Itron
Knowledge to Shape Your Future

Electric | Gas | Water
information collection, analysis and application

2006-2008 Evaluation Report for the Southern California Industrial and Agricultural Contract Group

Report Only

CALMAC Study ID: CPU0018.01

Submitted to:
California Public Utilities Commission
505 Van Ness Ave.
San Francisco, CA 94102

Submitted by:
Itron, Inc.
1111 Broadway, Suite 1800
Oakland, CA 94607

February 3, 2010

With assistance from:

ASW Engineering
Energy and Resources Solutions
Energy Metrics
Helios Resources
Jai J Mitchell Analytics
Michaels Engineering
PWP Inc.
Katin Engineering
SDV/ACCI
Warren Energy Engineering

Abstract

This report presents the evaluation results for the energy efficiency measures and programs within the scope of the California Public Utilities Commission's (CPUC) Southern California Industrial and Agricultural (SCIA) evaluation contract group. The evaluation addresses program impacts for the 2006-2008 energy efficiency program cycle.

The Itron team began the evaluation activities with an initial sample of projects from Southern California Edison's Industrial Energy Efficiency Program (SCE2509) and Agricultural Energy Efficiency Program (SCE2510). In July 2008, the CPUC Energy Division (ED) instructed the SCIA contract group to complete work for this initial sample, and redirected the remaining funds to the evaluation of two High Impact Measures (HIMs)[1] that were formerly in the Small Commercial contract group. These measures consisted of Steam Traps (SCG3507, SDGE3020 and PGE2080) and Pipe and Tank Insulation (SCG3507, SDGE3020, SDGE3012 and PGE2080). The ED also directed the SCIA contract group to evaluate the impacts of pump testing, a component of SCE's Agricultural Energy Efficiency Program (SCE 2510).

For this study, the SCIA contract group is divided into four measure groupings for reporting evaluation results. These groupings include two HIMs - pipe insulation and steam traps - as well as pump testing, and for the combination of SCE2509 Industrial measures and SCE2510 Agricultural measures that received incentives. This report presents impact evaluation results for each of these domains, including separate results for SCE2509 and SCE2510 where available.

The final net realization rate estimate for the Pipe Insulation HIM in SCG territory is 5.7 percent, with a relative margin of error at the 90 percent confidence level of 5.9 percent. Final net realization rate results for the PG&E service territory are 17.2 percent, with a relative margin of error at the 90 percent confidence level of 8.7 percent.

The ex-post net therms per commercial steam trap is estimated to be 8.63 therms for PG&E, 11.43 therms for SCG and 11.78 therms for SDG&E. These commercial steam trap evaluation values are significantly lower than the utilities' ex-ante estimates. For industrial steam traps, the net ex-post therm estimate is 2,630 therms for high pressure traps and 794 therms for low pressure traps. The industrial steam trap results are fairly close to the utility ex-ante estimates.

[1] The HIMs are defined as those efficiency measures common across IOU programs that contribute greater than one percent to the entire IOU savings portfolio for reductions in electrical consumption, electrical demand, or natural gas consumption.

For the SCE Pump Test program, evaluation results are roughly three-quarters of SCE's work paper estimate for energy savings and are practically equal in value for peak demand impact.

For the SCE Industrial program, SCE2509, the overall net realization rate is 57 percent of the SCE's claimed energy savings and 53 percent for peak demand.

For the SCE Agricultural program, only net-to-gross results were estimated, resulting in average NTG ratios for installed measures receiving an incentive of 0.59 (kWh) and 0.63 (kW).

Table of Contents

List of Tables

List of Figures

1

Executive Summary

This report presents the evaluation results for the energy efficiency measures and programs within the scope of the California Public Utilities Commission's (CPUC) Southern California Industrial and Agricultural (SCIA) evaluation contract group. The evaluation addresses program impacts for the 2006-2008 energy efficiency program cycle. This evaluation began in September 2007, midway through the 2006-2008 cycle.

The Itron team began the evaluation activities with an initial sample of projects (40 in total) from Southern California Edison's Industrial Energy Efficiency Program (SCE2509) and Agricultural Energy Efficiency Program (SCE2510). In July 2008, the CPUC Energy Division (ED) instructed the SCIA contract group to complete work for this initial sample, and redirected the remaining funds to the evaluation of two High Impact Measures (HIMs)[2] that were formerly in the Small Commercial contract group. These measures consisted of Steam Traps (SCG3507, SDGE3020 and PGE2080) and Pipe and Tank Insulation (SCG3507, SDGE3020, SDGE3012 and PGE2080). Consequently, no additional EM&V sample was drawn for programs SCE2509 and SCE2510. The ED also directed the SCIA contract group to evaluate the impacts of pump testing, a component of SCE's Agricultural Energy Efficiency Program.

For this study, the SCIA contract group is divided into four measure groupings for reporting evaluation results. These groupings include two HIMs - pipe insulation and steam traps, as well as pump testing, and for the combination of SCE2509 Industrial measures and SCE2510 Agricultural measures that received incentives. This report presents impact evaluation results for each of these domains, including separate results for SCE2509 and SCE2510 where available.

The impact evaluation results address verification findings, ex-post energy savings estimates, gross savings realization rates,[3] and the net-to-gross ratio (NTGR).

[2] The HIMs are defined as those efficiency measures common across IOU programs that contribute greater than one percent to the entire IOU savings portfolio for reductions in electrical consumption, electrical demand, or natural gas consumption.

[3] Realization rates are developed for each site and the program as a whole and are defined as the ratio of program ex-post savings estimated by the evaluation team divided by the ex ante savings.

1.1 Summary of Gross Realization Rate and Net-to-Gross Results

1.1.1 Pipe Insulation

Pipe insulation gross therm impact claims are heavily concentrated in the SCG service territory, making up 94 percent of the statewide therm impact claims over the 2006-2008 period. In support of the gross impact measurement objectives, evaluation activities included on-site engineering based measurement and verification of therm impact from 66 sites limited to the SCG service territory. The method for estimating free ridership was based on self-reported data gathered from the phone surveys which covered all relevant service territories, SCG, PG&E and SDG&E.

Final results for the Pipe Insulation HIM Evaluation are shown in Table 1-1 below. The final realization rate estimate for SCG territory is 5.7 percent, with a relative margin of error at the 90 percent confidence level of 5.9 percent. Final realization rate results for the PG&E service territory are 17.2 percent, with a relative margin of error at the 90 percent confidence level of 8.7 percent.

Table 1-1: Summary of Final Pipe Insulation Realization Rates for SCG and PG&E Service Territories

Final Realization Rate Results	SCG	PG&E
Gross Impact Realization Rate	**7.9%**	**35%**
Sample Size	66	38
90% Confidence Bounds	7.4 - 8.4%	33.5 - 36.4%
Relative Margin of Error	6%	4%
Net-to-gross Ratio	**72.2%**	**49.2%**
Sample Size	248	38
90% Confidence Bounds	72.0-72.4%	47.7-50.7%
Relative Margin of Error	0.3%	3.1%
Final Realization Rate	**5.7%**	**17.2%**
90% Confidence Bounds	5.4-6.1%	16.3-18.1%
Relative Margin of Error	5.9%	8.7%

The dry cleaner segment accounts for 78 percent of pipe insulation sites in the SCG tracking system and about 64 percent of total ex-ante gross impact therm claims. The dry cleaner segment has a lower gross impact realization rate than other business type segments (4.6 versus 15.3 percent). This is due to a combination of factors, including finding lower-than-assumed operating hours, and higher-than-assumed ambient air temperatures. In addition, there was a high likelihood of pre-existing insulation at these sites.

1.1.2 Steam Traps

Steam traps qualify as a HIM based on their ex ante estimated annual therm savings contributions to the overall statewide portfolio and to the portfolios of SCG, SDG&E, and PG&E. The ex-ante therm savings from steam trap replacement are equal to approximately

24.9 million therms annually, or 18% of the total annual statewide portfolio therm savings for the 2006 –2008 programs.

