the HIM evaluation plan for RCx, and subsequently implemented, called for the development of ex post savings estimates by IOU, rather than by program. The resulting cross-program HIM samples were designed to be large enough to provide utility-level estimates of RCx HIM performance at approximately 90/10 precision. These estimates can inform statewide RCx realization rates that may be applied generally. The decision to treat RCx as a single HIM led to changes to the original evaluation plan as described in section 1. These changes provided a more robust assessment of RCx performance because so much of the RCx savings in the 2006–08 portfolio occurred outside the programs that had originally been assigned to the Commercial RCx contract group.

As shown in the Table 1, the gross impact evaluation calculated gross realization rates for individual projects, and the impact approach was focused on project-level analysis. Site-specific M&V plans were developed for each project, using a combination of engineering analysis and building simulation methods. Both engineering and building simulation analyses were supported by extensive on-site data collection as specified in the site M&V plan, including, as appropriate, inspection, metering/trend logging and interviews with building operators and the commissioning (Cx) agent who performed the RCx study, thus allowing engineers to understand each site's systems, as well as the proposed and implemented measures.

1.1.2. Net Savings Evaluation

Consistent with the approach outlined in the original RCx Evaluation Plan, the net analysis of the RCx HIM utilized the same self-report approach developed and applied to all Large Nonresidential measures and programs. The self-report option was chosen because alternative methods of estimating the NTGR, such as discrete choice or billing analysis were not practical in light of the limited number of projects, the heterogeneity of participating customers, and the small impacts relative to overall energy usage. The self-report approach recommended by the NTGR working group, comprising ED staff, technical advisors, and evaluation contractors, was used, as described in greater detail later in this report. Two levels of rigor were used in the net analysis. The Standard - Very Large rigor level integrated information from other sources besides the customer interview, thereby allowing us to tell the full story behind each organization's decision to proceed with RCx recommendations and the role that the programs play in the causing the work to occur. The second, Basic rigor level used the same standard data collection instrument and algorithm to calculate the NTGR, but did not bring in the additional noncustomer viewpoints to support the analysis.

The original Commercial RCx M&V plan stated that all participants in the impact sample would be interviewed using the Standard rigor level NTG method, with the largest (i.e., those in the Large and Certainty strata) being subject to additional data collection and analysis anticipated by the Standard – Very Large NTG method. All other participants would be analyzed using the Basic NTG methodology. The HIM net impact evaluation improved on this analysis by interviewing all impact sample participants using the Standard – Very Large NTG method, rather than just the large sites. We randomly selected from among the remaining sites so that 50% of all sites were treated, in accordance with guidance from the Energy Division and its technical advisors. These randomly selected sites were interviewed using the Basic NTG method.

1.1.3. Effective Useful Life Evaluation

The goal of the EUL investigation was to examine selected measures from prior RCx programs, and determine if each measure was still in place, operational, and yielding savings based on available evidence. The results were then used to synthesize the data, to the extent possible, into estimated EULs for common classes of RCx measures / projects.

It must be emphasized that the EUL study was not designed to comply with the California Energy Efficiency Evaluation Protocols for EUL studies, but was rather undertaken with limited resources and time to address an immediate need for improved RCx measure life information, since assumed EULs for RCx measures vary widely across programs. In addition to enhancing understanding and serving as the basis for future investigations, the EULs that result from the study—whether at the RCx-project or measure level—can be used directly by the CPUC in assessing the cost-effectiveness of RCx projects implemented during the 2006–08 program cycle.

The intent of this study was to take a fresh look at EUL estimates for RCx measures through building owner/operator interviews and field inspection. This study sought to provide observation-based RCx measure persistence data based on measures from three third-party RCx programs from previous program cycles:

- Oakland Energy Partnership - Large Commercial Building Tune-Up Program – 2002–03
- Building Tune-Up Program – 2004–05
- UC-CSU-IOU Partnership RCx Projects – 2004–05

This relatively small retrospective EUL study provided valuable, if imperfect, empirical data. It was an important initial step, which can inform future, more comprehensive efforts to determine RCx EULs.

1.2. Identification of the RCx Study Population

As described above, both gross and net activities were conducted for RCx as a single measure. RCx was treated as a HIM by the ED from the start, both because it was likely to be a significant source of savings and because it was felt to be an important emerging delivery/technology type. As a result, the grouping of measures for the 2006–08 RCx evaluation has always been on the basis of the single RCx measure. What changed was the number of RCx projects that were completed through programs other than those initially within the Commercial RCx contract group. Once it was decided that all RCx projects should be treated in this evaluation, the task became one of identifying projects that were not included in the original Commercial RCx programs.

The Partnership Programs were relatively easily identified as having numerous RCx projects, which were included in the present evaluation as soon as the decision to conduct HIM-based analysis was made. Other programs took longer to identify, in part because many of these projects were not completed (and entered into the tracking database) until the second half of 2008, the third program year. We therefore carefully reviewed the fourth quarter 2008 program tracking data, both for programs that were explicitly identified as RCx programs (e.g. SCE2508, PGE2090), as well as for individual measure descriptions that contained any of a string of key words, such as commissioning, Cx, retro-commissioning, RCx, etc. This comprehensive review resulted in a population of 260 RCx projects.

Because many of these projects, and much of the RCx savings, in the 2006–08 portfolio occurred outside of the Commercial RCx contract group, this HIM approach led to a much broader sample that took in projects from a variety of utility and third party programs. Both gross and net samples were developed across these programs. Programs we identified as containing RCx projects are listed below.

1.3. Programs that Implemented RCx Projects

Each of the utilities implemented RCx projects in the context of one or more programs. Table 5 provides a listing of these programs, along with the number of RCx projects completed through each. Programs offered by different IOUs that had very similar design and markets are listed as a group. A more detailed table that includes program descriptions and key program elements is presented in Appendix 5.1 Overview of Programs Evaluated.

Table 5: Summary of evaluated Programs

Programs Included in this evaluation	# of RCx Projects*
PGE2001 Ag & Food Processing	1
PGE2002 Schools and Colleges	10
PGE2005 Hi-Tech Facilities	4
PGE2006 Medical Facilities	2
PGE2007 Office Buildings (Large Commercial)	28
PGE2015 LGP Association of Bay Area Governments	10
PGE2025 LGP Marin County	2
PGE2032 LGP Sonoma County	4
PGE2035 LGP Silicon Valley Leadership Group Energy Watch	3
PGE2036, SCE2530, SCG3520, SDG&E3029 UC-CSU-PG&E-SCE-SCG-SDG&E Partnership	54
PGE2052 Lodging Savers	9
PGE2056 Monitoring-Based Persistence Commissioning	3
PGE2070 Data Centers	5
PGE2071 Hospitality Energy Efficiency Program	3
PGE2072 Hospitals Pilot	7
PGE2088 State Leased Facilities	3
PGE2090 Airflow and Fume Hood Control Systems Re-Commissioning	1
PGE2091 Retrocommissioning Services and Incentives Program	15
PGE2094 Macy's Comprehensive Energy Management	18
SCE2508 Retro-Commissioning (RCx)	22
SCE2526, SCG 3518 California Community Colleges Partnership	2
SCE2528, SCG 3527 SCE-SCG County of Los Angeles Partnership	53
SDGE3010 SEnergy Savings Bids	1
Total	**260**

* Some projects were excluded from the evaluation because of insignificant savings or because they were retrofit only

2. Gross Savings Evaluation

2.1. Evaluation Objectives

Parameters examined by the RCx gross impact evaluation included kWh, kW and therm impacts. While the focus of the evaluation was initially on ten RCx programs operated by the four IOUs to improve the energy performance of existing plant and equipment for participating customers, the HIM evaluation plan for RCx called for the development of ex post savings estimates by IOU, rather than by program. Evaluation samples were drawn from all RCx projects completed by each IOU, regardless of the program under which they were completed.

The goal of the HIM gross impact evaluation was to maximize the value of EM&V expenditures within the CPUC policy framework to best meet ongoing informational needs by:

- Providing utility-level estimates of RCx HIM performance across all relevant contract groups at approximately 90/10 precision. These estimates can inform statewide RCx realization rates that may be applied generally. Given that much of the RCx savings in the 2006–08 portfolio occurred outside of the Commercial RCx contract group, this HIM approach was designed to provide a more robust assessment of RCx performance.
- Specifying large enough samples to make meaningful comparisons of gross realization rates between the largest programs. While the HIM sample design meant that we no longer had statistically valid evaluated gross savings for each RCx program, this was a moot point in many cases, since many programs had few participants. Other programs with significant RCx savings generally were part of a larger program that also included custom non-RCx projects as well (such as the UC/CSU/IOU Partnership in the LGP group, for instance).
- Obtaining key information on poorly-understood aspects of RCx performance, such as measure effective useful lives (EUL) as described in greater detail elsewhere, and which commonly-occurring RCx measures yield the most savings. Even though the EULs for RCx were not updated in this evaluation, Energy Division believedt the additional research on RCx EULs was warranted.

2.2. Sample Design and Selection

Although early attempts were made to assemble information on RCx projects likely to be completed directly from the program implementers, the final sampling for the RCx evaluation was assembled from the program tracking records submitted by the IOUs in March 2009, which found that seven Pacific Gas and Electric (PG&E) programs included a total of 57 RCx projects, which subsequently became part of the RCx evaluation population. These data were used to develop the HIM sample frame in support of the calculation of gross realization rates and net-to-gross ratios (NTGRs) at the IOU level.

The sampling plan for the RCx HIM aimed to provide 90/10 precision for gross impact estimates at the IOU level within the constraint of available budget. The sample design presented here provides such precision with a sample *n* of 50 out of 252 projects across all four IOUs that supported projects with RCx. The size of the gross impact sample was based on the expected relative precision for various sample sizes, strata and strata boundaries, given the number of completed projects and the assumption that ex ante savings estimates serve as a good predictor of the associated ex post savings. We developed stratified random samples of participants for each IOU service territory, with stratification based on the ex ante estimates of gross savings from each completed project, normalized to MMBTU, with the goal of achieving 90/10 precision at the IOU level.

Strata were defined using the Dalenius-Hodges method described in section 5A.7 of *Sampling Techniques*, 3rd Edition, by William G. Cochran. For each IOU, we sorted all projects by their ex ante impacts, then calculated all the first differences between each project's impact and the next smallest project's impact, and then took the square root of each first difference so that we could calculate their cumulative value. The final value of these cumulative square roots was divided by the number of strata, and the resulting value defined the upper bound of the first stratum. That same value was doubled to define the upper boundary for the second stratum and so on. In determining how many strata to employ, we balanced the greater precision made possible by added stratification against the uncertainty introduced by having too few observations in any individual stratum, taking care to avoid having any strata with just a single observation.

Before drawing the strata boundaries, we allocated the largest projects for each IOU territory to a certainty stratum. By reducing uncertainty from sampling to zero for the largest projects (that is, we know with certainty for this stratum that the mean savings for the sample are the same as the mean savings for the population), the certainty stratum allows us to attain greater overall relative precision with a smaller sample for the remaining sites. Similarly, projects with the very smallest ex ante savings were excluded from the population before we set strata boundaries using the procedure described above.

The resulting gross impact sample is presented in Table 6.

Table 6: Gross Impact Evaluation Sample

		Stratum Boundaries Ex Ante Savings1 (MMBTU)		RCx Projects			% of Ex Ante Savings in Sample		
Utility	**Stratum**	**Lower**	**Upper**	**Population**	**Sample**	**% in Sample**	**kW**	**kWh**	**Therms**
PG&E	1	264	1,414	22	2	9	2	4	3
PG&E	2	1,444	2,872	25	2	8	25	15	3
PG&E	3	3,044	6,057	12	2	17	13	11	22
PG&E	9	6,317	17,224	10	9	90	78	98	80
PG&E	Phase 2 - 1	78	1,928	45	2	4	0	2	0
PG&E	Phase 2 - 2	1,964	5,885	16	2	13	0	10	18
PG&E	Phase 2 - 9	9,944	26,651	5	5	100	100	100	100
PG&E	Excluded	5	55	2	0				
PG&E	Phase 2 - Excl.	0	61	14	0				
PG&E	**All Sampled**	**0**	**26,651**	**135**	**24**	**18**	**33**	**39**	**61**
SCE	1	83	793	29	3	10	19	14	0
SCE	2	800	1,939	13	3	23	19	20	6
SCE	3	2,048	3,344	12	3	25	0	26	11
SCE	9	4,290	10,607	4	4	100	100	100	100

		Stratum Boundaries Ex Ante Savings1 (MMBTU)		RCx Projects			% of Ex Ante Savings in Sample		
Utility	**Stratum**	**Lower**	**Upper**	**Population**	**Sample**	**% in Sample**	**kW**	**kWh**	**Therms**
SCE	Excluded	5	41	2	0				
SCE	Phase 2 - Excl	2,830	2,830	1	0				
SCE	**All Sampled**	**5**	**10,607**	**58**	**13**	**22**	**26**	**46**	**38**
SCG	1	240	1,744	13	3	23	NA	NA	18
SCG	2	1,753	3,621	10	2	20	NA	NA	21
SCG	9	4,524	18,147	5	5	100	NA	NA	100
SCG	Phase 2 - Excl	0	2,695	7	0				
SCG	**All Sampled**	**240**	**18,147**	**28**	**10**	**36**	**NA**	**NA**	**51**
SDG&E	1	1,575	2,034	2	1	50	63	63	55
SDG&E	9	4,319	24,959	2	2	100	100	100	100
SDG&E	Phase 2 - Excl	1,155	2,123	3	0				
SDG&E	**All Sampled**	**1,575**	**24,959**	**4**	**3**	**75**	**90**	**89**	**78**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

As shown in the table, the number of sampled (non-certainty) strata varies by IOU, ranging from a single sampled stratum plus the certainty stratum for SDG&E to three strata plus the certainty stratum for PG&E, while the number of projects in the certainty stratum ranges from two for SDG&E to 14 for PG&E. The resulting sample of 50 points was projected to provide relative precision of about 10% at the 90% confidence level for mean RCx project savings for each of the IOUs.

Changes to Sample Design from Original Plan

The original evaluation plan for the RCx contract group called for the development of program-specific realization rates, and in that context it made sense to sample from all of the projects completed under the RCx programs covered in the RCx contract group. With the shift to the HIM approach, which made SBW responsible for developing RCx realization rates by IOU, the 10 retrofit-only projects under program PGE2094 (Macy's Comprehensive Energy Management Program) were excluded. These 10 retrofit-only projects had only kWh savings, and their removal from the RCx population reduced the kWh savings total by about 6%. The three of these projects that had been drawn in the original sample were replaced by the next projects in line, using the random order established for each size stratum.

When estimating Program PGE2094's total impact, the RCx HIM realization rates were applied to this program's RCx projects, and the appropriate Non-Residential Custom Retrofit HIM realization rates were applied to the retrofit-0nly projects.

Sample Disposition

As described in section 3.3.2, steps were taken to minimize non-response bias in this research. Only one project was dropped from the gross savings sample. This was due to a major equipment failure that made collection of post-data impossible.

2.3. Methodology

In this evaluation, gross energy (kWh, therms) and peak kW impacts were evaluated for RCx projects completed between January 1, 2006 and December 31, 2008. The following sections describe (1) the overall approach and rationale for the chosen methods, including rigor level, and (2) the site specific approach to data collection and analysis.

2.3.1. Overall Approach, Rigor Level and Protocols

The gross impact evaluation calculated utility-level gross realization rates from the realization rates for individual projects, and the impact approach was therefore focused on project-level analysis of ex post kWh, kW and therms savings. This approach involved the following steps.

2.3.1.1 Baseline Assessment

Because these programs provide customized treatment for each participant, the relevant baseline information depended on which RCx measures were found to be applicable. For applicable measures, the EM&V team reviewed available baseline data the commissioning agent obtained as part of the diagnostic process used to identify these measures and to estimate for the building owner the costs and benefits associated with each measure. When it was both feasible and necessary, we collected additional baseline data, from visual inspection of affected systems, one-time measurements, short-term trend logging (customer control system trend logs or special metering), manufacturers' specifications, and self-reports from building operators and tenants.

2.3.1.2 Gross Energy (kWh and Therm) Assessment

All projects in the gross impact sample received an assessment of gross energy savings for measures implemented at the site of the project. Typically, the RCx diagnostic process identified numerous actions that the participant could have taken to improve the energy performance of the building, along with the cost and savings associated with each measure to motivate the owner's adoption of the recommended improvements. The program may have offered incentives to further motivate adoption. Measures paid for with an incentive created direct program impact, as did those measures taken as a direct consequence of the RCx study funded (or co-funded) by a program. Other program services may have identified actions that the customer took without receiving an incentive and without any specific recommendation from the program, but based on general information or techniques learned from the program (spillover).

For the RCx evaluation, we had the benefit of knowing which of a potentially long list of items recommended by the program to the owner/operator of a facility had been selected and implemented. We prioritized these items according to their respective contribution to the overall site savings, and developed data collection and analysis approaches that best addressed the uncertainties in the savings calculations.

The algorithms used in the evaluation to estimate savings for corrective actions that were taken varied with the particular circumstances encountered, and included: (1) spreadsheet engineering models using bin weather data, (2), other spreadsheet engineering models, (3) linear regression analysis of measured variables, (4) building simulation programs, (5) analysis of whole building or whole system metering records, or (6) a combination of these. The most appropriate analysis method was determined on a case-by-case basis based upon factors such as the information available from the ex ante estimation process, available performance data, ease of further data collection, complexity of determining system performance, the relative savings contribution of an action to the overall program package and budget constraints. Key sources of information for each analyzed measure included trend logs from energy management systems; spot or long-term measurements, using existing gauges, displays, or handheld equipment; manufacturer performance curves or equipment cut sheets; nameplate data; generic curves as provided in DOE-2 software; as-built construction drawings; visual observations; and interviews with facility staff.

2.3.1.3 Gross Peak Demand (kW) Assessment

The peak electric demand is currently defined in DEER (the CPUC-sponsored Database for Energy Efficient Resources) as the average grid-level impact for a measure between 2 p.m. and 5 p.m. during the three consecutive weekday periods containing the weekday with the hottest temperature of the year. We used the 2008 DEER update to identify these days for each of the 16 California climate zones, based on the typical weather data sets developed for the California Title 24 Building Energy Efficiency Standards. RCx project peak demand impacts were estimated accordingly.

For projects with significant kW demand savings, we estimated 8,760 values for peak kW demand savings. The procedures involved the extrapolation of post-retrofit short term time-of-use measurements to annual (8,760 hours/year) values by day type or an analytical relationship between post-retrofit site-specific short-term measurements of measure performance and outside air temperature. This relationship was then applied to typical 8760 hourly temperature values from the climate zone weather.

For each of the steps described above, site-specific M&V plans were developed for each project in the gross impact sample, using a combination of engineering analysis, building simulation, and billing analysis methods. Both engineering and building simulation analyses were supported by extensive on-site data collection as specified in the site M&V plans, including, as appropriate, inspection, metering/trend logging, and interviews with building operators and the Cx agent who performed the RCx study. This allowed evaluators to understand each site's systems, as well as the proposed and implemented measures.

2.3.1.4 Rigor Level

The Energy Division initially specified the evaluation type and rigor levels for RCx programs either in the RFP or in subsequent guidance. As per the original M&V plan developed for the Commercial RCx contract group and confirmed in the HIM addendum to that plan, we proposed that all sampled projects should be done with enhanced rigor rather than the standard rigor specified. We also suggested that all projects be assigned the protocol-guided direct impact evaluation type. This evaluation type was appropriate because these programs implement complex, controls-intensive changes to complex buildings and systems. Custom engineering analysis of savings was required to understand the impact of the program on participating facilities. Such projects did not lend themselves to a simplified verification-guided analysis.

Similarly, all gross kWh savings estimates also had enhanced rigor. Accurate estimation of RCx savings require measured data from pre and post configurations of affected systems. This is fundamentally consistent with International Performance Measurement and Verification Protocol (IPMVP) Option B and thus is consistent with enhanced gross savings impact rigor. Most of the RCx programs had been assigned basic kW rigor, but Energy Division agreed that these be changed to enhanced to be consistent with the treatment of gross kWh savings, and we conducted the HIM gross impact evaluation accordingly. We believe this was appropriate as it is consistent with their treatment with Option B.

Our HIM approach maintained this highest level of rigor for gross savings. Tthe site-specific M&V plans for each sampled site carefully assessed the measures (and by extension, the parameters for calculating savings) at that site, and determined the most appropriate data collection and analysis approaches to minimize savings uncertainty within budget and customer constraints.

2.3.2. Site-specific Approach

To implement the above approach, we generally developed an M&V plan consistent with IPMVP Option B (retrofit isolation) for each sampled RCx project. Note that in some instances, where data were scarce or where ex ante simulation models had already been developed, we determined that Options C (whole building) or D (calibrated simulation), respectively, would be more appropriate. Since the M&V plans were prepared at the project level, they contained site information, including contact information, IOU tracking database information and data collection and analysis methods for spillover measures. Each plan was reviewed and approved by Energy Division and its technical advisors prior to the scheduling of a site visit. A more detailed description of the contents of an M&V plan is provided in section 5.3.4.2..

The M&V plan proposed site-specific data collection and analysis procedures for each project that could be performed within the allotted budget. Specific guidance for protocol-guided direct impact data collection and analysis were provided in a detailed handbook developed prior to the start of field data collection, and included in section 5.3.

The specific analytic approach selected for each measure group within an RCx project depended upon a number of technical factors, the available information, and the evaluation budget. In all instances we considered alternative evaluation methodologies, weighed their reliability and accuracy advantages against their costs, selected the most appropriate evaluation method, and allocated the evaluation funds in the most cost-effective manner. The overall objective of the M&V plan for each sampled project was to minimize the uncertainty in the evaluation estimate of project-level savings. When they were demonstrated to have been induced by program participation, spillover measures were subject to the same enhanced rigor analysis methods.

In accordance with the approved plans, we conducted project- and measure-specific data collection and analysis of gross energy savings as follows for the sampled projects:

- As soon as a project was sampled, the program provided ex ante savings estimates, savings calculations, and supporting data for each measure implemented at the site. In addition, the program provided contact information for the lead project engineer and the site liaison.
- Based on the information provided by the program, we developed a project-specific savings verification plan that encompassed all or most of the measures selected by the owner to be installed. Depending on the nature of the measures and the magnitude of their expected energy savings, we sometimes selected only large saver measures, or a sample of measures that accounted for a significant or representative portion of the overall project savings. The development of the plan included a review of the ex ante savings calculations and supporting data for each measure, and conversations with the lead project engineer and site liaison about appropriate methods to collect additional site data. The plan also described either how the additional data would be integrated into the original model developed by the program implementer, or how a more appropriate method would be applied. This description included a review of the ex ante savings calculation algorithms, measurements to be taken, and acceptable stipulations.
- Energy Division reviewed the aggregate M&V plan for all selected measures for a site, and recommended needed enhancements. Sometimes this process exposed data collected or other information that the program possessed that was otherwise unknown to our team. However, we were careful not to allow the program to bias our approach and made sure that all communications on this topic with the program were fully transparent.
- We took the agreed-upon post measurements and performed supplemental data collection, including inspection of the building to confirm that the RCx improvements had been made.

Once post data collection was complete, we completed the evaluation savings analysis, and developed final evaluation estimates of peak kW, kWh/year, and therms/year savings for each implemented, sampled measure (or group of measures). As appropriate, we discussed any major discrepancies in the site savings estimates with the customer or commissioning agent to understand why the difference occurred. Savings estimates were developed using baseline assumptions adjusted for post conditions, to account for any changes in the operation or characteristics of the sampled sites that were not related to the efficiency improvements, such as changes in tenancy or business schedule.

The results from the gross savings (kWh, kW and therms) analysis were documented in project-specific reports, which also documented the impacts of any significant spillover measures encountered at a site. The reports also document the final data collection procedures that were applied in the analysis and the gross energy and peak demand savings that were achieved. Whenever possible, the reports compared the ex post gross savings estimate to the ex ante savings estimate prepared by the IOU and explained significant differences that were found.

2.4. Confidence and Precision of Key Findings

As noted previously, the sample for the RCx HIM evaluation was designed with the goal of attaining 90/10 relative precision; i.e., so that the 90% confidence interval around the estimated parameter is less than or equal to 10% of the mean value of the parameter estimate. This expected precision level was calculated for the sample using the Cochran (1977) approach to estimating the variance of a stratified random sample, where the ex ante estimates of MMBTU savings for sampled cases were used as the basis for the calculation. To the extent that ex ante estimates are in fact a good predictor of ex post savings, the resulting estimate of 50 gross impact sample points should yield a comparable attained precision for total BTU savings.

The actual precision attained was calculated using several different approaches taken from the California Evaluation Framework (2004), Cochran's Sampling Techniques (1977) and Taylor's *An Introduction to error analysis: The study of uncertainties in physical measurements* (1997), as cited in the International Performance Measurement & Verification Protocol (2002), with the specific techniques used depending on the type of impacts analyzed.

- For sites where all savings parameters had ex ante and ex post values (i.e., kW, kWh, and/or therms), precision was based on the estimation of a realization rate, defined as the ratio of ex post to ex ante savings. As described in the Framework (p. 358) and discussed in greater detail in appendix section 5.2, the sample realization rate was calculated both overall and by stratum, but only the overall realization rate was used in the calculation of the UES. The overall realization rate was multiplied by the mean population ex ante savings estimate to calculate the mean ex post UES value. The standard error and error ratio were calculated for the realization rate and used to determine the 90% confidence band around the realization rate and thus around the UES. The confidence bound was then divided by the sample realization rate to calculate the relative precision of the estimate. Results are presented in section 2.6.
- In cases where there was no ex ante claim— for example, gas savings impacts for an SCE project that only claimed electric savings—no realization rate could be calculated so a different approach was used. The selected approach is the same one taken from Cochran (p. 95) to calculate relative precision using the ex ante estimates for the sample. This approach involved using strata weights (the proportion of the total population accounted for by each stratum) to calculate a weighted mean and variance for a stratified random sample. The resulting mean and standard error were used to calculate the error band and relative precision for these cases.

The resulting confidence and precision levels for all sites are reported in section 2.6. Since the targets for confidence and precision for this project were 90/10 and sampling was based on this standard, all relative precision numbers are reported in those terms.

2.5. Validity and Reliability

In order to maximize the accuracy of the final program-level savings estimate, the evaluation team optimized both the sample size and the level of measurement and analysis effort spent on sites within the sample. As discussed further below, there are a number of significant sources of uncertainty associated with measurement of the key impact parameters of gross energy savings, gross demand savings, and net impacts. These sources of uncertainty are particularly challenging within large nonresidential program evaluations due to the heterogeneity of applications, processes, and energy efficiency measures.

There is an appropriate balance that can be achieved by trading off the sample size against the level of M&V activity (while holding a fixed budget constant) to ensure the best overall reliability of the program savings estimates. For the purposes of the gross impact evaluation, the sample size was essentially dictated by the goal of attaining 90/10 precision at the IOU level, so the analysis of uncertainty focused on elements that contribute to uncertainty in estimates of savings for each individual site. Examples of these elements of uncertainty include variations due to equipment scheduling and variations in occupant behavior and business levels; modeling errors; instrument error (if measurement is conducted); measure sampling error within a site; and planned and unplanned assumptions (according to IPMVP, this category "encompasses all the unquantifiable errors associated with stipulations, and the assumptions necessary for measurement and savings determination").

Every step of the gross impact analysis of the RCx HIM both recognized the need to minimize uncertainty and incorporated specific areas of engineering analysis and judgment to increase validity and reliability. This process can be thought of as a continuous and iterative effort to maximize available evaluation resources for valid, reliable results by targeting the evaluation effort to the areas of greatest uncertainty. Even after initial site level M&V plans were developed, it was often necessary to make real-time adjustments to the parameters measured and the techniques used to measure them.

At the highest level, the sampling plan allocated sample points so that the targeted precision of population estimates would be realized, as discussed in section 3.2 above. And given the resulting sample, the impact analysis approached each individual site as a separate case that required uncertainty to be minimized. This process is shown graphically in Figure 2.

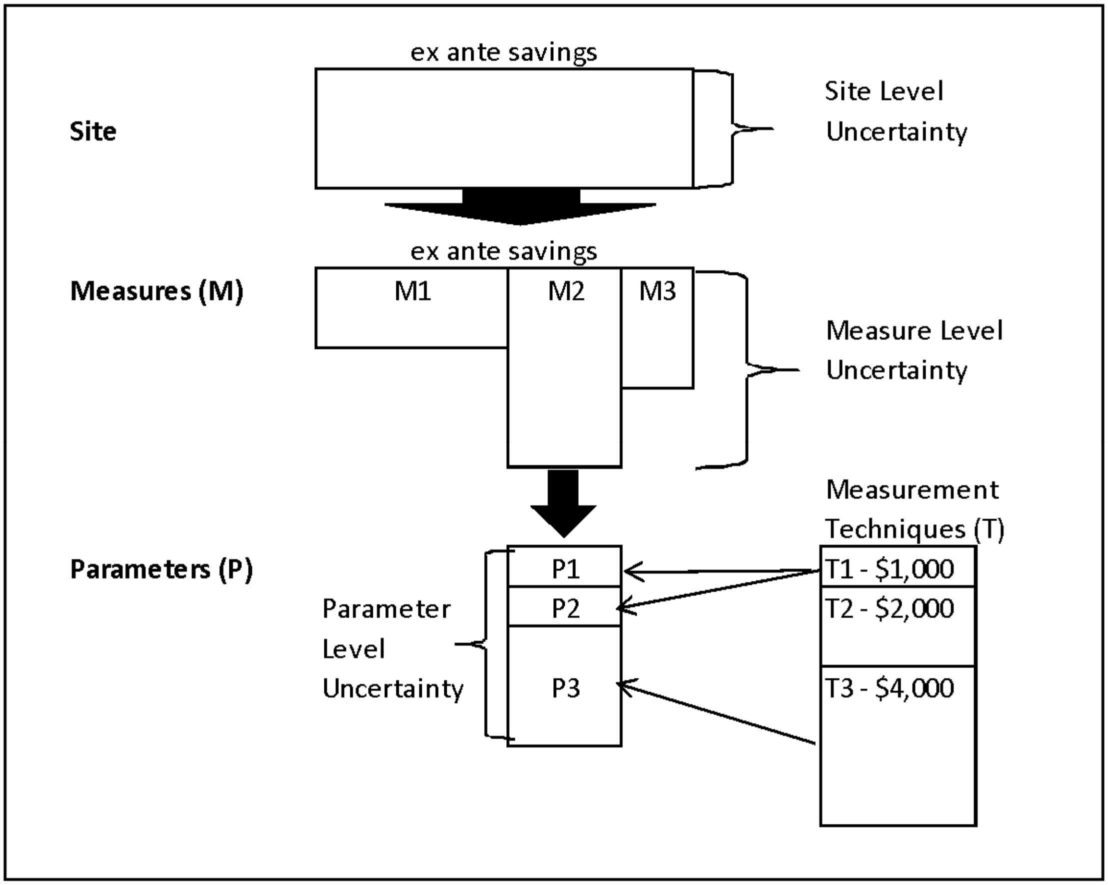


Figure 2: Allocating evaluation resources to reduce uncertainty

The first step in the site-level analysis was to rank the measures that contributed to the ex ante impact associated with that site. For most sampled RCx sites, evaluators had the advantage of a detailed assessment of pre-RCx energy usage as determined by the RCx study and master list of findings, which led to a better-than-average degree of certainty for pre- usage estimates and made it clear what portion of ex ante site-level impacts was associated with which measure. Having good pre data helped minimize uncertainty, meaning the site lead did not have to devote as many resources to the post- measurement. On the other hand, if the pre data were non-existent (but critical), then it would not make much sense to attempt to measure post usage with high precision, since the pre-post uncertainty is so large.

While it would seem that one might simply allocate evaluation resources according to that share of impacts (i.e., Measure 1 in the exhibit accounts for about half of impacts and would therefore receive about half of evaluation resources assigned to the site), the appropriate level of resources depended on the degree of uncertainty associated with each measure; the greater the uncertainty, the more resources were needed to obtain a reliable estimate of savings. It was through engineering judgment, based on the experience of the site lead and the impact evaluation team, that the degree of uncertainty associated with each measure was assessed, as shown graphically in the middle of the exhibit.

The engineering team's assessment of uncertainty, in combination with the ex ante impact, allowed us to assign resources in a way that would increase confidence in the results. For example, two measures—

supply air temperature reset control and lighting schedule change—might both be associated with a high proportion of site-level savings, but the impact of the lighting schedule change would generally be more certain and require fewer resources to measure than the changes to the operation of the HVAC system. By weighing the relative uncertainty associated with each measure along with the ex ante savings, we were able to increase the reliability of the site -level savings estimate.

The same procedure was applied at the measure level. The impact for each measure is determined by a variety of parameters, including for example connected load, set points, run time, occupancy, throughput and others. These parameters play varying roles in the algorithms used to calculate savings and they are subject to varying degrees of uncertainty. Connected load can be determined with minimal measurement error; occupancy or throughput may be subject to much greater variation over time and therefore require a greater allocation of resources to gather enough data to minimize uncertainty. The knowledge and judgment of the engineering team were used to make decisions regarding this allocation among parameters.

Finally, site leads had at their disposal a variety of measurement techniques to collect data on the selected parameters, and engineering judgment figured prominently in what kinds of measurements were undertaken or requested in the evaluation. Lighting run time, for example, could be measured using loggers placed at the level of banks controlled by a single switch, by loggers placed one to a floor, or simply by asking the building owner questions regarding hours of occupancy or operation. Similarly, data on the operation of a building's air handling units might be collected by monitoring one, several, or all of the units in the building. Lower cost measurement and analysis techniques might be appropriate for parameters that have a relatively lower degree of uncertainty, as illustrated by the hypothetical $1,000 techniques assigned to Parameters 1 and 2 in the exhibit. On the other hand, a more extensive, higher cost data collection approach, illustrated by the $4,000 techniques in the exhibit, might be needed to address the higher degree of uncertainty surrounding Parameter 3. In addition, site leads were acutely conscious of a broader definition of cost that included the customer hassle factor. Particular measurement techniques might yield highly accurate results, but require facility shutdown, extensive use of customer staff time, or other "costs" that could undermine the evaluation effort and cause the engineering team to ultimately lose access to the site. By balancing all costs and uncertainty factors in this manner and then applying both engineering and management judgment, the evaluation team was able to select measurement techniques consistent with available resources in a way that helped minimize uncertainty and increase the validity of the results.

2.6. Detailed Findings

This section presents quantitative findings from our gross saving sample concerning realization rates and unit energy savings.

2.6.1. IOU-Level Results

The gross savings evaluation yielded quantitative results for all 50 sampled projects. A detailed listing of the ex ante and ex post gross savings for each of these projects can be found in appendix section 5.5.

Table 7 below summarizes the analytical approaches used for the sample projects. For most sites, we relied upon extensive trend and/or logger data, coupled with a custom spreadsheet-based engineering analysis. For about a quarter of the projects, the program had created a calibrated building simulation (eQUEST®) model, which we subsequently modified with evaluator-collected post-implementation data. In one instance, the evaluation team created a new building simulation model.

Five projects required analyses structured around whole building or whole system (e.g., chilled water use, hot water use) metering data provided by the customer or utility. Of the remaining three projects, two had poor baseline data, so the analyses were simply verifications that the measures had been installed and were operational, and one project had been completely disabled, and thus required no analysis.

Table 7: Analysis approaches

Evaluation Gross Analysis Approach	# of projects
Detailed monitoring / custom analysis	28
Building simulation - updated program model	13
Building simulation - created new model	1
Whole building/system analysis	5
Verification	2
None needed	1
Total	**50**

The evaluation also found that all projects and measures in the gross sample could be considered “early replacement,” consisting of “add-on” measures, as opposed to major equipment replacements or retrofits. Consequently, the savings analyses assumed pre-implementation conditions to be the baseline.

Table 8 and Table 9 summarize gross realization rates and ex post unit energy savings (UES), respectively, for each utility and for the RCx HIM overall, where the realization rate is defined as the ratio of ex post to ex ante savings, and the UES is defined as the average savings per RCx project for all the projects in the population in kW, kWh or therms. These tables also provide estimates of the relative precision around these results where this could be calculated (in the case of SDG&E, for example, one of the two strata had only a single observation, making the calculation of a standard error meaningless.) Realization rates were not calculated for SCE gas savings or SCG electric savings.

As can be seen from the tables, the relative precision for the realization rate and ex post UES for all four sampling domains did not attain the 90/10 goal that we had targeted based upon the ex ante savings values. This was because there was greater variability in the ex post values than in the ex ante values, and suggests that sample designs based upon ex ante savings may underestimate the number of points required. In addition, the PG&E overall kWh realization rate of 0.45 meant that the confidence interval around the realization rate estimate that led to relative precision of 0.23 would have yielded a relative precision closer to ±0.1 if we had been dividing by a realization rate closer to 1.0.

Figure 3 provides a graphical comparison of claimed versus evaluated gross energy savings for each IOU and for the HIM overall. Quantities on the graph are expressed in millions of BTUs saved annually (MMBTU), combining both electrical energy and natural gas energy impacts.

Table 8: Gross Savings Realization Rates

			Gross Savings Realization Rate			90% Relative Precision		
Utility	Population	Sample	kW	kWh	Therms	kW	kWh	Therms
PG&E	135	24	0.31	0.45	0.53	0.51	0.23	0.24
SCE	58	13	2.07	0.94	N/A	0.66	0.17	N/A
SCG	28	10	N/A	N/A	0.93	N/A	N/A	0.23
SDG&E	4	3	2.60	1.23	0.21	0.04	0.06	0.01

Table 9: Gross Ex Post Unit Energy Savings

			Gross Ex Post Unit Energy Savings			90% Relative Precision		
Utility	Population	Sample	kW	kWh	Therms	kW	kWh	Therms
PG&E	135	24	13	178,355	7,334	0.51	0.23	0.24
SCE	58	13	30	383,712	462	0.66	0.17	1.56
SCG	28	10	6	28,781	23,735	0.00	0.00	0.23
SDG&E	4	3	129	606,849	11,454	0.04	0.06	0.01

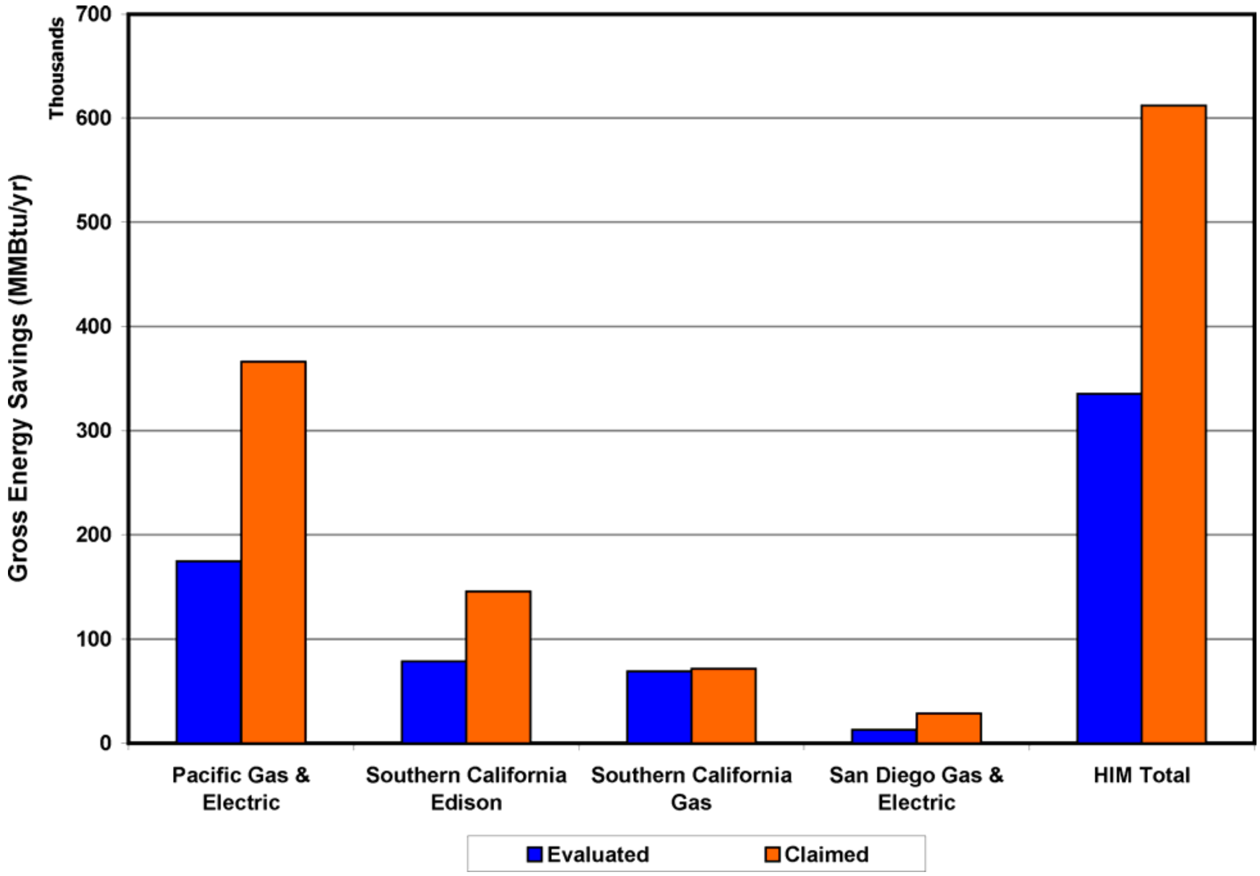


Figure 3: Comparison of evaluated and claimed gross energy savings

A number of key findings shown in these tables and graphs deserve mention:

- The number of RCx projects varied tremendously between IOUs. PG&E claimed 135 projects, nearly two-thirds of the RCx population of 225 projects. In contrast, SDG&E claimed four projects, all of which occurred at the same university campus. Figure 3 illustrates how, on an energy basis, the PG&E RCx projects accounted for about half of the claimed and evaluated savings.
- PG&E projects as a whole had relatively low realization rates for both electric and gas energy savings (0.45 and 0.53), compared to SDG&E, which had realization rates that varied between gas and electric savings; for example, 1.23 for kWh and 0.21 for therms. SCG's single realization rate for therms was 0.93. These differences may reflect the diversity of programs and program delivery models at PG&E. By comparison, the savings claims for SCE and SCG were dominated by two local government partnership programs, UC/CSU/IOU and Los Angeles County.
- IOU-level realization rates for peak kW demand vary widely, from 0.31 to 2.60. Two related factors likely explain this— program implementers using different definitions of peak period in their savings calculations, and a tendency by implementers to be conservative and claim zero peak demand reduction for measures where the actual reduction is very uncertain and/or difficult to estimate.
- The unit energy savings vary widely among IOUs. For single fuel IOUs, lower values are to be expected for any secondary fuel savings claimed, i.e., gas savings from an SCE project, and electric

savings from an SCG project would generally be an afterthought. But even for fuels across comparable utilities, such as electric savings at PG&E, SCE, and SDG&E, the UES vary by a factor of three. Much of this variation likely springs from the fact that some programs, such as the aforementioned local government partnership programs, often had larger, campus-style projects that yielded more claimed savings. This explains the high UES for SDG&E, which included only UC/CSU/IOU Partnership projects.

To determine the major reasons why the claimed and evaluated gross savings were different, we reviewed the site M&V reports and calculations for all gross impact sites. Our review classified differences between claimed and evaluated savings according to two main criteria:

- Whether the difference increased or reduced actual savings, compared to the claim.
- Whether the difference was driven primarily by actions the customer took (such as overriding a recommended fan schedule) or actions the program was responsible for (such as applying too high a chiller efficiency in the savings calculations).

Across the 50 projects in the gross sample, we determined 83 significant reasons for differences, with a relatively even split between customer-driven and program-driven reasons. Over 75% of these reasons worked to reduce savings, a percentage that was fairly uniform across all four IOUs. Critically, nearly half of these savings-reducing reasons were instances where the RCx measure was no longer operational. Put simply, the most common reason why savings fell short of the claim was that measures were not working anymore. Other common reasons for differences included discrepancies between program calculation assumptions and actual conditions, changes in building operation, and measures being only partially implemented. Detailed tables showing our reasons classifications can be found in Appendix 5.5.

2.6.2. IOU Strata-Level Results

Table 10 to Table 13 provide more detailed results for each of the four IOUs. These tables show the strata boundaries, sample counts, unit energy savings (UESs), and realization rates for each stratum's sampled projects (and not the stratum population). They also provide the standard error, confidence interval, and relative precision estimates for the overall realization rate and UES for each IOU sample.