There is considerable uncertainty surrounding the energy savings attributable to steam traps. The current California IOU programs have established deemed savings values for steam traps installed in commercial, low pressure industrial and high pressure industrial applications. This HIM evaluation of commercial steam trap applications incorporated site-specific data collection using a phone survey which helped to inform a commercial billing analysis and to determine a self reported NTGR. The phone survey data was analyzed separately for PG&E and the Sempra (SDG&E and SCG) utilities. The evaluation approach for industrial steam traps used a phone survey to collect site-specific self-report information to calculate a net to gross ratio, and on-site data collection to evaluate gross savings. The evaluation worked to minimize response bias for survey based results and recruitment and undertook uncertainty analyses both before and after the on-site visits were conducted.

Table 1-2 combines the results from the NTG and the gross analysis to produce the ex-post net savings per commercial steam trap. The gross ex ante savings per trap was 45.87 therms for PG&E and 139 therms for the Sempra Utilities. After applying the estimated realization rate and the NTG ratio, the ex-post net therms per trap were 8.63 therms for PG&E, 11.43 therms for SCG and 11.78 therms for SDG&E.

Table 1-2: Small Commercial Steam Trap Net Ex-Post Therm Savings

Strata	Total Gross Ex-Ante Therms	Number of Traps	Gross Ex-Ante Therms per Trap	Realization Rate	NTG Ratio	Ex-Post Net Therms per Trap
PG&E	993,544	21,660	45.87	0.30	0.62	8.63
SCG	4,646,492	33,428	139	0.12	0.70	11.43
SDG&E	514,022	3,698	139	0.12	0.72	11.78

Table 1-3 combines the gross ex-post realization rate with the NTG ratio to determine the net ex-post realization rate and savings per industrial trap by pressure level. The average realization rates for low and high pressure industrial steam traps is greater than 200%. The NTG ratio for high and low pressure traps was 0.52 and 0.57, respectively. The high realization rate combined with the reduced NTG ratio leads to a final net ex post therm estimate of 2,630 therms for high pressure traps and 794 therms for low pressure traps. The estimated net ex-post therms per trap are slightly higher than the work paper gross ex ante values due to the very high gross realization rate.

Table 1-3: Measure Level Realization Rates, NTG Ratios, and Net Ex-Post Therm Savings per Industrial Steam Trap

Measure Type	Total Gross Ex-Ante Therms	Number of Traps	Gross Ex-Ante Therms per Trap	Realization Rate	NTG Ratio	Net Ex-Post Therms per Trap
High Pressure	18,639,978	7,959	2,342	2.15	0.52	2,630
Low Pressure	1,159,884	1,818	638	2.19	0.57	794

1.1.3 Pump Testing

Pump testing is part of Southern California Edison's (SCE) Agricultural Energy Efficiency Program (SCE 2510). Gross savings claim impacts for the 2006-2008 program cycle total 32.5 million gross kWh, and 10,810 gross kW. Pump tests are currently not associated with claimed savings for any other PY2006-2008 California Energy Efficiency program or Contract Group. The main objectives of this evaluation were to calculate both the gross and net impact of a pump test.

A participant self-report survey was implemented to collect data used to evaluate the gross and net impacts of a pump test through SCE's Ag Program. The ex post gross impact analysis approach leveraged the participant tracking database and the participant surveys to true up the assumptions in the pump test work papers. The net impact approach collected and analyzed free-ridership self-reports from participant surveys.

The strata level results for the pump test findings are presented in Table 1-4 below, as are SCE's Work Paper assumptions and the overall results. As indicated in the table, the evaluation found that the percent of pump tests that result in a non-incented pump repair is 8.00%. The gross per-pump test kWh impact per pump test is 0.84 kW and 1,872 kWh. The resulting net impact per pump test is 0.53 kW and 1,182 kWh. When these numbers are applied to the population of 13,149 pump tests that were completed in PY 2006-2008, the gross impacts for the program were found to be 11,067 kW and 24,619,025 kWh.

Table 1-4: Evaluation Findings vs. Work Paper Assumptions

Savings Measurement	SCE Work Papers	PY '06-'08 Evaluation Findings				Realization Rate
		Total	Strata 3	Strata 2	Strata 1	
Gross kW/unit repaired	11.34	10.52	7.6	16.4	5.62	92.80%
Gross kW/test	0.82	0.84	0.94	1.13	0.24	102.40%
Gross kWh/unit repaired	34,092	23,392	16,223	33,456	23,320	68.60%
Gross kWh/test	2,472	1,872	2,013	2,311	994	75.70%
Percent of tested pumps that result in non-incented repairs	7.25%	8.00%	12.41%	6.91%	4.26%	110.30%
NTG	0.75	0.63	0.63	0.64	0.62	84.00%
Participant Population PY 06-08	13,149	13,149	4,380	4,382	4,387	
Total Program Gross kW	10,810	11,067				102.40%
Total Program Gross kWh	32,499,989	24,619,025				75.80%

1.1.4 SCE Industrial Measures and Agricultural Measures

As mentioned above, the Itron team began the initial evaluation of SCE2509 and SCE2510 with a proportional sample of projects obtained in March 2008 and the overall goals submitted by SCE for these two programs for the 2006-2008 cycle. The initial sample of 30 projects from SCE2509 and 10 projects from SCE2510 allowed the SCIA contract group to get started with evaluation activities using the limited information available. The CPUC refocused efforts in July 2008 and instructed the SCIA contract group to finalize work for the sampled projects and to take on the evaluation of two HIMs. This was done because the HIMs assigned represent a much larger share of the portfolio savings than SCE2509 or SCE2510. As a result of this shift, the impact evaluation results in this report are based on a very small sample size.

For SCE2509, a site-specific engineering approach was used for this evaluation, and included measurement and in-depth engineering analyses. The key steps involved in developing the overall savings estimate for the program were to independently verify reported measure installation records, develop ex-post estimates of the energy savings for each project in the sample, and statistically apply those findings to the full participant population.

Table 1-5 below presents a comparison of the evaluation verified net savings with the final program claimed net savings, as obtained from the final PY2006-2008 reports posted on EEGA. This table includes a calculation of the verified net savings as a percentage of the claimed net savings.

Table 1-5: Comparison of First-Year Evaluation-Based Net Savings with the Final Program-Claimed Net Savings: SCE2509 Industrial Projects

	Electric Savings	
	kWh/year	Avg. peak kW
Tracking		
a. Claimed Gross Savings	145,467,578	16,776
b. Claimed Realization Rate	0.89	0.89
c. Claimed Adjusted Gross Savings (c = a x b)	129,466,144	14,931
d. Claimed NTG Ratio	0.80	0.80
e. Claimed Net Savings (d = c x d)	103,572,915	11,945
Evaluation		
f. Evaluation Gross Realization Rate	0.72	0.65
g. Evaluated Gross Results (g = c x f)	93,438,719	9,710
h. Evaluation NTG Ratio*	0.63	0.65
i. Evaluated Net Results (i = g x h)	59,149,486	6,299
j. Evaluation Net Realization Rate (h = d x f)	0.46	0.42
k. Evaluated Net Savings as a Fraction of Claimed Net Savings (k = i / e)	**0.57**	**0.53**

* Consistent with current CPUC policy, the Net-to-Gross ratios in this evaluation reflect the effect of free ridership only and exclude any consideration of spillover.

As the table shows, the evaluated gross realization rate is 72% for kWh and 65% for kW for program SCE2509. There are fairly significant differences between claimed and evaluated NTGRs. The evaluated net savings as a percentage of program claimed net savings are 53% for kW and 57% for kWh. These values indicate that verified net program savings are on the order of one-half (for SCE2509) of claimed savings, far below program savings estimates.

1.2 Summary of Recommendations

1.2.1 Pipe Insulation

Controls should be instituted to ensure compliance with program guidelines.
Controls should be instituted to ensure that incented insulation is not installed on pipe with pre-existing insulation. In addition, controls should ensure that incented insulation is not installed in new construction applications, new pipe additions, and/or pipe replacements. At a minimum, verification of these characteristics should be provided by the installation contractor and the customer prior to distribution of incentive money. Another more stringent alternative would be to require IOU representatives to inspect sites prior to approving incentive applications.

Recommended revisions to pipe insulation work paper assumptions and ex-ante impact claims.
Assumptions about the environmental conditions (temperature surrounding pipes) in which the piping systems operate are generally inconsistent with data collected from participating sites. Another assumption contributing to an over-estimation of ex-ante therm impact was with respect to the diameter of piping systems insulated.

1.2.2 Steam Traps

The results from this evaluation lead to the recommendation that industrial steam traps be rebated as a custom measure.
The extremely high variability in per trap savings strongly supports the recommendation that industrial steam traps should not be rebated as a prescriptive measure.

It is also recommended that the utility closely monitor the installation of steam traps.
The onsite data collection effort found that a significant number of rebated traps were either not installed or not operational.

1.2.3 Pump Testing

The kW and kWh savings values should be verified.
The task of verifying the kW and kWh savings was not completed for pump testing. It was decided that such an effort was not warranted, given that the focus of this evaluation element was on measuring program influence on pumping system upgrades that occur in the absence of incentives. Use of deemed savings values is common for the evaluation of similar education program elements, such as audits. On future evaluation efforts it is recommended that on-site verification of pump operating efficiency be conducted. Another method would be to conduct a statistically adjusted engineering billing regression within a telephone survey sample.

Pump test frequency should be optimized.
Half of the participants surveyed reported that they test their pumps every one to two years and about half of those participants stated that if the program did not exist, they would still test their pumps on the same schedule. This implies that participants have a need to test pumps in some instances at a frequency greater than two years. Consideration should be given to optimizing any program rules surrounding the frequency of testing.

Examine the pump testing and repair practices in areas where free water pump efficiency testing is not offered.
This would provide a broader perspective on the value of the testing service, and may be a better foundation for estimating baseline conditions and net-to-gross ratios for the program than self-reported data from SCE participants.

Future evaluations should look for ways to get participants to follow through with repairs.
This evaluation found that only 20 percent of pumps in need of repair go on to be repaired. When a repair is recommended through an audit, SCE could dispatch a repair technician to provide participants with an estimate for the needed repair. In past audit programs, these

follow-up calls appear to have made a significant impact on the likelihood that the customer will implement audit recommendations.

1.2.4 SCE Industrial Measures and Agricultural Measures

Below are several recommendations aimed at improving the accuracy of savings claims and increasing the degree of program influence on rebated projects.

Recommendation: Improve baseline specification.
End the practice of using in situ baselines over the effective useful life (EUL) of the measure as the baseline for estimating savings and paying incentives. Identify projects explicitly in program files as replace-on-burnout, natural turnover, or early replacement. In the case of early replacement, provide evidence and documentation of the remaining useful life of the equipment replaced, the estimated time at which the equipment would have been replaced in the future, and the effect of the program in accelerating early replacement.

Recommendation: Clarify and enforce the definition of "industry standard practice".
This definition is used to set baselines for savings estimates and incentives (such that program savings estimates improve with better baselines and result in more accurate evaluation gross and net realization rates).

Recommendation: Be more conservative in estimating savings.
We recommend that the programs make more conservative assumptions for calculated projects in the future. Increased measurement could be used to address any customer concerns about savings estimates being too conservative; claims of higher savings would have to be substantiated with pre- and post-installation measurement (possibly with Energy Division review for the largest projects).

Recommendation: Empirically study the effective useful life of measures in an industrial setting.
Due to the uncertain nature of measure persistence associated with equipment operation in an industrial setting, it is recommended that empirically based persistence studies be completed to assess the appropriateness of EUL assumptions used by the programs and evaluators.

Recommendation: Incorporate greater levels of real-time measurement and pre- and post-installation measurement based verification.
Particularly for projects that are larger and have more uncertain savings, incorporate greater levels of either 'within-program' measurement or evaluation-based measurement, conducted in parallel with program measure installation. Pre- and post-installation measurement in particular can be useful in not only establishing more robust ex ante savings estimates but can also be incorporated when available by the relevant evaluation team. It is anticipated that these efforts would close the substantial gap that exists between current ex ante models and results and those based on evaluation M&V in this report.

Recommendation: Require a greater level of technical documentation for the largest and most complex projects.
Consider increases in the level of technical documentation required for the largest, most complex projects. There is a balance between keeping the application process and forms from being overly complex and costly to navigate, while at the same time providing adequate levels of documentation for verification and savings analyses. Application documentation should not be over-simplified, given the complexity of measures and range of site-specific characteristics in this program.

Recommendation: Require better documentation of pre-installation operating conditions.
Better documentation is needed regarding pre-installation or pre-retrofit operating conditions. In particular, large complex projects might be required to submit a greater level of site-specific application data than smaller projects, since (a) they contribute disproportionately to total program savings; (b) the large incentive payments increase the temptation for gaming or fraud; (c) measures implemented are often site-specific or industry-specific, and (d) savings may be very sensitive to baseline conditions.

Recommendation: Aggregate and approve fuel switching and distributed generation-related projects in one or more explicit programs or clearly identified program elements.
If the CPUC approves use of fuel switching, it should require all applications to follow the three-prong test set forth in the CPUC Policy Manual[4] and any other CPUC or other regulatory agency requirements (e.g., those related to GHG reduction goals).

Recommendation: Make application-level information readily accessible to evaluators.
Consideration should be given to making application-level information readily accessible to evaluators through electronic storage of all application files and possibly a retrievable on-line system made available to evaluators. Such a system might provide easy viewing access to the project tracking data plus downloading rights to project documentation in electronic format for each project. This documentation and storage and retrieval system would greatly facilitate the evaluation, while removing a step that commonly impedes evaluation progress, namely data requests. This level of access and documentation would represent best practice in this area for this type of custom program.

Recommendation: Tracking system ex ante impacts adjustments should be applied at the beginning of the program cycle.
The ex ante impacts stored in the SCE tracking system should not require mid-program cycle adjustments. The application of these adjustment factors were required on the part of the evaluators to match E3 claims, but should have been included in the tracking system format from the start of the program cycle. Extracts are provided to evaluators multiple times in a given program cycle. Such extracts should neither include changes to the impacts achieved during the program cycle, nor structurally change the database itself.

---

[4] See CPUC Decision 92-10-020, Conclusion of Law 5. The Three Prong Test requires that any fuel switching measures: (1) not increase source-BTU consumption; (2) have a TRC benefit-cost ratio of 1.0 or greater; and (3) not adversely affect the environment. The Three Prong Test does assess total fuel input, in addition to determining if the switch is cost-effective.

Recommendation: Improve the capability of program implementation staff to materially influence industrial efficiency improvements.

To move these customers further along the efficiency spectrum takes time and advanced levels of technical expertise, often requiring expertise in specific industry production practices and options for improvement. There is already significant industrial expertise available at the utility and third-party contractors. This expertise should be built upon and further increased. Development of the depth of technical expertise required to increase the net effects of the programs is a long term endeavor that requires both utility and regulatory support.

Recommendation: Influence and provide incremental energy efficiency options directly to end users at the earliest decision-making stages of major equipment or facility modifications.

Program involvement at an early stage to identify large equipment and facility changes helps ensure efficiency opportunities are appropriately considered and maximizes the chances of program influence. Utilization of sales or related tracking systems helps prevent projects from becoming lost opportunities.

Recommendation: Consider using early project NTG and baseline screening prior to the incentive being approved for the largest projects and those with significant policy issues such as fuel switching, self generation, and greenhouse gas impacts.