Table 10: PG&E – First Year Gross Savings Parameters

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBTU)		Gross Ex Post Unit Energy Savings2			Gross Savings Realization Rate2		
Utility	Stratum	Population	Sample	Lower	Upper	kW	kWh	Therms	kW	kWh	Therms
PG&E	1	22	2	264	1,414	4	96,081	2,112	0.52	1.44	3.12
PG&E	2	25	2	1,444	2,872	17	164,156	0	0.10	0.23	0.00
PG&E	3	12	2	3,044	6,057	42	218,563	1,216	0.99	0.61	0.07
PG&E	9	10	9	6,317	17,224	67	815,551	52,532	0.55	0.64	1.18
PG&E	Phase 2 - 1	45	2	78	1,928	1	44,898	0	1.16	0.80	0.53
PG&E	Phase 2 - 2	16	2	1,964	5,885	0	166,506	6,740	0.31	0.36	0.35
PG&E	Phase 2 - 9	5	5	9,944	26,651	47	400,001	44,357	0.60	0.33	0.34
PG&E	Excluded	2	0	5	55						

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBTU)		Gross Ex Post Unit Energy Savings2			Gross Savings Realization Rate2		
Utility	**Stratum**	**Population**	**Sample**	**Lower**	**Upper**	**kW**	**kWh**	**Therms**	**kW**	**kWh**	**Therms**
PG&E	Phase 2 - Excl.	14	0	0	61						
PG&E	**All Sampled**	**135**	**24**	**0**	**26,651**	**13**	**178,355**	**7,334**	**0.31**	**0.45**	**0.53**
					Statistics						
					Standard Error	**4**	**25,434**	**1,089**	**0.10**	**0.06**	**0.08**
					90% Confidence Interval	**7**	**41,840**	**1,791**	**0.16**	**0.11**	**0.13**
					Relative Precision	**0.51**	**0.23**	**0.24**	**0.51**	**0.23**	**0.24**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

2 Stratum level results are based on sampled sites only.

Table 11: SCE – First Year Gross Savings Parameters

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBtu)		Gross Ex Post Unit Energy Savings2			Gross Savings Realization Rate2		
Utility	**Stratum**	**Population**	**Sample**	**Lower**	**Upper**	**kW**	**kWh**	**Therms**	**kW**	**kWh**	**Therms**
SCE	1	29	3	83	793	18	275,414	923	1.34	1.84	
SCE	2	13	3	800	1,939	14	276,635	0	1.63	0.88	
SCE	3	12	3	2,048	3,344	56	526,471	0	-168.78	0.72	
SCE	9	4	4	4,290	10,607	1	1,319,758	0	0.02	0.71	
SCE	Excluded	2	0	5	41						
SCE	Phase 2 - Excl	1	0	2,830	2,830						
SCE	**All Sampled**	**58**	**13**	**5**	**10,607**	**30**	**383,712**	**462**	**2.07**	**0.94**	
					Statistics						
					Standard Error	**12**	**40,762**	**437**	**0.82**	**0.10**	
					90% Confidence Interval	**19**	**67,054**	**719**	**1.35**	**0.16**	
					Relative Precision	**0.66**	**0.17**	**1.56**	**0.66**	**0.17**	

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

2 Stratum level results are based on sampled sites only.

Table 12: SCG – First Year Gross Savings Parameters

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBTU)		Gross Ex Post Unit Energy Savings2			Gross Savings Realization Rate2		
Utility	**Stratum**	**Population**	**Sample**	**Lower**	**Upper**	**kW**	**kWh**	**Therms**	**kW**	**kWh**	**Therms**
SCG	1	13	3	240	1,744	0	0	7,169			0.81
SCG	2	10	2	1,753	3,621	0	0	15,193			0.57
SCG	9	5	5	4,524	18,147	31	161,173	81,464			1.27
SCG	Phase 2 - Excl	7	0	0	2,695						
SCG	**All Sampled**	**28**	**10**	**240**	**18,147**	**6**	**28,781**	**23,735**			**0.93**
					Statistics						
					Standard Error			**3,338**			**0.13**

	RCx Projects			Stratum Boundaries Ex Ante Savings1 (MMBTU)		Gross Ex Post Unit Energy Savings2			Gross Savings Realization Rate2		
Utility	Stratum	Population	Sample	Lower	Upper	kW	kWh	Therms	kW	kWh	Therms
					90% Confidence Interval			5,491			0.22
					Relative Precision	0.00	0.00	0.23			0.23

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

2 Stratum level results are based on sampled sites only.

Table 13: SDG&E – First Year Gross Savings Parameters

	RCx Projects			Stratum Boundaries Ex Ante Savings1 (MMBTU)		Gross Ex Post Unit Energy Savings2			Gross Savings Realization Rate2		
Utility	Stratum	Population	Sample	Lower	Upper	kW	kWh	Therms	kW	kWh	Therms
SDG&E	1	2	1	1,575	2,034	27	101,740	2,820	1.48	0.56	0.17
SDG&E	9	2	2	4,319	24,959	240	1,157,817	20,384	2.84	1.37	0.21
SDG&E	Phase 2 - Excl	3	0	1,155	2,123						
SDG&E	**All Sampled**	**4**	**3**	**1,575**	**24,959**	**129**	**606,849**	**11,454**	**2.60**	**1.23**	**0.21**
					Statistics						
					Standard Error	**3**	**20,604**	**92**	**0.07**	**0.04**	**0.00**
					90% Confidence Interval	**6**	**33,893**	**151**	**0.12**	**0.07**	**0.00**
					Relative Precision	**0.04**	**0.06**	**0.01**	**0.04**	**0.06**	**0.01**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

2 Stratum level results are based on sampled sites only.

2.6.3. System and Measure Class Results

To better understand the types of measures and systems that are yielding RCx savings, we developed a detailed classification scheme for all expected RCx and Major Commercial measures. This scheme standardized the designation of the building system and general strategy addressed by each measure. We examined the final data from all gross impact sample sites to determine which types of measures were most common. We then aggregated system and measure classes into broader groups (e.g., central plant for system; improve scheduling for measure) for analysis purposes. Section 5.5 in the appendix contains additional information about how this was accomplished.

Table 14 and Table 15 below present the percentage breakdown of ex post savings by system and measure class in the RCx HIM group (i.e., across all four IOUs). Note that measure counts are a fairly imprecise parameter for describing the level of RCx activity, since program implementers often define measures differently (e.g., at one site, an action affecting four air handlers might be considered one measure; at another, the same action might be considered four measures). Nonetheless, they do provide a general idea of where RCx efforts were focused.

Extrapolating the sample to the RCx HIM population, we estimated that the 225 projects comprised 623 measures, or slightly less than three per project. This includes unclassifiable projects, where individual measures could not be broken out for analysis purposes and the project was counted as having one unclassified measure. The most common system class was HVAC air distribution systems; the most common measure class was improving control strategies, with each accounting for more than one-third of measures installed. Generally, though, the measures and savings appeared distributed fairly evenly among the classes, so that no one class dominated.

Table 14: Percent Ex Post Savings by Building System Class

	% of			
Type of System	**Measures**	**kW**	**kWh**	**Therms**
Central plant	19	22	20	19
HVAC (general)	29	18	12	41
HVAC (air distribution system)	36	18	37	34
Other/unclassified	17	41	32	6

Table 15: Percent Ex Post Savings by Measure Class

	% of			
Type of Measure	**Measures**	**kW**	**kWh**	**Therms**
Improve control strategies	36	38	29	48
Improve outside air use	13	9	14	4
Improve scheduling	18	-3	11	28
Other/unclassified	17	40	29	12
Install/replace variable speed drive	16	15	16	7

We had hoped through this analysis to identify particular measure or system classes that might be particularly ripe targets for achieving RCx savings—if, for instance, we had observed that 60% of the kWh savings resulted from the 10% of central plant measures. The numbers in these tables, however, revealed no such obvious targets for future programs.

2.6.4. Spillover Results

During onsite visits, the site lead evaluators asked facility contacts for sampled projects about spillover while collecting data onsite. This process uncovered three sites with possible spillover, and interviews with decision makers confirmed that the actions in question had been significantly influenced by participation in the RCx program. For these cases, we analyzed energy savings using methods that were more simplified than the direct measures because spillover quantification was a secondary objective of the study. One of the three sites ultimately fell away when analysis showed no spillover savings.

The first measure took place at a hospital, which received RCx services through the Pacific Gas & Electric Hospital Pilot Program (PGE2072). At this site, the owner updated controls to be able to better schedule four packaged rooftop HVAC units, which originally operated 24/7. A simple evaluation engineering analysis estimated savings of 22,291 kWh/year and 2,617 therms/year from this action.

The second measure occurred at an office complex, which received RCx services through the Pacific Gas & Electric PECI Retrocommissioning Services and Incentives Program (PGE2091). At this site, the owner made an identical economizer repair to an air handler not treated by the original RCx effort. The evaluation engineering analysis estimated savings of 410 kWh/year from this action.

Our overall finding, then, was that spillover impacts appeared to be insignificant.

2.7. Recommendations

In developing recommendations for the Gross Savings Evaluation, we considered the implications of our findings for three aspects of the evaluation: program design and implementation, program evaluation, and future research needs. Each of these aspects is discussed below.

2.7.1. Program-Related Recommendations

While this was clearly not a process evaluation, and the focus was on measuring savings rather than assessing the effectiveness of program delivery, there were nevertheless some findings that have implications for the mechanism by which the RCx HIM is delivered. Note that these recommendations also tie into the evaluation-related recommendations presented below.

- **Provide program participants with adequate follow-up RCx services**. Once RCx service providers have identified RCx opportunities, maintaining the value of those findings requires sustaining a long-term relationship with customers to make sure the measures are implemented correctly and maintained properly over time. RCx is an incremental process that needs to be done over a longer period of time, rather than through a one-time process of dropping in, making recommendations and then moving on. This evaluation found frequent examples where measures failed soon after implementation. A particular example is economizer repair measures[1], which comprise a significant fraction of all RCx measures. Although these measures had been vetted by program staff, they often had failed, were only partly functional, or had faulty programming that meant they produced no savings a relatively short time after project completion. In many cases, more sustained follow-up from the program might have eliminated these problems. Program designs should provide latitude for RCx service providers to work with customers to "dial in" measures; that is, the original measure recommendation may need several tweaks and adjustments over many months of operations before facility staff consider it truly acceptable and sustainable. It may be appropriate to consider a payment schedule that withholds a fraction of final payment to the RCx provider until such follow-up activities are completed and documented for vulnerable measures such as economizer repair.
- **Reduce RCx service providers' burden for quantifying energy savings**. The corollary to the recommendation above is that programs should be designed to minimize the RCx service provider's

[1] Economizers on HVAC air handling units include dampers that permit the use of cool outside air directly to reduce indoor space temperatures, rather than relying on the mechanical cooling system to do so. Repairing malfunctioning dampers, as well as ensuring that the corresponding controls are working properly, can yield substantial electric savings.

responsibilities to perform rigorous calculations and analysis to back up utilities' claimed savings. Estimating savings to the level of rigor necessary for a utility claim is a complex, challenging endeavor that requires specialized analytical skills. While some RCx service providers might be able to meet this requirement, many may not, as it is not their primary skill set. It might make more sense to have utility staff or their consultants perform separate post-implementation M&V studies to back up their savings claims.

RCx service providers do need simple, straightforward tools that allow them to quantify costs and savings well enough to justify projects to customers. This is particularly true for well-established, clearly cost-effective measures such as changing HVAC setpoints or schedules. These tools might include broadly accepted methods for estimating complex parameters that are critical to good estimates of savings, such as boiler and chiller efficiencies, and the effect of cogeneration systems.

- **Give program staff primary responsibility for collecting baseline data**. Program staff and the RCx service providers are in the best position to collect all-important baseline data. In fact, if impact evaluations use a post-only design, RCx providers essentially become the only source of the data. Program implementers should bear primary responsibility for collecting and clearly documenting and archiving baseline information that will be necessary in the future to verify achieved savings. This data collection would not have to be onerous, but would mainly derive from information and situations that RCx service providers would naturally come across during their investigations. Such items might include photos, notes, seasonal trend data and screen shots from building energy management systems, and important one-time measurements. If program implementers and evaluators could agree upon the necessary elements, it could simplify the M&V process for all.

2.7.2. Evaluation-Related Recommendations

Several recommendations are based on specific issues encountered during the current evaluation, in the hope that future evaluations may be able to avoid some of these problems.

- **Improve baseline data collection**. Good baseline data is critical to accurate savings estimates, particularly for measures where the baseline is inherently uncertain, such as those involving broken dampers, stuck valves, or sequencing strategies. As noted above, it is generally most expeditious to have customers and/or RCx service providers collect this information. Equally important, baseline data should be consistently recorded and maintained in program tracking systems.
- **Specify post-only sample designs**. The 2006–08 program cycle once again demonstrated how, with RCx programs, it is nearly impossible until after the program ends to know which projects, and which measures within those projects, will be claimed as complete. The resources required for a pre-post sample design—including tracking projects throughout the cycle and collecting baseline data for projects that ultimately fall away—could better be spent increasing the sample size in a post-only design.
- **Balance the need for accurate first-year savings against the need to track savings over time.** While it is important to develop rigorous estimates of first-year gross savings for RCx projects, it is equally important to understand how those savings change over time. The allocation of evaluation resources should reflect this. Complex commercial facilities can be very dynamic, so that building

conditions and uses change regularly. Additionally, as both the gross impact and EUL portions of this evaluation have shown, a significant number of RCx measures fail within a year or two of implementation. Therefore, tracking how RCx savings degrade and how programs might be designed to minimize this degradation through appropriate interventions should be an important objective of future evaluations.

- **Maximize time allotted for onsite data collection**. The evaluation schedule should allow as much calendar time as possible for field data collection, to support seasonal analyses and the oftentimes iterative process of data collection, analysis, and quality control. For that to happen, up-front planning and approval activities should be kept to a minimum. Longer data collection periods can improve the quality of the evaluated savings estimates, as more representative data can be collected. Moreover, onsite conditions and constraints will often necessitate changes to original M&V plans, so it is important to allow adequate time for such variations. Finally, providing flexibility in scheduling with customers may help reduce the inconvenience to customers and improve the likelihood that they will cooperate during the M&V process.
- **Minimize the use of whole-building analysis.** Using billing records or interval data to estimate savings, per IPMVP Option C, can be appropriate in limited circumstances. But since this approach does not analyze how individual devices and systems are functioning, it makes it nearly impossible to determine whether particular measures are functioning well and the reasons why. Also, it is difficult to adjust with much confidence for the wide variety of external factors that could also be changing facility-level energy use; e.g., changes in building occupancy or usage, operational changes that the customer may not think significant, and meter change-outs, to name just a few. Factors such as these can lead to a high degree of uncertainty in the analyzed savings.

2.7.3. Future Research–Related Recommendations

- **Continue refining the measure classification scheme**. This evaluation developed a general scheme for grouping RCx measures. This scheme filled a need, since RCx measures are fundamentally different from conventional retrofit program measures for which other classification schemes already exist. Applying the scheme to the evaluation sample provided a way to assess what kinds of measures are most prevalent. Further analysis of these data could determine the amount of savings that a particular measure at a particular size of site might yield. This in turn could help program implementers and evaluators focus their resources on the most attractive measures. Additional work on the classification scheme, if undertaken in the near future, could be applied to the next round of RCx programs and evaluations.
- **Study the relative effectiveness of different programmatic approaches**. The population of 225 RCx projects spanned more than two dozen programs that used diverse approaches and delivery strategies. Some were strictly RCx, while for others RCx was a very small component. Some programs were general purpose and encompassed many types of commercial buildings, while others targeted narrow niches, such as hospitals or data centers. Still others were partnerships with public agencies, and others featured a monitoring-based approach to install permanent metering. A combined process and impact evaluation that compared results for different RCx approaches could

yield insights into best practices and effective designs for future RCx programs. Process evaluations for this program cycle have dealt with particular programs, so there is no overarching sense of how RCx overall fared from a process standpoint. It is also advisable to link the process and impact evaluations of RCx, when possible. The highly technical and complicated nature of RCx projects often requires process evaluators to work closely with gross impact evaluators with strong engineering backgrounds to assess the programs.

- **Compare Retro- and Monitoring-based Commissioning**. A related research issue is the comparative efficacy of monitoring-based commissioning (MBCx) and standard RCx in creating and maintaining savings. On the one hand, MBCx provides facility managers with powerful tools to sustain energy savings; on the other hand, utilizing those tools effectively may require more time and training than building professionals typically have available. A systematic comparison of the persistence of savings obtained through the two approaches should be included in EUL studies such as those discussed in section 4.7.

3. Net Savings Evaluation

3.1. Evaluation Objectives

The goal of the RCx net impact evaluation was to determine how much of the calculated gross savings should be attributed to the intervention of the program; that is, what fraction of the savings would not have been achieved if the program had not been in place and therefore represents the net savings attributable to the program. This was done by estimating project-specific net-to-gross ratios (NTGRs) for both the gross and net impact samples.

3.2. Sample Design and Selection

The sample design for the net impact evaluation followed the same basic approach used for the gross sample and described in section 2.2. For the net-only sample that received the Basic rigor level data collection treatment, additional sites were randomly selected from each stratum so that approximately 50% of the population was included. The resulting net impact sample is shown in Table 16.

Table 16: Net Impact Evaluation Sample

		Stratum Boundaries Ex Ante Savings[1] (MMBTU)		RCx Projects			% of Ex Ante Savings in Sample		
Utility	**Stratum**	**Lower**	**Upper**	**Population**	**Sample**	**% in Sample**	**kW**	**kWh**	**Therms**
PG&E	1	264	1,414	22	2	9	2	4	3
PG&E	2	1,444	2,872	25	2	8	25	15	3
PG&E	3	3,044	6,057	12	2	17	13	11	22
PG&E	9	6,317	17,224	10	9	90	78	98	80
PG&E	Phase 2 - 1	78	1,928	45	2	4	0	2	0
PG&E	Phase 2 - 2	1,964	5,885	16	2	13	0	10	18
PG&E	Phase 2 - 9	9,944	26,651	5	5	100	100	100	100
PG&E	Excluded	5	55	2	0				
PG&E	Phase 2 - Excl.	0	61	14	0				
PG&E	**All Sampled**	**0**	**26,651**	**135**	**24**	**18**	**33**	**39**	**61**
SCE	1	83	793	29	3	10	19	14	0
SCE	2	800	1,939	13	3	23	19	20	6
SCE	3	2,048	3,344	12	3	25	0	26	11
SCE	9	4,290	10,607	4	4	100	100	100	100
SCE	Excluded	5	41	2	0				
SCE	Phase 2 - Excl	2,830	2,830	1	0				
SCE	**All Sampled**	**5**	**10,607**	**58**	**13**	**22**	**26**	**46**	**38**
SCG	1	240	1,744	13	3	23	NA	NA	18
SCG	2	1,753	3,621	10	2	20	NA	NA	21
SCG	9	4,524	18,147	5	5	100	NA	NA	100

		Stratum Boundaries Ex Ante Savings[1] (MMBTU)		RCx Projects			% of Ex Ante Savings in Sample		
Utility	Stratum	Lower	Upper	Population	Sample	% in Sample	kW	kWh	Therms
SCG	Phase 2 - Excl	0	2,695	7	0				
SCG	**All Sampled**	**240**	**18,147**	**28**	**10**	**36**	**NA**	**NA**	**51**
SDG&E	1	1,575	2,034	2	1	50	63	63	55
SDG&E	9	4,319	24,959	2	2	100	100	100	100
SDG&E	Phase 2 - Excl	1,155	2,123	3	0				
SDG&E	**All Sampled**	**1,575**	**24,959**	**4**	**3**	**75**	**90**	**89**	**78**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

Sample Disposition. As described in section 3.3.2, steps were taken to minimize non-response bias, and NTGR interviews were completed with 90% of customers we attempted to reach.

3.3. Methodology

The net analysis of the RCx HIM utilized the same self-report approach developed and applied to all 2006–08 Large Nonresidential measures and programs. The self-report option was chosen because alternative methods of estimating the NTGR, such as discrete choice or billing analysis were not practical in light of the limited number of projects, the heterogeneity of participating customers, and the small impacts relative to overall energy usage for many projects.

3.3.1. Overall Approach, Rigor Levels and Protocols

Two levels of rigor were used in the net analysis. The Standard - Very Large rigor level integrated information from other sources besides the customer interview, thereby allowing us to tell the full story behind each organization's decision to proceed with the RCx study and recommendations and the role that the programs played in the causing the work to occur. The second, Basic, rigor level used the same standard data collection instrument and algorithm to calculate the NTGR, but did not bring in the additional noncustomer viewpoints to support the analysis.

The HIM net impact evaluation interviewed all impact sample participants using the Standard – Very Large NTG method, rather than just the large sites. We randomly selected from among the remaining sites (by stratum) so that approximately 50% of all sites were treated. These randomly selected sites were interviewed using the Basic NTG method.

The Nonresidential NTGR methodology used for the RCx net impact evaluation was developed to address the unique needs of large nonresidential customer projects developed through energy efficiency programs offered by the four California investor-owned utilities and third parties. This method relied exclusively on the Self-Report Approach (SRA) to estimate project and program-level Net-to-Gross Ratios (NTGRs), since other available methods and research designs are generally not feasible for large nonresidential customer programs. This methodology provided a standard framework, including decision rules, for integrating findings from both quantitative and qualitative information in the calculation of the net-to-gross ratio in a systematic and consistent manner. This approach was designed to fully comply with the *California Energy Efficiency Evaluation: Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals* (Protocols) and the *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches*[2] (Guidelines).

This approach preserved the most important elements of the approaches previously used to estimate the NTGRs in large nonresidential customer programs, and incorporated several enhancements designed to improve upon that approach, such as:

- The method introduced a 0 to 10 scoring system for key questions used to estimate the NTGR, rather than using fixed categories that were assigned weights (as was done previously).
- The method asked respondents to jointly consider and rate the importance of the many likely events or factors that may have influenced their energy efficiency decision-making, rather than focusing narrowly on only their rating of the program's importance. This question structure more accurately reflected the complex nature of the real-world decision making and helped to ensure that all non-program influences were reflected in the NTGR assessment in addition to program influences.

A detailed description of the nonresidential NTGR approach, including the theoretical basis for the self-report approach, the specific questions asked, and the algorithm used to calculate the NTGR, is provided in the appendix in section 5.4.5. It is important to note that the NTGR approach described there is a general framework, designed to address all large nonresidential programs. In order to implement this approach for the RCx net impact evaluation, it was customized somewhat to reflect the unique nature of RCx projects, as follows:

- Instead of a single question regarding the importance of the rebate in their decision, participants were asked about the influence of both the RCx study incentive and the implementation incentive on their decision to perform RCx at their facility.
- For high-cost measures (>$10,000) implemented through a RCx program, customers were asked if they had been considering this specific measure prior to the RCx study; if they had been, they were asked a separate battery of questions to ascertain the extent to which the RCx program influenced their decision to implement that particular measure (as distinct from their decision to retro-commission the whole facility), resulting in a separate NTGR for the savings associated with that measure. Where both overall and measure-specific NTGRs were calculated, the project NTGR was based on the individual NTGRs and associated ex post savings as well as on the overall NTGR.
- Decision makers who said they would have performed RCx on their facility even in the absence of the program were asked to specify how much they would have spent on that RCx effort as a means of assessing partial free ridership.

[2] California Public Utilities Commission Energy Division and the Master Evaluation Contractor Team, *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches*, October 15, 2007

3.3.2. Data Collection – Net Impact Evaluation

For the RCx net savings sample, data collection was exclusively through telephone interviews conducted by professional staff, with a number of steps taken to reduce non-response bias. A potential source of non-response bias is the inability or unwillingness of targeted decision makers to participate in the survey. This becomes a source of bias if we believe the program participants who did not respond would likely have responded differently to survey questions than participants who were sampled as replacements and did respond. If external factors (business conditions, staff turnover) lead to some respondents becoming unavailable, those factors may not lead to different responses regarding program influence. On the other hand, if a participant chooses not to respond, there may be underlying perceptions of utility or program influence that differ from those of more cooperative respondents, which could lead to biased results.

There are several reasons why the appropriate decision maker might not respond. These reasons are explained below, along with the methods that we deployed to minimize their contribution to the non-response rate.

- The firm that made the original decision has gone out of business or been sold, so that none of the individuals responsible for the decision to have the building retro-commissioned is available and the personnel of the new owner (if any) have no knowledge of the decision.
- More common is the situation where one of the individuals responsible for making a decision has since left that position, either through retirement, transfer within the company, or leaving for another firm. When this occurs, utility account representatives and third party program representatives can often help identify individuals who can provide informed responses to NTGR questions. We encountered several cases where the original decision maker was not available. In most of those cases we were able to identify a suitable substitute who was involved in the decision and knew about the decision criteria that led to a project being implemented, but there were a few instances where no one knowledgeable about the project was available so the interview could not be completed. One of the most effective techniques to counter this source of non-response is to conduct interviews as soon as possible after the participation decision is made; unfortunately, changes in evaluation plans and procedures led to significant delays before the decision maker was contacted, which is why we encountered some cases of personnel turnover.
- Another potential source of non-response is the unwillingness of decision makers to discuss their decision to RCx their facility, either because they consider this information confidential or because they do not have time to complete the interview. In the professional interviews conducted for the RCx net analysis, we minimized non-response from this source by: (1) providing maximum flexibility in scheduling, allowing respondents to pick a time up to three weeks out that best fit their needs, (2) working through the account or program representative to alert the decision maker to the coming call and encourage their participation (3) pointing out that by participating in a public-benefits-charge-funded program, they explicitly agreed to cooperate with subsequent program evaluations, (4) assuring the respondent that any confidential or financial data will not be publicly disclosed, and (5) having professional staff with extensive knowledge of the program and the customer's project conduct the interview. There were a few cases, however, in which customers refused to respond when our interviewer contacted them, even after they had been contacted by one or more interviewers for

other programs they had participated in. In addition, a few decision makers could not be reached, or failed to return calls or emails after repeated five or more attempts.

Through these efforts, we obtained a high response rate and minimized the risk of non-response bias in the results. The disposition of data collection calls for the NTGR interviews are presented in Table 17. Note that 90% of attempted customer contacts ultimately resulted in completed surveys.

Table 17: Disposition of NTGR Interview Calls

Disposition	Gross Sites	Net Only Sites	Total
Completed	**48**	**72**	**120**
Could not reach a knowledgeable respondent		4	4
Respondent refused to participate	1	2	3
No response after multiple attempts	1	5	6
Total Not Completed	**2**	**11**	**13**

3.4. Confidence and Precision of Key Findings

The mean NTGR was calculated based on the NTGR sample, which used the same stratification defined for the gross impact sample for this project. However, larger samples were taken within each stratum in order to attain the targeted 50 percent of the population in each stratum. As noted previously, the gross sample points were allocated to achieve 90/10 overall precision using the ex ante estimates of gross savings. Since there were no site-specific ex ante estimates of NTGR, the same approach could not be used for the expected precision of the net sample. We determined, however, that the 90/10 precision for a gross sample of 50 of 252 points was consistent with an error ratio of about 0.5, which would imply 90/5.4 precision for the net sample of 123 out of 252 points.

Calculation of the achieved precision for the NTGR was done using the Cochran (1977) method of calculating means and standard errors for a stratified random sample (see appendix section 5.2 for details). This approach involved using stratum weights (the proportion of the total population accounted for by each stratum) to calculate a weighted mean, variance and standard error. The resulting mean and standard error were used to calculate the error band and relative precision for the NTGR.

3.5. Validity and Reliability

As discussed in section 3.3, the self report approach (SRA) was selected for this evaluation as the best method for estimating net effects of the program or programs evaluated. The term "self-report" as applied to the nonresidential sectors understates the methodological complexity of this approach as applied in the current evaluation, which has deep roots in the social sciences and is widely used by the evaluation community. To suggest that it only involves asking one key decision maker to hypothesize about what equipment they would have installed in the absence of the program is misleading.

More specifically, the SRA as applied in this evaluation is a mixed method approach that involves asking key decision makers and decision influencers a series of structured and open-ended questions about their motivations for pursuing the retro-commissioning project and whether they would have pursued the same project in the absence of the program. A central component of this approach is to ask questions that attempt to rule out rival explanations for the retro-commissioning activity. In the case of large nonresidential customers participating in programs with a retro-commissioning component, the SRA was strengthened by the inclusion of additional quantitative and qualitative data sources, including, among others, in-depth, open-ended interviews, direct observation, review of customer and program records and analysis of industry and company data. Such qualitative data regarding the customer's decision and the decision process itself can be very useful in supporting or modifying quantitatively-based results and can help increase the reliability of results.

There have been a number of challenges to the SRA because we are not only asking participants to recall what has happened in the past, we are asking them, among other things, to report on a hypothetical situation; that is, what they would have done in the absence of the program. In many cases, the respondent may simply not know and/or cannot know what would have happened in the absence of the program. Even if the customer has some idea of what would most likely have happened, there is, of necessity, uncertainty about it. The situation just described creates potential for invalid answers (low construct validity) and answers with low reliability, where reliability is defined as the likelihood that a respondent will give the same answer to the same question whenever or wherever it is asked. Where we are asking for motivations and processes in hypothetical situations that occurred one or two years ago, there is room for bias. Examples include the following:

- Some respondents may believe that claiming no impact for the program is likely to cause the program to cease, thus removing future financial opportunities from the respondent, which would lead them to overstate the degree of program influence and raise the NTGR.
- On the other hand, some people may want to portray themselves in a positive light in that they would have installed energy-efficient equipment without any incentive (the socially desirable response), which could result in an artificially low net-to-gross ratio.
- The third hypothesized source of bias involves an interaction between the positive perception of taking energy efficiency actions, the often-observed difference between stated intentions and actual behaviors, and the fact that the counterfactual outcome cannot be viewed by the participant or outsiders.
- Another hypothesized source of bias arises when participants are asked to identify the reasons why they installed the energy efficient measures, since respondents might not always be able recall all the possible reasons and influences and rank each in terms of its importance.

In designing the approach to the calculation of the NTGR for the RCx HIM, we were very aware of these issues and took a number of actions to mitigate potential problems. Specifically, we followed the steps outlined in the *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches*, which was commissioned by the Energy Division of the CPUC to address the challenges listed earlier with respect to reliability and validity. Among the steps noted in the Guidelines were the following: building in consistency checks, using multiple questions, employing triangulation, ruling out of rival hypotheses, using a combination of quantitative and qualitative data, conducting sensitivity analysis, incorporating other documentation such as a company's procurement policies and standard practice in a particular industry, and, for situations when substantial savings are being claimed, using two analysts to independently review the data collected. These recommendations were all incorporated into the non-residential SRA as applied to the RCx HIM.

For the Standard-Very Large rigor cases that comprised the entire gross impact sample, we also incorporated interviews with program staff and/or account reps, reviewed retro-commissioning documents and talked with site engineers as needed both to ensure that the right respondent was reached and to provide context for the overall decision. This enabled us to develop an internally consistent, plausible "story" behind each project. Both by incorporating input from multiple other sources and by encouraging the decision maker to think about and weigh the full range of factors influencing their equipment installation decision and give them the opportunity to fully explain their situation, we believe we have minimized the likelihood of bias in survey responses and enhanced the reliability of the results.

3.6. Detailed Findings

The tables below document the findings of this research related to NTGR. Each table presents our findings for one of the four IOUs that implemented RCx projects.

Table 18: PG&E – NTGR Parameters

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBTU)		NTGR Weighted by…			
Utility	**Stratum**	**Population**	**Sample**	**Lower**	**Upper**	**Case Weight**	**kW**	**kWh**	**Therms**
PG&E	1	22	11	264	1,414	0.83	0.57	0.75	0.98
PG&E	Phase 2 - 1	45	22	78	1,928	0.77	0.78	0.80	0.78
PG&E	Phase 2 - 2	16	8	1,964	5,885	0.80	0.84	0.83	0.77
PG&E	Phase 2 - Excl.	14	0	0	61				
PG&E	Phase 2 - 9	5	5	9,944	26,651	0.87	0.89	0.88	0.88
PG&E	2	25	12	1,444	2,872	0.86	0.82	0.81	0.94
PG&E	3	12	6	3,044	6,057	0.73	0.72	0.71	0.83
PG&E	Excluded	2	0	5	55				
PG&E	9	10	9	6,317	17,224	0.88	0.82	0.90	0.87
PG&E	**All Sampled**	**135**	**73**	**0**	**26,651**	**0.81**	**0.76**	**0.80**	**0.86**
						Statistics			
						Standard Error	**0.02**	**0.02**	**0.01**
						90% Confidence Interval	**0.03**	**0.03**	**0.02**
						Relative Precision	**0.04**	**0.03**	**0.02**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

Table 19: SCE – NTGR Parameters

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBtu)		NTGR Weighted by…			
Utility	Stratum	Population	Sample	Lower	Upper	Case Weight	kW	kWh	Therms
SCE	1	29	12	83	793	0.82	0.81	0.82	0.89
SCE	2	13	7	800	1,939	0.90	0.64	0.90	0.93
SCE	3	12	6	2,048	3,344	0.86	0.83	0.86	0.93
SCE	Excluded	2	0	5	41				
SCE	Phase 2 - Excl	1	0	2,830	2,830				
SCE	9	4	4	4,290	10,607	0.95	0.93	0.94	0.93
SCE	**All Sampled**	**58**	**29**	**5**	**10,607**	**0.86**	**0.78**	**0.86**	**0.91**
						Statistics			
						Standard Error		**0.01**	**0.01**
						90% Confidence Interval		**0.02**	**0.01**
						Relative Precision		**0.02**	**0.01**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

Table 20: SCG – NTGR Parameters

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBTU)		NTGR Weighted by…			
Utility	Stratum	Population	Sample	Lower	Upper	Case Weight	kW	kWh	Therms
SCG	1	13	5	240	1,744	0.93			0.93
SCG	2	10	5	1,753	3,621	0.90			0.90
SCG	Phase 2 - Excl	7	0	0	2,695				
SCG	9	5	5	4,524	18,147	0.96			0.94
SCG	**All Sampled**	**28**	**15**	**240**	**18,147**	**0.92**			**0.92**
					Statistics				
					Standard Error				**0.01**
					90% Confidence Interval				**0.01**
					Relative Precision				**0.01**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

Table 21: SDG&E – NTGR Parameters

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBTU)		NTGR Weighted by…			
Utility	Stratum	Population	Sample	Lower	Upper	Case Weight	kW	kWh	Therms
SDG&E	1	2	1	1,575	2,034	0.98	0.98	0.98	0.98
SDG&E	Phase 2 - Excl	3	0	1,155	2,123				

		RCx Projects		Stratum Boundaries Ex Ante Savings1 (MMBTU)		NTGR Weighted by…			
Utility	Stratum	Population	Sample	Lower	Upper	Case Weight	kW	kWh	Therms
SDG&E	9	2	2	4,319	24,959	0.62	0.51	0.51	0.37
SDG&E	**All Sampled**	**4**	**3**	**1,575**	**24,959**	**0.80**	**0.75**	**0.75**	**0.68**
						Statistics			
						Standard Error	**0.00**	**0.00**	**0.00**
						90% Confidence Interval	**0.00**	**0.00**	**0.00**
						Relative Precision	**0.00**	**0.00**	**0.00**

1 For most of the strata, boundaries were determined using the best estimate of program tracking impacts available at the time the sample was designed rather than the actual claimed savings. Phase 2 strata were drawn when additional RCx projects were revealed at the end of 2008. The final claimed ex ante savings were used in assigning the Phase 2 sample to strata.

Overall, NTGR scores for the RCx HIM are consistently high, reflecting the continued influence of a variety of programs on the motivation and ability of organizations to pursue RCx projects. With only a single exception, NTGR scores averaged more than 0.50 for every fuel type in every stratum, and the overall mean was significantly higher for all IOUs.

While SDG&E had a somewhat lower average NTGR than the other IOUs, this was because there were only three projects in the sample, including the commissioning of a central plant that had already been planned and was submitted to the Partnership program only after a vendor had told the facility managers about the availability of incentives though that program. While the program may have accelerated the project and helped ensure its approval by campus funding sources, the overall degree of program influence was relatively low, with the resulting project NTGR of 0.26 contributing to the lower SDG&E average.

In contrast, the relatively higher overall mean SCE and SCG NTGR scores reflect the influence of the UC/CSU/IOU and Local Government partnership programs, which appear to strongly influenced facilities that participated in those programs.

Finally, the slightly lower mean NTGRs for PG&E sites are indicative of the more diverse mix of participating projects, since customers completed RCx projects through 19 different programs. More of these projects may have been influenced by non-program concerns such as occupant comfort, client/tenant/customer perceptions and equipment reliability.

As noted above, NTGR values are relatively high across IOUs. Based upon interviews with decision makers at participating organizations as well as review of other secondary data, it appears that several factors contribute to the high average NTGR.

- Several of the largest programs selected projects and submitted applications relatively early in the program cycle specifically as part of a coordinated effort to secure Program funding for their projects. Using a consistent process for all sites, these projects were triggered by and contingent on the availability of program funding, so that the NTGR scores for all projects submitted as part of this process would be expected to be high.

- From review of other data and questions to decision makers regarding RCx activities outside of utility or other programs either in California or other parts of the country, it appears that the level of such RCx activity independent of utility program continues to be very low.
- For example, a representative of one hotel chain said his company was developing a national RCx initiative, and when asked about how many of its facilities had been retro-commissioned, he mentioned six that had been done in the previous year. Upon further investigation, it was found that five of those six projects had been done through California utility programs, and the one that had been done in another state with no program had a budget about 3% as large as the average of the other projects.
- Similarly, the decision maker for a number of RCx projects for a department store chain reported that he has had very little success encouraging RCx in other areas unless thet activity is directly supported by other utility programs, as is the case in the Northeastern U.S..

Many of the organizations implementing RCx projects in the 2006–08 program cycle were public or quasi-public institutions (government agencies, hospitals, colleges) that are generally capable of accepting longer paybacks on energy efficiency projects in accordance with the cost and source of their funding (e.g., bond issues). This raises the question (asked of a number of decision makers) of why these projects would not be pursued without access to the program. Several responses help explain the continued reliance on utility programs and therefore the high NTGR.

- One reason is the uncertainty and risk associated with RCx projects. A thorough RCx study of a large, complex building typically costs tens of thousands of dollars, and there is no certainty that the study will pay for itself through the potential savings it identifies. As a result, the reduction in risk provided by incentives that cover all or part of the cost of the RCx study reduces the risk associated with an RCx project enormously, and leads many organizations (both public and private) to proceed with the project. The highest mean rating for all program influences cited by respondents was for the influence of incentives that cover the cost of the commissioning study; this was rated even higher than incentives associated with implementation of recommended measures.
- In addition to uncertainty, decision makers continue to be constrained by funding cutbacks, staffing shortages, and reductions in maintenance budgets, particularly in public institutions, but also by hard-hit private sectors such as office buildings and the hospitality industry. Both government agencies and universities emphasized the importance of the incentives in making possible projects that simply would not happen otherwise. Examples were provided by counties and universities of severe cutbacks in maintenance budgets, including both outsourced services and in-house staff. The Partnership programs provide these agencies/organizations with a way to pursue energy efficiency projects that address immediate concerns, while also putting in place the metering and control infrastructure to improve buildings management in the future.

The primary trend that has the potential to counter reliance on utility programs to facilitate RCx activity is the growing number of private and public sector organizations with explicit green, carbon reduction, energy efficiency or similar policies. Some policies, notably those of the Federal Government and the University of California system, explicitly call for buildings to be retro-commissioned at a given interval.

Other policies are more generic, but nonetheless strongly encourage actions like RCx that save energy and reduce the organization's carbon footprint.

Most of these policies, however, are neither enforced nor funded. A number of the decision makers interviewed for the NTGR survey complained about these "unfunded mandates" and stated that without these programs to support RCx, they would be unable to comply with their organization's policies.

There were only two cases with relatively low NTGR scores where the respondent said that organizational policies—one at the Federal and one at the local level—were likely to have led to the RCx project in the absence of the program. In one of these cases, the program still helped accelerate the implementation of the project or caused it, rather than an alternative project, to be selected by the organization. Among several municipalities that participated in an RCx program, only one said that they probably would have pursued the project anyway because of their green policy. In that case it was a question of using the program to implement that policy as cost-effectively as possible, but it was likely that they would have gone ahead in the absence of the program.

3.7. Recommendations

In the current economic climate, where businesses and non-profit organizations face very limited funds to pursue RCx projects in the absence of assistance from programs such as those that offered the RCx measure in 2006–08, customers will continue to need incentives and technical assistance to make these projects happen.

- Specifically, incentives to cover the cost of the RCx study and remove the risk associated with initiating such a project are critical to encouraging RCx activity and were rated higher than any other factor for their influence on the decision to RCx. Such incentives should remain the foundation of RCx programs.
- Requiring implementation of all measures that meet specific payback criteria (e.g. one year) with no additional incentive also helps ensure that recommended measures are actually implemented. This requirement could also be modified so that if more of the initial study cost is covered, measures with a somewhat longer payback period could be required.
- Partnership programs appear to have a powerful influence in promoting projects that otherwise would not happen; as such they should be continued in order to sustain high net savings for RCx projects. The current financial status of partner organizations should continue to be monitored, but it seems unlikely that budget concerns will disappear and enable universities and local government to pursue RCx projects using only their own resources in the near future.
- A project screening process before the RCx study is initiated is strongly recommended to ensure that the proposed project is not already scheduled for a similar review and analysis.
- Sustainability and green policies help encourage organizations to pursue RCx projects through utility sponsored programs, but may also represent a potential source of free ridership. Their growth and application by both private and public sector organizations should be monitored as it affects the RCx market, particularly if evidence arises—either in California or elsewhere—that such policies are causing organizations to pursue RCx outside of utility programs.

In addition, we recommend consideration of using other aspects of the interaction between decision makers and program staff to help establish a solid understanding of baseline practice with regard to RCx.

- As part of the program application process, customers could be asked to provide information on their knowledge of and experience with RCx, including corporate or organizational policies, payback and other investment decision criteria, and practices at facilities elsewhere in the country, particularly in areas without utility programs in place.
- Documenting this information at the time of the application would provide program managers and evaluators with a detailed picture of actual organizational practices and provide context to help judge the extent of program influence.

4. Effective Useful Life Evaluation

4.1. Evaluation Objectives

Effective Useful Life (EUL) for a measure is the average number of years that it is expected to deliver savings. In general, the 2006–08 impact evaluations did not include new data collection or analysis of EUL, but because little or no prior work has been done to investigate the likely EUL for the measures that comprise RCx projects, this evaluation sought to inspect and test RCx measures implemented at sites that participated in RCx programs in California as part of the 2002–03 or 2004–05 program cycles, in order to provide some empirical basis for estimating RCx measure EUL.

The goal of the EUL investigation was to examine selected measures from prior RCx programs, and determine if each measure was still in place, operational, and yielding savings based on available evidence. By taking a fresh look at EUL estimates for RCx measures through building owner/operator interviews and field inspection, this study sought to provide observation-based persistence data on measures from three third-party RCx programs from previous program cycles:

- Oakland Energy Partnership - Large Commercial Building Tune-Up Program (2002–03)
- Building Tune-Up Program (2004–05)
- UC-CSU-IOU Partnership RCx Projects (2004–05)

In addition, a goal of this EUL study was to provide the ED with recommendations concerning the scope and timing of additional studies that should be performed to further improve the estimates of RCx measure EUL.

It must be emphasized that the EUL study was not designed to comply with the California Energy Efficiency Evaluation Protocols for EUL studies, but was rather undertaken with limited resources and time to address an immediate need for improved RCx measure life information, since assumed EULs for RCx measures vary widely across programs. In addition to enhancing understanding and serving as the basis for future investigations, the EULs that result from the study—whether at the RCx-project or measure level—can be used directly by the CPUC in assessing the cost-effectiveness of RCx projects implemented during the 2006–08 program cycle.

4.2. Sample Design and Selection

We evaluated gross impacts for samples of projects and measures in the three programs mentioned above. We combined information from these samples, and eliminated measures that originally had yielded no savings, or were retrofit-type measures, similar in nature to conventional capital projects, e.g., installing a cooling tower. Collectively, the EUL sample spanned 32 projects and 101 measures.

Table 22 shows the most common types of measure, each of which constituted 8% or more of the total number of improvement measures identified. Three measures (improving outside air use, improving reset schedules and improving control strategies generally) accounted for almost half the measure count, while other measures each accounted for 8–11% of the total. The remaining less-frequent improvement measures were aggregated into a miscellaneous classification. Thus grouped, each measure name had a count of at least eight.

Table 22: Measures Investigated in EUL Study

Measure name	N	% of N
Improve outside air use	16	16%
Improve temperature / pressure reset schedule	16	16%
Improve control strategies – general	16	16%
Improve scheduling	11	11%
Improve sequence of operation	10	10%
Improve setpoints	8	8%
Improve valve performance	8	8%
Improvements, miscellaneous	16	16%
TOTAL	**101**	**100%**

Sample Disposition. Steps were taken to minimize non-response bias. The field team succeeded in inspecting all the sites selected for this research.

4.3. Methodology

Prior to beginning field data collection, the evaluation team developed a subset of RCx measures to study, classifying the measures according to the scheme shown in Table 22 and identifying approximate measure installation dates from reports and original database summaries. Site contacts and former EM&V engineers were identified, and contacts were coordinated with the RCx HIM evaluation site leads in cases of overlap with the gross impact sample. The evaluators also prepared a survey tool with measure-group-specific questions, as well as a recruitment script explaining the purpose of the visit and soliciting cooperation.