For the largest projects and those with significant policy issues, we strongly recommend that the CPUC consider implementing an Early Project NTG and Baseline Screening step. This step would involve having the CPUC evaluation team review the baseline claim and conduct NTG interviews just after the participant's implementation decision is made. The purpose of this screening is to obtain critical information regarding program influence that may lead to the project being re-defined or dropped. We feel that this early review is critical to the proper specification of the project baseline and the minimization of free-ridership for such large and complex projects. This approach would also have the advantage of capturing critical information on program influence just after the decision is made, while the information is still fresh in the mind of the decision maker(s).

Recommendation: Carefully review the list of qualifying measures for each program and eliminate eligibility for those that are standard practice.

Measures that are already extremely likely to be installed by the vast majority of the market should not qualify for incentives. Although identification of such measures can be difficult in practice in the industrial sector, a number of such measures can be identified through investigation of industry practices (e.g., interviews with manufacturers, distributors, retailers, and designers), analysis of sales data, and review of evaluation results. A balance must be struck, in determining which measures to retain and which to eliminate, between reducing free ridership and avoiding significant lost opportunities.

Recommendation: Put measures with inadequate empirical basis for savings estimates in the Emerging Technologies program until more reliable information is developed.

The CPUC and IOUs should develop more explicit criteria for determining whether new measures are included under resource programs or the Emerging Technologies program. Measures with highly uncertain savings in need of detailed research to establish validity, expected savings, and repeatable algorithms and measurement protocols should be included in emerging technologies.

Recommendation: Improve training of program implementation staff in several key areas. These areas are: proper baseline specification, enforcement of program and policy rules, reasonableness of claims, and increasing program influence on end user's efficiency-related decisions.

Recommendation: Conduct analysis of customer incentives by customer and industry type, further research on use of incentive caps.
Customer incentive caps have been utilized in various forms for many years. During times of low budgets and low goals, caps were set low to spread incentives to a broad pool of participants. More recently, as goals and budgets have significantly increased, caps have increased greatly as well. We are not aware of any systematic study of the effect of the incentives caps. Similarly, research is needed to explore how much total incentive dollars have been distributed across or concentrated within certain customers to determine whether these patterns are aligned and supportive of efficiency policy goals.

Recommendation: More information is needed on industrial project costs, non-energy costs and benefits, net present value analysis, and associated participant cost-effectiveness analysis.

Other recommendations to reduce free ridership. The following are overarching free-ridership-related recommendations that are also relevant to this contract group.

Recommendation: Consider limiting or excluding incentive payments to known free riders[5].
One obvious and simple approach to reducing free ridership is for program administrators to simply exclude projects from the program that they (or possibly the Energy Division) believe have a high probability of being free riders. Administrators in several other jurisdictions have used this approach.[6] In these cases, the administrator has the flexibility to determine total incentive amounts on a case-by-case basis, including zero incentives. We believe consideration should be given to implementation of a process by which projects considered

[5] From the California Public Utilities Commission Energy Efficiency Policy Manual, v. 4.0: "Free riders (Free Ridership) are program participants who would have installed the program measure or equipment in the absence of the program."

[6] Itron, 2005. National Energy Efficiency Best Practices Study. *Volume NR5 – Nonresidential Large Comprehensive Incentive Programs.* www.eebestpractices.com

to be very high likelihood free riders are excluded from participation (or, conversely, must go to higher efficiency levels than initially planned in order to participate).[7]

Recommendation: Consider incorporating a payback floor.
The use of a payback floor (minimum payback level based on energy savings alone) helps to ensure that project generates meaningful and significant energy savings. With a payback floor, the program avoids incenting projects that are primarily being done for reasons other than energy savings (modernization, production efficiency, environmental compliance, etc.)

Recommendation: Set incentive levels to maximize net not gross program impacts.
Free riders dilute the market impact of program dollars. Payback floors and increasing incentives with increasing payback levels are one approach. Another is to tie incentive levels to individual measures or types of measures that are known to have extremely high or low naturally occurring adoption levels.

Recommendation: Consider tying staff performance to independently verified net results.
Tying performance reviews and bonuses of program staff to verified net savings as reported through an independent M&V or impact evaluation process is likely to increase project quality and the accuracy of initial savings estimates. Marketing staff, in particular, should have any financial incentives tied to savings that are independently verified.

Evaluation Related Recommendations. There are also a number of recommendations related to improving the evaluation process.

Recommendation: Involve impact evaluators in large projects and a sample of projects on a real-time basis throughout the program cycle.
The timing of evaluation processes should be accelerated. Moving the evaluation process forward in time to occur just after the project is installed would ensure the decision maker is still available, and that their memory of the basis for the project is still fresh. This can be accomplished through earlier contracting and implementation of the evaluation, combined with improved utility tracking and early reporting of installations (as well as projects in the pipeline), more frequent sampling, and evaluation of projects throughout the program plan period.

Recommendation: Evaluation participation requirements should be strengthened.
Evaluation participation requirements need to be clearly explained to participants, both at the time they are paid incentives, and later, when evaluation activities commence. Evaluation participation should be clearly and obviously written into program participation and incentive payment agreements.

[7] If necessary, such a process could involve an advisory group that includes staff from the Energy Division (to address any customer concerns). This would offer IOUs appropriate protection from claims that such exclusions were unfounded or unfair.

Recommendation: Conduct a full complement of impact, process, and market evaluations. Large customer programs and markets are very dynamic and require regular assessment in order that they may be continuously improved by program managers and policymakers. Most of the effort for the 2006-2008 industrial evaluation focused on impact evaluation, in accordance with Energy Division's evaluation priorities. Future evaluations should consider more integration of process evaluation and market assessment to capture research economies and reduce customer and vendor interview burdens.

Recommendation: Stagger the timing of process and ex post impact tasks so that process evaluations can be conducted and results communicated on a relatively real-time basis. If process and impact evaluations are more integrated in future evaluations, care must be taken to schedule activities and deliverables appropriately. Because of the sometimes long project installation lag after program commitment in these programs, it is important to schedule process evaluation tasks to be conducted during or just after each program year so that results can be utilized to improve program processes for the subsequent program year (rather than producing results only late in the three-year program cycle for use in the next program cycle).

Recommendation: Conduct baseline research to establish standard industry practices for key measures in key industries. Significant research is needed to establish meaningful and defensible data, especially market share, for establishing industry standard practices for measures that are not completely site specific. Improved information on industry standard practices can then inform decisions about which measures should be eligible for incentives, which could in turn lead to reductions in free ridership.

Recommendation: Conduct persistence study of industrial sector savings. Few studies of the persistence of program savings in the industrial sector have been conducted, particularly within the last decade. Some participants have closed facilities or shut down processes associated with program measures due to economic factors. In addition, in some program years and cycles industrial production levels will be higher or lower depending on economic conditions. Research is needed to measure the persistence of savings over time under a range of economic conditions. Sufficient time needs to pass in order to maximize the information provided from such persistence studies.

2

Introduction and Purpose of Study

2.1 Introduction

This report presents the evaluation results for the energy efficiency projects and programs within the scope of the California Public Utilities Commission's (CPUC) Southern California Industrial and Agricultural Program evaluation contract group. The evaluation addresses program impacts for the 2006-2008 energy efficiency program cycle. This evaluation began in September 2007, midway through the 2006-2008 cycle.

Itron Inc. Consulting and Analysis group is the prime contractor for this evaluation. Itron was assisted by a team of subcontractors in this evaluation effort. This evaluation was managed and directed by the Energy Division of the CPUC. Assistance was provided to the Energy Division and Itron on study design and quality control by the CPUC's technical support contractors for this evaluation cycle (the Data Management and Quality Control (DMQC) contractor, the Master Evaluation Contract Team (MECT)[8], and the Database for Energy Efficiency Resources (DEER) [9]contractor).

The Southern California Industrial and Agricultural contract group is comprised of the two Southern California Edison Programs: the Industrial Energy Efficiency Program (SCE2509) and the Agricultural Energy Efficiency Program (SCE2510). These programs address the following market segments: industrial manufacturing (includes industrial, fabrication and process); water supply and treatment and wastewater treatment; oil and gas extraction; producers of agricultural products, including both farmers and food processing firms, facilities that process and store food; and golf courses. Each program offers one or more of the following interventions in order to encourage end-users to upgrade to energy efficiency measures: site specific facility assessments, feasibility studies, project incentives, facility audits, pump testing, and specialized training.

---

[8] A group of consultants with specialized expertise in important aspects of program impact evaluation that are technical advisors to ED staff and assist the evaluation contractors with development and execution of the verification and evaluation plans.

[9] The California Energy Commission and California Public Utilities Commission (CPUC) sponsors this database designed to provide well-documented estimates of energy and peak demand savings values, measure costs, and effective useful life (EUL) all with one data source. DEER has been has been designated by the CPUC as its source for deemed and impact costs for program planning.

Originally this contract group also included the Value/Energy Stream Mapping Programs administered by both SCG (SCG3535) and SDG&E (SDGE3044). These two programs had no savings claims and are not included in Table 2-1.

A list of the SCE programs included in this evaluation, and their basic program elements, is presented in Table 2-1.

Table 2-1: SCIA Programs, Descriptions and Key Elements

Programs Included in this Evaluation	Program Description	Key Program Elements (Note: As stated in original program filings)
SCE2509, Industrial Energy Efficiency Program	Third-party program implemented by Lockheed Martin Aspen Systems (general manufacturing sector), Global Energy Partners (oil and gas extraction facilities), Bacgen (water/wastewater sector), Martin Aspen Onsite Energy (offers energy management tools) and California Manufacturing Technology Consulting (emphasizes process improvements).	Program features vary by implementer, but may include incentives, technical assessments, and diagnostic tools (such as the energy management process, and software). Measures: highly varied, site-specific industrial measures; some POCs, some HVAC.
SCE2510, Agricultural Energy Efficiency Program	Third-party program implemented by Global Energy Partners and EnSave (comprehensive audits and other energy efficiency services to end-users), Center for Irrigation Technology, California State University, Fresno (pump testing services), and Staples Marketing Communications (pump tests and facility audits for golf courses.)	Designed to enhance adoption of energy efficient equipment and practices among agricultural customers. Measures: highly varied, site-specific agricultural measures; some HVAC and lighting.

2.2 Program Goals vs. Program Accomplishments

As noted above, this evaluation began in September 2007, midway through the 2006-2008 program cycle. At that time, the industrial and agricultural programs were still early in their implementation activities. As a result, the evaluation team began its evaluation planning work using the program goals that SCE, SCG and SDG&E submitted to ED at the outset of the program cycle. Beginning the evaluation activities during the program cycle was an important and effective aspect of the study and we recommend that evaluation activities begin even earlier in the next program cycle.

The Southern California Industrial and Agricultural Contract Group consists of the four programs listed in Table 2-2 below. The table includes the savings goals as compared to the actual accomplishments for these programs. SCE's Industrial Energy Efficiency Program (SCE2509) had savings claims of over 103 million kWh which amounts to 120% of their original goal of 86.6 million kWh. Demand savings claims totaled nearly 12,000 kW, falling short of the goal of 36,100 kW. For SCE's Agricultural Energy Efficiency Program, savings claims were roughly half of savings goals for both kWh and kW. The Value/Energy Stream Mapping Programs administered by both SCG (SCG3535) and SDG&E (SDGE3044) had no claimed savings and therefore will not be discussed further in this report.

Table 2-2: SCIA Program Goals and Program Accomplishments

		Goals			Claimed Accomplishments		
Program	**Program Name**	**kWh**	**kW**	**Therms**	**kWh**	**kW**	**Therms**
SCE2509	Industrial Energy Efficiency Program	86,611,501	16,222	0	103,572,915	11,945	0
SCE2510	Agricultural Energy Efficiency Program	129,368,274	36,100	0	73,820,866	18,365	0
SCG3535	Value/Energy Stream Mapping Program	0	0	1,195,680	0	0	0
SDGE3044	Value/Energy Stream Mapping Program	5,170,000	588	810,750	0	0	0

2.3 High Impact Measures (HIM) and Industrial Measure Groupings

Although the evaluation planning process initially took utility programs as a key organizational element, it was also emphasized by many evaluation teams that the portfolio should be examined from the perspective of key measures. In this evaluation, this approach is referred to as the high impact measure (HIM) approach. The philosophy behind the HIM approach organizes energy and demand impacts by measure groups and energy metrics (electric energy, electric demand, and gas energy) across programs at the utility level. The HIM approach sought to standardize the analytical methods and data collection approaches for key measures across programs and contract groups to increase consistency and accuracy.

The first step in the HIM process was to identify which measures or like group of measures contributed most to each of the energy metrics for each of the utilities. Many of the industrial sector efficiency projects involve complex energy systems and processes that are unique to a particular industry or even individual site.

The ED and its consultants developed a preliminary list of HIMs from the E3 calculators delivered by the IOUs[10] covering program savings claims through the end of the second quarter of 2008 (Q2-2008). A single Access database containing the E3 measure line items from the Input tab of the E3 calculator was created. Each of the measures was assigned to a measure name using a consistent measure naming scheme. The savings claims for each IOU

[10] These are the California investor owned utilities (IOU's), Pacific Gas and Electric, Southern California Edison, San Diego Gas and Electric, and Southern California Gas.

were tabulated for each named measure, and the contribution of each measure to the total IOU portfolio savings claim for kWh, kW and Therms was calculated. An initial list of HIMs was developed by identifying all measures that contributed more than one percent of the portfolio savings by IOU. There were no High Impact Measures in the Southern California Industrial and Agricultural Contract Group. Because of this, the ED requested that the contract group perform EM&V for two Small Commercial HIMs: Pipe and Tank Insulation [SCG3507 (Express Efficiency Program), SDGE3020 (Small Business Super Saver Program), SDGE3012 (Express Efficiency Program) and PGE2080 (Commercial Mass Market Program)], and Steam Traps (SCG3507, SDGE3020 and PGE2080).

2.4 Measure Groupings for SCIA Evaluation Report

The SCIA Evaluation Report is divided into four main measure groupings for reporting evaluation results. As noted above, these groupings include two HIMs: Pipe Insulation and Steam Traps, and two non-HIMs: pump testing (included in SCE 2510) and SCE2509 Industrial Measures. Information on these measure groupings and details regarding claimed savings impacts can be found in the following subsections.

2.4.1 Pipe Insulation

Pipe insulation is a significant gas savings measure for SCG with savings claimed of over 16 million therms. Pipe insulation is also found in PG&E's and SDG&E's service territory, as can be seen in Table 2-3 below. Pipe insulation was originally a measure within the Small Commercial Contract Group; however, as mentioned above, this measure was reassigned to the SCIA Contract Group. Chapter 3 of this report includes the details of the measurement and verification of gross and net therm savings achieved through the installation of pipe insulation through the IOU incentive programs over the 2006-2008 program cycle.

Table 2-3: Pipe Insulation Annual Claimed Therm Impacts

Utility	Annual Therm Savings from Pipe Insulation
SCG	16,400,122
PG&E	934,033
SDG&E	126,630
Total	**17,460,784**

2.4.2 Steam Traps

Steam Traps account for a total savings claim of nearly 25 million therms annually, as can be seen in Table 2-4 below. The majority of therm savings claims can be found in SCG and PG&E service territories. This measure was originally included in the Small Commercial Contract Group but was reassigned to the SCIA Contract Group. Chapter 4 of this report includes the details of the measurement and verification of gross and net therm savings achieved through the installation of steam traps through the IOU incentive programs over the 2006-2008 program cycle.