The evaluation team proceeded with the field data collection as follows:

- **Project file review:** The EUL study coordinator assigned sites to an appropriate investigator, according to their location and expertise. The EUL investigator then thoroughly reviewed all project documentation from the original implementer and the follow-up EM&V to become familiar with the site and measures implemented and the EM&V results. The investigator may also have contacted the original evaluation analyst for supplemental information.
- **Recruitment:** The investigator contacted project site contacts to schedule a brief site visit to inspect the condition and functionality of equipment and associated control strategies for the RCx project measures. All sites contacted allowed access to the sites for inspection.
- **Site visit:** The investigator visually verified the presence of measure equipment and its functionality. If practical, photographs and/or EMS screenshots were collected as documentation. Building operators were informally interviewed on the performance of the measure, probing for any difficulties or modifications made since the measure implementation. Particularly if the measure was not

functional or had been modified, building operator recollection of the events leading to measure modifications sometimes provided valuable anecdotal background. The focus was to determine the timing and rationale for any changes that may have significantly impacted the measure EUL.

- **Post site visit:** The investigator documented the site visit findings for each measure in as much detail as possible. The three potential outcomes for each measure were: (a) Measure is fully operational as described in the impact evaluation, (b) Measure has been modified with a description of the changes, or (c) Measure is disabled or otherwise not functioning to achieve the savings described in the project documents.

These procedures were tested during a pilot phase involving two sites, but no major changes were made as a result.

Once all field data collection was complete, data was aggregated, quality control checks performed, and inconsistencies resolved by working with field investigators. Subsequently, measures were grouped according to the observed findings, and recommendations developed for appropriate EULs for each group.

4.4. Confidence and Precision of Key Findings

The analysis of RCx measures conducted for the EUL study was an attempted census of 32 sites that had installed measures as a result of participation in from three third-party RCx programs from previous program cycles.

Since all of the sites contacted agreed to allow access by evaluation staff, the results represent a census of participants who had measures installed through the program that initially delivered energy savings, so that there is no sampling error and 100% precision. Moreover, while site -level sampling was done, this too was generally a census or close to a census, with investigators counting the number/percentage of a specific measure still operating as intended (e.g. the number of AHUs still functioning, the number of set points not overridden, etc.) In total, 96 measures were analyzed at the 32 sites.

Because the number of individual measures analyzed, as described in section 4.2, was never more than 16 for any of the eight measure groups studied, it was decided to consider all the measures as a single RCx measure. To the extent that these measures represent the broader population of measures implemented through RCx programs, the 90% confidence interval surrounding the mean EUL for the RCx measures is given by 1.645 times the standard error of the EULs calculated for the sample, with no finite population correction. To reject the current EUL assigned for RCx measures by each of the IOUs and consider adopting the revised value found by the study, the error band around the mean EUL for the studied population would have to exclude the current EUL value. The results of these calculations are presented in section 4.6.

4.5. Validity and Reliability

The EUL data collection and analysis involved inspecting the condition and functionality of equipment and associated control strategies for the RCx project measures. The primary source of uncertainty in the analysis was the development of criteria for determining functionality. In some cases, a measure clearly was not functioning; in others there was evidence that the functionality had been diminished but not completely eliminated, as when one of several setpoints had been overridden but others remained in their post-RCx state. There was, for many of EEMs that were found to have failed, significant uncertainty around the time of failure. We could state with a fair amount of certainty when a measure was installed, when our initial evaluation of the previous program had found the measure functional, and when we went back and discovered the measure was no longer working. We could ask facility staff when the failure occurred, but in many cases, they could not recall (sometimes because they were new to the facility) or could only recall in very vague terms (e.g., "We disabled that a long time ago.") We used our best judgment to establish a lifetime range, and then used the midpoint of that range in our analysis. For measures that were still operating, a decision was made to call equipment still delivering at least 50% of the original energy savings functional, while equipment delivering less than 50% of original savings were considered partly functional. The lack of detailed monitoring of measures at the study sites increased the level of uncertainty surrounding the percentage of savings still being achieved, which were estimated using engineering judgment.

There is also significant uncertainty surrounding the estimation of EUL from the observed failure rate. We used a simple linear function to project the past failure rate into the future to determine at what point in time half the measures would have failed and half would still be working, and used that as our estimate of the EUL for all the measures studied. While there is uncertainty associated with this estimate, it represents an indication of the validity of the EUL values assumed by each of the IOUs in claiming savings from RCx projects, and thus reduces the overall level of uncertainty associated with these values.

4.6. Detailed Findings

All 32 sites targeted for the study were successfully recruited for participation in the EUL investigation. The study team, expert field investigators with decades of mechanical and electrical systems experience, conducted on-site inspections from July through September 2009. The field work required diligence and ingenuity in assessing measure performance in a wide variety of facilities and with site contacts who had varying degrees of technical expertise and knowledge of past RCx projects.

Oakland Energy Partnership (OEP) sites in particular proved challenging because of the longer time lag since measures were implemented. The OEP was part of the 2002–03 program cycle, and even though many measures were actually completed in 2004, a number of sites had experienced sufficient staff turnover so that no one remembered the OEP retro-commissioning effort. Sites from other programs had also experienced personnel turnover, but usually to a lesser degree. Nevertheless, all consented to host a site visit with the EUL investigators, and all did their best to help with the measure investigation.

Comprising a heterogeneous mix of measures and actions, the measures addressed by the EUL study totaled 96 of 100 measures originally implemented at these sites through the three programs. The other four measures, located at two sites, were declared “indeterminate” due to lack of adequate data and high uncertainty as to the outcome. The remaining EEMs were classified by measure group and measure system.

Table 23: EUL Measure Classification

Measure Name	N	Still Working	%
Improve boiler performance	3	1	33%
Improve building warmup / cooldown	2	2	100%
Improve control strategies - general	10	8	80%
Improve damper performance	1	1	100%
Improve maintenance practices - general	1	1	100%
Improve outside air use	12	9	75%
Improve scheduling	10	6	60%
Improve sensor performance	5	5	100%
Improve sequence of operation	10	10	100%
Improve setpoints	8	3	38%
Improve temperature / pressure reset schedule	15	12	80%
Improve valve performance	8	7	88%
Install / replace variable speed drive - HVAC air handler	2	2	100%
Install lighting occupancy sensors	1	0	0%
Install miscellaneous efficiency improvement	3	2	67%
Install VFD	5	5	100%
All Measures	**96**	**74**	**77%**

The study results provided insights both into the timing of measure failures and the reasons for failures. The following were among the key findings:

- In all, there were 22 failures out of the 96 measures. Of these, three failed in the first year after measure installation, nine in the second year, seven in the third year, and for the small number of measures that had been installed four years ago, three more failed.
- Cumulatively, this represented failure rates of 3%, 13%, and 20% for the first three years, respectively. Using a simple linear extrapolation, these results lead to an EUL of eight years – defined as the point at which half the installed measures have failed. Note, however, that there is a large uncertainty band around this estimate.

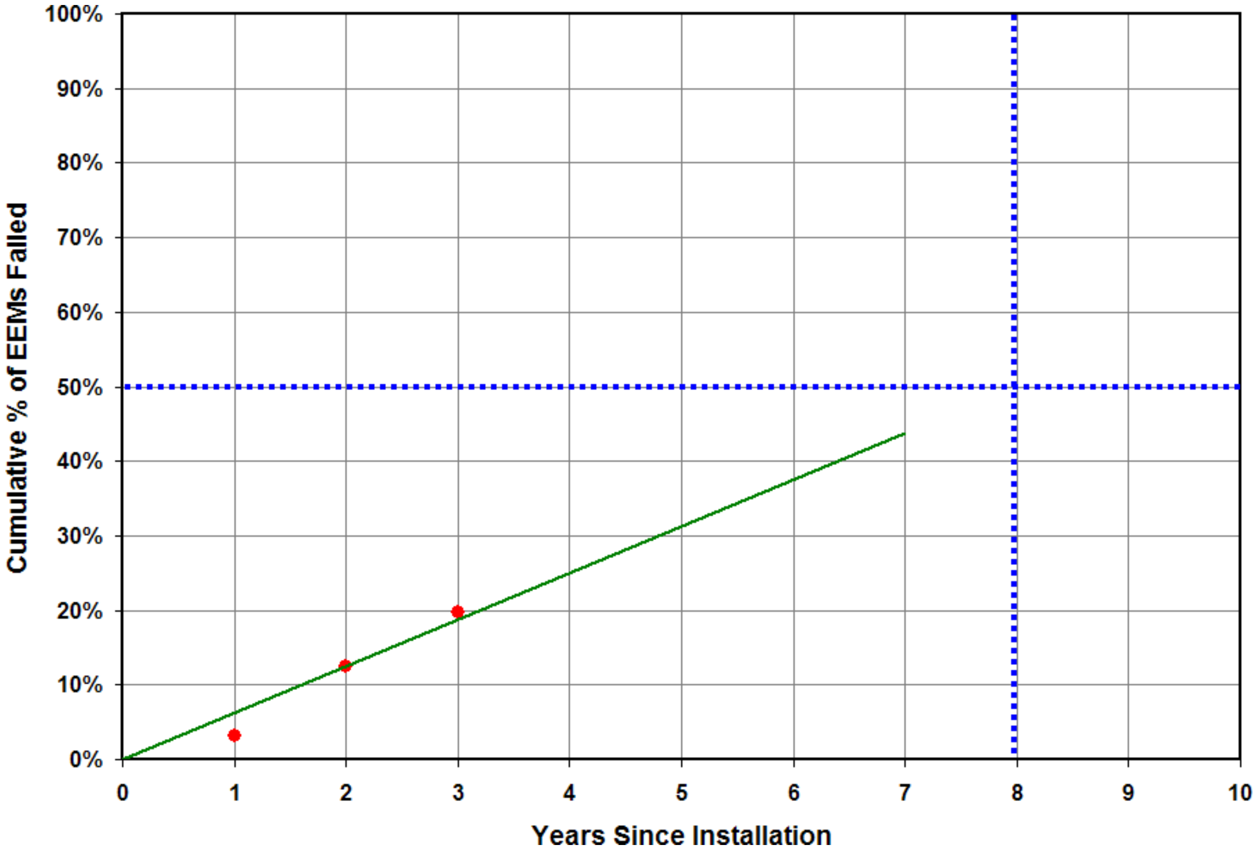


Figure 4: EUL Three-Year Failure Rate Projection

Although there were not enough measures in any category to provide statistically significant results on the comparative failure rates of different measure types, the results do provide the basis for a comparison to RCx measure EULs for program groups claimed by IOUs, summarized in Table 24.

Table 24: RCx Measure EULs Claimed

RCx Program Group		PG&E	SCE	SCG	SDG&E
Partnership-UC/CSU/IOU	Range	15	14	10	10
	Average*	**15**	**14**	**10**	**10**
Partnership-LA County	Range	-	10	15	-
	Average*	-	**10**	**15**	-
Partnership-Other	Range	3-15	14	20	-
	Average*	**6**	**14**	**20**	-
PECI	Range	3-12	10	-	-
	Average*	**6**	**10**	-	-
PG&E Core	Range	3-12	-	-	-
	Average*	**5**	-	-	-
Other	Range	3-15	-	-	10-15
	Average*	**8**	-	-	**11**

RCx Program Group		PG&E	SCE	SCG	SDG&E
All Programs	Range	3-15	10-15	10-20	10-15
	Average*	**7**	**11**	**14**	**10**

* Weighted by ex ante MMBtu savings.

Overall average EULs for RCx measures range from 7 for PG&E to 14 for SCG. Note that ranges are shown in cases where IOUs assign different values to different measure types. For example, PG&E applied EULs of 3 years for control-setting changes, 8 years for hardware repairs, and 12 years for hardware installations.

Similarly, different utilities may assume different EULs for the same programs and measures. Average EULs for UC/CSU/IOU Partnership MBCx projects range from 10 years for Sempra to 15 years for PG&E. Moreover, PG&E has substantially higher EULs for UC/CSU/IOU MBCx projects (15) than for the rest of its RCx portfolio (6–8 years), while the Southern California utilities do not have such a difference.

Generally, these inconsistencies between utility EUL claims point towards the need for a more uniform and defensible basis for future EULs. Additionally, they highlight the need for more research into whether RCx programs that feature installation of permanent monitoring (such as MBCx) substantially improve RCx measure EULs.

In addition to indications of measure life, the study also offered some interesting and potentially significant anecdotal findings regarding the reasons for measure failure. Of the 22 measures for which site contacts responded “no” to the question “Is the affected hardware still in place and operational?” the following reasons were offered for the equipment failure:

- Control sequence changed due to perception that the RCx measure compromised occupant comfort: 35%
- Control sequence changed – reason unknown: 22%
- Lack of maintenance compromised the measure: 22%
- Facility operating hours extended: 9%
- Miscellaneous: 12%

Despite the small sample size, particularly in terms of the breakdown of reasons for premature measure failure, it appears that human factors are generally more responsible for the failure of these measures than actual technical issues. Measures recommended by the RCx study and implemented as a direct result are usually selected based upon anticipated energy savings, but failure to consider occupant comfort issues appears to raise the risk of manual overrides of setpoints or other changes to control sequences.

Similarly, recommended RCx measures often require additional maintenance or at least adequate maintenance if savings are to persist. That one in five failed measures did so from lack of maintenance suggests that more explicit instructions regarding required maintenance actions should be provided as part of the recommendations for implemented measures.

In conclusion, this study has developed a rich data set for program planning purposes and identified area for future study, as discussed in the next section.

4.7. Recommendations

Given the uncertainty around the EUL estimated by this study, it does not appear that there are solid grounds for changing EULs claimed by IOUs for RCx measures. There are clear indications, however, that the average EULs claimed for Southern California utilities are higher than indicated by this study.

To confirm the EUL value developed by the current study on a relatively small sample, a much larger study will be required, and we recommend that such research be pursued aggressively and on a large scale.

A short-term option that could be completed by early 2010 would be to perform classic survival analysis to see if it comes up with firmer estimate to inform planning parameters for 2010–12 programs.

Two longer-term options also exist for solidifying the EUL estimate, as described briefly below. Both of these, however, would require a number of years to complete. One option could be selected, or both options could be pursued simultaneously.

- Continue 2002–05 panel study: Perform a similar measure failure investigation on the same panel after three more years have elapsed (end of 2012 or beyond). If the failure rates observed to date continue over the next few years, then around that time, the panel should be approaching a 50% failure rate, which would establish the EUL by definition.
- Establish new 2006–08 panel study: Results of a preliminary power analysis established that the EUL study would have to observe about 250 failures to obtain results with 90% confidence. Assuming the observed failure rate of 20% every three years, it would take a panel of over a thousand measures several years or more to reach a point when the requisite number of failures would be achieved. It is unclear at this point what the total number of observable measures in the 2006–08 RCx population is.

5. APPENDICES

5.1. Overview of Programs Evaluated

Table 25: Programs that Implemented RCx Projects

Programs Included in this Evaluation	Program Description	Key Program Elements
PGE2001 Ag & Food Processing	The Agricultural and Food Processing program coordinated a portfolio of products and services designed to enhance adoption of integrated demand side management among the diverse agricultural and food processing customers in PG&E's service area and to provide cost effective, comprehensive, relevant program elements to deliver the kWh, kW, therms, demand response (DR), and distributed generation (DG) goals for PG&E's energy procurement strategy.	Although RCx was not an explicit focus, a few RCx projects were completed through the Ag & Food Processing program.
PGE2002 Schools and Colleges	The Schools and Colleges program coordinated a diverse portfolio of products and services designed to enhance adoption of integrated demand side management among the educational institution customers in PG&E's service area. Third party offerings included Resource Solutions Group's (RSG) Campus Housing Efficiency Solutions (CHES) and School Energy Efficiency (SEE), and Low Income Investment Fund's (LIIF) California Preschool Energy Efficiency Program (CPEEP), while partnerships included the CCC/IOU Partnership and the UC/CSU/IOU Partnership	Although RCx was not an explicit focus, a few RCx projects were completed through the Schools and Colleges program.
PGE2005 Hi-Tech Facilities	The High Technologies Facilities program targeted high technology facilities and their unique energy utilization needs using both PG&E and third party industry specialists to deliver a range of energy efficiency services. The program addressed new construction and facility expansion and renovation as well as ongoing daily facility operation. The program incorporated statewide financial incentive elements and elements specifically targeted to and customized for the high technology customers in PG&E's service area.	Although RCx was not an explicit focus, a few RCx projects were completed through the Hi-Tech Facilities program.
PGE2006 Medical	This program targeted new and existing	Retro-commissioning was one of the key

Programs Included in this Evaluation	Program Description	Key Program Elements
Facilities	medical facilities using both PG&E and third party industry specialists to facilitate delivery of a portfolio of energy efficiency, demand response and distributed generation services.	energy saving measures promoted by the Medical Facilities Program.
PGE2007 Office Buildings (Large Commercial)	The Large Commercial market sector included large buildings where capital expansion, capital renewal, and/or operations and maintenance products and services are procured through contracts with manufacturers and/or distributors. The objective of the Large Commercial program was to provide the most cost effective and comprehensive portfolio of program elements for the targeted customers in order to deliver the kWh, kW, therms, demand response (DR), and distributed generation (DG) for PG&E's energy procurement strategy.	While most RCx projects targeted to large office buildings were delivered through PECI's PG&E 2091 program, some customers elected to complete RCx projects through PG&E's Core Large Commercial program.
PGE2015 LGP ABAG	This program promoted reduced energy use for local governmental agencies (cities, counties and special districts) in the Association of Bay Area Governments. It was designed to provide technical assistance and information services to assist cities, counties and special districts (local governments) in the ABAG membership areas.	Retro-commissioning was one of the energy saving measures promoted by the Association of Bay Area Governments Program.
PGE2025 LGP MARIN	This program brought together five elements: Marin Energy Management Team (MEMT), to act as "energy manager" for public agencies; Small Business Energy Alliance (SBEA), to provide energy audits and incentives for energy efficient retrofits; California Youth Energy Services (CYES), to provide hardware installation and energy assessments to targeted owners and renters in the Mass Market program; EnergyWise, to provide training and incentives to realtors and qualified home inspectors; and Building Tune-Up (BTU), to offer retro-commissioning and retrofit services and incentives to large commercial customers.	Retro-commissioning was one of the energy saving measures promoted by the Marin County Energy Watch (MCEW) Program.
PGE2032 LGP SONOMA	The Sonoma County Energy Watch (SCEW) was a joint project of QuEST, County of Sonoma and PG&E to realize energy savings by leveraging a whole community's public commitment to reduce greenhouse gas emissions. Working closely with the Climate	One of the elements of SCEW, Building Tune-Up (BTU) offered retro-commissioning and retrofit services to large commercial customers and provided incentives for implementing energy efficiency measures.

Programs Included in this Evaluation	Program Description	Key Program Elements
	Protection Campaign (CPC), SCEW reflected availability of services from third parties, and addressed the residential sector, office buildings and wastewater processing.	
PGE2035 LGP SVLGEW	This partnership between PG&E and Silicon Valley Leadership Group Energy Watch (SVLG-EW) promoted reduced energy use and energy savings targets for the SVLG members by providing energy efficiency information, commercial building energy assessments, energy efficient equipment and energy system metering and monitoring equipment to members of SVLG and Sustainable Silicon Valley (SSV)	Targeting the more than 240 Silicon Valley firms and supporting industries represented by SVLG, the program included retro-commissoning as one of the energy saving measures it promoted.
PGE2036, SCE2530, SCG3520, SDG&E3029 UC-CSU-PG&E-SCE-SCG-SDG&E Partnership	The University of California, California State University (UC/CSU) and Investor-Owned Utilities (IOUs) Energy Efficiency Partnership (UC/CSU/IOU Energy Partnership) was a statewide nonresidential program that continued in the 2006-2008 program cycle, with the goal of extending the reach and effectiveness of traditional utility programs by using the UC and CSU system communication and outreach channels to achieve broad penetration of energy efficiency services in the local campuses.	Retro-commissioning projects were among the options available through the 2006-2008 UC/CSU/IOU Partnership program, which comprised three elements that operate on a statewide, integrated basis to provide immediate energy savings while setting the foundation for a long-term program focused on sustainability and best practices, including Energy Efficiency Retrofits and Energy Efficiency Education and Best Practices Development as well as Monitoring Based Commissioning (MBCx).
PGE2052 Lodging Savers	Ecology Action's LodgingSavers program provided multi-measure comprehensive retrofits and retrocommissioning (RCx) to small, medium and large lodging facilities in PG&E's service area. LodgingSavers provided energy audits, job specification and design assistance, installation services and financial incentives to encourage measure adoption.	In addition to RCx, predominant measures included lighting, HVAC (heating, ventilation and air conditioning), controllers, refrigeration measures, and water saving measures.
PGE2056 Monitoring-Based Persistence Commissioning	MBPCx was a third party program operated by Enovity that combined traditional RCx activities with ongoing monitoring to improve savings persistence. The program targeted large public and private sector facilities with good potential for reducing energy use.	MBPCx initially focused on low-cost improvements and optimization of the building systems. A detailed engineering study quantified savings attributed to fixes done in the investigation phase, and recommended additional operation and maintenance improvements and capital retrofit measures. Additional program offerings included building operator training, improvements in system documentation and system monitoring. Long-term monitoring capability was

Programs Included in this Evaluation	Program Description	Key Program Elements
		achieved through the PACRAT diagnostic and tracking software system.
PGE2070 Data Centers	This third party program operated by QuEST was designed to improve energy efficiency at data center and server farm facilities, which are energy intensive, 24/7 operations that present significant opportunities.	Marketing concentrated on organizations with multiple facilities. The program began with a comprehensive auditing process to identify cooling system energy efficiency retrofits, with an emphasis on implementing measures which will not disrupt facility operations. Program funds were offered to buy measure costs down to a one year payback. QuEST, at the customer's discretion, provided turnkey service from auditing through installation and staff training.
PGE2071 Hospitality Energy Efficiency Program	HEEP was a third party program operated by QuEST to improve energy efficiency at hotel properties.	Marketing to large chains that manage their own properties, the program featured comprehensive audits of building HVAC systems, implementation of low-cost energy efficiency measures and deferred maintenance items, and development of a list of RCx activities and energy efficiency measures, together with detailed financial analysis and specifications. Measures were then selected by the owner.
PGE2072 Hospitals Pilot	HPP was a third party program operated by QuEST offering a comprehensive array of retro-commissioning services to best meet the variety of needs presented by energy-intensive hospital facilities.	HHP was implemented in four phases. Phase I included analysis of facility energy use, building systems inspection and identification of low-cost energy efficiency measures (EEMs). Additional areas for investigation and specific EEMs were detailed for follow up in Phase II, a targeted investigation resulting in a detailed EEM list including itemized cost-benefit analysis. The third phase, Implementation, was a joint effort with facility staff and QuEST personnel to complete RCx activities and install EEMs. The fourth phase involved a final as-built calculation of measure savings, documentation and staff training to help maintain savings persistence.
PGE2088 State Leased Facilities	This third party program offered by Enovity focused on buildings leased by the California Department of General Services that are 5,000 square feet or larger and offered benchmarking, energy audits, retro-commissioning, technical assistance, design advice, rebates, and direct implementation services.	While RCx was one of the measures covered by the program, most measures focused on HVAC, controls and lighting.
PGE2090 Airflow and Fume Hood Control	This third party program operated by Newmatic Engineering (NE) offered a	A detailed audit was conducted in participating labs, resulting in a list of

Programs Included in this Evaluation	Program Description	Key Program Elements
Systems Re-Commissioning	comprehensive suite of engineered options to optimize energy efficiency in laboratories through RCx activities focusing on controls (constant air volume to variable volume), installation of usage-based controls, air change rate reductions and static pressure setpoint reductions. The program sought to establish a long-term awareness in the laboratory-operating community of the benefits of bringing specialized technical expertise into the facility to conduct RCx.	RCx activities and capital EEMs, estimated savings, costs and incentives. The owner selected the measures for implementation.
PGE2091 Retrocommissioning Services and Incentives Program	The program was designed to expand building system optimization and RCx capabilities with program features that directly addressed market barriers, as well as to ensure the persistence of the program benefits.	The program provided screening of the candidate buildings, required approved candidates to enter into an agreement with the program; assisted owners in selecting a Cx provider, who completed the investigation, helped facility staff to select items for implementation, and aided implementation when necessary. For measures with outstanding paybacks (after incentive) of less than one year, the owner was expected to contract for this work and pay for the amount not covered by the incentive as a requirement of the program. For measures with outstanding paybacks over 3 years, the owner was given incentives to take advantage of the measure, but was not required to do so as part of the program contract.
PGE2094 Macy's Comprehensive Energy Management	MCEMP was a third party program operated by QuEST to deliver comprehensive energy efficiency services to approximately nine Macy's retail stores in the PG&E service territory. This program built on Macy's successful participation in earlier PG&E programs, and comprised both RCx and retrofit measures.	The program begins with an audit followed by the delivery of a preliminary report detailing RCx and retrofit measures, along with identification of deferred maintenance items requiring correction before proceeding. Upon approval by Macy's staff, QuEST proceeds to the Detailed Investigation Phase, where a comprehensive energy study on the items approved by Macy's is completed, including RCx and capital retrofit measures with associated costs, savings and incentives. Project implementation commences with RCx activities and retrofit installation, as directed by the customer. Staff training and periodic maintenance plans are provided at the end of the program to help maintain savings persistence.
SCE2508 Retro-Commissioning (RCx)	The program was designed to expand building system optimization and RCx capabilities with program features that	The program provided screening of the candidate buildings, required approved candidates to enter into an agreement

Programs Included in this Evaluation	Program Description	Key Program Elements
	directly address market barriers, as well as to ensure the persistence of the program benefits.	with the program; assisted owners in selecting a Cx provider, who completed the investigation, helped facility staff to select items for implementation, and aided implementation when necessary. For measures with outstanding paybacks (after incentive) of less than one year, the owner was expected to contract for this work and pay for the amount not covered by the incentive as a requirement of the program. For measures with outstanding paybacks over 3 years, the owner was given incentives to take advantage of the measure, but was not required to do so as part of the program contract.
SCE2526, SCG 3518 California Community Colleges Partnership	The 2006-08 CCC/IOU Energy Efficiency Partnership program was a partnership between the California Community Colleges (CCC) and the four Investor-Owned Utilities. Through the partnership, the campuses were able to build an infrastructure that facilitates project identification and implementation. It also allowed the utilities to focus on the varying needs of the community colleges and opportunities to tailor a program that was able to capture even greater energy savings while developing long term relationship.	RCx and MBCx projects were among the options available through the CCC/IOU Partnership Program, which also included the implementation of Retrofit and New Construction projects. The program also focused on training and education to expand existing vocational education programs while training faculty and staff on best practices on energy efficient technology implementation and energy management.
SCE2528, SCG 3527 SCE-SCG County of Los Angeles Partnership	The 2006-08 SCE/SCG/County of Los Angeles Energy Efficiency Partnership built on the success of the existing partnership program. Applying lessons learned from the existing program the Internal Services Department (ISD) contributed in- house staff to perform many of the tasks associated with the retro-commissioning process such as preliminary investigation, bench marking, determination of system deficiencies, and correction of those deficiencies.	This program focused mainly on retro-commissioning activities in County of LA facilities, but also included Retrofit, Technology Transfer/Feasibility Study and Public Housing Metering. The county also applied the lessons learned from current RCx projects to streamline the RCx process for future program cycles. Electric savings accrued to SCE; gas savings to SCG.
SDGE3010 SDGE3010_ESB_ESB-Energy Savings Bids	SDG&E's® Non-residential Energy Savings Bid (ESB) program provided financial incentives for qualifying retrofit projects of existing equipment and/or systems with new high-efficiency equipment and/or systems. A measured approach was used to determine the energy savings and incentive, with the incentive payment determined by the quantity of kilowatt-hour (kWh) savings resulting from installation of the new	Although RCx was not an explicit focus, a few RCx projects were completed through the Energy Savings Bids program.

Programs Included in this Evaluation	Program Description	Key Program Elements
	retrofit equipment and/or system.	

5.2. Statistical Estimation Methods

5.2.1. Gross Impact Evaluation Statistical Procedures

Two different approaches were required to estimate parameters for two different types of impacts analyzed.

- Impacts where ex ante savings (kW, kWh, and/or therms) were reported directly into the program tracking systems, where the savings are directly related to what was installed.
- Impacts where no ex ante claims were made for what was installed. This situation arises when one fuel constituted the only savings claimed, but where savings from another fuel were possible. In this situation, the utility typically did not report the other fuel benefits/savings, but the evaluators were required to include estimates of those savings in the ex-post procedures.

The actual precision attained was calculated using several different approaches taken from the *California Evaluation Framework* (TecMarket Works, 2004), Cochran (1977), Taylor (1997) and IPMVP (2002), with the specific techniques used depending on the type of impacts analyzed.

5.2.1.1 Procedures for Program Elements with Direct Savings

For sites where all gross savings parameters had *ex ante* and *ex-post* values (i.e., kW, kWh, and/or therms), precision was based on the estimation of a realization rate, defined as the ratio of gross *ex-post* to gross *ex ante* savings. The methods used were based on the *California Evaluation Framework* (2004, p. 358) and estimated and reported the following parameters:

- Overall and per-stratum realization rate
- Standard error of overall realization rate
- 90% confidence interval for overall realization rate
- Relative precision for realization rate
- Gross UES
- Standard error for Gross UES
- 90% confidence interval for Gross UES

Calculations followed the following steps:

Step 1: Calculate the sample-based realization rate using this equation:

$$b = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i x_i}$$

Where:

b = the realization rate

w_i = case weights, defined as the population count divided by the sample count for each stratum

y_i = sample ex-post savings for case *i*

x_i = sample ex ante savings for case *i*

Step 2: Calculate the standard error of b:

$$se(b) = \frac{\sqrt{\sum_{i=1}^{n} w_i (w_i - 1) e_i^2}}{\sum_{i=1}^{n} w_i x_i}$$

Where:

e = the *ex-post* value minus b times the *ex ante* value

Step 3: Calculate the error bound at the 90% level of confidence of the realization rate, *b*, by multiplying the appropriate t-statistic (1.645) by the standard error of the realization rate, *se(b)*. The upper and lower bound of the realization rate, *b*, were then calculated by adding and subtracting the 90% error bound from the realization rate.

CI=$b \pm (1.645 \text{ x } se(b))$

Step 4: Calculate the 90% relative precision of the realization rate by dividing the CI by the realization rate, *b*. Multiply by $\sqrt{1-\frac{n}{N}}$ to apply the finite population correction.

Step 5: Calculate the ex-post gross UES by multiplying the mean ex ante population savings by the realization rate.

Step 6: Calculate the upper and lower bound of the ex-post gross UES by multiplying the ex-post gross UES by the upper and lower bounds of the realization rate.

Step 7: For the purpose of estimating sample sizes for future studies, calculate an error ratio (êr) of the realization rate using the following equation:

$$\hat{er} = \frac{\sqrt{\left(\sum_{i=1}^{n} w_i e^2 / x_i^{\gamma}\right)\left(\sum_{i=1}^{n} w_i x^{\gamma}{}_i\right)}}{\sum_{i=1}^{n} w_i y_i}$$

As explained in the *Framework* (on page 358, step 4), in calculating the error ratio, the value of x_i is raised to the gamma (γ) power, usually assumed to be 0.8.

5.2.1.2 Procedures for Program Elements Where No Ex ante Savings Were Claimed

This situation arises when one fuel constituted the only savings claimed, but where savings from another fuel were possible. The evaluation team calculated *ex-post* savings in these cases, and these calculations were the basis for gross UES savings and their associated standard errors, as well as relative precision estimates. As indicated above, no realization rate is possible in these cases. The *California Evaluation Framework* (2004) does not provide guidance in this situation. Therefore, Cochran's (1977) approach to estimating stratified means and standard errors was followed. The steps involved in these calculations are as follows:

Step 1: Calculate the sample-based ex-post, stratified mean UES using the following equation (5.1):

$$\bar{y}_{str} = \frac{\sum_{h=1}^{L} N_h \bar{y}_h}{N}$$

Where:

y_{st}	= the stratified mean
N_h	= the stratum population
y_h	= the stratum mean
N	= the total population

Step 2: Calculate the standard error for a stratified sample mean by taking the square root of the variance (s^2), calculated using Cochran's equation 5.13, with the term on the right representing the finite population correction (fpc):

$$s^2\left(\bar{y}_{str}\right) = \sum_{h=1}^{L} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h s_h^2}{N}$$

where:

y_{st}	= the stratified mean

W_h = the stratum weight (N_h /N)

s_h^2 = the variance for stratum h

n_h = the sample size for stratum h

N = the total population

Step 3: Calculate the relative precision using the following equation:

$$rp = \frac{1.645 se\left(\bar{y}_{str}\right)}{\bar{y}_{str}}$$

The calculated UES from Step 1 was assigned to all claims in the population.

5.2.2. Net Evaluation Procedures

The core of the net evaluation procedures is the measurement and calculation of the net-to-gross ratio (NTGR). The procedures for deriving site-level and/or measure-level NTGRs were described in section 3.3. The parameters estimated by the following methods are:

1 Net-to-gross ratio (NTGR)

2 NTGR standard error

3 90% confidence interval for NTGR

4 Relative precision for NTGR

NTGRs were calculated for each fuel type for each sampling domain (IOU). Within each stratum, the project-specific NTGRs were weighted according to their contribution to stratum -level gross ex ante savings for each fuel type in calculation of the NTGR for that fuel type.

The procedures for calculating NTGR statistics involved several steps:

Step 1: Calculate a weighted mean NTGR for each stratum, with the NTGRs for individual cases weighted by their contribution to the total ex ante fuel savings for that stratum. For example, if case 1 of 10 cases in stratum *n* accounts for 130 kWh of the 1,000 kWh total ex ante savings in that stratum, the NTGR for that case is multiplied by 0.13. If Case 2 accounts for 80 kWh of the 1,000 kWh total, its NTGR is multiplied by 0.08, and so on. The sum of these products is the savings-weighted mean NTGR for the stratum.

Step 2: Calculate a stratified mean NTGR using Cochran's equation 5.1:

$$\bar{y}_{str} = \frac{\sum_{h=1}^{L} N_h \bar{y}_h}{N}$$

Where:

y_{st} = the stratified mean NTGR

N_h = the stratum population

y_h = the stratum mean NTGR, calculated as described in Step 1.

N = the total population

Step 3: Calculate the standard error for the NTGR by taking the square root of the variance (s^2), calculated using Cochran's equation 5.13, with the term on the right representing the finite population correction (fpc):

$$s^2\left(\bar{y}_{str}\right) = \sum_{h=1}^{L} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h s_h^2}{N}$$

Where:

y_{st} = the stratified mean NTGR

W_h = the stratum weight (N_h /N)

s_h^2 = the variance for stratum h, calculated as:

$$\sum_{i=1}^{n_h} w_i (x_i - \bar{x}_w)^2$$

where:

w_i = the savings weight for case *i* in stratum h

x_i = the NTGR for case *i* in stratum h

x_w = the savings weighted NTGR for stratum h

n_h = the sample size for stratum h

N = the total population

Step 4: Calculate confidence interval around the NTGR by multiplying the standard error by the critical t value for 90% confidence: 1.645.

Step 5: Calculate the relative precision of the NTGR using the following equation:

$$rp = \frac{1.645 se\left(\bar{y}_{str}\right)}{\bar{y}_{str}}$$

These NTGRs were applied to all claims in the tracking file.

References

Cochran, W.G. (1977). Sampling techniques (3rd ed.). John Wiley & Sons, A Wiley Publication in Applied Statistics.

Taylor, John R. (1997). An Introduction to error analysis: The study of uncertainties in physical measurements. Sausalito, CA: University Science Books.

TecMarket Works (2004). "California Evaluation Framework." Report to the *California Public Utilities Commission*, p. 327–359.

International Performance Measurement & Verification Protocol Committee (2002). International Performance Measurement & Verification Protocol , Concepts and Options for Determining Energy and Water Savings, Volume I, p. 77.

5.3. Gross Savings Evaluation Methodology

5.3.1. Overview

This methodology section provides an overview of the specific procedures and best practices for carrying out the tasks necessary to assess savings at each sampled project site. It is presented here in instruction style, as directions given by gross evaluation managers to the evaluation site leads who had primary responsibility for estimating savings.

5.3.1.1 Evaluation Goals

This evaluation examines the portfolio of programs and projects, sponsored by the State of California and funded through public benefits charges to investor-owned utility (IOU) ratepayers that comprise the 2006–08 Commercial Retro-Commissioning (RCx) high-impact measure (HIM) group. The ultimate goal of the evaluation is to rigorously quantify how much electric and natural gas energy this group saved for each IOU, as well as how much they reduced peak electric demand. The study was performed as an independent, third-party evaluation. The evaluation results will show how close this HIM came to matching IOU savings claims, and shed light on reasons why they did or did not meet them. These findings will also be used in determining the IOUs earnings from these programs.

5.3.1.2 Retro-commissioning Programs

Retro-commissioning is the systematic evaluation of major energy-using systems and equipment in existing buildings to find ways to ensure they operate both efficiently and in accordance with customer needs. More than with traditional retrofit programs, RCx focuses on lower-cost, O&M-type measures. A number of the programs serve small specialized markets such as data centers or target specific end uses, such as fume hoods. One program from each IOU serves a broad market and deals comprehensively with the energy end uses found at participating sites. All of them assist participants via investigations (audits or performance assessments) that identify opportunities for improving the operations or maintenance of existing assets, and provide incentives to reduce the barriers to adopting the program’s recommendations. Many provide other forms of educational and training support for the participants. In some cases, these programs also identify opportunities for conventional retrofit technologies, such as high efficiency chillers or new control systems, and either provide incentives directly to motivate adoption or refer customers to other efficiency programs.

5.3.1.3 Key Evaluation Perspectives

Impact evaluation, particularly for programs as complex as RCx, can lead to innumerable questions and situations where the proper course of action may not be obvious. Such cases require you to exercise your best judgment, keeping these general perspectives in mind:

Develop strong working relationships.

The success of RCx evaluation projects requires a cooperative working relationship with the facility operating engineers, facility managers, RCx engineers and vendors. They are essential for gaining the necessary access to equipment and data, and also can provide valuable insights on the history of building system operations and background on the project measures. Consciously strive to maintain their trust and cooperation, including making requests for assistance judiciously to avoid "burning out" the contact.

Optimize efforts to get the best possible answer.

For each site, you must assess how best to balance many factors—evaluation resources, the customer's resources, the idiosyncrasies of the particular site, and the analytical tools available—to come up with the "best answer," that is, the most accurate estimate of savings for the project, within the budget, time, and logistical constraints. You will probably have to prioritize, trade off, and modify your approach as the evaluation of a project goes on. When you do, keep in mind this overall goal of the "best answer."

Minimize influence on the project.

Customers may see you as a resource or an expert. They may ask you how well their RCx project is working, or if Recommendation X is a good idea, or how to solve some other operational problem. Even if you know the answer, resist the urge to tell them. As an evaluator, it is best if you play your cards close to the vest, and not say anything that could influence how the project turns out. While customers' inquisitiveness is totally understandable, you may have to gently and politely explain that to preserve the integrity of the evaluation, you have to remain tight-lipped.

Avoiding any kind of influence can be challenging. One plausible scenario has a facility engineer showing you system operation on the building energy management and control system (EMCS) display, when she notices that the temperature setback schedule that the RCx project had implemented has been overridden. Since she likely would not have noticed this had you not asked to see the EMCS, you would need to plead with her to keep the settings as they are while you collect data, and begin fixing the override only after you were done.

Consider all program effects.

Each project's evaluation should consider all program effects, regardless of whether they were specifically documented in the program's RCx report. For example, a customer may have signed up for an RCx study, and the RCx investigator mentioned offhandedly that broken dampers are a common problem in buildings. As a result, the customer repaired their damper actuators before the study began. In such an instance, we would want to investigate the energy implications of those damper actuator fixes as part of our M&V effort.

Spillover measures are another undocumented program effect. The evaluation procedures in later sections describe in more detail how to detect spillover.

Note that program effects only include measures and changes that were implemented and operational prior to the end of the program period, i.e., December 31, 2008. If measure performance or conditions have changed since that time, then we are only concerned with what was happening as of that date.

Ensure your work is transparent and well-documented.

Because of the substantial financial and policy implications of the evaluations, the CPUC Energy Division (ED) is holding the evaluation team to a high standard of transparency and documentation. The public ultimately will have the right to review all of our work. Depending on the evaluation results, our methodology and findings could be subject to intense scrutiny. Consequently, it is critical that we maintain clear records of what we did and how we did it, so that we can answer questions about our work long after it was originally completed. Because of the possibility of conflicts of interest or undue bias, the focus will especially be on how the evaluation team interacted with the investor-owned utility (IOU) programs. Practically speaking, be sure to keep the ED RCx evaluation project manager informed of important exchanges and events, such as customer recruitment or scheduled onsite visits.

5.3.2. Evaluation Questions for Each Site

Evaluating program energy savings correctly requires that we examine four essential, interrelated questions:

1 what happened because of the program (i.e., what were the gross savings, given appropriate baseline conditions),

2 what would have happened anyways if the program had not existed (i.e., free ridership),

3 what else happened, because of the program, beyond the documented activities (i.e., spillover), and

4 how long will any savings last (i.e., effective useful life).

How this evaluation assesses these three areas is discussed further in this section.

5.3.2.1 What Happened? (*Gross Savings / Baseline*)

Many measures included in RCx projects can be classified as *add-on measures*—that is, ones that can be "added on" to existing equipment to improve its operating efficiency and/or shift usage to off-peak periods. Examples include lighting or HVAC controls, variable frequency drives, programmable thermostats and system tune-ups or changes in control strategies. In these cases, your analysis should assume the conditions present before the measure was implemented as the baseline for calculating savings (also known as "early replacement"). An example might be a measure to restore proper controls to a VFD-controlled fan, so the fan does not run at full speed continually, as it was originally doing. The baseline usage would be based on full-speed operation, the post-implementation usage (assuming the measure was successful) would be based on the fan running at reduced speeds, and the savings would simply be the difference.

Some RCx measures, however, could be best categorized is *equipment replacement* (also known as "normal replacement"), since they involve changing of equipment because it was at the end of its effective useful life (EUL), i.e., it failed, became obsolete, or was not productive enough. The baseline might then be the alternative equipment that either codes or standard industry practice would have dictated. If this is the case for one or more measures in your project, then administer the baseline survey for these. It contains questions you should ask of knowledgeable facility personnel to assess the appropriate baseline.

With whatever baseline you ultimately select, you must clearly explain your rationale in the evaluation documentation.

5.3.2.2 What Would Have Happened If the Program Had Not Existed? (*Free ridership*)

Concurrent with our effort to evaluate the actual gross energy savings from the RCx projects, the evaluation's net/spillover lead will be estimating free ridership, that is, how much of the savings would have happened anyway had the RCx program not existed. For large, complicated non-residential projects, as RCx projects generally are, this effort is non-trivial. Every project sampled for the gross savings evaluation will receive the most detailed level of free ridership analysis, which means that in addition to decision maker and vendor interviews, other information will be reviewed regarding the customer's payback criteria, corporate policies and standard practice.

The net/spillover lead will be conducting the interviews and the bulk of the research for this effort. Your task will be set up this effort by prepping customers, identifying key contacts, and keeping tabs on when RCx projects are completed. The net/spillover lead will then examine information from program files and utility account reps. Next, he will conduct detailed surveys of the customer decision maker who elected to participate in the program, as well as the commissioning agents, contractors, and other "vendors" who were involved in the program.

Site Lead Free ridership Responsibilities

As site lead, you will need to support the free ridership survey effort in three ways, as listed here:

1. **Find out appropriate decision-maker:** During recruitment, identify the decision maker and get their contact information so the free ridership expert can contact them later. In many cases, there will be several individuals helping to make the decision. The facility engineer or manager may have initiated the project but may have needed approval from a corporate decision maker. Or, the decision may have come from someone higher in the organization telling the facility manager to participate. Any information on how different people contributed to the decision would be helpful, along with contact information from each.

   You can probe the site contact a bit if it is not obvious who the decision maker is. For example, if you find out the main contact started his job shortly after the RCx decision had been made, it would be helpful to ask "If you were not here, do you know who was responsible for the decision to participate?" Based on what you find out, the free ridership experts will do the follow-up investigation.

2. **Find out which vendors were involved in the project:** Either during recruitment or at the onsite visit, ask the site contact what vendors (i.e., service providers, such as RCx agents and/or controls

contractors) were involved in the project. If possible, collect the names of the vendor firms, the individuals involved, and phone numbers or e-mail addresses.

3 **Mention free ridership survey:** While recruiting, as well as during onsite visits, mention to the decision maker that others will be contacting him/her later, and explain why in simple terms. For example, you could say "For our evaluation, I will be focusing on trying to verify the actual energy savings, while others on our team will talk to you separately about the process by which your organization decided to participate in the SCE RCx program."

5.3.2.3 What Else Happened? (*Spillover*)

At any stage of the evaluation, you may encounter evidence of spillover measures. Spillover is a side-effect of RCx program participation, where participation in the RCx program influenced program participants to implement additional energy efficiency measures they would not have implemented otherwise. These spillover measures are implemented despite receiving neither rebates nor specific recommendations by the RCx program representatives. As such, they fall outside the scope of the direct impacts that are assessed for other program measures. Nonetheless, spillover measures must have been installed and operational before the end of the program cycle, that is, December 31, 2008 (or if the spillover measure is ongoing, then we are only concerned with its performance as of that date).

Your role is to ferret out potential spillover measures during the regular evaluation activities. While spillover is important, it is still secondary to assessing direct program impacts. Therefore, you should not spend an inordinate amount of time up front researching it—just enough so the evaluation management team can decide whether or not to formally evaluate the spillover measure(s). This will depend on whether (a) the measure has a sufficiently large impact on the evaluation, and (b) the utility program substantially influenced the action taken. Only if we decide to research the spillover further will you need to devote significant time to developing a spillover M&V plan, performing data collection and analysis, and reporting the results.