Table 2-4: Steam Traps Annual Claimed Therm Impacts

Utility	Annual Therm Savings from Steam Traps
PG&E	9,197,461
SCG	15,252,338
SDG&E	537,187
TOTAL	**24,986,986**

2.4.3 Pump Testing

The pumping measure in SCE's Agricultural Program (SCE2510) consists of both a free pump testing service and financial incentives for energy efficiency improvements to commercial, agricultural and industrial pumping systems. This report, however, only pertains to the Pump Testing Measure Impact Evaluation. As can be seen in Table 2-5 below, the total gross kWh savings claims are nearly 32.5 million kWh and over 10,810 kW. Chapter 5 of this report includes the details of the evaluation of the gross and net savings estimates for pump testing through the SCE Ag Program over the 2006-2008 program cycle.

Table 2-5: Pump Test Measure Claims for SCE2510

Program	Gross kWh	Gross kW	Net to Gross Ratio
SCE2510	32,499,989	10,810	0.75

2.4.4 SCE2509 Industrial Measures

The savings claims through Q4 2008 for SCE's Industrial Energy Efficiency Program (SCE2509) can be found in Table 2-6 below. These claims total over 129 million kWh and nearly 15,000 kW. Chapter 6 of this report includes the details of the measurement and verification of gross and net savings achieved through the installation of SCE2509 Industrial measures through SCE's Industrial Energy Efficiency Program over the 2006-2008 program cycle.

Table 2-6: SCE2509 Industrial Measures Savings Claims

Program	Program Name	kWh	kW
SCE2509	Industrial Energy Efficiency Program	103,572,915	11,945

2.5 Structure of Report

The overall organizational structure of this report can be found in Table 2-7 below.

Table 2-7: Structure of Report Sections

3. Pipe and Tank Insulation	**4. Steam Traps**
3.1 Evaluation Objectives	4.1 Evaluation Objectives
3.2 Methods used (on-site surveys, billing analysis where appropriate, self-report NTG surveys of participants and their vendors)	4.2 Methods used (on-site surveys, billing analysis where appropriate, self-report NTG surveys of participants and their vendors)
3.3 Validity and Reliability	4.3 Validity and Reliability
3.4 Detailed Findings	4.4 Detailed Findings
3.5 Discuss Findings & Recommendations	4.5 Discuss Findings & Recommendations

5. Pump Testing	**6. SCE Industrial and Agricultural**
5.1 Evaluation Objectives	6.1 Evaluation Objectives
5.2 Methods used (telephone surveys, self-report NTG surveys of participants and their vendors)	6.2 Methods used (on-site surveys, billing analysis where appropriate, self-report NTG surveys of participants and their vendors)
5.3 Confidence and Precision of Key Findings	6.3 Results
5.4 Validity and Reliability	6.4 Discussion of Findings & Recommendations
5.5 Detailed Findings	
5.6 Program Results	
5.7 Discuss Findings & Recommendations	

3

Pipe Insulation

3.1 Evaluation Objectives for Pipe Insulation

This section describes and outlines the evaluation objectives for the High Impact Measure (HIM) Pipe Insulation. The overarching objective of this evaluation was the measurement and verification of gross and net therm savings achieved by the installation of pipe insulation through the IOU incentive programs over the 2006-2008 program cycle.

This evaluation provides the unit energy savings (UES) per linear foot of installed pipe insulation for major ex-ante measure categories.[11] The evaluation also provides estimates of the net-to-gross ratio for pipe insulation installed in nonresidential applications in PG&E and SCG service territories[12] for the 2006 through 2008 program cycles.

Initially, this HIM included both pipe and tank insulation. During the planning phase for this HIM evaluation a decision was made to focus the evaluation on the pipe insulation component, which makes up over 94 percent of statewide pipe and tank insulation gross therm impact claims. Similarly, a decision was made during the planning phase to exclude the relatively minimal[13] electricity savings from the evaluation scope.

Pipe insulation gross therm impact claims are heavily concentrated in the SCG service territory, as shown in Table 3-1below. Pipe insulation claims in SCG territory make up 94 percent of the statewide therm impact[14] claims for pipe insulation over the 2006-2008 period. Within SCG service territory, pipe insulation has first year gross impact claims of 16,400,122 therms for installations during the 2006-2008 program cycle. All of these impact claims originate from the SCG nonresidential Mass Market Program (SCG3507) and make up 26 percent of total SCG first year therm impact claims for the 2006-2008 program cycle.

[11] Data collected can also support UES by ex-post verified temperature ranges, pipe diameters, pipe function, insulation thickness and more.

[12] Attempts were made to measure ex-post net-to-gross ratio for pipe insulation installed in SDG&E service territory, but insufficient sample was available to support such an estimate.

[13] Statewide first year total gross kWh impact claims for pipe insulation are negligible, at less than 20,000 kWh. Tank insulation first year kWh savings make up less than 0.2 percent of statewide impact claims. Neither pipe or tank insulation individually, nor the combination qualifies as a HIM for electricity impact claims for any individual IOU for the 2006-2008 program year cycle.

[14] This holds true on both a gross impact and a net impact basis.

Table 3-1: Pipe Insulation Gross Impact Summary

	Pipe Insulation Gross Therm Impact	Total IOU Gross Therm Impact (all measures)	Percent of IOU Therm Impact in Pipe Insulation	Percent of Pipe Insulation Impact through Express Efficiency/Mass Market	Percent of Pipe Insulation Delivered Through Small Business Super Saver
SCG	16,400,122	62,290,003	26.3%	100%	
PG&E	934,033	67,928,927	1.4%	100%	
SDG&E	126,630	7,667,793	1.7%	28%	72%
Total	**17,460,784**	**137,886,723**	**12.7%**	-	-

SDG&E pipe insulation claims are 126,630 therms, or 1.7 percent of total first year gross therm impact claims for the 2006-2008 period. Pipe insulation was delivered in SDG&E service territory through both the Small Business Super Saver (SDGE3020) and Express Efficiency (SDGE3012), with 72 percent and 28 percent of claims, respectively.

PG&E pipe insulation gross impact claims total 934,033 therms for the 2006-2008 program cycle, of which 99.99% (or 933,726) originate from the Nonresidential PG&E Mass Market Program (PGE2080). These claims make up 1.4 percent of all first year therm impact claims for PG&E over the 2006-2008 program cycle.

In support of the gross impact measurement objectives, evaluation activities included on-site engineering based measurement and verification of therm impacts from 65 sites. This on-site work was limited to the SCG service territory, due to the substantial concentration of statewide impact in this service territory. Results of the on-site work in the SCG territory informed ex-post adjustments to gross impact claims in the PG&E and SDG&E service territories.

Protocol compliant self report methods were utilized to measure net-to-gross ratios for pipe insulation installations in both SCG and PG&E service territories. Although the original evaluation plan called for a separate estimate of the net-to-gross ratio for the SDG&E service territory, the small population of participants[15] did not yield enough survey completes to produce a reliable estimate. Data in support of net-to-gross ratio estimation was collected using a CATI (Computer Aided Telephone Interviewing) system and an Energy Division (ED) approved survey instrument. The telephone survey instrument covers both net-to-gross data and some basic insulation characteristics[16].

[15] The SDG&E participant population is just 23 sites, and yielded just two successful survey completes.

[16] Telephone survey instruments are presented in Appendix B-1. Tables of participant response frequencies are presented in Appendix A-1.

3.2 Methods used for Pipe Insulation

This section describes the methodology used to determine the ex post gross and net therm impact resulting from the installation of pipe insulation through 2006-2008 California IOU energy efficiency programs.

3.2.1 Methodology Overview

As discussed above, estimation of gross energy savings was planned to be based on engineering analysis of data collected on-site from 65 sites where pipe insulation was installed through the SCG Express Efficiency Program. In fact, data was collected from 66 sites[17]. The site work provided the foundation for the estimate of a gross impact realization rate for pipe insulation installations in the SCG service territory, as well as verified unit energy savings (UES) across a variety of key parameters such as fluid temperature and pipe diameter.