Site Lead Spillover Responsibilities

As the site lead, you support the assessment of spillover in the following ways:

1 **Check with customer about possible spillover:** During the recruitment call and site visit, ask the facility manager about additional measures installed beyond those rebated or recommended by the program. Use the follow questions to determine whether there is a likelihood of spillover:

- a *Did you implement any additional energy efficiency measures at this facility since your participation in the 2006–2008 Program?*
- b *For each measure: Was this measure specifically recommended by the RCx report or the commissioning agent? IF YES, treat as direct impact. If NO, ask:*
- c *Did you receive or do you expect to receive a rebate for this measure?*

1. *If YES, through which program?*
2. *If NO, how significant was your experience in the 2006–2008 RCx Program in your decision to implement this Measure, using a scale of 0 to 10, where 0 is not at all significant and 10 is extremely significant?*

If potential spillover measures are identified, collect enough information to describe well and probe for the underlying decision-making context. Also, collect contact information for the key site personnel involved in the spillover implementation decision.

2 **Consult with the project RCx engineer** to get their perspective on the influence of the RCx program on the potential spillover measures. If the customer says the measure was recommended by the RCx agent, but it was not included in the RCx report, ask the RCx agent why the measure was not included.

3 **Confer with evaluation team:** Review data collected for steps 1 and 2 with the group manager. The group manager will relay the information to the net and spillover lead, who may conduct a spillover interview with those whom you have identified as key players in the spillover decision making process. The net and spillover lead and the evaluation team will decide whether or not to proceed with evaluation of the spillover measure.

4 **Develop spillover M&V plan:** If approved, you and the group manager will develop a data collection and analysis approach for the approved spillover measure(s). This approach will depend on the expected magnitude of the spillover effect, the feasibility of collecting necessary information, and budget available.

5 **Analyze and report spillover impact:** Proceed with the spillover evaluation to develop energy savings estimates. Include these results in the final site Evaluation Report.

5.3.2.4 How Long Will It Last? (*Effective Useful Life*)

Assessing the effective useful life of RCx measure savings is beyond the scope of the evaluation effort that the site leads will undertake. We will research RCx EULs on an evaluation-wide basis, based on past ED assumptions, informed by a new study of the actual persistence of RCx measures from the 2002–03 and 2004–05 evaluation samples. These revised EULs may be used for future program cycles, but the original claimed EULs will be retained for the 2006–08 cycle.

5.3.3. Working Safely

During this evaluation, and particularly when team members venture into the field, safety is a prime concern. Accidents have the potential to occur at any time at any site, and the fact that power measurements are an important part of evaluation work adds a large measure of risk.

We highlight the following key concepts for all evaluation team members to keep in mind when doing field work. Before you go into the field, make sure you thoroughly understand each concept, as it applies to evaluation work. *Also keep in mind that the ultimate responsibility for the safety of you and the others around you rests with you, so stay informed and make wise choices.*

- **Safety Is First:** Safety for evaluation team members, as well as people at the sites we visit, is a primary concern.
- **Follow Procedures:** Discuss evaluation safety issues with site contacts beforehand, and follow their procedures.
- **Protect Yourself:** Wear appropriate clothing and safety gear.
- **Avoid Hazards:** Watch out for known unsafe conditions.
- **Work Carefully:** Be sure of what you are doing, and work carefully and deliberately.
- **Get Help:** Avoid working alone when doing dangerous tasks. If at all possible, have customer electricians or staff help open electrical panels and make electrical connections.

5.3.4. Site Evaluation Procedures

This section provides detailed guidance on the core tasks essential to evaluating each site. The four subsections represent the major tasks associated with each RCx evaluation project, namely:

- Recruiting the site
- Preparing the M&V plan
- Collecting data and analyzing savings
- Preparing the M&V report

Figure 5 below shows the general work flow associated with each site to be evaluated. Throughout the process, the site lead works closely with the group manager, and keeps him or her apprised of progress and obstacles.

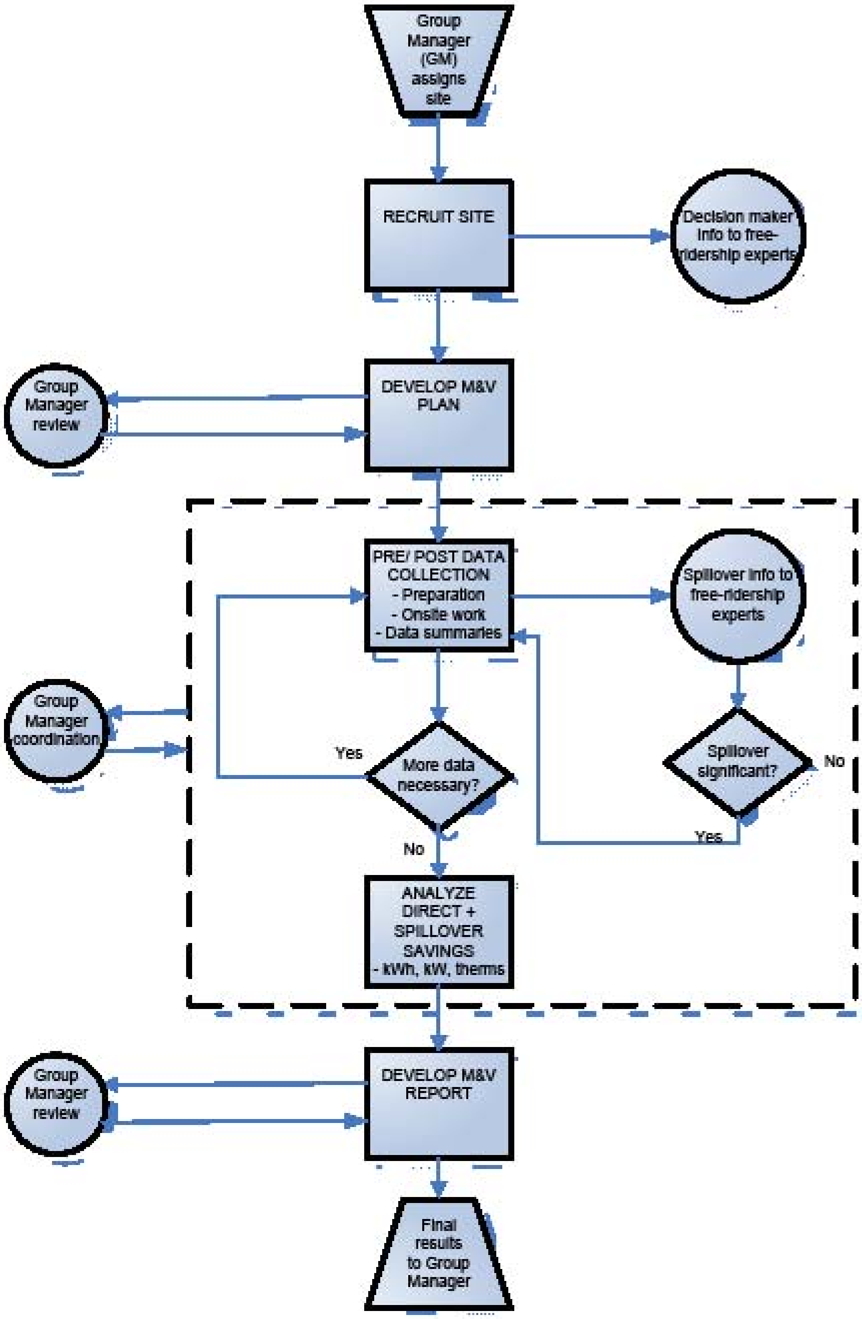


Figure 5: Site evaluation work flow

5.3.4.1 Recruiting the Site

You, as site lead, are responsible for getting the customer at the selected site to agree to participate in the evaluation. For each newly assigned project, you will receive a recruitment form as a worksheet in the site workbook. This form contains places to record important recruitment information.

The main steps for recruiting the customer are:

1 **Review the IOU project documents**, such as applications and engineering reports, to become familiar with the site participants and measures before contacting the customer. This review need not be exhaustive—just enough so you understand the site and the measures, and are comfortable discussing them with customer contacts.

2 **Contact the site authorization person.** When you've completed research on the project, contact the site authorization person by phone, email, or both (the specific approach is up to you, depending on preference or the expected preferences of the customer).

3 **Obtain authorization for participation in the evaluation.** Explain briefly to the site authorization contact what the evaluation process will involve for them. Key points include:

- When talking to the recruitment contact and the site authorization contact, emphasize that their participation in the evaluation will provide valuable information for their utility and the State of California to design future programs that produce cost-effective energy savings.
- When finalizing arrangements with the site authorization for the first site visit, explain that you'll need to be on-site for a full day typically, but you'll only need to meet with them for a short time at the start of the day.
- Evaluation field staff will be observing the equipment and its performance, but will strive to be as unobtrusive as possible. Data collection may include the installation of metering equipment and/or collection of trend logs.
- All data collected will be kept strictly confidential, and will not be shared with anyone outside of the utility or ED.
- This effort is an evaluation of a program, not of them. It will not affect any past program incentives they have received, nor their ability to participate in future programs.
- Mention that others will contact them separately regarding the process by which they decided to participate in the program (this is the free ridership part of the evaluation).
- Ask whom to schedule the site visit with. It may be possible to schedule the site visit at the time of recruitment but the M&V Plan must be written and approved before going on site. In most cases, it's preferable to say approximately when you'll be calling back to schedule the site visit to collect information necessary to evaluate the savings realized by the efficiency measures.

  If the customer needs additional information from the ED about the evaluation and the SBW Consulting team, then ED letters of introduction are available for your use.

4 **Send Confirmation Letter.** Immediately after recruitment is completed, prepare a confirmation letter and send it to the appropriate parties. As with all important communications, a cc should be sent to the ED contract manager for this evaluation. Keep the IOU representative and third party implementer informed as well.

5.3.4.2 Preparing the M&V Plan

With the project/site successfully recruited, you are ready to set up and write the M&V Plan. The Plan is a key document in the evaluation process as it summarizes what the program implementer did at the site, and how the evaluation is going to verify actual savings for all direct measures. The M&V Plan is also the first deliverable sent to the ED for review. It must be approved before proceeding with the field work.

5.3.4.2.1. Assemble and Review Project Information

Conduct a detailed review of all project documents and files provided. This review should be more thorough than the cursory review you did to prepare for scheduling. If the information is incomplete, describe the data you need as precisely as possible to the group manager, and the evaluation team will make a formal information request through the ED. Keep in mind that all IOU customer information, including spreadsheets and trend logs, are confidential.

5.3.4.2.2. Contact Project Personnel

After reviewing the project files, you'll probably still have questions about the project. The RCx project engineer is usually the best place to start. The RCx project engineer may be an employee of the third-party implementer or a sub-contractor. Additional sources for project-specific information include the third-party program manager or the IOU program manager. Check with the Evaluation Team for the communication and contact protocols that apply to each of the IOU programs.

If site contacts are willing to engage in technical discussions with you up front, then take advantage of this to enhance your plan. Otherwise, you can ask site personnel more questions after the M&V Plan is approved and you are on-site. It will be particularly important to try to reach the contacts if you need to administer a Baseline Survey to establish the savings baseline for each measure (either conditions before the measure was implemented or a threshold set by a code or standard).

5.3.4.2.3. Write the Plan

You'll be given a partially pre-filled, standardized M&V Plan with which to begin. The sample M&V Plan provides general tips to keep in mind when writing the M&V Plans and Reports. Additionally, your group manager can provide you with other relevant completed plans that can provide a starting point for your plan.

Each section of the Plan has detailed instructions to guide the content of your writing. In regards to writing style, strive to make your writing concise, particularly in the tables. Keep in mind the report audience is well-versed in engineering but not in the details of the particular site. Try to maintain the fine balance of providing just enough information. In addition to the section instructions embedded in the template, below are some comments to keep in mind while working in each of the major sections of the report.

Summary Information: Contains general site background and contact information. This information will be automatically imported from the site workbook.

Section 1 – Goals and Objectives: Describes the purpose of the evaluation at this site in general terms. It should only require a small amount of effort on your part to customize this section to reflect circumstances for the site.

Section 2 – Measure Description: Contains the measure descriptions and classifications you entered into the site workbook, as well as the 'best available' measure savings' from the database. You will also describe the equipment and operating conditions before and after the RCx effort. You must also use the guidelines in the Baseline Survey to establish the savings baseline for each measure (either conditions before the measure was implemented, or in rare cases, a threshold set by a code or standard).

Section 3 – Algorithms for Estimating Savings: The core of the evaluation planning process is the review of the methodologies used to estimate the IOU savings and, if necessary, the development of independent evaluation methodologies for re-estimating savings. As you review the IOU algorithms, identify and describe in the Plan the sources of data used for the key variables. Apply your professional judgment as to whether these data sources and methodologies are adequate for the evaluation algorithm. If not, propose an alternative methodology that you believe is superior.

Section 4 – Data Collection: Once you settle on the evaluation algorithms, list in the plan how each of the variables in the algorithms will be collected. Balance the desire for data quality and quantity with budgetary limits. Always feel free to talk it through with the group manager. You may also need to address sampling strategy if the measure has a large enough number of components to preclude a census. Consult the sampling discussion in section 5.3.4.6.1 for further guidance. You will also have a chance to identify uncertainties—that is, aspects of the plan that are not yet known, but could significantly affect it. In this section, you will also identify standard building characteristics data elements (also known as contextual data) relevant to the project. The site workbook contains tools to streamline this process.

5.3.4.2.4. Submit M&V Plan for Review

When you have completed the draft M&V plan, submit it to the group manager for review by uploading the site folder to the group manager. If you're in a remote location, you may use GoToMeeting and a phone line as a way to share and discuss visual information with the group manager in real time. Once you have revised the plan to the group manager's satisfaction, it will be submitted to the ED for their review and approval. Once the plan is approved, you may schedule your initial visit to the site and begin collecting data and analyzing savings.

5.3.4.3 Collecting Data and Analyzing Savings

Collecting and analyzing data from the customer site represents the majority of the impact evaluation effort, and involves many facets, techniques, and skills. This section presents an overview of key aspects of data collection and analysis, both general best practices for this kind of work plus requirements specific to this evaluation. It breaks out this phase of the work into four tasks:

1. Preparing for onsite work,
2. Structuring the onsite visit,
3. Summarizing data, and
4. Analyzing data.

The first two tasks pertain to onsite work, and they may be repeated in some manner as much as necessary for the specific site evaluation. For example, a pre-implementation site visit could be followed by first and second post-implementation visits, with possibly more visits in special cases. Each visit would involve some preparation and structuring.

5.3.4.3.1. Preparing for Onsite Work

Once the site M&V plan is approved, you can begin preparing for onsite work. Preparatory steps include:

1. **Schedule visit:** Coordinate well in advance with main site contact on an agreeable date. Make sure key personnel will be available—those who know the measures well, those who can navigate the EMCS, electricians or escorts if the customer requires them. Note that although the evaluation team will have notified the IOUs that the site was sampled, you are not required to have an IOU representative at the site when you are there.
2. **Sample (if necessary):** Certain measures might affect large numbers of similar pieces of equipment, such as AHUs, A/C units, or VAV boxes. If the program documentation includes inventories of such equipment, then it may be possible to select an evaluation sample in advance. Ideally, enough information exists about the population to permit an informed sampling approach that will balance the desire to minimize onsite efforts against the need for a reasonably robust sample. Once the sample is selected, it will dictate the amount of metering equipment required, as well as the time required to carry out the onsite work.
3. **Prepare toolkit:** Review the Data Collection section of M&V Plan carefully and inventory necessary metering equipment. Be sure to include some spares of critical components, as well as adequate equipment to handle all foreseeable contingencies. Obtain the metering gear per established procedures. Be sure you are familiar with the equipment before deploying it—time onsite is typically too limited to permit much on-the-job training.

   If you will be carrying equipment with you on a flight, be sure to pack it in accordance with airline weight and size regulations. Another alternative to consider is sending equipment via a commercial shipper (such as UPS or FedEx) to either your hotel or to a customer site, if the latter is amenable. If you ship equipment thusly, be sure to insure it adequately, and avoid sending tempting theft targets such as cameras, laptops, and power meters. A third option, during times of extensive field work, is to rent a centrally-located storage locker—preferably in conjunction with other site leads—and use this as a staging area for equipment and supplies.

4 **Get organized:** Have relevant documents, forms, maps, and plans ready before the trip. Confirm your appointment a day or two in advance, and find out exactly where and when you are to meet your contact. Review the M&V plan and program documents shortly before the visit, and if possible, obtain a site map to orient yourself. Make sure you have a clear idea of what you want to accomplish onsite, and how you plan to do it. While you must necessarily be flexible, having this plan will allow you to make smart choices about how to allocate your time while onsite.

Lastly, make sure you dress appropriately for the onsite visit. For the sake of both safety and decorum, wear long pants and sturdy, close-toed shoes. Your outfit should be business-like enough to comfortably meet facility managers or be seen amongst the facility clientele, but durable enough for data collection activities in dirty mechanical rooms and atop hot rooftops.

5.3.4.3.2. Structuring the Onsite Visit

While onsite, you will often need to deal with a wide range of facilities staff for the different elements of the site work. These individuals can run the gamut from the facility manager, chief engineer, energy manager, controls technician, mechanic, or electrician, to the custodian. Keep in mind that we are guests in their facility. Invariably these folks are busy, so be considerate of their time. If you anticipate a long site visit, ask your host if they need time for breaks or to attend to other business.

Throughout the visit, take good notes. The amount of information you receive during onsite visits can be bewildering, so having a written record of key points or observations can be very useful.

Below is a suggested order for site visits, but this is flexible and can be modified as needed. Be sure to develop back-up strategies beforehand should adjustments be necessary.

1 **Interviews:** Generally, it is best to talk to high-level managers early in the day before they get pulled away to other responsibilities. First, even if you have mentioned it to the managers before, it can be helpful to give them an overview of the evaluation process and how it fits in with the whole CPUC energy efficiency portfolio. Next, ask them open-ended questions about their project, and probe further as necessary. Helpful areas to discuss may include:

   a The full scope of the project (looking out for undocumented actions and spillover).

   b Measure completion status (whether the measures are fully done, for post-only sites, or when the measures are likely to be done for pre-post sites). Note that you may find out that measures that you thought were complete in actuality are not.

   c If the outcomes turned out differently than planned.

   d Any future plans that might affect the measure(s).

If they need to go, find out if there are good times later in the day to ask them follow-up questions.

2 **Quick walkthrough:** Take a cursory tour with the site contact of the measure-affected areas and equipment. If you do not already have one, ask the site contact for a simple map, so that you can maintain your bearings as you walk around. Some or all of the tour may be a virtual one, as reviewing the EMCS screens, controls, and capabilities might be important. The tour is an opportunity to get oriented, and to assess how difficult it will be to install metering equipment. While walking, chat with the escort and get additional information about the measures and general facility operations.

3 **Installing equipment:** Once you have settled on your strategy for installing monitoring equipment, inform the site contact in general terms of your intentions. As discussed previously in the safety section, be sure to check on safety procedures and housekeeping preferences (regarding, for instance, locking doors, opening electrical panels, leaving extension cords and metering equipment exposed, etc.). Once metering equipment is installed, download a sample of data to ensure the loggers are recording properly. Take pictures and/or write detailed descriptions of the locations of data loggers, particularly if there is a chance that someone else will be retrieving them. Clean up your work area, and affix equipment labels and warning signs as appropriate.

In some cases, it might make sense for site personnel to retrieve and send back data logging equipment. This can help you stretch your site evaluation budget. An example of a suitable site for this approach would be one where the loggers you left are few in number and easily retrieved, and skilled staff members are available to retrieve them. In such instances, make it convenient for them by leaving them with boxes and shipping numbers/labels. Do be clear exactly when you want the data loggers removed.

If you are requesting that the customer collect trend data for you, be sure to get information so you can specify exactly what points you want, and for how long. Clarify follow-up steps to ensure that the trend logs are being collected properly. Ideally, the customer will be able to provide early samples of the trend logs for you to check.

If time runs short, you may have to triage—accomplishing tasks essential to the visit, such as installing data loggers, while postponing other tasks, such as taking one-time power measurements, until later visits.

4 **Wrap-up:** Prior to leaving the site, check in with the site contact. Explain to them what you installed, where, and for how long. Ask any final questions, and explain what follow-up steps remain (e.g., evaluator needs to send data request to customer, and then customer needs to set up trend logs). Talk about the timing of future visits, if any. Answer any questions they may have, and thank them for their help. If they ask for their evaluation results, explain to them that we can provide them, but only after the full evaluation is complete. As discussed in prior sections, make every effort to minimize your influence on the project, whether pre or post implementation. Be cordial and tactful but minimize discussion of expected or actual measure outcomes.

Safeguard handwritten field notes, since these often contain information that can be difficult or impossible to replace. If time permits, make copies of them and stash them elsewhere in your luggage. In a similar vein, make sure important data files are stored in two places (e.g., your laptop as well as on a data stick).

Once you depart, you might follow up with thank you emails to those who assisted you during the site visit. If appropriate, copy their chain of command. This is also a good opportunity to reiterate follow-up needs such as trend logging or other information.

5.3.4.3.3. Summarizing Data

Data collection at an RCx site can involve dozens, and even hundreds of channels of data, with literally millions of data points. It can be challenging to assess these data for reasonableness, and then to distill them down for the savings analysis. Some general suggestions for handling these tasks are as follows:

1 **In the field:** Several software tools exist to allow you to quickly visualize time-series data files. These include (a) the graphing functions of the metering equipment software, (b) Excel's graphing capabilities, (c) LogTool[3], and (d) the Universal Translator. Examine the magnitude and the patterns of the data to see if they make sense for the equipment and the site. If you spot a significant discrepancy or odd finding, you can ask the site contacts if they have any explanations.

2 **After the visit:** As soon as possible after the visit, begin summarizing and filling in gaps in field notes, while information about the site is still fresh. Make sure you have a complete listing of the temporary metering that is in place. If you plan to have site personnel remove the equipment, then supply them with the listing as well.

Once all trend logs and metering files are in hand, perform a more careful and thorough summarization back in the office. Condition the data files and aggregate them using appropriate software (e.g., Excel, MS Access, LogTool, Universal Translator, and so on). As in the field, examine the data for reasonableness. Look for operating patterns based on time of day, day of week, outside air temperature, or other independent variables. If appropriate, set up data summaries by outside air temperature bins, so that equipment and system performance can be matched up to typical climate zone weather bins.

5.3.4.3.4. Analyzing Data

The manner in which you might analyze evaluation data to estimate energy savings can vary significantly, depending on the measures being studied. Some may be very simple, requiring just a handful of rows on a spreadsheet. Others may be fiendishly complex, using detailed summaries of millions of data points to set up multi-sheet Excel workbooks. Nonetheless, *the overarching standard for this evaluation is to structure and document our savings analyses completely*, so that an experienced energy analyst could open the evaluation files years later and understand and affirm what you did simply by looking at the M&V report, site workbook, and any auxiliary files.

While you have a great deal of leeway in structuring your analysis, two aspects need to conform to standardized procedures. First, the summarized results must ultimately flow into the Site Workbook. Second, the analysis should be consistent with the calculation approaches explained below. The latter section provides a mix of requirements, suggestions, and resources to both facilitate and standardize

---

[3] LogTool is a public domain software tool developed by SBW Consulting, Inc..

savings calculations. While it documents the evaluation team's understanding of current best practices, they are subject to revision as the evaluation proceeds. We welcome additions and refinements that will improve everyone's ability to do this work well.

Typical Weather

The ED has defined typical weather conditions for the state of California to be described though the California Climate Zone scheme[4]. In this scheme, the state is divided into 16 discrete zones. Each zone is assigned "typical" dry-bulb and wet-bulb temperatures for each hour of the year, based on historical data.

Peak Savings

The ED has defined the peak period for estimating electrical demand reduction to be the average kW reduction between 2:00 and 5:00 p.m. for the three consecutive weekdays containing the weekday with the hottest temperature of the year. These three days vary by climate zone, as shown in Table 26. The typical weather file for each site will identify the peak period for your site's location, and provide a means to create the corresponding temperature bins for analysis purposes.

Table 26: Peak Days by Climate Zone

Climate Zone	Reference City	Peak Days*
1	Eureka	Sep 30 – Oct 2
2	Napa	Jul 22–24
3	Oakland, San Francisco	Jul 17–19
4	San Jose	Jul 17–19
5	Santa Maria	Sep 3–5
6	Los Angeles (LAX)	Jul 9–11
7	San Diego	Sep 9–11
8	Long Beach	Sep 23–25
9	Los Angeles (Civic Center)	Aug 6–8
10	Riverside	Jul 8–10
11	Red Bluff	Jul 31 – Aug 2
12	Stockton	Aug 5–7
13	Fresno	Aug 14–16
14	Barstow	Jul 9–11
15	Brawley	Jul 30–Aug 1
16	Bishop	Aug 6–8

*Based on a 1991 reference year for defining weekdays.

[4] The official designation is: CZ2 (California Climate Zones Revision 2, 16 zones, 1992 supplied by the California Energy Commission).

Annual Energy Savings

Take great care in extrapolating short-term results—typically obtained over several weeks to several months—to a typical year. The appropriate means of doing so will depend on the particulars of the measure and the site. Explanatory factors may include outside air temperatures (dry bulb or wet bulb), building occupancy, production levels, or academic schedule, to name a few. *If the measure is weather-dependent, then be sure to use typical weather to estimate savings, as described above. In all other instances, results should be extrapolated to the first year after the measure was implemented*.

Interaction Between Measures

Often with RCx measures, the performance of one measure will affect another, and vice versa. An example might be a project where economizer performance was improved, and the chilled water system was optimized. The more efficient chillers would reduce the savings from the repaired economizers, and likewise, the improved economizers would reduce the load on the chillers. *These measure interactions need to be taken into account*. Our aim is to estimate the savings for the entire project, and not necessarily for the individual measures. Therefore, in many cases, it will be easier and more accurate to combine the effects of several interacting measures in one analysis rather than perform separate measure analyses.

Equipment Performance Curves

Much of impact analysis is determining how equipment performs under different conditions, and developing curves that define that performance. These curves can then be used to model pre- and post-implementation conditions. A wide range of techniques exists for developing these curves. The evaluation team has grouped them into four categories, which are listed below in descending order of rigor. You should incorporate these techniques, as appropriate to the situation, in your analyses.

- Regression analyses of metered data (linear, non-linear, multivariate)
- Manufacturer's curves
- Generic curves, such as those developed for DOE-2 and eQUEST modeling
- Affinity laws

Each of these techniques is described further in section 5.3.4.6.3. This section also provides reputable sources for generic curves.

Past Examples

The evaluation team possesses extensive libraries of past analyses of RCx measures at a wide variety of sites. If you feel you could benefit from seeing one or more past examples, then consult with your Group Manager, who will be able to select appropriate examples for your use.

Miscellaneous Issues

The following issues are presented for the sake of completeness. We expect them to be rare. If you come across a site where any of these situations applies, then consult with your group manager.

- **Building simulation programs:** In past RCx evaluations, program implementers occasionally used packaged simulation software, such as DOE-2, eQUEST or VisualDOE, to estimate savings from measures. In certain situations, the best way to evaluate the measure might be to rerun the simulation with enhanced inputs.

- **Interaction between lighting, heating, and cooling:** If an RCx project results in major lighting savings within a conditioned space (for example, commissioning of a daylighting system in an office, or improving lighting scheduling), then the reduced lighting energy use can (a) reduce cooling loads and/or (b) increase heating loads in the building.
- **"Retrofit" measures:** Sometimes RCx projects extend beyond the low/no-cost measures associated with RCx to include what are traditionally thought of as capital, or retrofit, measures. These might include:
  - Replacing major HVAC components, such as a chiller, cooling tower, boiler, pump, or motor.
  - Replacing lighting fixtures, lamps, or ballasts.

If you come across measures similar to these, confer with the group manager to determine whether the standardized procedures developed for the Major Commercial contract group might apply. If so, you can use those procedures as a starting point, rather than starting from scratch.

- **Cogeneration:** Some larger sites have cogeneration systems (also known as CHP, or combined heat and power, systems). These systems generally use natural gas to generate both electricity and heat. Often this heat is passed through absorption chillers to generate cooling. If such a system is in operation at the site, allocating electric and gas savings to RCx measures can become very complicated. For example, reduced fan use might actually save gas, since the reduced electric load would also reduce the gas needed to produce that electricity in the cogenerator.

5.3.4.4 Preparing the M&V Report

The final step in the project evaluation process is preparation of the M&V Report, which summarizes all findings from the evaluation process. When the analysis is complete and the results compiled, assemble the supporting documentation and relevant background information.

You will use the approved M&V Plan as a basis for the M&V Report. This report should essentially update the M&V Plan to reflect how the evaluation analysis actually took place, with updated Project Summary, Measure Description and Algorithms sections. In addition, it contains a new Findings section for providing evaluation results in both narrative form and tabular form. It may be appropriate to discuss where the evaluation process diverged for the M&V Plan, if that occurred. Spillover measures, which were unknown when the M&V Plan was written, may be discussed in the report.

The group manager will review and approve the report. For this review, as well as other subsequent external reviews, you can anticipate challenges to the conclusions. Consequently, all results need to be rigorously reasoned and documented.

5.3.4.5 Data Quality Assurance Guidelines

Throughout the evaluation process, you will need to assess, and take steps to ensure, the quality of the data you use to estimate savings. These data sources could include, but are not limited to:

- Trend logs from the customer's EMCS.
- Time-series data from temporary data loggers.
- Other time-series data from external sources, such as actual hourly weather data, utility interval data, or production records.
- One-time measurements of key parameters.
- Any data of these types that the program implementer collected as part of their engineering study.

By paying attention to how the data were collected, how they might be in error, and what the data trends reveal, you can better understand and find ways to minimize the uncertainty in your savings analysis. Below are several best practices that you should apply during the data collection process.

Be aware of potential sources and impacts of error: For data that you did not collect yourself, consider the source, and how reliable it might be. If it is from a customer EMCS, take stock of—as much as is reasonable—the age of the system, how well maintained and calibrated it might be, and the location and type of sensors. Pay particular attention to situations where the accuracy of a given measurement has significant impacts on savings estimation. An example might be a measure where reducing a temperature setpoint by two degrees yielded significant savings. In such a case, a temperature sensor accurate within ±1°F could conceivably introduce error in the savings estimate of 50%.

High accuracy is always important for outside air temperatures (OAT), because they are critical in bin and regression analyses. OAT measurements are also often confounded by poor sensor placement, such as a sensor placed in direct sunlight, or next to a very hot surface. Whenever possible, note whether or not sensors appear to be placed in suitable locations. One example might be an air handler mixed air sensor placed in the middle of the mixing chamber, or is it off to the side, close to the outside air intake, so that the mixed air temperature will be biased towards the OAT. Another example is a flow sensor that is located too close to a series of piping elbows, so that the sensor sees turbulent, rather than laminar, flow.

Another source of error is extrapolation from the short-term metering period to the year. Depending on the time of year and the variation in energy use at the facility, short-term data might not necessarily provide a representative snapshot of long-term operating conditions.

Use redundant metering if appropriate: In certain cases, you may opt to employ redundant metering to ensure that you get useable data. If you have asked a customer to trend numerous data points, but you have some doubts as to whether it will get done, you could also install your own loggers in selected key areas for "insurance." Keep in mind that this strategy has limits, as an EMCS can usually record far more points than you can with temporary loggers.

Check initial data sets: It is generally good practice to perform checks of time-series data right after they are set up to be sure that data points are being captured properly. Once a customer sets up a trend log, try to have them send you an early example so you can confirm that it was set up properly (although be warned that it can be difficult to convince customers to do this for you). If you are setting up data loggers, download the initial records and re-examine the configuration settings so that you can be sure the data are accurate, and the logger will operate for a sufficiently long period. Use your judgment to prioritize which loggers and data channels should be checked. For instance, it is a fairly safe bet that a dozen singlechannel temperature loggers measuring a sample of room temperatures will provide adequate data, so this could be a low priority. On the other hand, a logger programmed onsite to capture a single critical power measurement should be very carefully checked.

Check final data sets: Ask yourself these questions to help ensure that the data you have collected are of good quality. Note that some of these questions may apply for one-time measurements:

1 *Are the values and range for the data reasonable?* You should have some general sense of what range of values are appropriate for a given measurement. For example, one would not expect a 5-hp motor to yield readings of 67 amps. Nor would one expect power data to be negative. Points to compare against could include design values, one-time readings from handheld meters, EMCS readings, local gauge readings, nameplate values, or even past experience. Ideally, each of these points should reinforce each other, so that they paint a consistent picture. For instance, if a facility engineer states that a fan always runs at half speed, you could check to see if the VFD display, trend logs, one-time measurements, and visual/aural observations of the fan all support the statement.

2 *Do the data fit the expected profile?* You should have some sense of building operations, such as whether certain spaces or systems run continuously, are turned on and off according to occupancy schedules, or vary automatically according to some independent variable, such as outside air temperatures. Check if the data you have collected are consistent with the expected profiles, and if not, investigate for reasons why. For instance, one would expect cooling loads in a hotel to be higher when it is warmer outside. In another example, based on the description of an implemented RCx measure, you might predict that condensing water temperature will vary according to outside air temperature. A flat temperature profile might indicate that the measure is performing differently than predicted.

3 *Are there outliers?* If so, can these be explained?

4 *Are related data values and streams consistent with each other?* For example, are mixed air temperature readings between the outside and return air temperatures? (This should be the case, since mixed air is a combination of the two).

5 *Are there missing data?* If so, is there a reasonable scheme for filling the missing points?

If the responses to one or more of the questions above indicate problems with the data, then possible reasons could be one or more of the following:

1 *Data channels mislabeled.* This can happen either in the logger configuration file, or in the field notes. If caught in time, some onsite detective work can rectify this problem.

2 *Logger scaling factors incorrect.* An example of this would be a sensor with a range of 0–100 amps, inadvertently programmed in the logger as having a 0-200 amp range, so that recorded values are double what they should be.

3 *Logger failed.* For instance, sensors can be disconnected or damaged, or logger batteries can fail.

4 *Sensor poorly located.* For example, an outside air temperature sensor located in direct sunlight will register much higher readings than actual.

5 *Sensor poorly calibrated*, as might be found in a very old EMCS where sensors had not been checked or adjusted for accuracy in decades.

6 *Anomalous events occurred.* Data can be affected by unusual customer events, such as power outages, equipment failures, maintenance outages, or major changes in operations.

In some cases, you may be able to control and correct data problems (for instance, Problems 1 or 2). In others, you may have little or no leeway, and will have to use your best judgment about whether the data is sufficiently robust to carry forward to the analysis. In any case, the data flow should be clearly documented, from the raw data source through any data cleaning and adjustments to the final summarized data products.

5.3.4.6 Analysis Techniques

5.3.4.6.1. Sampling

During onsite work, you may only have time or resources to investigate a fraction of the affected equipment. In such instances, there are some fairly simple methods you can apply to randomly select a sample, and to estimate the number of sample points you need to characterize what is happening with reasonable accuracy. Provided below is a real-world example, in which the program calculations included a spreadsheet listing all affected air handlers. The general steps you can follow to create a sample are listed below, along with a description of how they play out in the example shown in Figure 6.

	Symbol	VAV	EconOpp/1000*Hrs/day	Cum%	Strata	Random #	Service	Hours Operated	CFM Total	CFM OSA	CFM Econ opportunity	CFM %
1	AH-10		746	12%	**Certainty**	58	Bldg "H" Prefunction (West)	24	34,330	3,250	31,080	9%
2	AH-25		611	22%	**Certainty**	99	Bldg "A" Atrium (SW)	24	27,700	2,250	25,450	8%
3	AH-24		605	31%	**Certainty**	41	Bldg "A" Atrium (SE)	24	27,440	2,250	25,190	8%
4	AH-22		453	38%	**Certainty**	40	Bldg "A" Atrium (NW)	24	21,140	2,250	18,890	11%
5	AH-14		418	45%	**Certainty**	16	Bldg "E" Main Kitchen	24	28,500	11,100	17,400	39%
6	AH-23		357	51%		29	Bldg "A" Atrium (NE)	24	17,140	2,250	14,890	13%
7	AH-16		303	55%	Random	17.0	Bldg "E" Level 252 Public Corr.	24	14,645	2,000	12,645	14%
8	AH-2		284	60%		19	Bldg "H" Ballroom Salon 7	15	23,800	4,900	18,900	21%
9	AH-3		284	64%		47	Bldg "H" Ballroom Salon 8	15	23,800	4,900	18,900	21%
10	AH-9		273	69%		51	Bldg "H" Prefunction (North)	24	12,430	1,060	11,370	9%
11	AH-17		181	72%	Random	7	Bldg "E" Multipurpose Restaurant	19	10,775	1,225	9,550	11%
12	AH-15	1	179	75%		69	Bldg "F-1" Level 252 Meeting Rooms	15	14,800	2,850	11,950	19%
13	AH-7		176	77%	Random	5	Bldg "H" Exhibit Hall Salon G	15	15,700	3,950	11,750	25%
14	AH-6		171	80%		41	Bldg "H" Exhibit Hall Salon F	15	15,350	3,950	11,400	26%
15	AH-21		149	82%		74	Bldg "C-1" Bar & Grille	14	12,340	1,720	10,620	14%
16	AH-1	1	147	85%		93	Bldg "H" Ballroom Salons 1-6	15	12,300	2,500	9,800	20%
17	AH-4	1	147	87%	Random	3	Bldg "H" Ballroom Salons 9-14	15	12,300	2,500	9,800	20%
18	AH-26		128	89%		58	Tuscany's Restaurant	14	11,255	2,100	9,155	19%
19	AH-11		122	91%		86	Bldg "H" Service Corridor	15	8,600	500	8,100	6%
20	AH-12		103	93%		98	Bldg "H" Finish Kitchen	15	10,500	3,650	6,850	35%
21	AH-13		102	94%		57	Bldg "E" Laundry	14	23,600	16,300	7,300	69%
22	AH-8	1	102	96%		17.1	Bldg "H" Exhibit Hall Salons H-L	15	8,850	2,050	6,800	23%
23	AH-5	1	89	97%		51	Bldg "H" Exhibit Hall Salons A-E	15	7,950	2,050	5,900	26%
24	AH-20		63	98%		55	Bldg "C-1" Entertainment Lounge	8	10,090	2,200	7,890	22%
25	AH-19	1	58	99%		93	Bldg "D" Japanese Restaurant	14	20,150	16,000	4,150	79%
26	AH-18		49	100%		29	Bldg "E" Seafood Restaurant	14	14,500	11,000	3,500	76%

Sum of all	**6,299**	**9 cases selected**
Sum of non certainty	**3,467**	
Mean	**165**	
Std deviation	**88**	
Coeff of variation	**53%**	
n (sample)	**4**	
N (population, non-certainty)	**21**	
Relative precision (for non-certainty part)	**39%**	<-- rp=1.645*sqrt(1-(n/N))*(cv/sqrt(n))
Total relative precision	**22%**	

Figure 6: Site sampling example.

1. If a list exists (of rooms, fixtures, AC units, etc.), then figure out a size variable that correlates well to energy use (such as square feet, installed watts, operating hours, horsepower, or tons). In the example, subtracting the OSA CFM from the total CFM yielded an estimate of how many more CFM the air

handler economizer could handle if working properly. Multiplying this by Hours Operated provided the "EconOpp/1000×Hrs/day" variable, which we assumed correlated well to energy savings.

2 Sort the list in descending order by the size variable, and calculate the cumulative percentage of the size variable total that each case contributes. Decide if a certainty stratum makes sense—that is, if a relatively small number of cases account for a large percentage of the total (in this example, it does, since five (20%) of the 26 items account for 45% of the total). If so, select certainty cases.

3 Using Excel functions, calculate the sample and population counts, mean, and standard deviation for the non-certainty stratum. Then calculate the coefficient of variation and relative precision for the non-certainty stratum. To calculate the overall relative precision, multiply the relative precision for the non-certainty stratum by the total of the "EconOpp/1000×Hrs/day" variable for this stratum. Then divide the result by the overall total (since the certainty stratum has no sampling error). For the example, this yields an overall relative precision of 22%.

4 Select enough random sample points so that the overall sample size is feasible, but yields reasonable relative precision (note that this may not always be possible, so some judgment is necessary here). In the example, we chose to sample 4 of the 21 non-certainty cases.

5 Select some replacement cases, should some cases prove impossible to treat (i.e., equipment is inaccessible or was taken out of service). Assuming 3 replacement cases, a total of 7 non-certainty cases would be required, so every third case would be randomly selected.

More details on this approach can be found in the California Evaluation Framework[5].

5.3.4.6.2. Temperature Data

The climate zone weather data file for each site contains typical hourly weather data (dry-bulb and wet-bulb temperatures) for the location, based on the appropriate California Climate Zone. The file contains easy-to-use tools for summarizing these data into temperature and hour bins for analysis. It also contains hourly conditions for the nine ED-defined peak hours.

In many cases, you will also need actual hourly weather data (generally, dry bulb temperatures) over the same time period as other data obtained from loggers or EMCS trends. For this, you have several options:

- Data logger installed temporarily at a suitable location onsite.
- EMCS trend data (although be warned that miscalibrated sensors and poor sensor location can introduce errors into these data).
- Data subscription from the NOAA National Climatic Data Center (http://www.ncdc.noaa.gov/oa/ncdc.html). For past projects, we have purchased annual subscriptions to hourly data for a particular weather station for $29. Both edited and unedited data streams are available. The latter, while it may

[5] The California Public Utilities Commission's 2004 California Evaluation Framework has a discussion of simple random sampling in Chapter 13, pp. 1318-327). It is available for download at http://www.cee1.org/eval/CEF.pdf.

contain missing values, has the advantage of being available within weeks of the actual data collection times.

- Data purchase from Weather Bank, Inc. (http://www.weatherbank.com/archive.html). This private firm maintains hourly and daily historical data records from every National Weather Service reporting station in the United States. They currently archive weather data on a real-time basis, and even update data for some weather stations hourly.

5.3.4.6.3. Developing equipment performance curves

Much of impact analysis is determining how equipment performs under different conditions, and developing curves that define that performance. These curves can then be used to model pre- and post-implementation conditions.

This section briefly explains four techniques for developing performance curves. These are listed below in descending order of rigor. The most accurate approach is linear regression with metered data. This is followed by manufacturer's curves, and then by generic curves, such as those developed for DOE-2 and eQUEST modeling. The simplest approach applies affinity laws.

1) Regression Analyses with Metered Data

Metered data can provide the most accurate means of characterizing equipment performance. If time-series data for independent and dependent variables associated with equipment is available (for instance, for a chiller, kW and tons of cooling, or tons and outside air temperature), then you can develop a correlation between the variables using built-in tools in MS Excel. The general rule of thumb is that an R^2 statistic exceeding 0.7 indicates a reasonable correlation.

The simplest way is to apply the "trend line" option when drawing an X-Y plot in Chart Wizard. More sophisticated is "Tools-Data Analysis-Regression" choice on the menu. Example outputs from each of these methods are shown in Figure 7.

2) Manufacturer's Curves

Certified performance curves issued by the equipment manufacturer can be quite useful, even though they reflect design performance, rather than performance under actual operating conditions. Using the equipment make and model number, you can often obtain curves from manufacturers online. Customers themselves may have the curves available in their onsite archives.

3) Generic Curves

If information specific to the installed equipment is unavailable, the next best thing is to apply generic curves developed for the DOE-2/eQUEST simulation model. The various references below describe curve parameters and the inputs necessary for them (for example, to determine part-load efficiency for a boiler, the equation would be:

$$\text{Efficiency}_{\text{part-load}} = \text{Efficiency}_{\text{full-load}} \times (a + b \times \text{part-load-ratio} + c \times \text{part-load-ratio}^2),$$

where a=0.082597, b=0.996764, and c=-0.079361.