The method for estimating free ridership among the pipe and tank insulation participants is based on self-reported data gathered from the phone surveys. While the site-specific M&V work was focused in the SCG service territory, the net-to-gross evaluation spanned all relevant service territories, SCG, PG&E and SDG&E.

Across these service territories, pipe insulation was installed both in small commercial establishments as well as in larger corporate and industrial facilities. For this reason, the pipe insulation net-to-gross evaluation drew from both the "*Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches and the Algorithm for the Residential and Small Commercial Consistent Free Ridership Method*" designed for small commercial and residential customers, as well as the "*Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers*" designed to accommodate larger corporate and industrial customers. These sources provide a standard framework for combining findings from qualitative and quantitative data sources to calculate the net-to-gross ratios for energy efficient measures in a systematic and consistent manner across ED contract groups evaluating nonresidential customers. Details regarding these net-to-gross methods are presented in Appendix A-2 and Appendix A-3. Related survey instruments are presented in Appendix B-1.

3.2.2 Sampling Plan

Gross Impact Evaluation Sample Plan

A primary objective of the study was an estimate of the pipe insulation realization rate, or the percentage of ex-ante therm impact claims achieved through measure installation. The goal of the sampling plan was to estimate the therm realization rate for pipe insulation installations in SCG service territory with 10 percent precision with a 90 percent confidence interval.

[17] One extra site was completed. There was uncertainty near the end of this project regarding the successful recruiting of the last required large site. Thus, a smaller site was substituted, and then the large site was successfully recruited and the analysis completed.

The sampling approach begins with model-based stratification by size, implemented in accordance with the California Evaluation Framework[18]. At the time of this sample design there was little data available on the dispersion of tank and pipe insulation realization rates. Consistent with guidance provided in the Evaluation Framework, an error ratio of 0.5 was assumed for sample design purposes. Though the Framework provides a range of acceptable values from 0.4 to 1, we selected 0.5 due to the relative simplicity of this technology and its applications.

The sampling unit was defined as a site, consistent with the overall goal of a therm realization rate for the pipe insulation measure. The population from which the sample was drawn was made up of 1,725 sites. Sites were ordered from greatest ex-ante savings claims to the smallest, and then divided into strata so that each group represented roughly one-fifth of total savings claims.

An assumed error ratio of 0.5 and a population of 1,725 sites yielded a required sample of 65 sites to achieve 90/10 precision. The 65 site sample was divided evenly across the five strata. The resulting sample stratification and associated quota objectives are shown below in Table 3-2. Also shown in Table 3-2 are actual completed on-site analyses, which match goals, with the exception of strata 2 where goals are exceeded by one.[19].

Table 3-2: Pipe Insulation M&V Sample Design

Stratum	Total Sites	On-Site Quota Planned	On-Site Quota Achieved
1	44	13	13
2	175	13	14
3	291	13	13
4	410	13	13
5	805	13	13
Total	**1,725**	**65**	**66**

[18] http://www.calmac.org/publications/California_Evaluation_Framework_June_2004.pdf. The Evaluation Framework Study states on page 335 "In most impact evaluation studies, the error ratio can be expected to be in the range 0.4 to 1.0…. If the tracking system is expected to provide quite accurate estimates of the actual savings of most sample projects in the evaluation study, then the error ratio is likely to be relatively small, e.g., near 0.4. This might be the case, for example, if the program provides energy efficiency retrofits to large commercial buildings, and the tracking estimates of savings are based on a fairly detailed analysis of each project that is undertaken in the program. If the tracking system is expected to provide rather poor estimates of the actual savings of most sample projects in the evaluation study, then the error ratio is likely to be larger, e.g., near 1.0. This might be the case, for example, if the program is an express-style program that requires only a simple application and does not provide any site-specific analysis as part of the program delivery."

[19] One extra site was completed within strata two. There was uncertainty near the end of this project regarding the successful recruiting of the last required strata 1 site. Thus, a strata 2 site was substituted. Then, the last strata 1 site was successfully recruited yielding an extra completed site for analysis.

Net-to-Gross Sample Design

Net-to-gross interviews were conducted across all three IOUs offering pipe insulation through the 2006-2008 program years. Available samples of pipe insulation customers within PG&E and SDG&E territory were somewhat limited, at 195 and 21 respectively, but larger within the SCG service territory, at 1,725 sites. Consistent with the *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*[20] objective samples were set at the minimum of 300 or 50 percent of the population. In accordance, a sample goal of 300 completes was set for the SCG service territory, and 50% of participation within the PG&E and SDG&E service territories (97 and 10, respectively). Although initially a stratified sampling plan approach was planned for the SCG territory, a census approach was required to reach the planned target number of completes.

Note on Steam Trap Participation among Pipe Insulation Participants

Steam traps were often installed in conjunction with pipe insulation. Eighty-seven percent of pipe insulation participant sites also had a record of steam trap participation. To ensure that both the Steam Trap HIM evaluation and the Pipe Insulation HIM evaluation had sufficient opportunity to meet sampling goals, a combined survey sample was developed and fielded with an instrument designed to meet both HIM evaluation objectives.

3.2.3 Gross Impact Evaluation Approach

This section addresses the gross impact estimation approach and algorithms developed for site-specific M&V. The methods presented in this section were developed in part based on a review of previous studies and publications. Impact estimates were initially calculated for seven pilot M&V points. Through the pilot process, data collection forms[21] were developed, the field measurement and verification process was fine-tuned, heat losses were calculated, and uncertainties were analyzed for various sites (dry cleaners, food processing, and commercial laundry). The field staff training, quality control processes, data entry system, and data management were all developed to handle the variety of sites, equipment, piping systems and environments that would likely be encountered in the field for this project.

The natural gas savings due to insulating pipes come from the reduced heat transfer out of the piping system after new insulation is added. Our methodology for calculating annual heat loss and total estimated energy savings comes from Chapter 25 of the 2001 American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHRAE) Fundamentals Handbook. Additionally, other adjustments were made to further improve the accuracy of our findings and enhance the on-site evaluation process. The primary calculation method and specifics regarding other adjustments follow.

[20] TecMarket Works, April, 2006

[21] On-site data collection forms are presented in Appendix A-4.

The energy savings due to the reduction in heat loss at the boiler(s) is calculated as the difference in heat transfer from the unit with and without the insulation using the American Society of Testing and Materials (ASTM) Standard C680 Practice for Determination of Heat Gain or Loss and the Surface Temperature of Insulated Pipe and Equipment Systems.

The equation utilizes the following variables to determine the annual energy impact in therms per year: hours per year (Hrs), surface and ambient temperatures (°F), and production levels.