TWO EASY WAYS TO DO REGRESSIONS

(1) TOOLS - DATA ANALYSIS - REGRESSION

DATA

a	b
1	1.1
4	6
2	4
3	4.5

y

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.956751
R Square	0.915372
Adjusted R Square	0.873059
Standard Error	0.730753
Observations	4

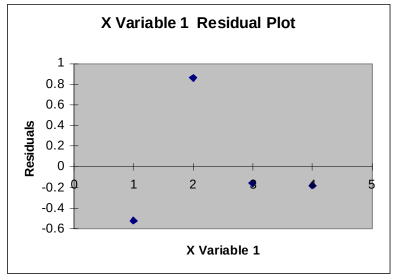
ANOVA

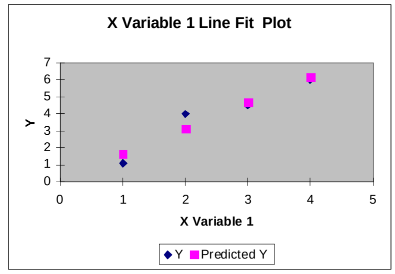
	df	SS	MS	F	ignificance F
Regression	1	11.552	11.552	21.63296	0.043249
Residual	2	1.068	0.534		
Total	3	12.62			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.1	0.894986	0.111734	0.921238	-3.750814	3.950814	-3.750814	3.950814
X Variable 1	1.52	0.326803	4.651124	0.043249	0.113882	2.926118	0.113882	2.926118

RESIDUAL OUTPUT

Observation	Predicted Y	Residuals	Standard Residuals
1	1.62	-0.52	-0.871522
2	6.18	-0.18	-0.301681
3	3.14	0.86	1.441363
4	4.66	-0.16	-0.268161





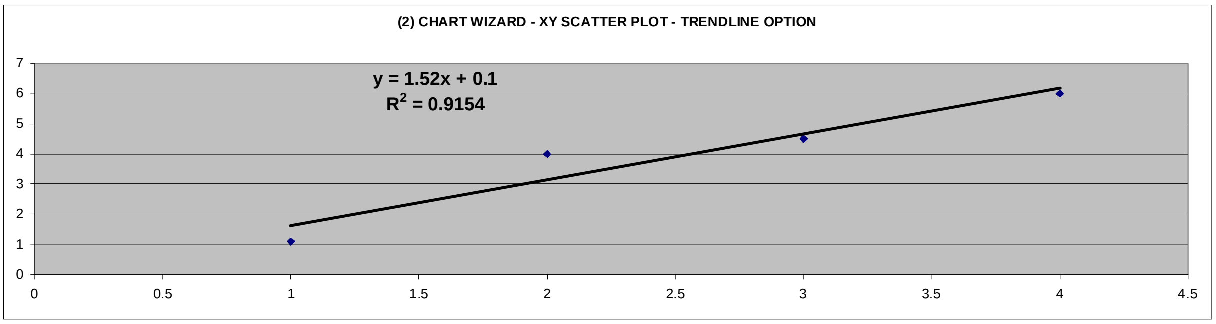


Figure 7: Simple regression examples using Excel tools.

DOE-2.1 California Compliance Supplement – Fan Speed Controls – Power Relationships (4/8/94)

This document (DOE21 California Compliance Supplement.pdf) contains information on:

- Variable flow fan curves
- Fan efficiency
- Motor efficiency
- Drive efficiency
- Speed control efficiency

DOE-2/eQUEST Manuals

Some versions contain information on part load curves for:

- Absorption chillers
- Compression chillers
- Double-bundle chillers

- Heating equipment (domestic hot water heaters, furnaces, boilers)
- Cooling towers

The file DOE-2EngineersManualVersion2.1A.pdf is available online, but does not contain the curve coefficients. The 2.1D version contains them, as likely does the 2.1E version. The DOE-2.2/eQUEST manual should have the most up-to-date numbers, but the current edition does not include Mechanical Equipment curve coefficients. Please notify your group manager if you believe you could use information from this source.

2007 Standard Performance Contract Procedures Manual – Appendix C

This appendix of the SPC Manual (SCE_C_Min_Equipment_Efficiency_041207.pdf) contains minimum equipment efficiency standards developed for the California Standard Performance Contracting program. Among the topics it covers are baseline efficiencies for:

- Motors
- Packaged HVAC units
- Heat pumps
- Chillers
- Furnaces and boilers
- Condensers

It also contains part-load curves for air- and water-cooled chillers.

4) Affinity Laws

For the simplest analyses, the affinity law relationships may suffice. The affinity laws express the relationship between several variables involved in pump or fan performance (such as head, volumetric flow rate, shaft speed, and power). They apply to pumps, fans, and hydraulic turbines, and for both centrifugal and axial flows.

The affinity laws state that:

- Flow is proportional to shaft speed.
- Pressure or Head is proportional to the square of shaft speed.
- Power is proportional to the cube of shaft speed.

These laws assume that the pump/fan efficiency remains constant. Some analysts have found that using an exponent of 2.7 in the third affinity law yields values that closely approximate the more rigorous generic curves developed for DOE-2 (described above). As of yet, we have not found an official reference for this “rule-of-thumb” value.

5.4. Net Savings Evaluation Methodology

5.4.1. Overview of the Large Nonresidential Free Ridership Approach

The methodology described in this section was developed to address the unique needs of Large Nonresidential customer projects developed through energy efficiency programs offered by the four California investor-owned utilities and third-parties. This method relies exclusively on the Self-Report Approach (SRA) to estimate project and program-level Net-to-Gross Ratios (NTGRs), since other available methods and research designs are generally not feasible for large nonresidential customer programs. This methodology provides a standard framework, including decision rules, for integrating findings from both quantitative and qualitative information in the calculation of the net-to-gross ratio in a systematic and consistent manner. This approach is designed to fully comply with the *California Energy Efficiency Evaluation: Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals* (Protocols) and the *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches* (Guidelines), as demonstrated in Supplement D.

This approach preserves the most important elements of the approaches previously used to estimate the NTGRs in large nonresidential customer programs[6]. However, it also incorporates several enhancements that are designed to improve upon that approach, for example:

- The method introduces a 0 to 10 scoring system for key questions used to estimate the NTGR, rather than using fixed categories that were assigned weights (as was done previously).
- The method asks respondents to jointly consider and rate the importance of the many likely events or factors that may have influenced their energy efficiency decision making, rather than focusing narrowly on only their rating of the program's importance. This question structure more accurately reflects the complex nature of the real-world decision making and should help to ensure that all non-program influences are reflected in the NTGR assessment in addition to program influences.

It is important to note that the NTGR approach described in this document is a general framework, designed to address all large nonresidential programs. It was customized to reflect the requirements of the retro-commissioning HIM evaluation.

5.4.2. Basis for SRA in Social Science Literature

The social sciences literature provides strong support for use of the methods used in the SRA to assess program influence. As the *Guidelines* notes,

> More specifically, the SRA is a mixed method approach that involves asking one or more key participant decision-makers a series of structured and open-ended questions about whether they

[6] Such as, for example, the NTGR method used to evaluate NTGRs for the California Standard Performance Contracting Program.

would have installed the same EE equipment in the absence of the program as well as questions that attempt to rule out rival explanations for the installation (Weiss, 1972; Scriven, 1976; Shadish, 1991; Wholey et al., 1994; Yin, 1994; Mohr, 1995). In the simplest case (e.g., residential customers), the SRA is based primarily on quantitative data while in more complex cases the SRA is strengthened by the inclusion of additional quantitative and qualitative data which can include, among others, in-depth, open-ended interviews, direct observation, and review of program records. Many evaluators believe that additional qualitative data regarding the economics of the customer's decision and the decision process itself can be very useful in supporting or modifying quantitatively-based results (Britan, 1978; Weiss and Rein, 1972; Patton, 1987; Tashakkori and Teddlie, 1998).[7]

More details regarding the philosophical and methodological underpinnings of this approach are in Ridge, Willems and Fagan (2009), Ridge, Willems, Fagan and Randazzo (2009) and Megdal, Patil, Gregoire, Meissner, and Parlin (2009). In addition to these two articles, Supplement A provides an extensive listing of references in the social sciences literature regarding the methods employed in the SRA.

5.4.3. Free Ridership Analysis by Project Type

There are three levels of free ridership analysis, two of which were used in the RCx evaluation. The most detailed level of analysis, the **Standard – Very Large Project** (S-VL) NTGR, is applied to the largest and most complex projects (representing 10 to 20% of the total) with the greatest expected levels of gross savings[8] The **Standard** NTGR, involving a somewhat less detailed level of analysis, is applied to projects with moderately high levels of gross savings. The least detailed analysis, the **Basic** NTGR, is applied to all remaining projects. For the RCx evaluation, the S-VL analysis was applied to all projects in the Gross impact sample, while the Basic analysis was applied to the remaining projects.

5.4.4. Sources of Information on Free Ridership

There are five sources of free ridership information in this study. Each level of analysis relies on information from one or more of these sources. These sources are described below.

1 **Program Files** As described in previous sections of this report, programs often maintain a paper file for each paid application. These can contain various pieces of information which are relevant to the analysis of free ridership, such as letters written by the utility's customer representatives that document what the customer had planned to do in the absence of the rebate and explain the customer's motivation for implementing the efficiency measure. Information on the measure payback with and without the rebate may also be available.

2 **Decision-Maker Surveys** When a site is recruited, one must also determine who was involved in the decision-making process which led to the implementation of measures under the program. They

[7] *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches*, October 15, 2007, pg. 3.

[8] Note that we do not refer to an Enhanced level of analysis, since this is defined by the Protocols to involve the application of two separate analysis approaches, such as billing analysis or discrete choice modeling.

are asked to complete a Decision Maker survey. This survey obtains highly structured responses concerning the probability that the customer would have implemented the same measure in the absence of the program. First, participants are asked about the timing of their program awareness relative to their decision to purchase or implement the energy efficiency measure. Next, they are asked to rate the importance of the program versus non-program influences in their decision making. Third, they are asked to rate the significance of various factors and events that may have led to their decision to implement the energy efficiency measure at the time that they did. These include:

- the age or condition of the facility or equipment,
- information from a feasibility study or facility audit
- the availability of an incentive or endorsement through the program
- a recommendation from an equipment supplier, auditor or consulting engineer
- their previous experience with the program or measure,
- information from a program-sponsored training course or marketing materials provided by the program
- the measure being included as part of a major remodeling project
- a recommendation from program staff, a program vendor, or a utility representative
- a standard business practice
- an internal business procedure or policy
- stated concerns about global warming or the environment
- a stated desire to achieve energy independence.

In addition, the survey obtains a description of what the customer would have done in the absence of the program. If they would have performed some sort of retro-commissioning activity in the absence of the program, the decision maker is asked to description what would have been done and how much would have been spent. This is used to adjust the no-program score in the calculation of the NTGR for partial free ridership, as discussed in section 5.4.5.2.

This survey contains a core set of questions for **Basic** NTGR sites, and several supplemental questions for both **Standard and Standard – Very Large** NTGR sites For example, if a Standard or Standard-Very Large respondent indicates that a financial calculation entered highly into their decision, they are asked additional questions about their *financial criteria* for investments and their rationale for the current project in light of them. Similarly, if they respond that a *corporate policy* was a primary consideration in their decision, they are asked a series of questions about the specific policy that led to their adoption of the installed measure. If they indicate the installation was a *standard practice*, there are supplemental questions to understand the origin and evolution of that standard practice within their organization. These questions are intended to provide a deeper understanding of the decision making process and the likely level of program influence versus these internal policies and procedures. Responses to these questions also serve as a basis for consistency checks to investigate conflicting answers regarding the relative importance of the program and other elements in influencing the decision. In addition, **Standard – Very Large** sites may receive additional detailed probing on various aspects of their installation decision based on industry- or technology-specific issues, as determined by review of other information sources. For Standard-Very Large sites all these data are used to construct an internally consistent "story" that supports the NTGR calculated based on the overall information given.

3 **Vendor Surveys** A Vendor Survey is completed for all **Standard** and **Standard- Very Large** NTGR sites that utilized vendors, and for **Basic** NTGR sites that indicate a high level of vendor influence in the decision to implement the energy efficient measure. For those sites that indicate the vendor was very influential in decision making, the vendor survey results enter directly into the NTGR scoring. Since the retro-commissioning high impact measure (HIM) investigated for this evaluation does not rely on vendors for measure delivery, vendor interviews were not conducted for this study.

4 **Utility and Program Staff Interviews** For the Standard and Standard-Very Large NTGR analyses, interviews with utility staff and program staff are also conducted. These interviews are designed to gather information on the historical background of the customer's decision to install the efficient equipment, the role of the utility and program staff in this decision, and the name and contact information of vendors who were involved in the specification and installation of the equipment.

5 **Other information** For **Standard – Very Large Project** NTGR sites, secondary research of other pertinent data sources is performed. For example, this could include a review of standard and best practices through industry associations, industry experts, and information from secondary sources (such as the U.S. Department of Energy's Industrial Technologies Program, Best Practices website URL, http://www1.eere.energy.gov/industry/bestpractices/). In addition, the Standard- Very Large NTGR analysis calls for interviews with other employees at the participant's firm, sometimes in other states, and equipment vendor experts from other states where the rebated equipment is being installed (some without rebates), to provide further input on standard practice within each company.

Table 27 below shows the data sources used in each of the three levels of free ridership analysis. Although more than one level of analysis may share the same source, the amount of information that is utilized in the analysis may vary. For example, all three levels of analysis obtain core question data from the Decision Maker survey.

Table 27: Information Sources for Three Levels of NTGR Analysis

	Program File	Decision Maker Survey Core Question	Vendor Surveys	Decision Maker Survey Supplemental Questions	Utility & Program Staff Interviews	Other Research Findings
Basic NTGR	√	√	√[1]		√[2]	
Standard NTGR	√	√	√[1]	√	√	
Standard NTGR-Very Large Projects	√	√	√[3]	√	√	√

1 Only performed for sites that indicate a vendor influence score (N3d) greater than maximum of the other program element scores (N3b, N3c, N3g, N3h, N3l).

2 Only performed for sites that have a utility account representative

3 Only performed if significant vendor influence reported or if secondary research indicates the installed measure may be becoming standard practice.

Supplement B provides the full battery of Decision Maker interview questions along with notes, for both NTGR levels used in the retro-commissioning evaluation, regarding which questions are asked (denoted by an "X"), and the intended uses of the information in the NTGR analysis. A copy of the complete survey forms (with lead-in text and skip patterns) are contained in *Final Large Nonresidential NTGR Survey Instruments.XLS* that is available upon request.

5.4.5. NTGR Framework

The Self-Report-based Net-to-Gross analysis relies on responses to a series of survey questions that are designed to measure the influence of the program on the participant's decision to implement program-eligible energy efficiency measure(s). Based on these responses, a NTGR is derived based on responses to a set of "core" NTGR questions. The NTGR includes the effects of deferred free ridership (i.e., accelerated adoption).

5.4.5.1 NTGR Questions and Scoring Algorithm

A self-report NTGR is computed for all NTGR levels using the following approach. Adjustments may be made for **Standard – Very Large** NTGR sites, if the additional information that is collected is inconsistent with information provided through the Decision Maker survey.

The NTGR is calculated as an average of three scores. Each of these scores represents the highest response or the average of several responses given to one or more questions about the decision to install a program measure.

1 A **Timing and Selection** score that reflects the influence of the **most important** of various program and program-related elements in the customer's decision to select the specific program measure at this time.

2 A **Program Influence** score that captures the perceived importance of the program (whether rebate, recommendation, training, or other program intervention) relative to non-program factors in the decision to implement the specific measure that was eventually adopted or installed. This score is determined by asking respondents to assign importance values to both the program and non-program influences so that the two total 10. The program influence score is adjusted (i.e., divided by 2) if respondents say they had already made their decision to install the specific program qualifying measure or, in the case of retro-commissioning, to pursue the retro-commissioning project, before they learned about the program.

3 A **No-Program** score that captures the likelihood of various actions the customer might have taken at this time and in the future if the program had not been available (the counterfactual). This score also accounts for deferred free ridership by incorporating the likelihood that the customer would have installed program-qualifying measures at a later date if the program had not been available.

When there are multiple questions that feed into the scoring algorithm, as is the case for the **Timing and Selection** score, the maximum score is always used. The rationale for using the maximum value is to capture the most important element in the participant's decision making. Thus, this score is always based on the strongest influence indicated by the respondent. However, high scores that are inconsistent with other previous responses trigger consistency checks and can lead to follow-up questions to clarify and resolve the discrepancy.

The calculation of each of the above scores is discussed below. For each score, the associated questions are presented and the computation of each score is described. For a detailed explanation of the scoring algorithm, including examples, see Supplement C.

Timing and Selection Score

For the Decision Maker, the questions asked are:

I'm going to ask you to rate the importance of the program as well as other factors that might influence your decision to implement this Retro-commissioning project. Think of the degree of importance as being shown on a scale with equally spaced units from 0 to 10, where 0 means not at all important and 10 means very important, so that an importance rating of 8 shows twice as much influence as a rating of 4.

Now, using this 0 to 10 rating scale, where 0 means "Not at all important" and 10 means "Very important," please rate the importance of each of the following in your decision to retro-commission this facility at this time.

- Availability of the PROGRAM rebate to cover the cost of the retro-commissioning study
- Availability of the PROGRAM rebate to help cover the cost of implementing recommended measures
- Information provided through a recent feasibility study, energy audit or other types of technical assistance provided through PROGRAM
- Information from PROGRAM training course
- Information from other PROGRAM marketing materials
- Endorsement or recommendation by PROGRAM or UTILITY staff
- Endorsement or recommendation by the UTILITY Account Representative
- Corporate policy or guidelines
- Standard practice in the business or industry
- Payback on the investment
- Recommendation from a vendor/supplier

The Timing and Selection Score is calculated as:

The highest of the responses to the first six decision maker questions above, although these questions may be asked in a different order.

Program Influence Score

The questions asked are:

1. Did you first learn about the Program BEFORE or AFTER you first began to think about implementing this project?
2. Did you learn about PROGRAM BEFORE or AFTER you decided to implement the project?
2. Now I'd like to ask you a last question about the importance of the program to your decision as opposed to other factors that may have influenced your decision. Again using the 0 to 10 rating scale we used earlier, where 0 means "Not at all important" and 10 means "Very important," please rate the overall importance of PROGRAM versus the most important of the other factors we just discussed in your decision to implement this retro-commissioning specific project. This time I would like to ask you to have the two importance ratings—the program importance and the non-program importance—total 10.

The Program Influence score is calculated as:

The importance of the program, on the 0 to 10 scale, to question 2. This score is reduced by half if the respondent learned about the program after the decision had been made.

No-Program Score

The questions asked are:

1. Regarding this retro-commissioning project, if the PROGRAM had not been available, using a likelihood scale from 0 to 10, where 0 is "Not at all likely" and 10 is "Extremely likely" how likely is it that you would have retro-commissioned this facility?
2. IF 1>0. You indicated that there was an "***X***" in 10 likelihood that you would have retro-commissioned this facility if the PROGRAM had not been available. When do you think you would have done this? Please express your answer in months

a	_____ _____	within 6 months?	(Deferred NTG Value=0)
b	_____ _____	7 to 47 months later	(Deferred NTG Value=(months-6) ×0.024)
c	_____ _____	48 or more months later	(Deferred NTG Value =1)
d	_____ _____	Never	(Deferred NTG Value=1)

Note: The value 0.024 is 1 divided by 41 (41 is calculated as 47 – 6). This assumes that the deferred NTG value is a linear function beginning in month 7 through month 47, increasing 0.024 for each month of deferral.

The No-Program Score is calculated as:

10 minus (the likelihood of retro-commissioning the facility multiplied by one minus the *deferred net-to-gross value* associated with the timing of the project).

Partial Free Ridership

In cases where the decision maker is able to state how much they would have spent on retro-commissioning in the absence of the program, the ratio of that expenditure to the amount actually spent to retro-commission the facility through the program is multiplied by the no-program likelihood in calculating the no-program score.

The Core NTGR

The self reported core NTGR is simply the average of the Program Influence, Timing and Selection, and No-Program Scores, divided by 10. The exception to this is when the respondent says there is a 10 in 10 likelihood that the facility would have been retro-commissioned in the absence of the program and that it would have been done within 6 months; in this case the core NTGR is calculated as the average of the no-program score (which is zero in this case) and the relative program influence score.

5.4.5.2 Data Analysis and Integration

The calculation of the Core NTGR is fairly mechanical and is based on the answers to the closed-ended questions. However, the reliance of the Standard NTGR – Very Large on more information from so many different sources requires more of a case-study level of effort. The SRA Guidelines point out that a case study is one method of assessing both quantitative and qualitative data in estimating a NTGR. A case study is an organized presentation of all these data available about a particular customer site with respect to all relevant aspects of the decision to install the efficient equipment. In such cases where multiple interviews are conducted eliciting both quantitative and qualitative data and a variety of program documentation has been collected, one will need to integrate all of this information into an internally consistent and coherent story that supports a specific NTGR.

The following data sources should be investigated and reviewed as appropriate to supplement the information collected through the decision maker interviews.

- Account Representative Interview
- Utility Program Manager/Staff Interview
- Utility Technical Contractor Interview
- Third party Program Manager Interview
- Evaluation Engineer Interview
- Gross Impact Site Plan/Analysis Review
- Corporate Green/Environmental Policy Review (if mentioned as important)
- Corporate Standard Practice Review (if mentioned as important)
- Industry Standard Practice Review (if mentioned as important)
- Corporate payback review (if mentioned as important)
- Review relevant codes and standards, including regulatory requirements

- Review industry publications, websites, reports such as the Commercial Energy Use Survey, historical purchase data of specific measures etc.

As detailed in the Self-Report NTGR Guidelines, when complementing the quantitative analysis of free ridership with additional quantitative and qualitative data from multiple respondents and other sources, there are some basic concerns that one must keep in mind. Some of the other data – including interviews with third parties who were involved in the decision to install the energy efficient equipment – may reveal important influences on the customer's decision to install the qualifying program measure. When one chooses to incorporate other data, one should keep the following principles in mind: 1) the method chosen should be balanced. That is, the method should allow for the possibility that the other influence can either increase or decrease the NTGR calculated from the decision maker survey responses, 2) the rules for deciding which customers will be examined for potential other influences should be balanced. In the case of Standard –Very Large interviews, all customers are subject to such a review, so that the pool of customers selected for such examination will not be biased towards ones for whom the evaluator believes the external influence will have the effect of influencing the NTGR in only one direction, 3) the plan for capturing other influences should be based on a well-conceived causal framework. The onus is on the evaluator to build a compelling case using a variety of quantitative and/or qualitative data for estimating a customer's NTGR.

Establishing Rules for Data Integration

Before the analysis begins, the evaluation team should establish, to the extent feasible, rules for the integration of the quantitative and qualitative data. These rules should be as specific as possible and be strictly adhered to throughout the analysis. Such rules might include instructions regarding when the NTGR based on the quantitative data should be overridden based on qualitative data, how much qualitative data are needed to override the NTGR based on quantitative data, how to handle contradictory information provided by more than one person at a given site, how to handle situations when there is no decision-maker interview, when there is no appropriate decision-maker interview, or when there is critical missing data on the questionnaire, and how to incorporate qualitative information on deferred free ridership.

One must recognize that it is difficult to anticipate all the situations that one may encounter during the analysis. As a result, one may refine existing rules or even develop new ones during the initial phase of the analysis. One must also recognize that it is difficult to develop algorithms that effectively integrate the quantitative and qualitative data. It is therefore necessary to use judgment in deciding how much weight to give to the quantitative versus qualitative data and how to integrate the two. The methodology and estimates, however, must contain methods to support the validity of the integration methods through preponderance of evidence or other rules/procedures as discussed above.

For the **Standard-Very Large** cases in the large Nonresidential programs, the quantitative data used in the NTGR Calculator (which calculates the "core" NTGR), together with other information collected from the decision maker regarding the installation decision, form the initial basis for the NTG "story" for each site. Note that in most cases, supplemental data such as tracking data, program application files and results of interviews with program/IOU staff and vendors, will have been completed before the decision maker is contacted and will help guide the non-quantitative questioning in the interview. In practice, this means that most potential inconsistencies between decision maker responses and other sources of information should have been resolved before the interview is complete and data are entered into the NTGR Calculator. For example, if a company has an aggressive "green" policy widely promoted on its website that is not mentioned by the decision makers, the interviewer will ask the respondent to clarify the role of that policy in the decision. Conversely, if the decision maker attributes the decision to install the equipment to a new company wide initiative rather than the program, yet there is no evidence of such an initiative reported by program staff, vendors, or the company's website, the decision maker will be asked to explain the discrepancy so that his or her responses can be changed if needed.

In some cases, however, it may be necessary to modify or override one of the scores contributing to the overall NTGR or the NTGR itself. Before this is done all quantitative and qualitative data will be systematically (and independently) analyzed by two experienced researchers who are familiar with the program, the individual site and the social science theory that underlies the decision maker survey instrument. Each will determine whether the additional information justifies modifying the previously calculated NTGR score, and will present any recommended modifications and their rationale in a well-organized manner, along with specific references to the supporting data. Again, it is important to note that the other influences can have the effect of either increasing or decreasing the NTGR calculated from the decision maker survey responses, and one should be skeptical about a consistent pattern of "corrections" in one direction or another.

Sometimes, *all* the quantitative and qualitative data will clearly point in the same direction while, in others, the *preponderance* of the data will point in the same direction. Other cases will be more ambiguous. In all cases, in order to maximize reliability, it is essential that more than one person be involved in analyzing the data. Each person must analyze the data separately and then compare and discuss the results. Important insights can emerge from the different ways in which two analysts look at the same set of data. Ultimately, differences must be resolved and a case made for a particular NTGR. Careful training of analysts in the systematic use of rules is essential to insure inter-rater reliability[9].

Once the individual analysts have completed their review, they meet to discuss their respective findings and present to the other the rationale for their recommended changes to the Calculator-derived NTGR. Key points of these arguments will be written down in summary form (e.g., Analyst 1 reviewed recent AQMD ruling and concluded that customer would have had to install the same measure within 2 years, not 3, thereby reducing NP score from 7.8 to 5.5) and also presented in greater detail in a workpaper so that an independent reviewer can understand and judge the data and the logic underlying each NTGR estimate. Equally important, the CPUC will have all the essential data to enable them to replicate the results, and if necessary, to derive their own estimates.

The outcome of the reconciliation by two analysts determines the final NTGR for a specific project. Again, the reasoning behind the "negotiated" final value must be thoroughly documented in a workpaper, while a more concise summary description of the rationale can be included in the NTGR Calculator

[9] Inter-rater reliability is the extent to which two or more individuals (coders or raters) agree. Inter-rater reliability addresses the consistency of the implementation of a rating system.

workbook (e.g., Analyst 1 and Analyst 2 agreed that the NTGR score should have been higher than the calculated value of 0.45 because of extensive interaction between program technical staff and the customer, but they disagreed on whether this meant the NTGR should be 0.6 or 0.7. After discussion, they agreed on a NTGR of 0.65 as reflecting the extent of program influence on the decision).

In summary, it has been decided that supplemental data from non-core NTG questions collected through these surveys should be used in the following ways in the California Large Nonresidential evaluations:

- Qualitative and quantitative information from other sources (e.g., industry data, vendor estimates of sales in no-program areas, and other data as described above) may be used to alter core inputs only if contradictions are found with the core survey responses. Since judgments will have to be made in deciding which information is more compelling when there are contradictions, supplemental data are reviewed independently by two senior analysts, who then summarize their findings and recommendations and together reach a final NTGR value.
- Responses will also be used to construct a NTGR "story" around the project; that is they will help to provide the context and rationale for the project. This is particularly valuable in helping to provide guidance to program design for future years. It may be, for example, that responses to the core questions yield a high NTGR for a project, but additional information sources strongly suggest that the program qualifying technology has since become standard practice for the firm or industry, so that free ridership rates in future years are likely to be higher if program rules are not changed.
- Findings from other non-core NTGR questions (e.g., Payback Battery, Corporate Policy Battery) are also be used to **cross-check the consistency** of responses to core NTGR questions. When an inconsistency is found, it is presented to the Decision Maker respondent who is then be asked to explain and resolve it if they can. If they are not able to do so, their responses to the core NTGR question with the inconsistency may be overridden by the findings from these supplemental probes. These situations are handled on a case-by-case basis; however consistency checks may be programmed into the CATI survey instrument used for the Basic and Standard cases.

5.4.5.3 Accounting for Partial Free Ridership

Partial free ridership can occur when, in the absence of the program, the participant would have installed something more efficient than the program-assumed baseline efficiency but not as efficient as the item actually installed as a result of the program. In the context of the retro-commissioning HIM, partial free ridership occurs if the participant would have conducted a smaller scale version of retro-commissioning or a building tune-up for the same facility.

Data Collection Procedures Information is gathered on partial free ridership using the following questions asked as part of the decision maker NTGR survey.

1 If you had retro-commissioned this facility, how much would you have spent if the program had not been available?

The amount that would have been spent is expressed or calculated as a proportion of the cost of the retro-commissioning effort.

Data Analysis and Integration Procedures In cases where partial free ridership is found and it is determined that the adjustment should be made to the net-to-gross ratio, the following procedure is used. Where the decision maker is able to state how much they would have spent on retro-commissioning in the absence of the program, the ratio of that expenditure to the amount actually spent to retro-commission the facility is multiplied by the no-program likelihood in calculating the no-program score.

5.4.6. NTGR Interview Process

The NTGR surveys are conducted via telephone interviews. Highly-trained professionals with experience levels that are commensurate with the interview requirements should perform these interviews. Basic- and Standard-level interviews should be conducted by senior interviewers, who are highly experienced conducting telephone interviews of this type. Standard - Very Large interviews should be completed by professional consulting staff due to the complex nature of these projects and related decision making processes. These often involve interviews of several entities involved in the project including the primary decision maker, vendor representatives, utility account executives, program staff and other decision influencers, as well as a review of market data to help establish an appropriate baseline.

5.4.7. Compliance with Self-Report Guidelines

The proposed NTGR framework fully complies with all of the CPUC/ED and the MECT's Guidelines for Estimating Net-to-Gross Ratios Using the Self-Report Approach, as demonstrated in Supplement D.

5.4.8. Supplemental Information

5.4.8.1 Supplement A: References

Blalock, H. (1970). Estimating measurement error using multiple indicators and several points in time," *American Sociological Review*, 35, pp. 101–111.

Bogdan, Robert and Steven J. Taylor. (1975). *Introduction to qualitative research methods*. New York: John Wiley & Sons.

Britan, G. M. (1978). Experimental and contextual models of program evaluation. *Evaluation and Program Planning,* 1: 229–234.

Cochran, William G. (1977). *Sampling techniques*. New York: John Wiley & Sons.

Crocker, L. and J. Algina. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.

Cronbach L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

DeVellis, R.F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage Publications, Inc.

Duncan, O.D. (1984). Notes on social measurement: Historical and critical. New York: Russell Sage.

Guba, E. G. (1978). Toward a methodology of naturalistic inquiry in educational evaluation. *CSE Monographic Series in Evaluation No. 8*. Los Angeles: Center for the Study of Evaluation.

Hall, Nick, Johna Roth, Carmen Best, Sharyn Barata, Pete Jacobs, Ken Keating, Ph.D., Steve Kromer, Lori Megdal, Ph.D., Jane Peters, Ph.D., Richard Ridge, Ph.D.,

Francis Trottier, and Ed Vine, Ph.D. (2007). California Energy Efficiency Evaluation: Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals. Prepared for the California Public Utilities Commission.

Lyberg, Lars, Paul Biemer, Martin Collins, Edith De Leeuw, Cathryn Dippo, Norbert Schwarz, and Dennis Trewin. (1997). *Survey measurement and process quality*. New York, NY: John Wiley & Sons.

Madow, William G., Harold Nisselson, Ingram Olkin. (1983). *Incomplete data in sample surveys*. New York: Academic Press.

Maxwell, Joseph A. (2004). Using Qualitative Methods for Causal Explanations. *Field Methods*, *Vol. 16, No. 3*, 243-264.

Megdal, Lori, Yogesh Patil, Cherie Gregoire, Jennifer Meissner, and Kathryn Parlin (2009). Feasting at the Ultimate Enhanced Free ridership Salad Bar. *Proceedings of the International Energy Program Evaluation Conference*.

Mohr, Lawrence B. (1995). *Impact analysis for program evaluation*. Thousand Oaks, CA: Sage Publications, Inc.

Netemeyer, Richard G., William O. Bearden, and Subhash Sharma. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: SAGE Publications.

Patton, Michael Quinn. (1987). *How to use qualitative methods in evaluation*. Newbury Park, California: SAGE Publications.

Ridge, Richard, Philippus Willems, and Jennifer Fagan. (2009). Self-Report Methods for Estimating Net-to-Gross Ratios in California: Honest! *Proceedings from the 19th National Energy Services Conference*.

Ridge, Richard, Philippus Willems, Jennifer Fagan and Katherine Randazzo. (2009). The Origins of the Misunderstood and Occasionally Maligned Self-Report Approach to Estimating the Net-To-Gross Ratio. *Proceedings of the International Energy Program Evaluation Conference*.

Rogers, Patricia J., Timothy A. Hacsi, Anthony Petrosino, and Tracy A. Huebner (Eds.) (2000). *Program theory in evaluation: Challenges and opportunities*. San Francisco, CA: Jossey-Bass Publishers.

Rossi, Peter and Howard E. Freeman. (1989). *Evaluation: A systematic approach.* Newbury Park, California: SAGE Publications.

Sayer, Andrew. (1992). *Method in social science: A Realist Approach*. New York: Routledge.

Sax, Gilbert. (1974). *Principles of educational measurement and evaluation*. Belomont, CA: Wadsworth Publishing Company, Inc.

Schumacker, Randall E. and Richard G. Lomax. (1996). *A beginner's guide to structural equation modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Scriven, Michael. (1976). Maximizing the power of causal explanations: The modus operandi method. In G.V. Glass (Ed.), *Evaluation Studies Review Annual, Vol. 1*, pp.101-118). Bevery Hills, CA: Sage Publications.

Shadish, Jr., William R. and Thomas D. Cook, and Laura C. Leviton. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage Publications, Inc.

Stone, Arthur A., Jaylan S. Turkkan, Christine A. Bachrach, Jared B. Jobe, Howard S. Kurtzman, and Virginia S. Cain. (2000). *The science of the self-report: Implications for research and practice*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Tashakkori, Abbas and Charles Teddlie. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE Publications.

TecMarket Works, Megdal & Associates, Architectural Energy Corporation, RLW Analytics, Resource Insight, B & B Resources, Ken Keating and Associates, Ed Vine and Associates, American Council for an Energy Efficient Economy, Ralph Prahl and Associates, and Innovologie. (2004). *The California evaluation framework.* Prepared for the California Public Utilities Commission and the Project Advisory Group.

Velleman, P. F., and Wilkinson, L. (1993), Nominal, ordinal, interval and ratio typologies are misleading. *American Statistician, 47(1)*, 65-72.

Weiss, Carol H. (1998). *Evaluation*. Upper Saddle River, New Jersey: Prentice Hall.

Weiss, R. S. and M.Rein. (1972). The Evaluation of broad-aim programs: Difficulties in experimental design and an alternative. In C. H. Weiss (ed.) *Evaluating action programs: Readings in social action and education*. Boston: Allyn and Bacon.

Wholey, Joseph S., Harry P. Hatry and Kathryn E. Newcomer. (1994). *Handbook of practical program evaluation*. San Francisco, CA: Jossey-Bass, Inc.

Yin, Robert K. (1994). *Case study research: Design and methods*. Newbury Park, California: SAGE Publications.

5.4.8.2 Supplement B: Net-to-Gross Questions and Uses of Data by Level of NTGR Analysis

RCx Decision Maker Survey

	Question Text	Basic	Standard – Very Large
	Introduction Hello, my name is ______ calling from COMPANY on behalf of the California Public Utilities Commission. I am calling about the retro-commissioning of FACILITY NAME through the PROGRAM NAME retro-commissioning program. Are you the person who was most involved with the decision to retro-commission this facility through PROGRAM NAME? [IF YES, CONTINUE]. We are interviewing firms that participated in the PROGRAM NAME in 2006 through 2008 to discuss the factors that may have influenced your decision to participate in the program. The interview		

	Question Text	Basic	Standard – Very Large
	will take about 20 minutes.		
	WARM-UP QUESTIONS		
A1	First, according to our records, you participated in PROGRAM NAME on (approximate date). [READ: Program Description. PROGRAM NAME promotes retro-commissioning of commercial/institutional facilities. The program offers incentives to cover the cost of retro-commissioning/MBCx studies as well as incentives to help cover the cost of implementing recommended actions/measures identified by the study.] Is that correct? Yes, No, DK, Refused	X	X
A2	Next, I'd like to confirm the following information regarding the measures you implemented through the program: (READ: PROJECT DETAILS INCLUDING SERVICES RECEIVED, MEASURES INSTALLED, KEY DATES, ETC. AS REQUIRED) Does that sound right? Yes, No, DK, Refused	X	X
A3	In your own words, please tell me why you decided to retro-commission this FACILITY? Were there any other reasons? a. Record VERBATIM b. DK/Refused	X	X
	NET-TO-GROSS BATTERY		
N1	When did you first learn about PROGRAM? Was it BEFORE or AFTER you first began to think about retro-commissioning FACILITY? a. Before (Skip to N3) b. After c. DK/Refused	X	X
N2	Did you learn about PROGRAM BEFORE or AFTER you decided to retro-commission FACILITY? a. Before b. After c. DK/Refused	X	X
	READ: Program Description: *As I mentioned earlier, [PROGRAM] promotes retro-commissioning/monitoring-based commissioning of commercial buildings and facilities. I'm going to ask you to rate the importance of the program as well as other factors that might have influenced your decision to retro-commission this facility. Think of the degree of importance as being shown on a scale with equally spaced units from 0 to 10, where 0 means not at all important and 10 means very important, so that an importance rating of 8 shows twice as much influence as a rating of 4.*		
N3	Now, using this 0 to 10 rating scale, where 0 means "Not at all important" and 10 means "Very important," please rate the importance of each of the following in your decision to retro-commission this facility when you did. [ROTATE PRESENTATION OF ITEMS. FOLLOW UP WITH "And is there anything else that I may have missed?" Record as p. Other (SPECIFY)]		
a	The age or condition of the facility	X	X
b	Availability of incentives to cover the cost of the study	X	X
c	Availability of incentives to cover the cost of implementing recommendations	X	X
d	Information provided through study, audit or other technical assistance provided through the Program or UTILITY	X	X

	Question Text	Basic	Standard – Very Large
e	The age or condition of the facility	X	X
f	Previous experience with Retro-Commissioning	X	X
g	Previous experience with the Partnership Program	X	X
h	Information from a UTILITY or program training course	X	X
i	Information from UTILITY or program marketing materials	X	X
j	A recommendation from an auditor or consulting engineer	X	X
k	Standard practice in your industry	X	TRIGGER
l	Recommendation from Partnership Program staff	X	X
m	Endorsement or recommendation by UTILITY Account Rep	X	X
n	Corporate/organizational policy or guidelines	X	TRIGGER
o	Payback on the investment	X	TRIGGER
p	Other (SPECIFY)	X	X
N4	Now I'd like to ask you a last question about the importance of the program to your decision. Again using the 0 to 10 rating scale we used earlier, where 0 means "Not at all important" and 10 means "Very important," please rate the overall importance of the Program versus the non-program factors in your decision to retro-commission this facility. This time I would like to ask you to have the two importance ratings—the program importance and the non-program importance—total 10		
	a. ________rating of the importance of PROGRAM NAME	X	X
	b. ________rating of the importance of Other Factors	X	X
	Now I would like you to think about the action you would have taken with regard to the retro-commissioning of this facility if PROGRAM had not been available.		
N5	Regarding the retro-commissioning of this facility, if the PROGRAM had not been available, what is the likelihood that you would have retro-commissioned this facility, using a 0 to 10 likelihood scale, where 0 is not at all likely and 10 is extremely likely?	X	X
	IF N5>0. You indicated in your previous response that there was a X in 10 likelihood that you would have retro-commissioned this facility if the PROGRAM had not been available.	X	X
N6	When do you think you would have undertaken this retro-commissioning project? (Please answer in months)________	X	X
	a. _____ ____ within 6 months?		
	b. _____ ____ 6 – 47 months later		
	c. _____ ____ 48 or more months later		
	g. _____ ____ Never		
N7	IF N6 <4 YEARS LATER: How much do you think you would have spent to retro-commission the facility at that time? (record in any of the formats below)	X	X
	In absolute terms (enter value) _________		
	Per square foot (enter value) _________		
	As % of program funded cost (enter estimate or calculate %)		
N8	Did you consider any alternatives to the Retro-Commissioning of &FACILITY that you would have implemented in the same time frame if the program had not been available?	X	X
N9	Please describe the alternative which you most likely would have pursued if the [PROGRAM NAME] Program had not been available. Please be as specific as	X	X

	Question Text	Basic	Standard – Very Large
	possible, including about how much you would have spent on an alternative tune-up/building study/retro-commissioning effort.		
N10	In the absence of the [PROGRAM NAME] Program, is it more likely that you would have done nothing or is it more likely that you would have taken the action that you just described?	X	X
	Additional Decision Maker Questions		
	PAYBACK BATTERY (If payback importance >5)		
P1	What financial calculations does your company make before proceeding with a project like this one?		X
P2	What is the cut-off point your company uses before deciding to proceed with the investment?		X
P3	Based on what you knew before you undertook this project, what was the result of the calculation for this project: a) with the rebate? b) without the rebate?		X
	INVESTIGATE INCONSISTENT RESPONSE		
	CORPORATE POLICY BATTERY (If corporate policy importance >5)		
CP1	Does your organization have a corporate policy that requires or encourages commissioning or retro-commissioning of facilities?		X
CP2	What specific corporate policy influenced your decision to retro-commission this facility?		X
CP3	Had that policy caused you to retro-commission this facility before participating in this program?		X
CP4	Had that policy caused you to retro-commission other facilities before participating in this program? When and where?		X
CP5	Did you receive an incentive for a previous retro-commissioning projects? If so, please describe.		X
	STANDARD PRACTICE BATTERY (If standard practice importance >5)		
SP1	How long has retro-commissioning been standard practice in your industry?		X
SP2	Does your company ever deviate from the standard practice? If yes, under what conditions?		X
SP3	How did this standard practice influence your decision to retro-commission FACILITY?		X
	OTHER INFLUENCES		
O1	Please state, in your own words, any other factors that influenced your decision to go ahead on this retro-commissioning project.	X	X

5.4.8.3 Supplement C: NTGR Scoring Algorithm and Example

The calculation of the self-report-based core NTGR is described below. The NTGR is calculated as an average of three scores representing responses to one or more questions about the decision to install a program measure.

1 A **Timing and Selection** score that captures the influence of the most important of various program and program-elated elements in influencing the customer to select the specific program measure at this time. Program influence through vendor recommendations is also captured in this score.

2 An overall **Program Influence** score that captures the perceived importance of the program (whether rebate, recommendation, or other information) in the decision to implement the specific measure that that was eventually adopted or installed. The overall program influence score is reduced by half if the respondent says they learned about the program only after they decided to install the program qualifying measure.

3 A **No-Program** score that captures the likelihood of various actions the customer might have taken at this time and in the future if the program had not been available. This score accounts for deferred free ridership by capturing the likelihood that the customer would have installed program qualifying measures at a later date if the program had not been available. It also accounts for partial free ridership by scaling down the score by the ratio of what would have been spent on retro-commissioning compared to what was actually spent through the program.

Calculation of each of the above scores is discussed below. For each score, the questions contributing to the calculation are presented, the calculation is described, and an example is provided.

Timing and Selection Score

For the decision maker, the questions asked are:

Using a 0 to 10 rating scale, where 0 means not at all important and 10 means very important, please rate the importance of each of the following in your decision to retro-commission this facility at this time:

- Availability of the PROGRAM rebate to cover the cost of the retro-commissioning study
- Availability of the PROGRAM rebate to help cover the cost of implementing recommended measures
- Information provided through a recent feasibility study, energy audit or other types of technical assistance provided through the PROGRAM
- Information from PROGRAM training course
- Information from other PROGRAM marketing materials
- Endorsement or recommendation by PROGRAM or UTILITY staff
- Endorsement or recommendation by the UTILITY Account Representative
- Corporate policy or guidelines
- Standard practice in the business or industry
- Payback on the investment
- Recommendation from a vendor/supplier

The Timing and Selection Score is calculated as:

The highest of the responses to the first six decision maker questions.

Example:

The decision maker provides responses of 6 for the importance of the rebate, 5 for an audit or feasibility study, 3 for training and 2 for program marketing material.

The timing and selection score is the maximum of the decision maker responses (6, 5, 3, and 2), so the timing and selection score is 6.

Program Influence Score

The questions asked are:

1. Did you first learn about the Program BEFORE or AFTER you first began to think about implementing this project?
2. Did you learn about PROGRAM BEFORE or AFTER you decided to implement the project?
3. Now I'd like to ask you a last question about the importance of the program to your decision as opposed to other factors that may have influenced your decision. Again using the 0 to 10 rating scale we used earlier, where 0 means "Not at all important" and 10 means "Very important," please rate the overall importance of PROGRAM versus the most important of the other factors we just discussed in your decision to implement the retro-commissioning project. This time I would like to ask you to have the two importance ratings—the program importance and the non-program importance—total 10.

The program influence score is calculated as:

The program importance response, on the 0 to 10 scale, to question 3. This score is reduced by half if the respondent became aware of the program only after having decided to retro-commission the facility.

Example:

The decision maker says they became aware of the program before deciding to implement the retro-commissioning project, and provides a response of 7 to question 3 which becomes the program influence score.

No-Program Score

The questions asked are:

1. Regarding the installation of this equipment, if the PROGRAM had not been available, using a likelihood scale from 0 to 10, where 0 is "Not at all likely" and 10 is "Extremely likely" how likely is it that you would have retro-commissioned this facility?
2. IF 1>0. You indicated that there was an "***X***" in 10 likelihood that you would have retro-commissioned this facility if the PROGRAM had not been available. When do you think you would have done this? Please express your answer in months

a	_____ ____	Within 6 months?	(Deferred NTG Value=0)
b	_____ ____	7 to 47 months later	(Deferred NTG Value=(months-6) ×0.024)
c	_____ ____	48 or more months later	(Deferred NTG Value =1)

d _____ ____ Never (Deferred NTG Value=1)

Note: The value 0.024 is 1 divided by 41 (41 is calculated as 47 – 6). This assumes that the deferred NTG value is a linear function beginning in month 7 through month 47, increasing 0.024 for each month of deferral.

The No-Program Score is calculated as:

10 minus (the likelihood of retro-commissioning the facility multiplied by one minus the *deferred net-to-gross value* associated with the timing of the project).

Partial Free Ridership

In cases where the decision maker is able to state how much they would have spent on retro-commissioning in the absence of the program, the ratio of that expenditure to the amount actually spent to retro-commission the facility is multiplied by the no-program likelihood in calculating the no-program score.

Example

The respondent says there is an 8 in 10 likelihood that they would have retro-commissioned the facility. In response to question 5, the decision maker says they would done this 18 months later, which has a NTGR value of (18–-6) × 0.024, or 0.29 associated with it. The also say they would have spent half as much on the retro-commissioning project in the absence of the program, which means the no-program likelihood is reduced by 50% to 4.

The No-Program likelihood score is 10 minus (4× (1–0.29)), which is 10 minus 4×0.71 or 7.16.

Core NTG Ratio

The self reported NTGR ratio is simply the average of the Program Influence, Timing and Selection, and No-Program Scores, divided by 10. The exception to this is when the respondent says there is a 10 in 10 likelihood that the facility would have been retro-commissioned in the absence of the program and that it would have been done within 6 months; in this case the core NTGR is calculated as the average of the no-program score (which is zero in this case) and the relative program influence score.

Example

The NTGR is the average of 6, 8 and 7.2, or 7.1 divided by 10 = 0.71.

5.4.8.4 Supplement D: Demonstration of Compliance with the CPUC/ED and MEC’s Guidelines for Estimating Net-to-Gross Ratios Using the Self-Report Approach

1 **Timing of the interview**

To minimize problems of recall, every effort should be made to conduct the NTGR interview as close to project completion as possible.

2 Identifying the correct respondent

The survey form includes some initial probing on the respondent's role in the completed project, to confirm their involvement in the decision to implement the energy efficiency measures. In addition, both the utility or third party representative and any trade allies involved should be asked to confirm they are the correct contact. If multiple decision makers are identified, each one should be interviewed and the results pooled.

In the unfortunate circumstance where the key decision maker has left the company, that sample point should be discarded and replaced with a respondent from within the same stratum in the backup sample.

3 Set-up questions

The survey includes a series of warm-up questions that serve to remind the respondent about the circumstances and motivations surrounding the project, the project scope (including installed measures), incentives paid, and the project schedule. This information also helps to build the "story" to substantiate the NTGR responses given.

4 Use of multiple questions

The NTGR scoring algorithm relies on responses from several questions to determine the final NTGR score. The scoring is a function of:

- The timing of their program awareness relative to their decision to implement the installed measure
- The importance of program versus non-program influences in their decision making
- The importance of specific influences in the participant's general decision to implement the measure and that led them to implement the specific measure at the time they did rather than an alternative
- Without the program, the probability of alternative actions to implementing the selected measure

5 Validity and reliability

The proposed NTGR method is designed to produce valid and reliable NTGR results, based on the use of:

- *"Tried and true" question wording*. Many of the core questions used in NTGR scoring are substantially the same as those that have been used extensively in previous large C&I program evaluations, such as the last several rounds of evaluation for the California Standard Performance Contracting Program. While the question construct is somewhat different from in the past, the wording used is essentially the same as has been used previously.
- *Information from supplemental questions and multiple data sources to corroborate and triangulate on the NTGR "story"*. In addition to self-reported information, the NTGR findings for Standard – Very Large NTGR sites include responses to a number of supplemental questions surrounding the project (e.g., corporate policy, standard industry practice and payback), and the results from an interview with the vendor(s) involved in the project. These findings will be used

to converge on a plausible estimate of the NTGR and to help tell the "story" behind the project and its context.

- *Multiple reviewers. Standard - Very Large customer projects are reviewed by two experienced analysts*. The two reviewers seek to develop a NTGR consensus on the project, and resolve any differences of opinion.
- *Identification and explicit consideration of alternate hypotheses*. Respondents are asked about the relative influence of a variety of program and non-program factors.

During the pre-test of the NTGR survey instrument, reliability tests should be conducted and any problem areas detected should be corrected.

6 Consistency checks

Questions within the NTGR battery that are more likely to produce inconsistent responses have been flagged. These include questions regarding the program's reported importance in the decision to implement the specified measure, alternative actions in the program's absence, questions reporting the motivations for doing the project, as well as any closely related supplemental questions. The CATI software should be specifically programmed to flag any inconsistencies, and include follow-up prompts when they are found. Interviewers should be instructed how to administer these follow-up questions to resolve these inconsistencies. Interviewers should make every effort to resolve any inconsistencies before concluding the interview.

7 Making the Questions Measure-Specific

In general, most projects involve one type or class of measure. However, there are a few instances where the project consists of multiple types of measures, but usually, one measure predominates. In such cases, the interview should be conducted around the dominant measure with the greatest share of savings. If there are projects with multiple types of measures and no one measure class predominates, the NTGR sequence should be repeated for each significant measure class (e.g., once for lighting and once for process measures). At the beginning of each interview, there is a prompt with a description of the measure class that the questions pertain to so that it is clear in the minds of the respondent which measures they are being asked about.

8 Partial free ridership

Questions on the amount that would have been spent on retro-commissioning in the absence of the program are designed to collect the information necessary to adjust for any partial free ridership.

9 Deferred free ridership

Question N6 addresses deferred free ridership, and provides specific adjustment factors for each response category. The NTGR algorithm (See section 5.4.5 and Supplement C) text fully explains the specifics of this adjustment.

10 Scoring algorithms

The methodology includes a specific algorithm for developing a NTGR based on responses received. The results of the 0 to 10 scoring are used to develop specific values for each question used to score the NTGR. A description of the scoring algorithm is provided in section 5.4.5 and in Supplement C.

11 Handling unit and item non-response

Every effort should be made to discourage non-responses (i.e., refusals and terminates). For example, in California, the interviewer points out that the energy efficiency program requires the project to be evaluated as a condition of participation. Absent such a requirement, interviewers should stress such things as the importance of evaluation in improving program design and delivery. In some cases, incentives can be offered to respondents. In the event various strategies are not successful, the non-responding customer should be replaced by another customer within the same stratum. While efforts to minimize item non-response ("don't knows" and "refusals") should be made using a variety of available techniques, one should recognize that forcing a response can distort the respondent's answer and introduce bias.

12 Weighting the NTGR

The mean NTGR for a given measure, end use or program should be weighted to take into account the size of the ex-post gross impacts.

13 Ruling out rival hypotheses

The core NTGR questions, particularly question 4 of the Decision Maker survey, have been carefully constructed to try to rule out rival hypotheses. The method asks respondents to jointly consider and rate the importance of the many likely events or factors that may have influenced their energy efficiency decision making, rather than focusing narrowly on only their rating of the program's importance. This question structure more accurately reflects the complex nature of the real-world decision making and should help to ensure that all non-program influences are reflected in the NTGR assessment in addition to program influences.

14 Precision of the NTGR

The calculation of the achieved relative precision of the NTGRs (for program-related measures and practices and non-program measures and practices) is expected to be straightforward. However, the inclusion of more complicated situations involving multiple participant and vendor interviews as well as the inclusion of additional qualitative information means that the NTGR standard errors may underestimate the uncertainty surrounding the NTGR estimate.

15 Pre-testing the questionnaire

The NTGR survey should be carefully and extensively pre-tested and adjusted in response to pre-test findings before it is fielded.

16 Incorporation of additional qualitative and quantitative data in estimating the NTGR (data collection, rules for data integration, analysis)

Specific rules have been established for data integration and these are described in section 5.4.3.

17 Qualified interviewers

The NTGR surveys should be fielded by highly experienced interviewers. High-level professional interviewers should be used for the largest and most complex projects, while less experienced professional interviewers should be used for smaller, simpler projects. A CATI approach may be used for all but the very largest and most complex projects.

5.4.8.5 Supplement E: Summary of Revisions to Approach and Survey Instruments for Retro-Commissioning HIM

As noted previously and summarized below, several revisions have been made to the implementation of the net-to-gross ratio estimation methods for the Retro-Commissioning HIM. The revised Retro-Commissioning NTG survey instruments (both Basic and Standard - VL) address both the decision to retro-commission the facility and the decision to select specific measures for implementation. A measure specific NTG battery may be asked of retro-commissioning program participants for measures with the following characteristics:

- Measures identified as Retrofit measures
- Measures with an implementation cost greater than $10,000
- Measures with a payback greater than 1 year (for some RCx programs, implementation of measures with payback of one year or less is a condition of participation)

In addition to asking decision makers about the retro-commissioning decision, the NTG survey asks them about the decision to implement the two largest impact measures that meet the above criteria. (We believe that two measures in addition to the RCx decision is the maximum possible without overwhelming the respondent.)

For each measure, a single screening question will be used to determine whether the measure-specific battery is needed. Respondents will be asked: *Were you considering this measure before it was recommended by the RCx study?*

- If the measure was not under consideration before it was recommended, it is considered purely program driven and assigned the same NTGR as the overall RCx effort, and the interviewer goes on to the next measure.
- If the measure was considered previously, a battery of questions similar to those asked for the retro-commissioning decision is asked and a separate NTGR score is calculated, and the overall project NTGR is calculated based upon the proportion of savings associated with each measure-specific NTGR.

Other revisions to the retro-commissioning survey instruments include the following:

- When decision makers say there was a non-zero likelihood that they would have commissioned their facility even with the RCx program, they are asked to estimate the amount they would have spent on that activity, either in absolute terms, in dollars per square foot, or as a percentage of the program-subsidized R-Cx cost. The reasoning is that firms may say that they pursue a R-Cx or R-Cx-like policy, but unless the amount they would have spent is comparable to the cost of the RCx investigation conducted through the program, the resulting energy savings would not be of the same magnitude.

NTGR Scoring

Measure and site -level NTGRs will be calculated using the NTGR algorithm described in the general write-up and modified as follows.

- For the overall RCx NTGR, if the amount that would have been spent on the R-Cx effort is less than the cost of the study provided by the program, the no-program score will be adjusted. The customer's self-reported likelihood that they would have retro-commissioned the facility in the absence of the program is decreased in proportion to the no-program RCx budget vs. cost of the program-provided RCx study. For example, if the respondent says there is a 7 in 10 likelihood that they would have retro-commissioned the facility without the program and that the "no program" RCx budget would have been 60% of the program RCx cost, we would multiply the no-program likelihood of 7 by 60% 60% × 7 = 4.2) before applying it to the deferred free ridership value associated with the time frame when RCx would have been done.
- The overall RCx NTGR is used for all measures that are not specifically asked about, including both measures that cost less than $10,000 and those that are not in the top two measures rated as described above (i.e., retrofit measures and measures with implementation costs greater than $10,000).
- For individual measures that were not considered before they were recommended by the RCx study, the NTGR is set equal to the overall RCx NTGR.
- For individual measures that were under consideration before they were recommended by the RCx study, the NTGR is calculated in accordance with the NTGR algorithm.
- If none of the recommended measures had been under consideration before they were identified by the RCx study, the overall NTGR is used for all project impacts.
- The project -level NTGR is the ex ante impact-weighted average of the individual measure NTGRs (including the RCx NTGR applied to "all other" measures.). If there are no retrofit measures and no measures with implementation cost greater than $10,000 the overall RCx NTGR is used for all measures and for the project.

5.5. Detailed Findings for the Gross Sample Projects

Table 30 through Table 37 show case-level results for each of the 50 projects in the gross sample. There are four sets of two tables, with one set for each IOU (PG&E, SCE, SCG, and SDG&E, in that order). The first table in each set shows general information about implemented measures, along with ex ante and evaluated ex post gross and net savings results. The second table in each set summarizes and classifies key reasons for differences between ex ante and ex post savings.

Below are explanations of the coding schemes used in the tables to impart information about each project.

System and Measure Classification

Originally, we developed a detailed classification scheme for all expected RCx and Major Commercial measures (see table below). This scheme standardized the designation of what building system and general strategy each measure addressed. For RCx projects, the vast majority of evaluated measures fell into the "Improve…" categories, indicating that they consisted of adjustments and minor improvements to existing equipment, rather than wholesale replacements of components. After the evaluation data were finalized, we examined which types of measures were most common and aggregated system and measure classes to group these measures more broadly for analysis purposes. Values in these classes are shown in Table 28 and Table 29 below:

Table 28: System Classes

System Class	System(s)
Central plant	Boilers / heating hot water / steam Chillers / chilled water Cooling towers / condensing water
HVAC – general	HVAC - general HVAC – packaged
HVAC - air distribution	HVAC - air distribution
Other/ unclassified	Combined RCx measures (could not disaggregate measures) Compressed air Domestic water Lighting – general Lighting - interior Refrigeration

Table 29: Measure Classes

Measure Class	Measure Name(s)
Improve scheduling	Improve scheduling (general) Improve building warmup/cooldown
Improve control strategies	Improve control strategies (general) Improve sequence of operations Improve setpoints Improve temperature/pressure reset schedule
Improve outside air use	Economizer operation Improve outside air use
Install/repair variable speed drive (VFD)	Install / replace variable speed drive (cooling tower fan, HVAC air handler, HVAC pump, non-HVAC pump, other)

Measure Class	Measure Name(s)
Other/ unclassified	Combined RCx measures (could not disaggregate measures)
	Improve damper performance
	Improve maintenance practices - general
	Improve sensor performance
	Improve system performance monitoring
	Improve valve performance
	Install / replace controls - central system
	Install / replace lamps / ballasts - linear fluorescent (non-high bay)
	Install / replace miscellaneous efficiency improvement
	Other

The flags for system and measure classes below indicate the presence in the project of one or more measures that fall within the class. For example, a "1" in the "Central Plant" system class indicates that of the various measures at the project, at least one directly addressed some aspect of the central plant systems (boilers, chillers, cooling towers, etc.). Note that the "Other/Unclassified" class for both systems and measures most often meant that the various measures associated with the project could not be disaggregated well enough for us to assign a percentage of project savings.

Reasons for Differences

The tables containing this information list up to four major reasons why the evaluated savings differed from the claimed savings for each project. In addition, each reason has a three-part code that answers the following questions about the reason:

- Did this reason tend to INCREASE (I) or REDUCE (R) evaluated compared to claimed savings?
- Did this reason occur primarily because of CUSTOMER (C) or PROGRAM IMPLEMENTER (P) actions?
- Was this reason the MAIN (M) explanation for the differences?

As an example, "R C M Fans operating during unoccupied periods" can be interpreted to mean that the main reason why evaluated savings for this project fell short was that the customer likely took some action to change or override the fan schedule changes that the RCx program had implemented." Do keep in mind that the listed reasons represent the best judgment of the evaluation team after a brief examination of available evidence collected during the study, and are not meant to be authoritative or exhaustive

Table 30: PG&E – Gross and Net Savings Results for Sampled Projects

				System Class				Measure Class					Ex Ante Gross Savings			Ex Post 1st Year Gross Savings			Gross Realization Rate			NTGR	Ex Post 1st Year Net Savings		
SBW ID	IOU Program	IOU ID	Case Weight	Central plant	HVAC - general	HVAC - air distribution	Other/ unclassified	VFD	Improve scheduling	Improve control strategies	Improve outside air use	Other/ unclassified	kW	kWh	Therms	kW	kWh	Therms	kW	kWh	Therms		kW	kWh	Therms
P00090	PGE2094	20	1.1	1		1		1		1	1	1	0.0	1,952,906	N/A	0.0	1,188,854	1,397	N/A	0.61	N/A	0.87	0	1,034,303	1,215
P00096	PGE2072	TAA0001648	1.1			1	1			1	1	1	586.0	603,684	28,548	88.2	471,450	-4,784	0.15	0.78	-0.2	0.78	69	367,731	-3,732
P00097	PGE2072	TAA0000947	1.1			1		1		1	1		86.7	1,401,794	20,157	-61.5	814,150	32,850	-0.71	0.58	1.6	0.78	-48	635,037	25,623
P00098	PGE2072	TAA0000947	6.0		1	1				1	1	1	14.0	209,013	35,769	10.6	28,407	2,432	0.76	0.14	0.1	0.78	8	22,157	1,897
P00107	PGE2091	3001-01	11.0			1	1		1	1	1		14.1	43,932	1,352	0.0	16,374	4,224	0.00	0.37	3.1	0.90	0	14,736	3,802
P00125	PGE2091	3015-10	11.0	1				1					0.0	89,724	0	7.4	175,788	N/A	N/A	1.96	N/A	0.81	6	142,388	N/A
P00136	PGE2056	2	1.1	1	1	1	1	1	1	1		1	62.0	3,156,290	107,575	161.2	2,770,266	59,407	2.60	0.88	0.6	0.93	150	2,576,348	55,248
P00140	PGE2070	2	12.5		1	1		1			1	1	222.0	1,152,386	0	25.0	328,312	0	0.11	0.28	N/A	0.62	16	203,553	0
P00145	PGE2071	1	6.0	1		1		1			1	1	71.1	502,439	0	73.3	408,718	0	1.03	0.81	N/A	0.97	71	396,456	0
P00150	PGE2090	1	1.1		1		1					1	133.0	1,161,967	68,791	0.0	-261,268	-34,906	0.00	-0.22	-0.5	N/A	N/A	N/A	N/A
P00278	PGE2036	2K6UCR002	1.1			1		1				1	44.0	255,190	92,200	47.0	549,704	98,618	1.07	2.15	1.1	0.83	39	456,254	81,853
P00281	PGE2036	2K0700071	1.1	1						1			69.4	607,533	11,528	80.0	597,538	261,330	1.15	0.98	22.7	0.90	72	537,784	235,197
P00285	PGE2036	2K6UCR013	1.1				1					1	103.0	1,412,350	25,275	29.8	359,764	0	0.29	0.25	0.0	1.00	30	359,764	0
P00289	PGE2036	2K0701231	1.1				1					1	0.0	852,825	45,405	255.8	849,505	58,877	N/A	1.00	1.3	1.00	256	849,505	58,877
P00292	PGE2036	2K6UCR012	12.5				1					1	105.0	258,962	6,207	9.2	0	0	0.09	0.00	0.0	1.00	9	0	0
P01137	PGE2002	2K08007653	22.5		1	1			1	1			0.0	23,000	0	0.0	9,116	0	N/A	0.40	N/A	0.90	0	8,204	0
P01770	PGE2007	2K08004335	1.0				1		1				0.0	0	197,610	0.0	0	0	N/A	N/A	0.0	0.88	0	0	0
P01982	PGE2015	2K07001074	1.0			1			1	1	1		60.5	823,858	88,738	3.9	110,708	14,232	0.06	0.13	0.2	0.86	3	95,209	12,240
P01987	PGE2015	2K0806875	22.5				1	1					2.5	88,924	0	2.9	80,679	0	1.16	0.91	N/A	0.18	1	14,522	0
P04407	PGE2032	TAA0001235	8.0		1		1					1	N/A	542,592	28,286	0.0	333,012	13,479	N/A	0.61	0.5	0.81	0	269,740	10,918
P04434	PGE2035	2K0809839	8.0		1				1	1			0.0	379,106	10,749	0.0	0	0	N/A	0.00	0.0	0.87	0	0	0

				System Class				Measure Class					Ex Ante Gross Savings			Ex Post 1st Year Gross Savings			Gross Realization Rate			NTGR	Ex Post 1st Year Net Savings		
SBW ID	IOU Program	IOU ID	Case Weight	Central plant	HVAC - general	HVAC - air distribution	Other/ unclassified	VFD	Improve scheduling	Improve control strategies	Improve outside air use	Other/ unclassified	kW	kWh	Therms	kW	kWh	Therms	kW	kWh	Therms		kW	kWh	Therms
P04435	PGE2035	2K07000456	1.0	1						1		1	294.9	2,913,574	0	6.0	601,541	0	0.02	0.21	N/A	0.90	5	541,387	0
P04665	PGE2052	TBA0001604	1.0			1				1		1	30.8	1,282,719	138,227	192.0	812,483	97,720	6.23	0.63	0.7	0.79	152	641,862	77,199
P04671	PGE2052	TBA0001779	1.0			1	1	1	1		1		1.5	1,107,349	228,717	31.9	475,275	109,834	21.24	0.43	0.5	0.93	30	442,006	102,146

Table 31: PG&E – Reasons for Difference Between Ex Post and Ex Ante

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P00090	20	1.1	R C Economizers not fully functional	R P M Program calculations overestimated air handler VFD and chilled water valve savings	R P	
P00096	TAA0001648	1.1	R P M Minimum outside air baseline assumptions overestimated savings.	R P Cogged fan belt savings overestimated	R C Several VFDs failed causing AHU fans to run at full speed	
P00097	TAA0000947	1.1	I P M Error in economizer savings calculation underestimated savings.			
P00098	TAA0000947	6.0	R C M Economizers not operational			
P00107	3001-01	11.0	I P M Deadband measure saving more than anticipated by program calculations			
P00125	3015-10	11.0	I P M Incorrect power equation underestimated savings for VFDs on cooling tower fans			

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P00136	2	1.1	R P Boilers operating at low loads	I P Simultaneous heating/cooling larger than estimated	R P Heating loads lower than estimated	I P M Chiller efficiency improvement higher than estimated
P00140	2	12.5	R P Supply air temperature control strategy for fan speed was not successful	I P Server room curtians to separate supply and return air saved more than anticipated by program calculation	R P M Savings estimated for economizer operation were unrealistic	
P00145	1	6.0	R C M Economizer not operational			
P00150	1	1.1	R P M Occupancy sensor control of fume hoods uses more energy than previous manual sash operation			
P00278	2K6UCR002	1.1	R P Filter changes were disallowed as regular maintenance	I P M Modeling error underestimated the fan energy savings		
P00281	2K0700071	1.1	I P M The steam production efficiency factor was omitted from program calculation resulting in greater natural gas savings			
P00285	2K6UCR013	1.1	R P M Enhanced post data in whole building analysis showed greatly reduced savings			
P00289	2K0701231	1.1	R P M Enhanced post data in whole building analysis showed greatly reduced savings			
P00292	2K6UCR012	12.5	R P M Enhanced post data in whole building analysis showed greatly reduced savings			
P01137	2K08007653	22.5	R C HVAC rescheduling did not occur	R P M Natural gas savings eliminated because HVAC units had electric resistance heat	I P Higher efficiency heat pump units installed than reported by program resulting in greater savings	
P01770	2K08004335	1.0	None			

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P01982	2K07001074	1.0	R P M Economizer operation unchanged	R P Discharge air reset unchanged from pre-implementation condition for 3 AHUs	R P AHU schedule changes partially implemented	
P01987	2K0806875	22.5	R C M Nighttime setback strategy defeated during part of the post-implementation monitoring period			
P04407	TAA0001235	8.0	R P M Program calculations overestimated savings			
P04434	2K0809839	8.0	R C M HVAC rescheduling did not occur			
P04435	2K07000456	1.0	R C Improved chiller sequencing not implemented.	R C Fan speed and temperature setpoint changes not implemented.		
P04665	TBA0001604	1.0	R C Reductions in HVAC hours less than estimated	R C Supply air temperature reset did not occur	R C Economizers not working	
P04671	TBA0001779	1.0	R C M Reductions in HVAC and lighting hours less than estimated	R C Economizers not fully working	R C Fan speeds not reduced as much as estimated	

Table 32: SCE – Gross and Net Savings Results for Sampled Projects

				System Class				Measure Class					Ex Ante Gross Savings			Ex Post 1st Year Gross Savings			Gross Realization Rate			NTGR	Ex Post 1st Year Net Savings		
SBW ID	IOU Program	IOU ID	Case Weight	Central plant	HVAC - general	HVAC - air distribution	Other/ unclassified	VFD	Improve scheduling	Improve control strategies	Improve outside air use	Other/ unclassified	kW	kWh	Therms	kW	kWh	Therms	kW	kWh	Therms		kW	kWh	Therms
P00007	SCE2508	1000-01	4.3	1	1					1			0.0	278,085	0	25.0	131,914	0	N/A	0.47	N/A	0.80	20	105,531	0
P00016	SCE2508	1006-01a	4.0	1		1				1	1		0.0	747,007	0	36.8	314,573	0	N/A	0.42	N/A	0.66	24	207,618	0
P00020	SCE2508	1006-05	9.7	1		1			1	1	1	1	10.0	138,114	0	10.0	287,873	2,769	1.00	2.08	N/A	0.76	8	218,783	2,104

				System Class				Measure Class					Ex Ante Gross Savings			Ex Post 1st Year Gross Savings			Gross Realization Rate			NTGR	Ex Post 1st Year Net Savings		
SBW ID	IOU Program	IOU ID	Case Weight	Central plant	HVAC - general	HVAC - air distribution	Other/ unclassified	VFD	Improve scheduling	Improve control strategies	Improve outside air use	Other/ unclassified	kW	kWh	Therms	kW	kWh	Therms	kW	kWh	Therms		kW	kWh	Therms
P00033	SCE2508	1015-01	9.7				1			1			11.9	142,766	0	19.3	285,308	0	1.62	2.00	N/A	N/A	N/A	N/A	N/A
P00243	SCE2528	V3057901	1.0	1		1			1	1	1		51.0	3,783,049	171,633	45.0	2,487,355	0	0.88	0.66	0.0	0.93	42	2,313,240	0
P00244	SCE2528	V3057902	1.0	1	1	1	1	1	1	1			59.0	722,223	48,289	-15.0	455,840	0	-0.25	0.63	0.0	0.93	-14	423,931	0
P00250	SCE2528	V3057901	4.0	1	1	1			1	1	1		-1.0	845,179	25,023	-6.0	626,706	0	6.00	0.74	0.0	0.93	-6	582,837	0
P00253	SCE2528	V3096901	1.0	1	1	1	1		1	1			59.0	1,670,298	N/A	-26.0	1,078,829	0	-0.44	0.65	N/A	0.93	-24	1,003,311	0
P00266	SCE2528	V3127902	4.3		1	1			1	1	1		-27.0	367,170	8,341	-14.0	238,428	0	0.52	0.65	0.0	0.93	-13	221,738	0
P00297	SCE2530	1023	1.0				1					1	0.0	1,257,006	0	0.0	1,257,006	0	N/A	1.00	N/A	1.00	0	1,257,006	0
P00298	SCE2530	1	4.0				1					1	0.0	613,146	0	138.0	638,135	0	N/A	1.04	N/A	0.90	124	574,322	0
P00302	SCE2530	5	9.7				1					1	19.0	169,243	0	25.3	253,061	0	1.33	1.50	N/A	0.83	21	210,041	0
P00303	SCE2530	1061	4.3				1					1	53.0	295,593	0	31.3	459,563	0	0.59	1.55	N/A	0.83	26	381,437	0

Table 33: SCE – Reasons for Difference Between Ex Post and Ex Ante

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P00007	1000-01	4.3	R C M Pre and post chiller staging strategies were similar	R P Program calculations overestimated savings		
P00016	1006-01a	4.0	R C Revised chiller staging strategy	R C M Less condensing water supply temperature reduction than recommended	R P Economizers not working as recommended	I C Less efficient chillers

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P00020	1006-05	9.7	I P M Program calculations underestimated fan savings			
P00033	1015-01	9.7	I P M Anti-sweat heater control had lower duty cycle	I C Changes to lighting operation		
P00243	V3057901	1.0	R C M Fans operating during unoccupied periods			
P00244	V3057902	1.0	R C M Fans operating during unoccupied periods	R P Lower baseline for static pressure control		
P00250	V3057901	4.0	R C Change in scheduled occupied times	R C M Economizers not operational	R C Chiller/boiler lockout not functioning	
P00253	V3096901	1.0	R C M Scheduled lighting hours increased	R C Chiller sequence optimization did not occur	R C Less condensing water supply temperature reduction than recommended	R C Less aggressive discharge air temperature reset than recommended
P00266	V3127902	4.3	R C M HVAC rescheduling did not occur	R C Cold deck temperature reset did not occur	R C Chilled water temperature reset did not occur	
P00297	1023	1.0	None			
P00298	1	4.0	None			
P00302	5	9.7	I P M Program calculations underestimated savings			
P00303	1061	4.3	R P M Program calculations underestimated savings			

Table 34: SCG – Gross and Net Savings Results for Sampled Projects

				System Class				Measure Class					Ex Ante Gross Savings			Ex Post 1st Year Gross Savings			Gross Realization Rate			NTGR	Ex Post 1st Year Net Savings		
SBW ID	IOU Program	IOU ID	Case Weight	Central plant	HVAC - general	HVAC - air distribution	Other/ unclassified	VFD	Improve scheduling	Improve control strategies	Improve outside air use	Other/ unclassified	kW	kWh	Therms	kW	kWh	Therms	kW	kWh	Therms		kW	kWh	Therms
P00235	SCG3520	1	1.0		1					1			0.0	0	16,500	81.3	237,295	39,104	N/A	N/A	2.4	1.00	81	237,295	39,104
P00268	SCG3527	1	1.0	1	1	1			1				0.0	0	141,167	0.0	0	151,942	N/A	N/A	1.1	0.93	0	0	141,306
P00270	SCG3527	3	1.0	1	1	1			1	1	1	1	0.0	0	36,482	0.0	0	44,486	N/A	N/A	1.2	0.93	0	0	41,372
P00274	SCG3527	7	5.0		1	1			1	1	1		0.0	0	25,169	0.0	0	27,023	N/A	N/A	1.1	0.93	0	0	25,131
P00276	SCG3520	2	1.0		1					1			0.0	0	49,750	76.1	568,572	148,353	N/A	N/A	3.0	1.00	76	568,572	148,353
P00324	SCG3527	9	4.3		1	1			1	1			0.0	0	10,289	0.0	0	18,102	N/A	N/A	1.8	0.93	0	0	16,835
P00325	SCG3527	2	1.0	1		1		1	1	1			0.0	0	77,208	0.0	0	23,433	N/A	N/A	0.3	0.93	0	0	21,793
P00327	SCG3527	11	5.0		1				1	1			0.0	0	27,680	0.0	0	3,363	N/A	N/A	0.1	0.93	0	0	3,128
P00336	SCG3527	20	4.3	1	1	1			1	1	1		0.0	0	7,653	0.0	0	6,356	N/A	N/A	0.8	0.93	0	0	5,911
P00339	SCG3527	23	4.3		1	1			1	1	1		0.0	0	8,566	0.0	0	-2,952	N/A	N/A	-0.3	0.93	0	0	-2,745

Table 35: SCG – Reasons for Difference Between Ex Post and Ex Ante

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P00235	1	1.0	None			
P00268	1	1.0	R C M Fans operating during unoccupied periods			
P00270	3	1.0	R C M HVAC running continuously with no setbacks	I C Economizer and fan VFD savings increased because of continuous operation.		
P00274	7	5.0	I C M Slight change in operating schedule	R C One of 7 economizers not working		

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P00276	2	1.0	I P M Cold and hot deck reset saved more than program calculation estimated			
P00324	9	4.3	I C M Improved occupancy scheduling	R C Outside air lockout not working		
P00325	2	1.0	R C M Fans operating during unoccupied periods			
P00327	11	5.0	R C M HVAC rescheduling did not occur			
P00336	20	4.3	R C M Change in scheduled occupied times			
P00339	23	4.3	R C M HVAC rescheduling did not occur	R C Cold deck temperature reset did not occur	R C Hot water pump lockout did not occur	

Table 36: SDG&E – Gross and Net Savings Results for Sampled Projects

				System Class				Measure Class					Ex Ante Gross Savings			Ex Post 1st Year Gross Savings			Gross Realization Rate			NTGR	Ex Post 1st Year Net Savings		
SBW ID	IOU Program	IOU ID	Case Weight	Central plant	HVAC - general	HVAC - air distribution	Other/ unclassified	VFD	Improve scheduling	Improve control strategies	Improve outside air use	Other/ unclassified	kW	kWh	Therms	kW	kWh	Therms	kW	kWh	Therms		kW	kWh	Therms
P00316	SDGE3026	1	1.0				1					1	58.8	587,965	29,595	86.0	1,000,639	40,768	1.46	1.70	1.4	0.98	84	980,626	39,953
P00317	SDGE3026	2	2.0			1					1		18.1	181,269	16,147	26.9	101,740	2,820	1.48	0.56	0.2	0.98	26	99,705	2,764
P00319	SDGE3026	4	1.0	1	1			1		1			110.0	1,100,000	162,100	394.1	1,314,994	0	3.58	1.20	0.0	0.26	102	341,898	0

Table 37: SDG&E – Reasons for Difference Between Ex Post and Ex Ante

SBW ID	IOU ID	Case Weight	Reason 1	Reason 2	Reason 3	Reason 4
P00316	1	1.0	R P M eQuest model overestimated reheat savings			
P00317	2	2.0	R P M eQuest model overestimated reheat savings			
P00319	4	1.0	I Original estimate not available, so reason(s) not known.			

5.6. Detailed Findings for the Net Sample Projects

Project specific results of all the NTGR data collection and analysis activity are presented in the table below. As detailed elsewhere in this appendix, decision makers were asked a series of questions to ascertain the degree of program influence and the likelihood that a retro-commissioning (RCx) project would have been implemented in the absence of the program. The interview results in three scores: a Timing and Selection Score, a Program Influence Score, and a No-Program Score, which are averaged to calculate the NTGR.

The table shows each of the components of the NTGR score for every completed interview. In the few cases where respondents said they had been considering specific measures prior to the RCx project, measure-specific results were used to calculate the NTGR, and the value is shown in the table with an asterisk. For those cases, all in PG&E's service territory, measure specific Timing and Selection, Program Influence and No-Program scores are presented on a separate page.

For customers subject to the Standard-Very Large rigor level, the table also shows the results of the review by two analysts. The columns under "Adjusted or Consensus" show whether this review resulted in individual components of the NTGR score being adjusted and whether the scoring was overridden by analyst judgment based on information available from other sources. The right-most column shows the final NTGR used in the net analysis.

Table 38: PG&E – NTGR Scoring Components and Final Score for Sampled Measures

			Decision Maker Survey(s)				Adjusted or Consensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P00074	4	12.5	8.0	8.0	10.0	0.87	N/A	N/A	N/A	N	N	0.87
P00078	8	12.5	8.0	8.0	10.0	0.87	N/A	N/A	N/A	N	N	0.87
P00086	16	11.0	8.0	8.0	10.0	0.87	N/A	N/A	N/A	N	N	0.87
P00090*	20	1.1	8.0	8.0	10.0	0.88	8.0	8.0	10.0	Y	Y	0.87
P00094	24	11.0	8.0	8.0	10.0	0.87	N/A	N/A	N/A	N	N	0.87
P00096	TAA0001648	1.1	10.0	4.0	9.4	0.78	10.0	4.0	9.4	Y	N	0.78
P00097	TAA0000947	1.1	10.0	4.0	9.4	0.78	10.0	4.0	9.4	Y	N	0.78
P00098	TAA0000947	6.0	10.0	4.0	9.4	0.78	10.0	4.0	9.4	Y	N	0.78
P00099	TAA0000947	12.5	10.0	4.0	9.4	0.78	N/A	N/A	N/A	N	N	0.78
P00100	TAA0001521	12.5	10.0	4.0	9.4	0.78	N/A	N/A	N/A	N	N	0.78
P00107	3001-01	11.0	10.0	7.0	10.0	0.90	10.0	7.0	10.0	Y	N	0.90
P00108	3001-03	11.0	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P00111*		6.0	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P00116*	3006-01	12.5	10.0	8.0	10.0	0.92	N/A	N/A	N/A	N	N	0.92
P00125	3015-10	11.0	10.0	6.0	8.3	0.81	10.0	6.0	8.3	Y	N	0.81
P00126*	3019-01	11.0	7.0	7.0	2.0	0.67	N/A	N/A	N/A	N	N	0.67
P00135	1	6.0	10.0	2.5	3.0	0.52	N/A	N/A	N/A	N	N	0.52
P00136	2	1.1	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00137	3	6.0	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00140	2	12.5	8.0	7.0	3.5	0.62	8.0	7.0	3.5	Y	N	0.62
P00141	3	12.5	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P00145	1	6.0	10.0	10.0	9.2	0.97	10.0	10.0	9.2	Y	N	0.97
P00239*	2K07001692	6.0	10.0	3.5	0.0	0.28	N/A	N/A	N/A	N	N	0.28
P00240*	2K07001689	11.0	10.0	3.5	0.0	0.20	N/A	N/A	N/A	N	N	0.20
P00242	2K07002031	12.5	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P00277	2K6UCR007	12.5	8.0	8.0	9.0	0.83	N/A	N/A	N/A	N	N	0.83

			Decision Maker Survey(s)				Adjusted or Consensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P00278	2K6UCR002	1.1	8.0	8.0	9.0	0.83	8.0	8.0	9.0	Y	N	0.83
P00279	2K6UCR004	12.5	7.0	6.0	9.6	0.75	N/A	N/A	N/A	N	N	0.75
P00281	2K0700071	1.1	10.0	7.0	10.0	0.90	10.0	7.0	10.0	Y	N	0.90
P00283	2K6UCR009	11.0	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P00285	2K6UCR013	1.1	10.0	10.0	10.0	1.00	10.0	10.0	10.0	Y	N	1.00
P00286	2K0700010	11.0	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P00287	2K0700008	11.0	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P00288	2K0700009	12.5	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P00289	2K0701231	1.1	10.0	10.0	10.0	1.00	10.0	10.0	10.0	Y	N	1.00
P00290	2K0700007	11.0	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P00292	2K6UCR012	12.5	10.0	10.0	10.0	1.00	10.0	10.0	10.0	Y	N	1.00
P00294	2K0700212	1.1	10.0	5.0	8.9	0.80	N/A	N/A	N/A	N	N	0.80
P01130	2K08007648	22.5	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P01131	2K08007718	22.5	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P01132	2K08007655	22.5	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P01133	2K08007640	22.5	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P01135	2K08007651	22.5	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P01137	2K08007653	22.5	10.0	7.0	10.0	0.90	10.0	7.0	10.0	Y	N	0.90
P01138	2K08007656	22.5	10.0	7.0	10.0	0.90	N/A	N/A	N/A	N	N	0.90
P01309	2K08012093	8.0	10.0	3.0	10.0	0.77	N/A	N/A	N/A	N	N	0.77
P01311	2K08008459	22.5	8.0	8.0	10.0	0.87	N/A	N/A	N/A	N	N	0.87
P01312	2K07002036	22.5	10.0	4.0	9.4	0.78	N/A	N/A	N/A	N	N	0.78
P01764	2K07003854	22.5	8.0	2.0	10.0	0.67	N/A	N/A	N/A	N	N	0.67
P01766	2K08012079	22.5	10.0	2.5	10.0	0.75	N/A	N/A	N/A	N	N	0.75
P01769	2K07003856	8.0	8.0	4.0	1.4	0.45	N/A	N/A	N/A	N	N	0.45
P01770	2K08004335	1.0	10.0	8.0	8.3	0.88	10.0	8.0	8.3	Y	N	0.88
P01776	2K08006602	22.5	10.0	4.0	8.9	0.76	N/A	N/A	N/A	N	N	0.76
P01777	2K08008474	22.5	9.0	7.0	7.2	0.77	N/A	N/A	N/A	N	N	0.77

			Decision Maker Survey(s)				Adjusted or Consensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P01780	2K07003849	22.5	10.0	8.0	8.3	0.88	N/A	N/A	N/A	N	N	0.88
P01781*	2K07003851	22.5	10.0	4.0	1.4	0.55	N/A	N/A	N/A	N	N	0.55
P01783	2K08006880	22.5	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P01784	2K08010177	22.5	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P01982	2K07001074	1.0	10.0	10.0	5.7	0.86	10.0	10.0	5.7	Y	N	0.86
P01987	2K0806875	22.5	6.0	3.5	0.0	0.18	6.0	3.5	0.0	Y	N	0.18
P01988*	2K08004138/ 2K07001093	8.0	10.0	7.0	2.0	0.76	N/A	N/A	N/A	N	N	0.76
P01990	2K07000624	22.5	8.0	5.0	10.0	0.77	N/A	N/A	N/A	N	N	0.77
P04407	TAA0001235	8.0	10.0	6.0	8.3	0.81	10.0	6.0	8.3	Y	N	0.81
P04433*	2K0809480C	8.0	10.0	6.0	10.0	0.87	N/A	N/A	N/A	N	N	0.87
P04434	2K0809839	8.0	9.0	7.0	10.0	0.87	9.0	7.0	10.0	Y	N	0.87
P04435	2K07000456	1.0	10.0	7.0	10.0	0.90	10.0	7.0	10.0	Y	N	0.90
P04474	2K0806921C	8.0	10.0	10.0	10.0	1.00	N/A	N/A	N/A	N	N	1.00
P04664*	TBB0001604	22.5	10.0	2.5	0.0	0.13	N/A	N/A	N/A	N	N	0.13
P04665	TBA0001604	1.0	10.0	6.0	7.7	0.79	10.0	6.0	7.7	Y	N	0.79
P04666	TBA0001779	22.5	10.0	2.5	3.0	0.52	N/A	N/A	N/A	N	N	0.52
P04671	TBA0001779	1.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P04674	TAA0001197	22.5	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P04676	TAA0001197	8.0	10.0	5.0	10.0	0.83	N/A	N/A	N/A	N	N	0.83

* Measure-specific values were incorporated into the Standard Scoring NTGR.

Table 39: PG&E – Measure-Specific NTGRs

		MEASURE 1					MEASURE 2					PROJECT		
SBW ID	IOU Project ID	Timing and Selection	Program Influence	No-Program	NTGR	Measure 1 Ex Ante MMBTU savings	Timing and Selection	Program Influence	No-Program	NTGR	Measure 2 Ex Ante MMBTU savings	RCx Decision NTGR	Total Project Ex Ante MMBTU savings	Ex Ante Savings Weighted Project NTGR
P00090	20	10.0	7.0	10.0	0.90	1678	10.0	7.0	10.0	0.90	727	0.87	6678	0.87

		MEASURE 1					MEASURE 2					PROJECT		
SBW ID	IOU Project ID	Timing and Selection	Program Influence	No-Program	NTGR	Measure 1 Ex Ante MMBTU savings	Timing and Selection	Program Influence	No-Program	NTGR	Measure 2 Ex Ante MMBTU savings	RCx Decision NTGR	Total Project Ex Ante MMBTU savings	Ex Ante Savings Weighted Project NTGR
P00111		10.0	7.0	10.0	0.90	910	10.0	7.0	10.0	0.90	367	0.90	3199	0.90
P00116	3006-01	10.0	4.0	10.0	0.80	168	10.0	4.0	10.0	0.80	66	0.93	1761	0.92
P00126	3019-01	10.0	3.5	10.0	0.78	226	N/A	N/A	N/A	N/A	N/A	0.53	419	0.67
P00239	2K07001692	10.0	7.0	0.0	0.35	1463	10.0	7.0	0.0	0.35	476	0.18	3268	0.28
P00240	2K07001689	10.0	7.0	0.0	0.35	136	N/A	N/A	N/A	N/A	420	0.18	878	0.20
P01781	2K07003851	10.0	5.0	1.4	0.55	364	0.0	N/A	N/A	N/A	0	0.51	364	0.55
P01988	2K08004138/ 2K07001093	10.0	7.0	10.0	0.90	327	10.0	7.0	10.0	0.90	1419	0.63	3552	0.76
P04433	2K0809480C	10.0	6.0	9.4	0.85	16	N/A	N/A	N/A	N/A	0	0.87	2373	0.87
P04664	TBB0001604	10.0	2.5	0.0	0.13	89	10.0	2.5	0.0	0.13	476	0.13	252	0.13

Table 40: SCE – NTGR Scoring Components and Final Score for Sampled Measures

			Decision Maker Survey(s)				Adjusted or Concensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P00007	1000-01	4.3	9.0	5.0	10.0	0.80	9.0	5.0	10.0	Y	N	0.80
P00009	1000-05	9.7	9.0	7.0	6.4	0.75	N/A	N/A	N/A	N	N	0.75
P00010	1001-01a	9.7	10.0	8.0	8.6	0.89	N/A	N/A	N/A	N	N	0.89
P00014	1002-03	9.7	10.0	5.0	5.5	0.68	N/A	N/A	N/A	N	N	0.68
P00016	1006-01a	4.0	8.0	4.0	7.9	0.66	8.0	4.0	7.9	Y	N	0.66
P00020	1006-05	9.7	10.0	6.0	6.8	0.76	10.0	6.0	6.8	Y	N	0.76
P00026	1008-01b	9.7	10.0	5.0	4.0	0.63	N/A	N/A	N/A	N	N	0.63
P00243	V3057901	1.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00244	V3057902	1.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00245	V3096901	4.0	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00250	V3057901	4.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00252	V3057901	4.3	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93

			Decision Maker Survey(s)				Adjusted or Concensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P00253	V3096901	1.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00255	V3096901	4.0	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00256	V3096901	4.3	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00260	V3096901	9.7	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00261	V3096901	4.3	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00264	V3127902	4.3	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00265	V3057901	9.7	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00266	V3127902	4.3	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00267	V3127903	9.7	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00297	1023	1.0	10.0	10.0	10.0	1.00	10.0	10.0	10.0	Y	N	1.00
P00298	1	4.0	10.0	8.0	9.2	0.90	10.0	8.0	9.2	Y	N	0.90
P00299	2	9.7	10.0	8.0	5.5	0.78	N/A	N/A	N/A	N	N	0.78
P00302	5	9.7	8.0	7.0	10.0	0.83	8.0	7.0	10.0	Y	N	0.83
P00303	1061	4.3	8.0	7.0	10.0	0.83	8.0	7.0	10.0	Y	N	0.83
P00304	1054	4.0	8.0	7.0	10.0	0.83	N/A	N/A	N/A	N	N	0.83
P00305	1058	9.7	8.0	7.0	10.0	0.83	N/A	N/A	N/A	N	N	0.83
P00321	V3127903	9.7	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93

Table 41: SCG – NTGR Scoring Components and Final Score for Sampled Measures

			Decision Maker Survey(s)				Adjusted or Concensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P00235	1	1.0	10.0	10.0	10.0	1.00	10.0	10.0	10.0	Y	N	1.00
P00268	1	1.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00270	3	1.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00274	7	5.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93

			Decision Maker Survey(s)				Adjusted or Concensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P00276	2	1.0	10.0	10.0	10.0	1.00	10.0	10.0	10.0	Y	N	1.00
P00309	5	5.0	10.0	8.0	5.5	0.78	N/A	N/A	N/A	N	N	0.78
P00324	9	4.3	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00325	2	1.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00327	11	5.0	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00329	13	5.0	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00331	15	4.3	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00333	17	4.3	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00334	18	5.0	10.0	8.0	10.0	0.93	N/A	N/A	N/A	N	N	0.93
P00336	20	4.3	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93
P00339	23	4.3	10.0	8.0	10.0	0.93	10.0	8.0	10.0	Y	N	0.93

Table 42: SDG&E – NTGR Scoring Components and Final Score for Sampled Measures

			Decision Maker Survey(s)				Adjusted or Concensus					
SBW ID	IOU Project ID	Case Weight	Timing and Selection	Program Influence	No-Program	Standard Scoring NTGR	Timing and Selection	Program Influence	No-Program	Was Standard Very Large (Y/N)	Scoring Overridden by other information (Y/N)	Final NTGR
P00316	1	1.0	10.0	10.0	9.5	0.98	10.0	10.0	9.5	Y	N	0.98
P00317	2	2.0	10.0	10.0	9.5	0.98	10.0	10.0	9.5	Y	N	0.98
P00319	4	1.0	7.0	2.0	0.7	0.32	5.0	2.0	0.7	Y	Y	0.26

5.7. Site Reports

Reports that document the gross savings analyses for the completed sample are provided in the file called "SiteReports.exe" (a self-extracting archive created with 7-Zip) that accompanies this document. Contained within this file are files for each site, identified by the SBWID. These reports have been redacted to protect the identities of the businesses that participated in this study.

5.8. Glossary of Acronyms

Acronym	Definition
BTU	Building Tune-Up
CCC	California Community Colleges
CFM	cubic feet per minute
CHES	Campus Housing Efficiency Solutions
CHP	combined heat and power
CPC	Climate Protection Campaign
CPEEP	California Preschool Energy Efficiency Program
CPUC	California Public Utilities Commission
CSU	California State University
CYES	California Youth Energy Services
DOE	Department of Energy
EEM	energy efficiency measure
EMCS	Energy Management Control System
EMS	Energy Management System
ESB	Energy Savings Bid
EUL	Effective Useful Life
HIM	High Impact Measure
HVAC	Heating, Ventilation, and Air Conditioning
IOU	Investor-Owned Utility
ISD	Internal Services Department
LGP	Local Government Partnership
LIIF	Low Income Investment Fund
MCEMP	Macy's Comprehensive Energy Management Program
MCEW	Marin County Energy Watch
MEMT	Marin Energy Management Team
NOAA	National Oceanic and Atmospheric Administration
NTGR	Net-To-Gross Ratio
OAT	Outside Air temperature
OEP	Oakland Energy Partnership
OSA	Outside Air

Acronym	Definition
RSG	Resource Solutions Group
SBEA	Small Business Energy Alliance
SCEW	Sonoma County Energy Watch
SCG	Southern California Gas
SDG&E	San Diego Gas & Electric
SEE	School Energy Efficiency
SRA	Self-Report Approach
SSV	Sustainable Silicon Valley
UES	Unit Energy Savings
VAV	Variable Air Volume
VFD	Variable frequency Drive

5.9. Responses to Public Comments on the Draft Evaluation Report

During the public review period from December 7-31, 2009, the CPUC received 23 sets of comments on the draft evaluation report. These sets ranged from a single comment contained within one paragraph, to complicated sets of nested comments addressing general issues, as well as site- and measure-specific issues. We consolidated these comments to eliminate redundancy, and organized them so that similar questions could be addressed in a like manner. As necessary, we consulted with the primary analysts for particular sites evaluated, as well as other members of the evaluation team, to develop responses to address the issues raised.

Table 43 lists the 35 discrete public comments, sorted into four groups—general, gross (site-specific), net, and EUL—and summarizes each comment and the corresponding evaluator response. The remainder of this appendix contains the full verbatim text of the comments, comment documents, and responses. This text is ordered by the Comment ID number we assigned to each comment.

Table 43: Summary of Public Comments and Evaluator Responses

Grouping	Comment ID*	Author	Subject	Section/ Page	Date	Comment summary	Response summary
General	1	Keith Rothenberg	Commercial RCx	--	12/23/09	Provide more information on how to improve programs, particularly savings estimates.	Evaluation scope limited what we could determine in this area. Ex ante estimates were generally performed credibly, but perhaps change in focus, baseline data collection is needed.
General	2	Lia Webster	Commercial Retro-Commissioning	p 115 - p 124	12/30/09	Add program numbers to site tables.	Have been added.
General	3	Lia Webster	Commercial Retro-Commissioning	--	12/30/09	Show how program design differences affected realization	Evaluation's focus on savings, not so much program processes.

Grouping	Comment ID*	Author	Subject	Section/ Page	Date	Comment summary	Response summary
						rates, or at least warn of the dangers of lumping all programs together.	Premature to conclude that RCx is "high risk."
General	10	PG&E Company	HIM	General	12/31/09	Inappropriate to draw conclusions about programs and measures from HIM analysis. Why was it done this way?	General focus on RCx as a measure and delivery method at the IOU level. These results will be applied per ED framework for updating IOU claims. RCx bundling in IOU records made detailed analysis difficult.
General	11	Alison Watson	Commercial Retro-Commissioning - M&V Data	--	12/31/09	Request all supporting files.	Provided through IOU channels.
General	12	PG&E Company	Review not completed	General	12/31/09	Insufficient site-level info, time, and resources to do full review.	Agree that time was limited, but feel that site info was complete.
General	13	PG&E Company	Review not completed (part 2 of 2)	General		See Comment #12.	See Comment #12
General	14A	PG&E Company	PG&E Comments	General	12/31/09	Very similar to Comment #12.	See Comment #12
General	15A	Athena Besa	Overarching Comments	--	12/31/09	No specific reasons for non-operational measures given--concerned that facilities down for maintenance, for instance, may have been unfairly zeroed out.	Provide several examples illustrating how we were careful not to penalize programs for anomalous events.
General	17	Pierre Landry	Realization Rates	--	12/31/09	Impact evaluations benefit from closer connections with the programs and the process evaluations, and might be best managed by the IOUs.	Agree that impact-process evaluation coordination can be beneficial.
General	18	Reuben Deumling	TURN Comments	--	1/1/10	6 pages reiterating key evaluation results, expressing concern over poor program performance, and prescribing various policy goals.	Comments duly noted but outside evaluation scope.
General	23A	Partnership	General comments	--	12/22/09	Individual measure analysis used in evaluation less accurate than whole facility basis.	Whole building vs. individual measure analysis decided on case-by-case basis. Former is not necessarily better.
General	23B	Partnership	General comments	--	12/22/09	Discrepancies between actual IOU claims and tracking DB claims used in evaluation for some sites.	Only the final IOU database numbers provided to the CPUC were used for realization rates.
General	23C	Partnership	General comments	--	12/22/09	Discrepancies between savings shown in evaluation	See Comment #23B.

Grouping	Comment ID*	Author	Subject	Section/ Page	Date	Comment summary	Response summary
						site report and main report appendix.	
Gross (site-specific)	15B	Athena Besa	Overarching Comments	--	12/31/09	3 measure-specific technical issues.	Further review yielded insufficient grounds for changing analysis.
Gross (site-specific)	15C	Athena Besa	Overarching Comments	--	12/31/09	1 measure-specific technical issue.	Further review yielded insufficient grounds for changing analysis.
Gross (site-specific)	19	Rachel Christenson	Response to NG Realization Rate P00136	Site Report P00136	12/30/09	Boiler efficiencies and loads used in evaluation are too low and should be revised.	Comment is valid. Analysis has been revised accordingly, improving the gas savings.
Gross (site-specific)	20	PG&E Company	Site report P01137	Site report PO 1137	12/31/09	Evaluation info about school occupancy/HVAC use contradicts that found during the RCx evaluation.	Information collected during evaluation came from a credible source, and seems to best reflect the latest conditions.
Gross (site-specific)	21	PG&E Company	Site report P01137	Site report PO 1137	12/31/09	Evaluation info about HVAC controls contradicts that found during the RCx evaluation.	Information collected during evaluation came from a credible source, and seems to best reflect the latest conditions.
Gross (site-specific)	23D	Partnership	1) SBW ID: P00281, IOU ID: 2K0700071 [CSU San Jose, MBCx --- Central Chilled Water Plant]	--		Discrepancy between application, claimed savings.	See Comment #23B.
Gross (site-specific)	23E	Partnership	2) SBW ID: P00285, IOU ID: 2K6UCR013 [UC Davis, MBCx - Life Sciences Addition]	--	12/22/09	Limited pre-implementation data used in analysis.	Pre data used was only source available from customer.
Gross (site-specific)	23F	Partnership	3) SBW ID: P00292, IOU ID: 2K6UCR012 [UC Davis, MBCx - Plant and Environmental Sciences]	--	12/22/09	Limited pre-implementation data used in analysis.	Pre data used was only source available from customer.
Gross (site-specific)	23G	Partnership	1) SBW ID: P00303, IOU ID: 1061 [UC Santa Barbara, MBCx - Life Sciences]	--	12/22/09	Discrepancies between (a) savings shown in evaluation site report and main report appendix, (b) measures included in eval and in project.	See Comment #23B for (a); for (b), site report has been revised to better reflect measures implemented.
Gross (site-specific)	23H	Partnership	1) SBW ID: P00317, IOU ID: 2 [UC SD, MBCx - Engineering Building Unit 2]	--	12/22/09	See Comment #15B.	See Comment #15B.
Gross (site-specific)	23I	Partnership	2) SBW ID: P00319, IOU ID: 4 [UC SD, MBCx Hybrid - Hillcrest Medical Center]	--	12/22/09	See Comment #15C.	See Comment #15C.
Net	6	PG&E	NTG		12/31/09	Lack of non-	Such a lack tells

Grouping	Comment ID*	Author	Subject	Section/ Page	Date	Comment summary	Response summary
		Company				program-driven RCx in Calif. should make NGTR = 1.	nothing about the "naturally-occurring" level of RCx in the absence of programs.
Net	7	PG&E Company	NTG	Table 3, p 3	12/31/09	How were NTGR results extrapolated?	Project-level NTGRs led to savings-weighted strata NTGRs for each fuel, then to strata-weighted IOU-level NTGRs.
Net	8	PG&E Company	NTG	Table 3, p 3	12/31/09	Repeat of Comment #7.	See Comment #7.
Net	9	PG&E Company	NTG	p 36, Sec 3.5	12/31/09	Good write-up of NTG self-report bias.	Glad this was noticed, as we strove to minimize bias per CPUC guidelines.
EUL	4	PG&E Company	Evaluation	General	12/31/09	How did/will evaluators handle short-lived measures?	First-year savings were based on as-found conditions, with some latitude to establish a typical year. This is consistent with CPUC direction. Measures that failed early were assigned a zero, as attempts to establish partial savings would be highly speculative.
EUL	5	PG&E Company	Evaluation	General	12/31/09	Were savings based on conditions when visited?	See Comment #4
EUL	14B	PG&E Company	PG&E Comments	General	12/31/09	Assigning zero savings to short-lived measures ignores any initial savings that occurred before the measures failed.	See Comment #4
EUL	16A	Pierre Landry	EULs	--	12/31/09	EUL study is not protocol-compliant and is fairly uncertain, so its results should not be used for this program cycle.	The CPUC does not plan to use these EUL results for this program cycle's cost effectiveness calculations.
EUL	16B	Pierre Landry	EULs	--	12/31/09	Agree that future RCx EUL studies are necessary and should be larger, protocol-compliant.	Comment noted.
EUL	22	PG&E Company	Site Report P01170	Site Report P01170	12/31/09	Despite recent failure, RCx measure actually lasted longer than claimed EUL.	See Comment #4

* Comments were posted between 12/23/09-1/1/10 on www.energydataweb.com/cpuc under "Commercial Retro-commissioning". Comments 1-18 were filed under the document "RCx 2006-08_Final_DRAFT_EMV Report v12061800_DISTR"; Comments 19-22 under the document "SiteReports v20091207.exe". Comment 23 was received separately from the Local Government Partnership contract group, on whose section this set of comments was posted. The evaluation team assigned values to the Grouping, Comment ID, and Comment/Response summary fields. Other fields came verbatim from www.energydataweb.com.

Comment 1

In general it would be helpful if the evaluation team can expand on their recommendations to improve the RCx programs. During the webinar, the presenter stated that the ex ante values are not a reliable predictor of ex post savings. Please request that the gross impact team expand on any recommendations to improve this situation. For instance, did they find that there was gaming? Were the ex ante technical reviews performed by the IOUs (if any) inadequate? Was the ex ante approach not based on accepted engineering methods? Were total meter approaches inappropriate? Were there not enough baseline and post implementation measurements to verify impacts? Comments to improve in this area could be very useful for the program managers.

Response to Comment 1

We understand the desire for program managers for feedback to improve their programs, and as much as the scope of the evaluation allowed, we tried to do so. The recommendations provided in Sections 2.7, 3.7, and 4.7 reflect advice that we felt we could support, given the constraints of the research conducted to date. The IOU-managed process evaluations for these programs might be able to shed more light on potential improvements to program designs and practices.

We can, however, speak in general terms about some of the questions raised in this comment. In light of the caveats above, these statements should be taken with a grain of salt.

Gaming: Among the sampled projects, we found little evidence of systematic attempts to doctor the numbers or mislead customers.

Ex ante technical reviews: We noticed that some programs did seem to expend a great deal of resources checking and refining ex ante estimates of savings. Whether these additional reviews provided enough value to justify the cost and the effort is unknown, but there did seem to be undue emphasis on refining the savings calculations, when in the end, the realized savings depend on many factors outside the program implementer's control, such as occupant feedback, changes in building operation, or adjustments to optimize a measure.

Ex ante calculations: Savings calculations ran the gamut from very simple spreadsheet calculations to extremely involved building simulation models. While evaluation analysts found errors and questionable assumptions in places, overall the quality of the ex ante calculations seemed reasonably good. It is unclear whether more involved calculations for a given project necessarily resulted in higher realization rates for that project. As we alluded to in the gross impact recommendations in Section 2.7, serious thought should be given to streamlining the savings calculation process and the resulting burdens on service providers, so that more of their efforts can be directed to better customer follow-up.

Total meter approaches: Using whole-building metering to estimate savings can be an appropriate and reasonable approach in certain cases, particularly when the RCx impacts are large and the facility is known to be quite stable in terms of other operational changes. But we found that it can also be risky. If confounding factors enter the picture, such as unexpected occupancy or equipment changes, or meter malfunctions, then savings estimates from this approach can become very unreliable. See the response to Comment 23A for further discussion of this topic.

Baseline and post data: The gross impact recommendations in Section 2.7 make clear the need for good baseline data, preferably collected by program staff. Lack of adequate baseline data was a major source of uncertainty in our evaluation, and logistically, it is nearly impossible for evaluators to collect much of it. The distinction needs to be made between collecting the data, and analyzing it, though: program implementers should have primary responsibility for data collection, but the subsequent analysis could be done by them or by others at some later date.

Comment 2

Please include program numbers in tables 30 through 37

Response to Comment 2

These tables have been updated to include IOU program ID numbers.

Comment 3

The report is thorough in its analysis and presentation of gross savings realization results from retro-commissioning programs in California from 2006-2008, but less so in its development of the context within which the results should be considered. The set of sampled projects includes a diverse array of programmatic approaches to retro-commissioning. Within each program for which a project was analyzed, any number of factors would have an impact on savings realization, including: service provider qualifications, owner incentives and/or measure maintenance obligations, energy calculation tools utilized, rigor of energy savings analysis and quality control methods, measure installation verification, operator training delivered, persistence management approach, etc., etc. Lacking an in-depth discussion or analysis of the factors that lead to the range of realization rates, the conclusions drawn risk misleading the audience toward the simple assumption that retro-commissioning through California utility programs, in all of their forms, is at best a high-risk endeavor. We strongly recommend that the rigorous quantitative analysis of this study should be complimented with either an analysis of how programmatic factors influenced gross realization rates or a detailed discussion that provides this context as a clear caveat to all of the hard numbers and key findings presented.

Response to Comment 3

We acknowledge that the RCx portfolio included a wide range of program sizes, designs, and target markets. The primary purpose of this evaluation, however, was to assess the portfolio's energy impacts, not those of individual programs. We would hope that readers of this report would not reach the hasty conclusion, based on its contents, that RCx is "high-risk." Without the examining other critical elements of the portfolio, such as (1) program costs, incorporated into economic models to assess cost-effectiveness of various programs and programmatic approaches, (2) customer and provider satisfaction, as measured in IOU-managed process evaluations, or (3) savings lifetimes, which the EUL portion of this evaluation concluded are still not well-known, it would be premature to conclude thusly. We agree that the aggregate gross and net savings numbers presented in our study are only part of the story.

Comment 4

How do the evaluators recommend conducting RCx evaluations in the future for short-lived measures, such as RCx scheduling measures? If the EUL is only three years and the evaluators are on site one to three years after installation, we would expect a significant portion of measures to have failed. Have the evaluators accounted for and/or quantified the effect of this issue on their findings?

Response to Comment 4

[Note: this response addresses Comments 4, 5, 14B, and 22 together.]

We developed our estimates of gross first-year savings based on current conditions at the time we visited and monitored the site. This is consistent with the direction given to all CPUC evaluators for this program cycle. So the case of a rescheduled cogeneration system that was taken out of service over a year before evaluators arrived would be no different than the case of an efficient lighting system installed in a building that was demolished prior to evaluation. In both cases, current policy dictates that the first-year savings be set to zero.

Evaluation analysts had some latitude to determine what constituted "typical" first-year conditions, based on metered data and customer reports: for example, at one site, a government building, a county election led to around-the-clock operation for several weeks during the site inspection and metering period. The analyst chose to use data from prior to the election to establish a typical year.

Any attempts by evaluators to assess when RCx measures stopped yielding savings would likely be highly speculative. Generally, detailed past trend data that would allow an analyst to pinpoint exactly when a measure failed is extremely rare. Also, measure performance tends to vary, particularly soon after it is implemented, as facility personnel make adjustments to "dial in" measures. The EUL portion of our study encountered just these kinds of problems in our efforts to establish actual measure lives. Future EUL research, as we recommended in Section 4.7, could shed more light on how to quantify and properly account for short-term RCx measures. Such research needs to consider not only outright measure failure, but savings degradation, where measures are slowly tweaked and adjusted away over time.

Comment 5

Is it true that site savings estimates were based on current conditions at the time of site visit?

Response to Comment 5

Refer to the response for Comment 4.

Comment 6

For PG&E, the NTG findings range from 0.76 to 0.86 (p. 38) based on the self report methodology. However, in response to a question during the webinar, the evaluators were unable to cite a single RCx project performed in California outside of a utility program. This is consistent with the utilities' assessment of current California retro-commissioning practices. These findings are mutually exclusive. While some respondents may say they would have performed some portion of RCx work without program influence, given that none actually exist, isn't the appropriate NTG 1.0?

Response to Comment 6

The fact that no RCx was encountered (and we were not explicitly looking) outside a program in California when there were numerous RCx programs offered means nothing in terms of the extent of RCx in the absence of a program. One would have been hard pressed to find a commissioning agent who was not fully booked with RCx program work to even do a project outside the program, and if asked they would most likely have told their client about the programs. In fact, we encountered a good example of the latter in the case of a university medical center that asked a building automation company to do a retro-commissioning-like effort and was then steered to the Partnership program by that automation company. Similarly, the fact that some level of RCx appears to take place in areas of the country where no RCx programs are offered indicates that the NTGR, while relatively high as indicated by self-report results, is not equal to 1.0.

Comment 7

Were NTG values estimated at the project level and then weighted by savings type? We see in section 1.1.2 that the self report method was used, but how were the individual survey results expanded to the population? The "projects" NTG for PG&E is shown in Table 3 as 0.81. How was this calculated?

Response to Comment 7

Yes, NTG values were estimated at the project level and weighted by savings type to calculate savings-type-specific NTGRs for each IOU. As described in Appendix section 5.5.2, we calculated a weighted mean NTGR for each stratum, with the NTGRs for individual cases weighted by their contribution to the total ex ante fuel savings for that stratum. For example, if case 1 of 10 cases in stratum n accounts for 130 kWh of the 1,000 kWh total ex ante savings in that stratum, the NTGR for that case is multiplied by 0.13. If Case 2 accounts for 80 kWh of the 1,000 kWh total, its NTGR is multiplied by 0.08, and so on. The sum of these products is the savings-weighted mean NTGR for the stratum. Results for each stratum were then weighted by that stratum's weight (the stratum population divided by the total population) to come up with the savings-type-specific NTGR for the IOU. In the case of the "Projects" NTGR, we simply assigned a savings value of 1 to each project so that all projects within the stratum were weighted equally in their contribution to the NTGR.

Comment 8

Were NTG values estimated at the project level and then weighted by savings type? We see in section 1.1.2 that the self report method was used, but how were the individual survey results expanded to the population? The "projects" NTG for PG&E is shown in Table 3 as 0.81. How was this calculated?

Response to Comment 8

This comment duplicates Comment 7.

<u>Comment 9</u>

We appreciate the evaluators' good write up of the NTG self report areas of potential bias (p. 37).

<u>Response to Comment 9</u>

We are glad that you noted our awareness of potential bias in the NTG self-report and would only point out that we did our best to minimize those sources of bias in our application of the self report approach by complying with the CPUC/ED and MEC's Guidelines for Estimating Net-to-Gross Ratios Using the Self-Report Approach, as described in Appendix 5.4. Net Savings Evaluation Methodology, Section 5.4.8.4 Supplement D.

Comment 10

Treating all RCx measures as a single High Impact Measure (HIM) is quite surprising. In this evaluation, there were many different measures comprising the population of projects, delivered through different channels, programs, and implementers. Furthermore, some projects included in the population are not well characterized as "retro-commissioning" and other projects were a mix of RCx measures and retrofit measures. While the HIM approach may have some value merely to account for the savings, we believe it is inappropriate to draw conclusions from the HIM analysis in terms of overall RCx measures and programs. Please explain if and why this was done.

Response to Comment 10

RCx is a HIM because all the RCx measures add up to greater than 1% of one or more savings parameters (kWh, kW, Therms). This evaluation focused on RCx as a measure and delivery method (a possible proxy for program strategy). The analysis approach is based on project-level sampling pooled across the RCx programs, excluding projects that had a large retrofit component. Samples were designed to obtain results calculated at the IOU level, across the programs that included RCx projects. We intend to draw conclusions from the RCx samples in accordance with the Energy Division Decision Framework for Updating IOU savings claims, distributed to the energy efficiency service list by Carmen Best on December 2, 2009. Furthermore, RCx is frequently a bundled measure in the utilities' own records and filings, which made it difficult to segregate when planning the HIM projects.

Comment 11

Quantum Energy Services & Technologies, Inc. appreciates the hard work that has gone into compiling this draft final report. We respectfully request all supporting files for the site-specific measurement and verification reports be made available prior to the issuance of the final report. We would like to receive the data in ample time to review and provide comments to the final report.

Response to Comment 11

Soon after the CPUC posted the draft final evaluation report on December 7, 2009, PG&E requested supporting files for all site-specific M&V reports prepared for projects in their service territory. Shortly thereafter, the evaluation team and CPUC provided these materials to PG&E. PG&E's RCx service providers should have been able to obtain these materials through their existing channels with the PG&E program managers.

Comment 12

PG&E has not been able to conduct a meaningful review at the site level of the projects. The site reports didn't contain enough information. We requested and received detailed site data (over 1GB) hand delivered 12/22 pm. It appears that this data is incomplete, as it does not include ex post savings analysis or calculations -- the key files are missing. For example, site report P00292, Table 5, shows three supporting files, but we only received two, the Msword Final (site) Report and the Evaluation Site Workbook (excel), but not the file PES Savings Analysis.xls.

Response to Comment 12

We acknowledge that the final report review period was short, but note that it was consistent with the tight timeframe for all phases of the evaluation work. In light of the difficulties that PG&E faced finding appropriate reviewers, the CPUC granted them a one-week extension to January 7, 2010 for providing comments. We received no additional comments from PG&E during the extension period.

We were puzzled by the example of missing data cited in this comment. We re-checked the files for the referenced project and found that the key files, including the ex post savings analysis and calculations mentioned in the Supporting Documents section of the M&V report, were indeed present. To the best of our knowledge, all key supporting documents identified in these reports were indeed provided to PG&E per their request.

Comment 13

We sent the data to our RCx consultant on 12/23 for arrival on 12/24. However, on 12/23 our consultant informed us that they must discontinue work as a result of the conflict of interest letter sent by ED to all consultants on 12/17. As a result of this withdrawal and the lack of key data files, PG&E has not been able to even begin a meaningful review of the detailed site data. PG&E made significant efforts to retain another qualified consultant, consistent with the CPUC's 12/17 letter and available during the Xmas/New Year's week. To date, we have not been able to do so and therefore no meaningful evaluation of the core data in this report -- the specific sample site findings -- has been possible.

Response to Comment 13

Refer to the response for Comment 12.

Comments 14 A & B

Pacific Gas and Electric Company®

Sandy Lawrie
Principal Regulatory Case Manager

Pacific Gas and Electric Company
77 Beale St., Mail Code B9A
P.O. Box 770000
San Francisco, CA 94177
Phone 415-973-2494

December 31, 2009

Ms. Zenaida Tapawan-Conway
Energy Division
California Public Utilities Commission
505 Van Ness Avenue,
San Francisco, CA 94102

Comments posted on www.energydataweb.com/cpuc

RE: *Draft Final Report 2006–08 Retro-Commissioning Impact Evaluation*

Dear Ms. Tapawan-Conway:

Pacific Gas and Electric Company (PG&E) respectfully submits the following questions and comments to the *Draft Final Report 2006–08 Retro-Commissioning Impact Evaluation.*

PG&E provides its comments in two sections: (1) an Executive Summary that provides a high-level discussion of the strengths and weaknesses of the report; and (2) PG&E's detailed comments to specific items in the report, presented in spreadsheet format. PG&E posted its detailed comments on www.energydataweb.com/cpuc and also attaches them as "Attachment A" hereto for your convenience.

A. EXECUTIVE SUMMARY OF COMMENTS

Comment 14A

1. Despite the significant efforts of the Evaluators and Energy Division, the comment period was insufficient to conduct a meaningful review of the Draft Final Report findings.

PG&E acknowledges the effort required to author a Draft Final Report such as this and appreciates the opportunity to submit these comments. PG&E also appreciates Energy Division's (ED) efforts to present evaluation findings through webinars and to respond quickly to PG&E's data requests that followed. Despite ED's and the evaluators' significant efforts in this regard, PG&E's analysis of this report was hindered by an abbreviated review period, incomplete data to support certain evaluation findings and by the loss of PG&E's evaluation consultant in the middle of its review.

In addition to a very tight review and comment schedule, the site reports did not contain enough information to support a meaningful site-level review of the projects. PG&E requested the detailed site data, which it received on the afternoon of 12/22. However, the data provided was incomplete, as it generally did not include the ex-post savings analysis and calculations. This is discussed in more detail in the attached comments.

Though incomplete in some respects, PG&E attempted to review the data provided. PG&E sent the data to its RCx consultant on December 23, 2009. On the same day, the consultant notified PG&E that it would discontinue work as a result of a conflict of interest assertion made by ED. PG&E made significant efforts to replace the consultant, but has not yet found another qualified, non-conflicted consultant who is immediately available during the year-end holidays.

To date, PG&E has not received the complete site-level data. As a result of incomplete data, as well as the issues referenced above, a meaningful and comprehensive site-level review was not possible.

Comment 14B **2. <u>The Evaluators did not account for all energy savings from short-lived measures</u>**

The evaluation's executive summary states that the gross statewide realization rate was 55% with "nearly half of savings reductions attributable to the measures not working any more." [p.3] Many of the RCx measures and savings claimed by PG&E were short-lived measures with EULs of three years. Given that evaluators were generally on-site between one and three years after measure installations, it is not surprising that many short-lived measures were no longer working at the time of evaluation. However, in those cases, the evaluators assigned zero energy savings to the sites, as if they never yielded any savings. While we appreciate the evaluation challenge of assigning savings to a currently, non-functioning measure, short-lived measures that yielded savings should not be given

2

zero credit. This problem is further compounded when the sample is expanded to the population.

B. CONCLUSION

Pacific Gas and Electric Company appreciates the opportunity to submit these questions and comments to this Draft Final Report. PG&E also appreciates the effort of Energy Division in attempting to respond quickly to PG&E's data requests and to manage a tight schedule for review and comment. As such, PG&E looks forward to the opportunity for further discussion of these topics with the evaluators and the Commission's Energy Division. Nevertheless, for the reasons stated above, PG&E feels that the review period was inadequate, and that certain information and data, critical to properly evaluating these studies, remains outstanding. As such, PG&E reserves the right to supplement these comments as necessary.

PG&E's comments have been posted on www.energydataweb.com/cpuc.

Respectfully Submitted,

By: /s/
SANDY LAWRIE

Principal Regulatory Case Manager
Pacific Gas and Electric Company
Mail code B9A
P.O. Box 770000
San Francisco, CA 94177
Telephone: (415) 973-2494
Facsimile: (415) 973-7131
E-Mail: slda@pge.com

.

Pacific Gas and Electric Company
Comments on Draft Final Report 2006–08 Retro-Commissioning Impact Evaluation (12/31/09)

Attachment A

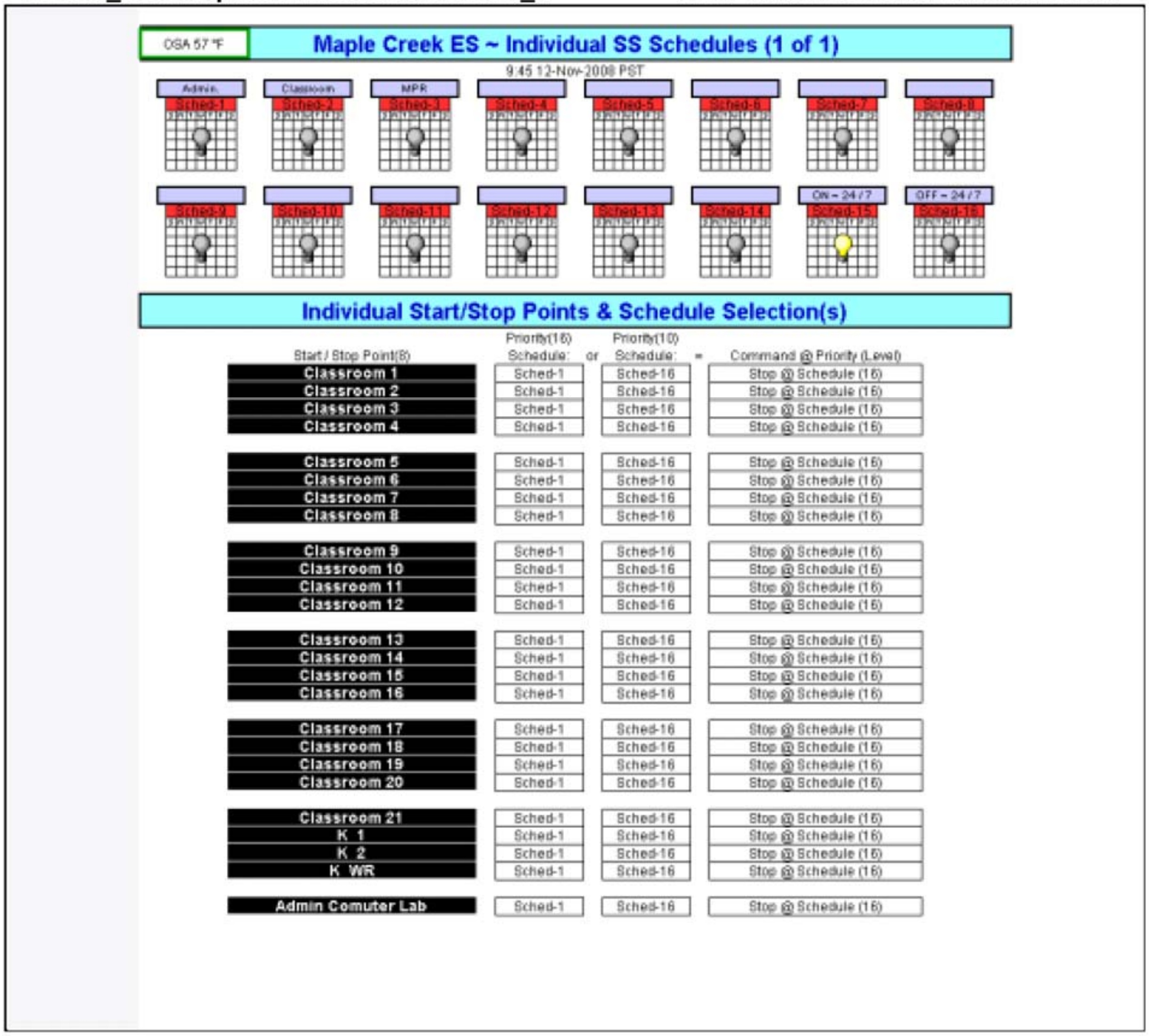
Subject:	Section/ Page:	Type (Question or Comment):	Comment or Question:
Review not completed	General	Comment	PG&E has not been able to conduct a meaningful review at the site level of the projects. The site reports didn't contain enough information. We requested and received detailed site data (over 1GB) hand delivered 12/22 pm. It appears that this data is incomplete, as it does not include ex post savings analysis or calculations -- the key files are missing. For example, site report P00292, Table 5, shows three supporting files, but we only received two, the Msword Final (site) Report and the Evaluation Site Workbook (excel), but not the file PES Savings Analysis.xls. We sent the data to our RCx consultant on 12/23 for arrival on 12/24. However, on 12/23 our consultant informed us that they must discontinue work as a result of the conflict of interest letter sent by ED to all consultants on 12/17. As a result of this withdrawal and the lack of key data files, PG&E has not been able to even begin a meaningful review of the detailed site data. PG&E made significant efforts to retain another qualified consultant, consistent with the CPUC's 12/17 letter and available during the Xmas/New Year's week. To date, we have not been able to do so and therefore no meaningful evaluation of the core data in this report -- the specific sample site findings -- has been possible.
HIM	General	Comment	Treating all RCx measures as a single High Impact Measure (HIM) is quite surprising. In this evaluation, there were many different measures comprising the population of projects, delivered through different channels, programs, and implementers. Furthermore, some projects included in the population are not well characterized as "retro-commissioning" and other projects were a mix of RCx measures and retrofit measures. While the HIM approach may have some value merely to account for the savings, we believe it is inappropriate to draw conclusions from the HIM analysis in terms of overall RCx measures and programs. Please explain if and why this was done.
NTG	p 36, Sec 3.5	Comment	We appreciate the evaluators' good write up of the NTG self report areas of potential bias (p. 37).
NTG	Table 3, p 3	Question	Were NTG values estimated at the project level and then weighted by savings type? We see in section 1.1.2 that the self report method was used, but how were the individual survey results expanded to the population? The "projects" NTG for PG&E is shown in Table 3 as 0.81. How was this calculated?
NTG	n/a	Question	For PG&E, the NTG findings range from 0.76 to 0.86 (p. 38) based on the self report methodology. However, in response to a question during the webinar, the evaluators were unable to cite a single RCx project performed in California outside of a utility program. This is consistent with the utilities' assessment of current California retro-commissioning practices. These findings are mutually exclusive. While some respondents may say they would have performed some portion of RCx work without program influence, given that none actually exist, isn't the appropriate NTG 1.0?

Pacific Gas and Electric Company
Comments on Draft Final Report 2006–08 Retro-Commissioning Impact Evaluation (12/31/09)

Attachment A

Evaluation	General	Question	Is it true that site savings estimates were based on current conditions at the time of site visit?
Evaluation	General	Question	How do the evaluators recommend conducting RCx evaluations in the future for short-lived measures, such as RCx scheduling measures? If the EUL is only three years and the evaluators are on site one to three years after installation, we would expect a significant portion of measures to have failed. Have the evaluators accounted for and/or quantified the effect of this issue on their findings?
P01770	Site Report P01770	Comment	Apparently, the evaluation policy estimated energy savings based on current conditions at the time of site visit. At this site, the PG&E recommended RCx measure (re-scheduling of the co-generation system) was implemented in January 2006 with a claimed three-year EUL. The facility manager of the site (contact information available) confirmed that the absorption chiller and co-gen system subsequently failed in March 2009. Therefore the RCx measure actually exceeded its three year EUL, and real energy savings were achieved. However, since the system failed prior to on-site EM&V, the evaluators assigned zero savings to this site. Given the foregoing, we believe PG&E should be given appropriate credit for the claimed energy savings.
Site report PO 1137	Site report PO 1137	Comment	Section 2.6, Evaluation Measure ID:1 states that "The 23 unit ventilators with cooling and heating coils were equipped with a single control that did not allow individual temperature reset and scheduling." Enovity agrees that there is no ability to do re-set and there were no savings claimed or calculated for re-set. Contrary to the assertion of the evaluation however, the units can be scheduled individually, permitting unoccupied classrooms to be unconditioned, thereby resulting in energy savings. A screen shot is attached that shows the individual schedules.
Site report PO 1137	Site report PO 1137	Comment	Section 2.6, Evaluation Measure ID: 1 states, "The classrooms are not used during vacations and summer break periods; hence the unit ventilators remained off during that period." This is incorrect and conflicts with information given to Enovity during the investigation phase. The savings claimed by Enovity were for reducing operating hours of the units' ventilators during vacations and summer breaks. The school informed Enovity that some, but not all, classrooms are used during vacations and summer breaks; the existing baseline condition was one master schedule that turned all the ventilators on regardless of whether the classrooms were being used. Enovity therefore proposed controls that would only bring on the classrooms that were in use for a given day during vacations and summer breaks. During verification Enovity confirmed that all the controls were in place and functional. Additional savings should be credited to account for this information. To verify this, ECM data needs to be collected for the unit ventilators during vacation/summer break.

/CUSD_BAS/MapleCreekES/Sched/Sched_1 12:44 12-Nov-2008 America/New York



Start / Stop Point(8)	Priority(16) Schedule:	or	Priority(10) Schedule:	=	Command @ Priority (Level)
Classroom 1	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 2	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 3	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 4	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 5	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 6	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 7	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 8	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 9	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 10	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 11	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 12	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 13	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 14	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 15	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 16	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 17	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 18	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 19	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 20	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 21	Sched-1		Sched-16		Stop @ Schedule (16)
K 1	Sched-1		Sched-16		Stop @ Schedule (16)
K 2	Sched-1		Sched-16		Stop @ Schedule (16)
K WR	Sched-1		Sched-16		Stop @ Schedule (16)
Admin Comuter Lab	Sched-1		Sched-16		Stop @ Schedule (16)

Response to Comment 14A

Refer to the response for Comment 12.

Response to Comment 14B

Refer to the response for Comment 4.

Comments 15 A, B, & C

SAN DIEGO GAS & ELECTRIC COMPANY
SOUTHERN CALIFORNIA GAS COMPANY

Comments on 2006–08 Retro-Commissioning Impact Evaluation

Overall Comments: Comment 15A

The report discusses how a big percentage of the measures studied were not operational, at page 24, the report states, "Critically, nearly half of these savings-reducing reasons were instances where the RCx measure was no longer operational." However, the report does not discuss specific reasons why they were not operational and may have unduly reduced the savings. Several facilities were down for maintenance which would then have the installed equipment not fully operational. Therefore, the study timeframe may have been when the equipment would not be fully operational due to other conditions at the facilities not necessarily related to the equipment itself.

SDG&E UC/CSU Project Specific Comments Comment 15B

1) SBW ID: P00317, IOU ID: 2
The specific MBCx project evaluated is EEM # 1, Re-Commission Laboratory and Fume Hood Controls

a) *Issue: The original savings calculations for the project utilized a discharge air temperature of 55F for the air handlers while the reviewer's calculations used a discharge air temperature of 64.3F, resulting in a reduction in gas savings associated with reheat. Page 24 of the site report for SBW ID P00317 states that the average discharge air temperature of 63.4F was used to correct the baseline model's assumption where it had been set to 55, however a summary table of discharge air temperature supplied from the BMS trended from May 6, 2008 to September 17 2009 states an average discharge air temperature of 59F. It is recommended that the reviewer revise the calculations using the correct trended average discharge air temperature which appears to be 59F, not 63.4F.*

b) *Issue: Page 22 of the site report for SBW ID P00317 states that, "Slight corrections were also made to the model to correct assumptions regarding the central plant kW per ton rating used", but no initial values or altered values are provided. The report also lacks any specific justification for the revision of the central plant efficiency.*

c) *Issue: The energy savings estimation method used assumes a fixed 5,135 CFM air flow reduction, as provided by Phoenix Controls. The CFM reduction will vary depending upon sash position, and affect energy savings. A fixed discharge air temperature (DAT) was used for the energy savings calculation, based on short term monitoring of several air handlers. All 8 air handlers DAT are automatically reset and vary throughout the day. To more completely examine the heating, cooling, and fan power energy changes provided by this measure building power, chilled water, and hot water consumption pre and post should be compared. UCSD records this information monthly.*

2) SBW ID: P00319, IOU ID: 4 Comment 15C

The specific MBCx project evaluated at is EEM #1, VSDs on CT fans, Optimize Chiller/Boiler Sequencing.

a) *Issue: It appears as though the reviewer did not account for new boiler controls implemented as part of the overall MBCx RCx project. The HW system measures are mentioned on page 18 of the site report for SBW ID P00319 under section 3.3 "Evaluation Algorithms- Energy Savings", however section 7.2. "Key Findings" of the report does not include any information about the HW system measure. The associated gas savings for the MBCx project were reduced to zero therms yet there is no post implementation trending data, calculations, or supporting justification for reducing the gas savings. It is recommended that the reviewer verify the original gas savings calculations or eliminate the ex ante gas savings associated with the boiler controls as the resulting therm realization rate of zero is inaccurate.*

Response to Comment 15A

When circumstances were encountered in the EM&V investigation which temporarily affected the savings, a good faith effort was made to make accommodation for the situation and to give the site the benefit of the doubt. For example,

- One measure was a setback mode for a swimming pool circulation pump speed. The EM&V engineer found the setback mode was not enabled per the measure description. The site contact explained it must have been accidently disabled recently in a maintenance operation. The pool area was on a separate electrical meter which allowed the EM&V engineer to verify the recent control sequence change and to not penalize them for the inadvertent and temporary change.
- A site with refrigeration measures had a major equipment failure in the refrigeration system and was undergoing lengthy repairs when contacted by the EM&V engineer and was not likely to be completed in time to meet the EM&V completion schedule. We dropped the site from the sample and did not penalize the utility.
- A number of projects in the sample had chiller and boiler control sequences which locked out the equipment at certain outside air temperatures. Given the compressed timeline for site evaluation, we weren't able to observe or trend actual equipment performance but the measure was approved if the control sequence was enabled in the energy management system.

The individual EM&V reports supply sufficient background on non-operational measures to show that these measures were not achieving savings on a more permanent basis than a maintenance shutdown.

Response to Comment 15B

Overall, we found insufficient grounds for modifying the evaluation analysis. Specific responses to the issues raised are as follows:

Issue (a) Response: The analysis is correct as it stands. The data which supports the 59F average DAT is NOT for P00317 (EBU-2) but is for P00316 (CMMW).

Issue (b) Response: The evaluation team put considerable effort into trying to estimate the efficiencies of the Central Plant at the campus. The central plant is sufficiently complex with a mix of electric chillers, steam-driven chillers, natural gas boilers, a large cogeneration system, heat recovery from the cogeneration and a chilled water thermal energy storage system. Data was collected on the major systems. Facility personnel indicated that they do not have a study, data or documentation on the Central Plant performance.

The RCx Provider had used a Central Plant heating efficiency of 75% for EBU-2, but used 80% for CMMW. Our evaluation report used 80% for both facilities and gives all the heat savings credit to the natural gas boilers.

The RCx Provider had used a Central Plant cooling efficiency of 0.90 kW/ton for EBU-2, but used 1.17 kW/ton for CMMW. Our evaluation report used identical kW/ton input values for both facilities based on the chiller efficiencies, pump sizes and configuration, and qualitative information on equipment sequencing. It was estimated that when the Central Plant operated on all electric chillers, the efficiency is

0.998 kW/ton. The average annual efficiency (electric chillers and steam-driven chillers) was estimated to be 0.62 kW/ton.

Issue (c) Response: A fixed DAT was used in the evaluation but the value is an average that was a derived value based on actual trended data for about a 16-month period. Chilled Water and Hot Water data on the building was not available. Furthermore, electric data for this building was collected from the BMS, but data trending stopped after Dec 2008.

Response to Comment 15C

Overall, we found insufficient grounds for modifying the evaluation analysis. The RCx Provider provided no basis of savings relative to sequencing of the boilers. In fact, on page 14 of their Final MBCx Report they state, "No sequence description is provided for the boiler plant as monitoring is the primary function provided for this system by the APOGEE system." Facility operators indicated that after the RCx project they are able to see boiler operating information on the APOGEE computer screens. Although being able to see operation by plant personnel will likely result in maintaining efficient operation of the equipment, there is no basis for calculating natural gas savings for this monitoring upgrade.

According to page 12 of the Final MBCx Report, "The boiler plant was significantly upgraded in 1996 including two new 500 HP boilers. The new boilers, B-1 and B-2, are three-pass, Mohawk Model 2506 by Superior Boilers Works and the third boiler, B-3, is an older model Clever Brooks CB-500. All three boilers have forced draft, low NOx, flue gas recirculation burners by Industrial Combustion, and a rated fuel-to-steam efficiency of 80%. The boiler plant upgrade also included heat recovery systems, condensate and feed water pumps, deaerator, chemical treatment system and controls. The heat recovery systems include flue gas-to-feed water heat recovery economizers for boilers B-1 and B-2, and a blow down-to-make-up water heat exchanger that is connected to all three boilers."

No pre-installation or post-installation data was included with the Final MBCx Report to support the savings (Appendix C). It does not appear that the baseline boiler efficiency calculations were based on detailed trended data (no monitoring system in the pre-installation phase), while the post-installation analysis was. Due to the lack of pre-installation data, there was no real apples-to-apples comparison, and we concluded that there was no basis to prove that the savings were achieved.

Comments 16 A& B

SCE COMMENTS ON THE DRAFT FINAL REPORT 2006–08 RETRO-COMMISSIONING IMPACT EVALUATION — EXPECTED USEFUL LIFE

Comment 16A *1. The EUL Study was an interesting exercise, but it did not generate usable estimates, only recommendations for further research. We agree with this recommendation.*

One intended purpose of the study was to update the EUL for RCx programs: "the EULs that result from the study—whether at the RCx-project or measure level—can be used directly by the CPUC in assessing the cost-effectiveness of RCx projects implemented during the 2006–08 program cycle" (p. 44). However, "the EUL study was not designed to comply with the California Energy Efficiency Evaluation Protocols for EUL studies" (p. 44), so it should not serve as the basis for changing parameters.

Furthermore, the report states: "Given the uncertainty around the EUL estimated by this study, it does not appear that there are solid grounds for changing EULs claimed by IOUs for RCx measures...To confirm the EUL value developed by the current study on a relatively small sample, a much larger study will be required" (p. 51).

Therefore, there are no valid reasons for the CPUC to use these values in assessing the cost-effectiveness of RCx projects implemented during the 2006–08 program cycle

Comment 16B *2. Two other intended purposes of the study were partially met: "enhancing understanding and serving as the basis for future investigations" (p. 44). We agree with the report's conclusion: "a much larger study will be required, and we recommend that such research be pursued aggressively and on a large scale" (p. 51). In addition, we would expect next time that the CPUC would conduct such research in compliance with the California Energy Efficiency Evaluation Protocols for EUL studies.*

Response to Comment 16 A

In line with this comment, the CPUC has already decided that the results of the EUL study will not be applied to the cost-effectiveness calculations for the 2006-08 program cycle.

Response to Comment 16 B

This comment concurs with our conclusion that additional EUL study is desirable, and is duly noted.

Comment 17

One drawback to having the CPUC conduct the impact evaluation is the lack of information available to explain why a particular phenomenon is observed. In the case of the disparity in realization rates across IOUs, the impact evaluation study looked at the differences in measures, customers, etc., with limited success. However, the SCE M&E staff and RCx program managers understand that in-house collaboration was the key reason for the relatively good results reported in this study for the SCE RCx program. SCE's process evaluation for the 2006-08 RCx program recommended several modifications that were adopted by the program manager during the course of the program.

Therefore, the CPUC should reconsider its decision to take the impact evaluations away from the organizations that best understand the programs and the environments in which they are implemented.

Response to Comment 17

The first part of this comment echoes the contention voiced in our gross savings recommendations (Section 2.7.3, Recommendation 2), that better coordination between process and impact evaluations might provide useful information about the reasons for success (or lack thereof) for particular programs and programmatic approaches.

The second part of the comment concerns future CPUC evaluation policy, and is duly noted.

Comment 18

TURN Comments on the 2006-08 Commercial Retro-commissioning Draft Evaluation Report December 31, 2009

TURN appreciates this opportunity to comment on the Commercial Retro Commissioning draft Energy Division EM&V report (Draft Final Report 2006-2008 Retro-Commissioning Impact Evaluation). The Commercial Retro Commissioning Program includes 225 projects from more than two dozen programs offered by the four IOUs. The program is considered a high impact measure (HIM). TURN has conducted a limited review of the documents and has significant concerns about the M&V findings on program performance. TURN's comments are organized into the following three sections:

- Comparison of the IOUs' ex ante reported accomplishments and ED's ex post measured and verified savings and net benefits.
- Changes to how the EM&V was conducted.
- Savings toward goals and net benefits, freezing ex ante data for the 2010-2012 portfolio cycle, and the Commission's Long Term Energy Efficiency Strategic Plan.

The draft EM&V report points to the following shortcomings in the Commercial Retro Commissioning program: (1) savings are dramatically lower than the ex ante values predicted, (2) reasons identified to explain ex ante and ex post discrepancy point to a number of failures on the part of program.

The gross impact analysis found an overall realization rate of 0.55 on a BTU basis. In other words, measured and verified ex post program energy savings were only slightly more than half of what the IOUs had claimed. Although levels of free ridership were found to be lower than for some programs (NTG ≥0.8), savings shortfalls of this magnitude threaten the cost effectiveness not only of the program, but of the portfolio as a whole.

I. Comparison of the IOUs' ex ante reported accomplishments and ED's ex post measured and verified savings and net benefits.

Seventy-five per cent of the reasons identified in this report for why the ex ante and ex post values diverged corresponded to reduced savings. While an extensive list of reasons was identified, "nearly half of these savings-reducing reasons were instances where the

Retrocommissioning measure was no longer operational."[10] The report further notes that human factors are more often responsible for premature measure failure than technical issues. More specifically, both a failure to consider occupant comfort issues in the context of retro commissioning, and inadequate attention to maintenance were identified as causes of failure.[11] These matters are (or should be) the focus of the program. Retro commissioning concerns itself with evaluating and improving both operation and maintenance. When nearly one-quarter (22%) of the interventions result in failure because of inadequate maintenance one has to ask whether someone else should be running the program?

The report distinguishes two types of failure: "actions the customer took (such as overriding a recommended fan schedule) or actions the program was responsible for (such as applying too high a chiller efficiency in the savings calculations)."[12] While this distinction may be analytically useful, it is important to remember that (within reason) it is the responsibility of the program to anticipate actions such as in the example above and strive to reduce their probability. The interaction of people and building systems is, after all, the subject of this program.[13] Other common reasons for differences in ex ante reported accomplishments and ex post measured and verified savings and net benefits included discrepancies between program calculation assumptions and actual conditions, changes in building operation, and measures being only partially implemented.

Table 1 Summary of Commercial Retrocommissioning findings[14]

---

[10] *Draft: 2006-08 Retro-Commissioning Impact Evaluation*, p. 24.

[11] Ibid., p. 50.

[12] Ibid. p. 24.

[13] "Retro-commissioning is the systematic evaluation of major energy-using systems and equipment in existing buildings to find ways to ensure they operate both efficiently and in accordance with customer needs. … All of them assist participants via investigations (audits or performance assessments) that identify opportunities for improving the operations or maintenance of existing assets, and provide incentives to reduce the barriers to adopting the program's recommendations. Many provide other forms of educational and training support for the participants. In some cases, these programs also identify opportunities for conventional retrofit technologies, such as high efficiency chillers or new control systems, and either provide incentives directly to motivate adoption or refer customers to other efficiency programs." (*Draft: 2006-08 Retro-Commissioning Impact Evaluation* p. 63).

[14] Excerpt from *Application 4, TURN Protest Utility Advice Letter, December 14, 2009*

APPLICATION 4. 2006-08 EM&V Results

Date Report Posted	Program Name	Net Realiz. Rates	Findings - Discussion of Gross & Net Realization Rates, NTG, etc.	Explanations offered in report for diff ex post vs ex ante
7-Dec-09	Commercial Retrocommissioning	0.55, 0.45, 0.55, 1.23, 0.21, 0.93	The gross impact analysis found an overall realization rate of 0.55 when measured in MMBTU. PG&E projects as a whole had relatively low realization rates for both electric and gas energy savings (0.45 and 0.50), compared to SDG&E, which had realization rates that varied between gas and electric savings; for example, 1.23 for kWh and 0.21 for therms. SCG's single realization rate for therms was 0.93.	Across the 50 projects in the gross sample, we determined 83 significant reasons for differences between the claimed and evaluated gross savings, with a relatively even split between customer-driven and program-driven reasons. Over 75% of these reasons worked to reduce savings, a percentage that was fairly uniform across all four IOUs. Critically, nearly half of these savings-reducing reasons were instances where the Retrocommissioning measure was no longer operational.

II. Changes to how EM&V was conducted.

TURN welcomes the change in this program cycle toward greater oversight of the EM&V process by ED. The Commission recognizes the need for "a clearer separation between 'those who do' (the program administrators and implementers) and 'those who evaluate' the program performance."[15] As noted earlier, there is a need to continue to improve on the earlier adopted independent EM&V structure. While the present set of reports reflect movement in this direction, EM&V contractors can and must strive toward even greater independence from the IOUs when evaluating programs.

III. Savings toward goals and net benefits, freezing ex ante data for the 2010-2012 portfolio cycle, and the Commission's Long Term Energy Efficiency Strategic Plan.

A. Savings toward goals and net benefits.

[15] D.05-01-055, pp.10-11 (cited in the December 8, 2009 Comments of The Utility Reform Network on 2009-2011 Portfolio EM&V, p. 2).

The overall findings in this evaluation report were that gross savings were significantly below those estimated by the IOUs, although the variation among the IOUs was large. PG&E's program, which yielded gross realization rates of 0.31 (kW), 0.45 (kWh), and 0.50 (therms) was by far the worst. The other IOUs had significantly smaller programs and higher realization rates. SCE: 2.07 (kW), 0.94 (kWh); SDG&E 2.60 (kW), 1.23 (kWh), and 0.21 (therms). SCG's gross realization rate was 0.93 (therms).

B. Freezing ex ante data for the 2010-2012 portfolio cycle

As noted in TURN's recent Protest,[16] Decision 09-09-047 requires the IOUs to use "ex ante values … that are based upon the best available information at the time the 2010-2012 activity is starting."[17] The present series of EM&V reports that have been published in draft form in December,[18] include updated data, at times significantly different from the ex ante data. To execute the proposed set of programs for the next program cycle with the discredited ex ante data when an updated and field-verified set is now available would be a grave mistake.

C. The Commission's Long Term Energy Efficiency Strategic Plan

The Strategic Plan calls for reducing overall energy consumption in commercial buildings (zero net energy by 2030 for all new and a substantial proportion of existing buildings).[19] While energy efficiency savings is a key output of the EM&V process, EE savings are, in and of themselves, not equivalent to reductions in total energy consumption. The IOUs' 2006-08 EE portfolio has resulted in EE savings[20] while the state's total energy consumption has continued to rise.[21] Absolute reductions in energy consumption, if they are to be brought about

---

[16] *TURN Protest to Utility Advice Letters Implementing 2010-2012 Energy Efficiency Portfolio Budgets And Other Directives Pursuant to D.09-09-047.*

[17] Ordering Paragraph 48 of D.09-09-047, and quoted in TURN's Protest of December 14th, 2009, p.2.

[18] And presumably available to the IOUs in draft form prior to December.

[19] *California Long Term Energy Efficiency Strategic Plan*, p. 30.

[20] Albeit at a much lower level than reported by the IOUs. See Commission Resolution E-4272 October 15, 2009 adopting Energy Division's 2006-2008 Interim Verification Report.

[21] For the trend in statewide natural gas consumption over the program period (an increase of 6.5% between 2005 and 2007, the latest year for which data are available) see http://tonto.eia.doe.gov/dnav/ng/ng_cons_sum_a_EPG0_VC0_mmcf_a.htm; for statewide electricity consumption (an increase of 1.7% between 2005 and 2008) see

through utility EE, must exceed the sum of the forces that presently encourage increased energy consumption. If EE is to play a role in achieving the goals of the Strategic Plan then the programs that make up the current and future program portfolio must deliver results that exceed IOU forecasts rather than lag behind them. This program, along with most of the others which have now been evaluated, have fallen short of their goals.

IV. Key Measures and the importance of human factors

As noted above, the majority of 'measures' in the Commercial Retro Commissioning program involve the systematic evaluation of energy using systems and the identification of improvements to operation and maintenance that can be expected to yield efficiency improvements and either enhance or maintain customer satisfaction. This can be as much a social as a technical undertaking. While the program goals are not in question, the means of achieving them, the strategies and qualifications of those charged with executing the program are. While the personnel in charge of the facilities participating in the Retro Commissioning program can be expected to know the operations of their buildings intimately, it is incumbent upon the program staff to motivate buy-in to the principles and goals of the program. The nature of this kind of intervention requires collaboration and a commitment from both parties to follow up and follow through. The results of the EM&V report under review suggest this is not happening to the degree anticipated by the IOUs or expected by the author's of reports such as California's Long Term Energy Efficiency Strategic Plan. A 55% realization rate and the analysis of the failures which explain it suggest the need for a reprioritization around what the report calls human factors.

Table 2 List of Measures in Commercial Retrocommissioning program[22]

Improve boiler performance
Improve building warmup / cooldown
Improve control strategies - general
Improve damper performance

http://www.eia.doe.gov/cneaf/electricity/epm/epm_sum.html. Per capita consumption of natural gas has increased 5% over this period, while per capita electricity consumption has fallen by 1%, see http://quickfacts.census.gov/qfd/states/06000lk.html

[22] Ibid,. Table 23.

Improve maintenance practices - general
Improve outside air use
Improve scheduling
Improve sensor performance
Improve sequence of operation
Improve setpoints
Improve temperature / pressure reset schedule
Improve valve performance
Install / replace variable speed drive - HVAC air handler
Install lighting occupancy sensors
Install miscellaneous efficiency improvement
Install VFD

Response to Comment 18

These comments are duly noted.

Comment 19

enovity inc

MEMORANDUM

To:	Matthew Smizer, PG&E
Cc:	Jonathan Soper, Principal, Enovity Henry Summers, Project Manager, Enovity Rachel Christenson, Program Manager, Enovity
Date:	December 30, 2009
Subject:	Comments to CPUC Final Evaluation Report, P00136
From:	David Chan, Project Engineer, Enovity

This memo is in response to CPUC's Final M&V report for P00136, following a retro-commissioning project under Enovity's PG&E third party program, Monitoring Based Persistent Commissioning (MBPCx). There were a total of 19 measures that were implemented at this site. The total electricity saving for the 19 measures implemented were estimated to be 3,156,300 kWh per year and the natural gas savings was estimated to be 107,000 therms per year. CPUC final evaluation report concluded that the realized savings for electricity is 81%, realized demand saving is over 200% and realized natural gas savings is 12%.

<u>It is Enovity's recommendation that CPUC re-evaluate the calculations and resulting energy savings for EEM 14 Supply Air Reset in the post installation CPUC Retrocommissioning Evaluation Report.</u>

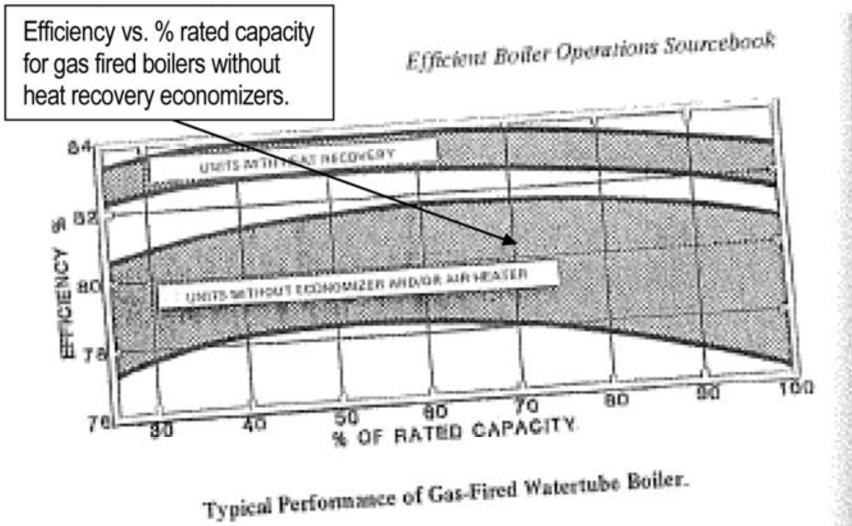
Enovity believes that the natural gas savings has been underestimated in CPUC's evaluation report. The majority of the gas savings comes from EEM 14, "Supply air reset & space temperature set back" after hours for the lab's AHUs which has an estimated natural gas savings of 79,200 therms. The lab air handlers are constant volume 100% outside air units that operate 24/7. These air handlers feed constant volume boxes with hot water reheat coils in all zones. The total cfm for these air handlers is about 400,000 cfm. Heating is provided by five natural gas force draft boilers. The gas savings from this measure comes from the reduction in reheat energy as the zone temperature gets set back at night (from 6 pm to 6 am) Monday through Friday and all day on weekends. The setback is typically 10 to 12 deg F in winter (unoccupied zone setpoint = 60 deg F) depending on occupied setpoint which is adjustable by tenants from 68 to 78 deg F. A reduced space temperature setpoint at night during periods of low outside air temperature will result in a significant gas savings as the re-heat coils will re-heat the air much less to maintain the lower space temperature setpoint.

CPUC's M&V evaluation report concluded that "there were significant natural gas reheat savings". In addition, the CPUC report claims that as a result of the drop in heating load, there is a much lower boiler efficiency and the drop in boiler efficiency is the primary reason for the low realization rate for this measure. Enovity agrees that boiler part load efficiency was not analyzed in detail as part of the post installation measurement and verification, however real

100 Montgomery Street, Suite 600, San Francisco, CA 94104-4331 (415) 974-0399 (f) (415) 974-0390 (v)

boiler efficiency does not drop off to such a significant amount during part load to cause such a drastic reduction in energy savings that is claimed in the CPUC report. Figure 1 shows a typical performance curve of a gas fired watertube boiler as a function of part load; this curve was taken from Efficient Boiler Operation Sourcebook, 4th Edition, F. William Payne and Richard E. Thompson. This chart shows that even as the load drops below 30%, the efficiency of the boiler is still between 77% to 80%. Even when a boiler is cycling at very low loads, Enovity's experience is that cycling losses account for about a 5 to 10% reduction in system efficiency. For the reasons stated above, Enovity believes that the realization rate for this measure should be significantly higher than 12%.

Figure 1: Typical Performance of Gas Fire Watertube Boiler



Typical Performance of Gas-Fired Watertube Boiler.

Enovity have re-evaluated this measure and reviewed trend data. The data re-evaluated for this measure includes the number of boilers operating. Enovity compared the number of boilers in operation for 2007 and 2009. Enovity selected the month of February for this comparison. Although this one month does not represent the entire year, it was chosen to show that this measure has a significant impact in heating energy reduction in this facility. The result of Enovity's re-evaluation shows that 70% of the time three boilers where operating during the month of February in 2007, compared to 17% during February 2009. In addition, during 2009 this facility required only one boiler to meet its heating demand 30% of the time, while in 2007 the facility ran only one boiler 2% of the time (See histogram in Figure 2). The five boilers in this facility are staged on and off based on heating demand and fewer boilers running suggested that the facility heating demand have reduced. The average heating degree days for February 2007 and 2009 is 13 and 14 respectively; this suggested that the heating reduction is from temperature reset and set back controls and is not weather influenced.

Figure 2: Number of Boilers Operating

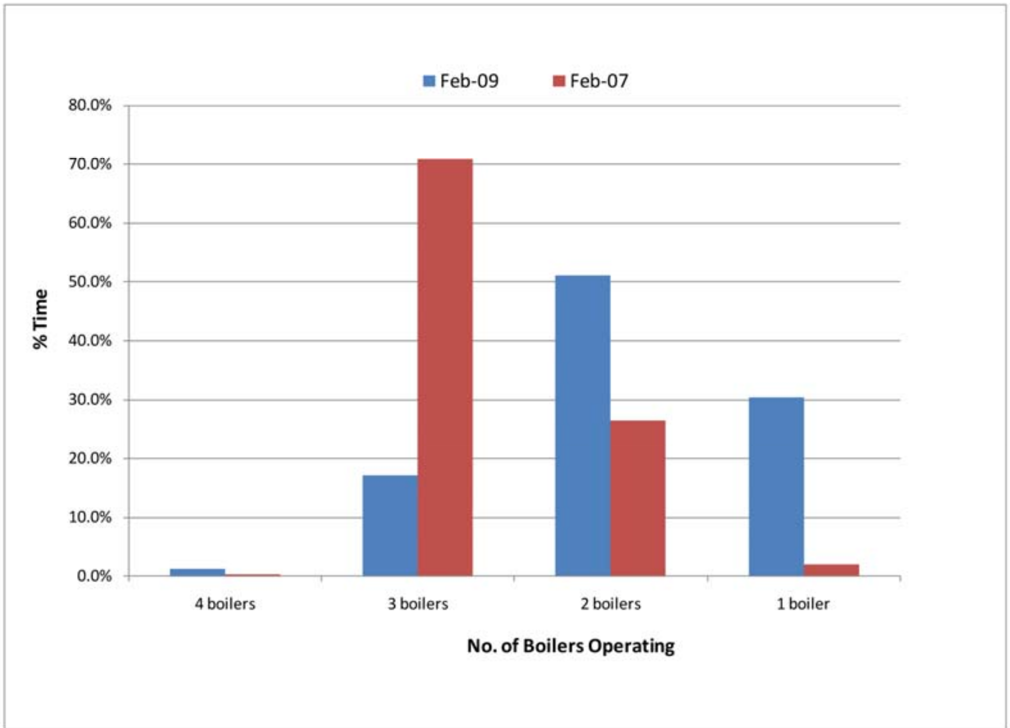


Figure 3 and Figure 4 shows the number of boilers running during a typical weekday in February in 2007 and 2009. Figure 3 shows that two boilers are running continuously and a third boiler stages on as needed. Figure 4 shows data from 2009 where only one boiler is needed to satisfied the heating load during the setback period, (between 6pm to 6am).

The original energy savings claimed for this measure was 79,200 gross therms. Enovity recalculated the energy savings for this measure using a spread sheet bin analysis, 5 degree bins were used. This analysis compares the difference in heating load during unoccupied hours and assumes that the heat load during occupied hours remains the same in both before and after scenarios. Using the assumption that the heating load has been reduced by a conservative 20% during unoccupied hours yields, the measure results in a total gross natural gas savings of 74,000 therms per year, a reduction of 5,200 therms (roughly 6.5%) when the boiler efficiency reduction due to lower load have been taken into account. Note, the trend data does not show significant boiler cycling. There is a significant difference between the energy savings that were calculated with the additional reduction of the boiler efficiency losses and the 12% realization rate that is being reported in the CPUC Retrocommissioning Evaluation Report.

Figure 3: Number of Boilers Running (2/26/07)

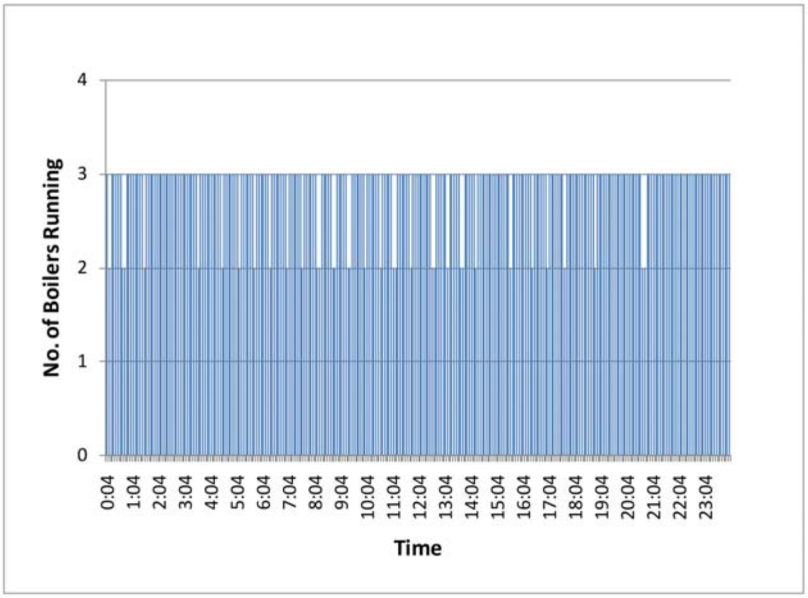
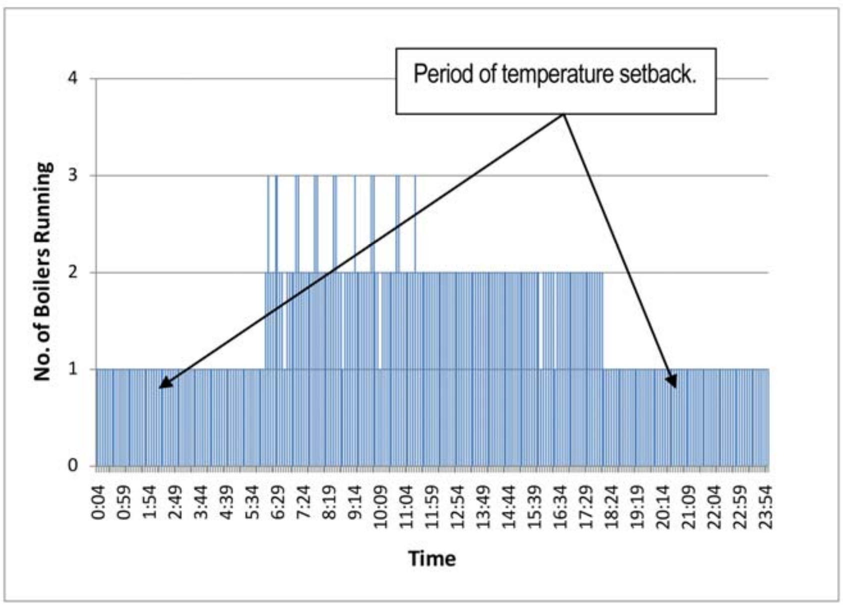


Figure 4: Number of Boilers Running (2/23/09)



Response to Comment 19

We carefully considered the reviewer's critique of the evaluation methodology for calculating boiler efficiency. After conferring with other senior engineers, including experts in the eQUEST algorithms and curves used in the analysis, we determined that the reviewer's assertion is correct. We corrected the algorithms to reflect much smaller reductions in boiler efficiency at part loads. This increased the natural gas realization rate for this project from 12% to 55%, and increased the overall PG&E gas realization rate by several percentage points.

<u>Comment 20</u>

Section 2.6, Evaluation Measure ID: 1 states, "The classrooms are not used during vacations and summer break periods; hence the unit ventilators remained off during that period." This is incorrect and conflicts with information given to Enovity during the investigation phase. The savings claimed by Enovity were for reducing operating hours of the units' ventilators during vacations and summer breaks. The school informed Enovity that some, but not all, classrooms are used during vacations and summer breaks; the existing baseline condition was one master schedule that turned all the ventilators on regardless of whether the classrooms were being used. Enovity therefore proposed controls that would only bring on the classrooms that were in use for a given day during vacations and summer breaks During verification Enovity confirmed that all the controls were in place and functional. Additional savings should be credited to account for this information. To verify this, ECM data needs to be collected for the unit ventilators during vacation/summer break.

<u>Response to Comment 20</u>

Our analyst for this site used the latest information available at the time of the evaluation from a knowledgeable site contact. The site contact is the energy manager for the school district. He is told when the school needs extra class rooms during the summer/winter break periods, and it was his information that led us to conclude that the controls on the unit ventilators have not yielded savings over the latest school year (2008-2009).

Comment 21

/CUSD_BAS/MapleCreekES/Sched/Sched_1 12:44 12-Nov-2008 America/New York

Start / Stop Point(8)	Priority(16) Schedule:	or	Priority(10) Schedule:	=	Command @ Priority (Level)
Classroom 1	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 2	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 3	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 4	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 5	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 6	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 7	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 8	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 9	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 10	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 11	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 12	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 13	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 14	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 15	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 16	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 17	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 18	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 19	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 20	Sched-1		Sched-16		Stop @ Schedule (16)
Classroom 21	Sched-1		Sched-16		Stop @ Schedule (16)
K 1	Sched-1		Sched-16		Stop @ Schedule (16)
K 2	Sched-1		Sched-16		Stop @ Schedule (16)
K WR	Sched-1		Sched-16		Stop @ Schedule (16)
Admin Comuter Lab	Sched-1		Sched-16		Stop @ Schedule (16)

Response to Comment 21

We agree with the comment's assertion that the unit ventilators (UVs) can indeed be individually controlled. Our analyst never claimed that they could not be controlled independently. Our finding, however, was based on information provided by the site contact (school district energy manager) that individual controls on the UVs did not benefit them much because the classrooms were fully occupied.

Comment 22

Apparently, the evaluation policy estimated energy savings based on current conditions at the time of site visit. At this site, the PG&E recommended RCx measure (re-scheduling of the co-generation system) was implemented in January 2006 with a claimed three-year EUL. The facility manager of the site (contact information available) confirmed that the absorption chiller and co-gen system subsequently failed in March 2009. Therefore the RCx measure actually exceeded its three year EUL, and real energy savings were achieved. However, since the system failed prior to on-site EM&V, the evaluators assigned zero savings to this site. Given the foregoing, we believe PG&E should be given appropriate credit for the claimed energy savings.

Response to Comment 22

Refer to the response for Comment 4.

Comments 23 A, B, C, D, E, F, G, H, & I

2006-08 Retro-Commissioning Impact Evaluation Report
Partnership Comments – 12/22/09

The 2006-2008 Retro-Commissioning Impact Evaluation Report was reviewed specifically for analysis of the UC/CSU/IOU Partnership MBCx projects. The comments contained herein, contributed by the Partnership team, highlight areas of concern which should be addressed to establish validity of the final evaluated savings results as well as overall program evaluation rates. General comments and issues associated with the overall analysis are provided. Additionally, we discovered some inconsistencies and discrepancies on a project by project basis, which follow the general comments.

GENERAL COMMENTS:

Comment 23A

1. Individual Measure Approach. The evaluator's approach to pull select measures from overall MBCx projects at some campuses is not an accurate analysis of MBCx project impact. The MBCx program is by nature a more holistic approach to energy savings than a specific measure by measure program. Several MBCx projects analyzed for the Retro-Commissioning Impact Evaluation did not lend themselves to individual measure savings analysis; instead savings should be evaluated on a whole facility basis. Attempts to disaggregate individual measures from MBCx projects lead to inaccurate savings measurement and by extension, inaccurate realization rates.

Comment 23B

2. Documented Savings Database. The report calculated realization rates for each project by dividing ex post 1st year gross savings (report verified savings) by ex ante savings which were taken to be IOU claimed savings. The ex ante savings were crosschecked with actual IOU claimed savings through the program database and several inconsistencies were discovered. It is recommended that ex ante savings in the report be revised to match actual IOU claimed savings to ensure that proper realization rates and savings values are being reported. A summary of the discrepancies for each project are indicated in red in Table 1 at the end of these comments.

Comment 23C

3. Appendix Results vs. Main Report. For many evaluated MBCx projects, savings values in the individual site reports contained in the report Appendix do not match values reported in main report. No documentation could be found explain the discrepancy between values found in the report and values pulled from the appendix calculations.

Specific MBCx PROJECT SPECIFIC COMMENTS:

PG&E UC/CSU Projects:

Comment 23D

1) SBW ID: P00281, IOU ID: 2K0700071 [CSU San Jose, MBCx - Central Chilled Water Plant]

The specific MBCx project evaluated at CSU San Jose was the improvement of control strategies for chillers and chilled water system.

a) Issue: Reviewer used proposed savings from project application for ex ante savings as opposed to IOU claimed savings, resulting in inaccurate realization rates.

Comment 23E

2) SBW ID: P00285, IOU ID: 2K6UCR013 [UC Davis, MBCx – Life Sciences Addition]

The "combined" retro-commissioning measure at UC Davis in the Life Sciences Addition consisted of improvements to the AHU operation, chilled water system, and hot water system.

a) Issue: Reviewer only analyzed limited pre retrofit utility data (two months worth) for measurement and verification purposes which resulted in the conclusion that no savings could be determined. Previous report documentation from program tracking reveals at least one year's worth of pre implementation baseline utility data which was used by the initial provider to determine realized savings. It is recommended that the reviewer revise savings estimates using appropriate pre and post implementation data which will allow for an accurate verification of savings.

Comment 23F

3) SBW ID: P00292, IOU ID: 2K6UCR012 [UC Davis, MBCx – Plant and Environmental Sciences]

The "combined" retro-commissioning measure at UC Davis in Plant and Environmental Sciences Building consisted of improvements to the AHU operation, chilled water system, and hot water system.

a) Issue: Reviewer only analyzed limited pre retrofit utility data for measurement and verification purposes which resulted in the conclusion that no savings could be determined. The site report on page 12 states, "The evaluation savings relied on the pre-implementation data available in the program documentation. Only three months worth of pre-implementation chilled water and hot water energy use data was available", while previous report documentation from program tracking reveals at least one year's worth of pre implementation baseline utility data for electricity, steam, and chilled water which was used by the initial provider to determine realized savings. It is recommended that the reviewer revise savings estimates using appropriate pre and post implementation data which will allow for an accurate verification of savings.

SCE UC/CSU Projects:

Comment 23G

1) SBW ID: P00303, IOU ID: 1061 [UC Santa Barbara, MBCx – Life Sciences]

The "combined" retro-commissioning measure at UC Santa Barbara in the Life Sciences Building consisted of adding VSD's to four exhaust fans, fixing ambient air make up damper, adding booster fans, and controls setpoints.

b) Issue: Appendix values do not match main report values

c) Issue: The results summary for the SBW ID P00303 site report clearly states on page 16 second paragraph that the VSD's and booster fans have annual savings of 459,563kWh, with a tracking estimate of 509,725kWh. The next paragraph states that tracking savings include 44,131kWh that are linked with "non-applicable air-side measures", which represent savings associated with the central plant per program tracking documentation. The calculation of a realization rate using an ex ante savings estimate that includes savings from what the reviewer considers to be a different measure is inaccurate. It is recommended that the reviewer either remove savings associated with the "non-applicable air-side measures" from the ex ante savings in order to calculate the proper realization rate, or preferably, measure and verify the savings associated with the central plant as these are valid savings directly resulting from the project.

SDG&E UC/CSU Projects:

Comment 23H

1) SBW ID: P00317, IOU ID: 2 [UC SD, MBCx – Engineering Building Unit 2]

The specific MBCx project evaluated at UCSD engineering Building Unit 2 is EEM # 1, Re-Commission Laboratory and Fume Hood Controls

a) Issue: The original savings calculations for the project utilized a discharge air temperature of 55F for the air handlers while the reviewer's calculations used a discharge air temperature of 64.3F, resulting in a reduction in gas savings associated with reheat. Page 24 of the site report for SBW ID P00317 states that the average discharge air temperature of 64.3F was used to correct the baseline model's assumption where it had been set to 55, however a summary table of discharge air temperature supplied from the BMS trended from May 6, 2008 to September 17 2009 states an average discharge air temperature of 59F. It is recommended that the reviewer revise the calculations using the correct trended average discharge air temperature which appears to be 59F, not 64.3F.

b) Issue: Page 22 of the site report for SBW ID P00317 states that, "Slight corrections were also made to the model to correct assumptions regarding the central plant kW per ton rating used", but no initial values or altered values are provided. The report also lacks any specific justification for the revision of the central plant efficiency.

Comment 23I

2) SBW ID: P00319, IOU ID: 4 [UC SD, MBCx Hybrid – Hillcrest Medical Center]

The specific MBCx project evaluated at UCSD Hillcrest Medical center is EEM #1, VSDs on CT fans, Optimize Chiller/Boiler Sequencing.

a) Issue: It appears as though the reviewer did not account for new boiler controls implemented as part of the overall MBCx project. The HW system measures are mentioned on page 18 of the site report for SBW ID P00319 under section 3.3 "Evaluation Algorithms- Energy Savings", however section 7.2. "Key Findings" of the report does not include any information about the HW system measure. The associated gas savings for the MBCx project were reduced to zero therms yet there is no post implementation trending data, calculations, or supporting justification for reducing the gas savings. It is recommended that the reviewer either verify the original gas savings calculations or eliminate the ex ante gas savings associated with the boiler controls as the resulting therm realization rate of zero is inaccurate.

Table 1: IOU Claimed vs. Ex Ante Savings

PG&E Project Summary from Report (IOU claimed savings from program database in red)

SBW ID	IOU ID	Campus	Project Title	Case Weight	IOU Claimed Savings kW	IOU Claimed Savings kWh	IOU Claimed Savings Th	Ex Ante Gross Savings kW	Ex Ante Gross Savings kWh	Ex Ante Gross Savings Th	Ex Post 1st Year Gross Savings kW	Ex Post 1st Year Gross Savings kWh	Ex Post 1st Year Gross Savings Th	Gross Realization Rates kW	Gross Realization Rates kWh	Gross Realization Rates Th	NTGR
P00278	2K6UCR002	CSU East Bay	Science Building Complex MBCx	1.1	44.0	255,190	92,200	44	255,190	92,200	47.0	549,704	98,618	107%	215%	110%	0.83
P00281	2K0700071	CSU San Jose	MBCx - Central Chilled Water Plant	1.1	47.0	523,075	86,210	69.4	607,533	11,528	80.0	597,538	261,330	115%	98%	2270%	0.9
P00285	2K6UCR013	UC Davis	MBCx - Life Sciences Addition	1.1	91.0	1,412,350	25,275	103	1,412,350	25,275	29.8	359,764	0	29%	25%	0%	1
P00289	2K0701231	UC Davis	MBCx - Shields Library	1.1	-	852,825	45,045	0	852,825	45,405	255.8	849,505	58,877	N/A	100%	130%	1
P00292	2K6UCR012	UC Davis	MBCx - Plant and Environmental Sciences	12.5	75.0	258,962	6,207	105	258,962	6,207	9.2	0	0	9%	0%	0%	1

SCE Project Summary from Report (IOU claimed savings from program database in red)

SBW ID	IOU ID	Campus	Project Title	Case Weight	IOU Claimed Savings kW	IOU Claimed Savings kWh	IOU Claimed Savings Th	Ex Ante Gross Savings kW	Ex Ante Gross Savings kWh	Ex Ante Gross Savings Th	Ex Post 1st Year Gross Savings kW	Ex Post 1st Year Gross Savings kWh	Ex Post 1st Year Gross Savings Th	Gross Realization Rate kW	Gross Realization Rate kWh	Gross Realization Rate Th	NTGR	Ex Post 1st Year Net Savings kW	Ex Post 1st Year Net Savings kWh	Ex Post 1st Year Net Savings Th
P00297	1023	CSU Long Beach	Central Plant MBCx	1	0	1,257,006	0	0	1,257,006	0	0	1,257,006	0	N/A	100%	N/A	1	0	1,257,006	0
P00298	1	CSU Pomona	Central Chiller/TES Plant MBCx	4	0	638,335	0	0	613,146	0	138	638,135	0	N/A	104%	N/A	0.9	124	574,322	0
P00302	5	UC Santa Barbara	Bren MBCx	9.7	19	169,243	0	19	169,243	0	25.3	253,061	0	133%	150%	N/A	0.83	21	210,041	0
P00303	1061	UC Santa Barbara	Life Sciences MBCx	4.3	53	509,725	0	53	295,593	0	31.3	459,563	0	59%	155%	N/A	0.83	26	381,437	0

SCG Project Summary from Report (IOU claimed savings from program database in red)

SBW ID	IOU ID	Campus	Project Title	Case Weight	IOU Claimed Savings kW	IOU Claimed Savings kWh	IOU Claimed Savings Th	Ex Ante Gross Savings kW	Ex Ante Gross Savings kWh	Ex Ante Gross Savings Th	Ex Post 1st Year Gross Savings kW	Ex Post 1st Year Gross Savings kWh	Ex Post 1st Year Gross Savings Th	Gross Realization Rate kW	Gross Realization Rate kWh	Gross Realization Rate Th	NTGR	Ex Post 1st Year Net Savings kW	Ex Post 1st Year Net Savings kWh	Ex Post 1st Year Net Savings Th
P00235	1	CSU LA	MBCx - Admin Building	1	0	0	16,500	0	0	16,500	81.3	237,295	39,104	N/A	N/A	240%	1	81	237,295	39,104
P00276	2	CSU LA	MBCx - Salazar, Simpson Tower, Phys Sci	1	0	0	49,750	0	0	49,750	76.1	568,572	148,353	N/A	N/A	300%	1	76	568,572	148,353

SDG&E Project Summary from Report (IOU claimed savings from program database in red)

SBW ID	IOU ID	Campus	Project Title	Case Weight	IOU Claimed Savings kW	IOU Claimed Savings kWh	IOU Claimed Savings Th	Ex Ante Gross Savings kW	Ex Ante Gross Savings kWh	Ex Ante Gross Savings Th	Ex Post 1st Year Gross Savings kW	Ex Post 1st Year Gross Savings kWh	Ex Post 1st Year Gross Savings Th	Gross Realization Rate kW	Gross Realization Rate kWh	Gross Realization Rate Th	NTGR
P00316	1	UCSD	MBCx - Center for Molecular Medicine West	1	65	587,965	23,127	58.8	587,965	29,595	86	1,000,639	40,768	1.46	1.7	1.4	0.98
P00317	2	UCSD	MBCx - Engineering Building Unit 2	2	22	181,269	14,153	18.1	181,269	16,147	26.9	101,740	2,820	1.48	0.56	0.2	0.98
P00319	4	UCSD	Hillcrest Medical Center - MBCx Hybrid	1	315	1,100,000	150,000	110	1,100,000	162,100	394.1	1,314,994	0	3.58	1.2	0	0.26

2006-2008 Retro-Commissioning Impact Evaluation Report Comments 4

Response to Comment 23A

The decision to rely on the more holistic whole-building savings analysis vs. individual measure analysis was made on a case-by-case basis by the EM&V and CPUC technical advisers. When possible, projects with individual measures accounting for a large portion of the ex ante savings were evaluated as individual measures if large uncertainty due to interactive effects was not a factor. When this was not possible, the more holistic whole-building approach was taken.

Response to Comment 23B

The IOU claimed savings provided to the CPUC were always used in realization rate calculations. When preparing site M&V reports, however, evaluation analysts based their assessment of reasons for differences between ex ante and ex post savings on utility-supplied program documentation, which sometimes differed from the utility-claimed savings. However, realization rates were only calculated from the utility savings claim data sent to the CPUC. Addenda have been added to all site M&V reports in the gross sample to indicate whether or not gross savings reported in the documentation were the same as the utility claimed savings.

Response to Comment 23C

Refer to the response for Comment 23B.

Response to Comment 23D

Refer to the response for Comment 23B.

Response to Comment 23E

Comments for both sites cited, P00285 and P00292, are identical, highlighting the use of short pre data set in the analysis. Both the evaluation analyst and the CPUC technical advisers had similar concerns, but ultimately concluded that the approach taken was the only feasible one. The primary site contact was concerned with the time commitment as they felt burdened with numerous evaluation requests. As per the primary site contact's request a list of required data (for analysis) was provided in the interest of saving the site staff's time during the site visit. We received one year worth of post data. However, based on discussion with the site staff, our understanding was that the pre data was recorded by a different system and could not be conveniently retrieved. In such a case the pre data available in the project documentation was used. The following pre data was available from project documentation for each site.

P00285: LSA CHW Pre Baseline.XLS (Chilled water data from 6/22/07 to 9/14/07); LSA HW Pre Baseline.XLS (Hot water data from 6/22/07 to 9/14/07); LSA Electrical Pre Baseline (Electricity data from 12/3/06 to 11/30/07)

P00292: PES Thermal Energy.XLS (This spreadsheet has the chilled and hot water monthly values. The spreadsheet points to another spreadsheet "PES baseline – 9-25-06 to 9-25-07.xls" but that file is not included in the project documentation. Instead the data that was available was from 7/3/07 to 12/2/07); PES Electrical Pre Baseline (Electricity data from 6/1/07 to 5/31/08)

Response to Comment 23F

Refer to the response for Comment 23E.

Response to Comment 23G

Response to Issue b (there was no 'a'): Similar to comment 23B, utility-supplied program documentation sometimes reported savings that differed from the IOU-claimed savings. Addenda were added to all gross sample reports to indicate whether or not report gross savings are the same as the utility claimed savings.

Response to Issue c: Per the reviewer's recommendation, savings associated with the "non-applicable air-side measures" associated with the central plant from the ex ante savings were removed from the ex ante savings in the report.

Response to Comment 23H

Refer to the response for Comment 15B.

Response to Comment 23I

Refer to the response for Comment 15C.