The heat transfer equations presented below are used in the ASTM C680 standard:

hc = C * (1/d)^0.2 * (1/avg.temp)^0.181 * (Ts-Ta)^0.266 * (1+1.277(wind))^0.5

hr = ε * σ * (Ts^4 - Ta^4)/ (Ts-Ta)

where:

hc = convection surface heat transfer coefficient, Btu/h-sq.ft-°F

hr = radiation heat transfer coefficient, Btu/h-sq.ft-°F

d = diameter of cylinder, in. For flat surface and large cylinders (d>24), use d=24

avg. temp = average temperature of the air film, °F

Ts = unit surface temperature, °F

Ta = ambient air temperature, °F

wind = air speed, mph

C = constant depending on shape and heat flow condition, = 1.235 for longer vertical cylinders (closest match to current case), = 1.016 for horizontal cylinders (horizontal pipes)

ε = surface emittance, 0.12 for steel

σ = Boltzman's constant, 0.172E^-8, Btu/h-sq.ft-°F

From this, annual heat losses (AHL) resulting from the pipes were estimated using the heat transfer coefficients evaluated above, hc and hr, as follows:

AHL = A * (hc + hr) * (Ts-Ta) * H * UF * C1, (in MMBtu/yr)

where:

A = effective surface area, sq. ft

H = annual hours of operation

UF = fraction of operational time with flow in pipes (no units, 1.00 for 24/7 operation)

C1 = conversion constant, 10^-6 MMBtu/Btu

MMBtu = million British thermal units

<u>Adjustments Made to the Algorithm</u>

For steam and hot water piping systems heat transfer primarily occurs through convection and conduction. Either for simplicity or to err on the conservative side, both ASHRAE and the NAIMA 3E plus software tool assume negligible radiation effects. This assumption seemed accurate for indoor or sheltered piping systems, however, and the evaluation team wanted to see just how much solar gain and nighttime sky losses occur through radiative effects. We calculated that the solar gain or *insulation* (solar flux) represents up to 5%

additional heat transfer to the pipe or insulation surfaces depending on exposure time. Therefore, we added a solar gain component (SG) into the equation as follows:

SG = emissivity * surface area exposed to the sun * sun exposure factor * solar flux

The solar flux is a measure of how much energy from the sun is hitting the surface of the earth; it is an average exposure over the course of the year based on location. The solar flux data was obtained by using an online solar flux calculator, which takes the zip code for the location of the outdoor pipe system to obtain the global position of the pipes as it relates to the sun. The surface area that is exposed to the sun is assumed to be exactly half of the total surface area of the pipe or of the insulation, in other words a maximum of one half of the surface can be directly hit by the sun at any one time. The sun exposure factor took into account time periods where the pipes would be shaded or not directly see the sun as well as at what times the piping system was usually operational.
Therefore, the final annual heat loss (AHL) calculation that we utilized is shown below:

AHL = (A * (hc + hr) * (Ts-Ta) * H * UF * C1) - SG, (in MMBtu/yr)

To obtain annual therms saved, the following conversion was used:

AHL (therms/yr)= AHL (MMBtu/yr) *10

AHL was calculated for both the un-insulated or pre-existing condition (AHLpre) and the newly insulated condition (AHLpost). Then, using the efficiency of the boiler (EFFboiler), the estimated energy savings (ES) was calculated as follows:

ES = (AHLpre - AHLpost) / EFFboiler (therms/yr)

Field Data Collection Methodology

Table 3-3 below presents the key engineering parameters used to complete gross impact calculations and the data sources from which the information was collected. The measurement approach used for parameters is a combination of field observations, logger data, flue gas analysis, self-reported data, application data and independent third party sources. Descriptions of these parameters and how they were collected are presented in Appendix A-5.

Table 3-3: Key Measured Parameters Used in Gross Impact Calculations

Parameter	On-site Survey	Logger Data	Flue Gas Analysis	Independent Sources	Telephone Survey	Incentive Application
Temperatures	X	X				
Operating Hours	X	X				
Pipe sizes/lengths	X					X
New Pipe/Pre-existing Insulation	X				X	
Boiler efficiency	X		X			
Emissivity	X			X		
Wind/solar flux	X			X		

Baseline Selection

Program guidelines require that incented pipe insulation be installed on existing (i.e. not new) and bare pipe. The SCG program requirements state that:

"A minimum of 1inch of pipe insulation must be added to existing bare commercial or industrial steel pipe system applications. The bare pipe size must be at least ½ inch or larger. The following applications are not eligible: new construction, new pipe system replacement, fuel switching, residential."

Consistent with these guidelines, the baseline condition for all installations is bare pipe. Adjustments for pipe insulation projects that did not meeting program requirements are discussed next.

Adjustments to Gross Impact Claims Related to Program Qualifying Status

Gross therm impact claims associated with non-program qualifying pipe insulation installations are assigned an ex-post impact of zero. All installations –or portions of installations—found not to adhere to the terms of the program were assigned an ex-post gross impact of zero.

Data regarding the age of the pipe insulated through the program was collected as part of the telephone survey, and noted as part of the on-site M&V effort. The on-site data collection forms also recorded the sections of pipe that were insulated prior to the retrofit, based on interviews with on-site staff.

In addition to on-site data, telephone survey data was also collected regarding the age of pipe retrofit through the program and the presence of insulation prior to retrofit. In the event that data collected on-site do not match those collected through the phone interview, follow up inquiries were made to appropriately resolve any inconsistencies.

Gross Impact Estimation for PG&E and SDG&E Service Territories

On-site M&V activities were not completed for participating sites in PG&E or SDG&E service territories. Application of an SCG based UES result to PG&E or SDG&E would require specific data regarding the diameters, and temperatures of the pipes insulated through the program within these other service territories.

Similar to SCG, both PG&E[22] and SDG&E programs exclude from rebate eligibility *all pipes which*:

a. have insulation already present at the time of retrofit,
b. are new pipes or part of a new construction project.

Telephone surveys of pipe insulation participants within PG&E and SDG&E were asked to report on these characteristics in regards to the pipe that was insulated through the programs. Survey responses indicate that these program rules were often overlooked during program implementation. Adjustments to gross impact were made based on these data and the degree to which they indicated that these program rules were violated during the program year 2006-2008 cycles.

Two approaches were considered for adjusting PG&E and SDG&E ex-ante gross impact claims based on data regarding implementation of non-program qualifying measures. While two approaches were considered, only one was used.

The first approach leveraged telephone survey data regarding program qualifying status. Adjustments were made to the PG&E gross impact claim consistent with the self-reporting of non-qualified installations. Sites that self-reported during the telephone survey to have installed insulation *only on new pipes* were assigned a gross impact of zero.

Sites that reported having installed a portion of the total insulation on new pipe were assigned a corresponding portion of ex-ante gross impact. The reduction corresponded to the percentage of the insulation that was reported to have been installed on new pipe.

Sites reporting that insulation was present on the pipes prior to the installation of program incented insulation were also assigned a proportional downward adjustment based on an estimate of the percent of pipes at the site that had pre-existing insulation.

The second approach to adjusting PG&E and SDG&E gross impact claims leveraged data collected on-site in SCG service territory. The second method used the ratio of SCG program qualifying and verified impact to SCG verified impact (both program qualifying and not-qualifying). This ratio represents a *non-program qualifying penalty*, which can be applied to PG&E gross impact claims. This penalty does not adjust for gross impact algorithms, and assumptions, but only for non-program qualifying status.

[22] PG&E Nonresidential Mass Market, SDG&E Express Efficiency, and SDG&E Small Business Super-Saver.

This second method is attractive due to its foundation of on-site inspection data. Because of this, the underlying data are more detailed and reliable. If there is no perceivable difference in the probability of non-program qualifying installations in SCG territory versus PG&E/SDG&E service territory, than the SCG based estimate is more robust because of its foundation in on-site based investigation. For this reason we prefer (and ultimately select) the second method. However, we were careful to first use telephone survey data to examine evidence that rates of non-program qualifying installations are comparable between SCG and PG&E/SDG&E service territories.

3.2.4 Net-to-Gross Evaluation

Two separate methodologies were applied to the estimation of the net-to-gross ratio for the pipe insulation measure. Small commercial participants were subject to an approach consistent with *the "Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches and the Algorithm for the Residential and Small Commercial Consistent Free Ridership Method*[23]*"*. The large commercial and industrial customers were subject to an approach consistent with *"Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers.*[24] These two methodologies are discussed below:

Small Commercial Net-to-Gross Methodology

One objective of the California energy efficiency program evaluations is to identify the portion of savings directly attributable to the program effort, and to properly account for those effects that would have occurred in the absence of the program. California reporting protocols for the 2006-2008 program require the discounting of savings by a "free-ridership factor" in the estimation of net program savings by applying a net-to-gross ratio (NTGR). The 2006 Evaluation Protocols allow for the use of a participant self-report approach (SRA) to estimate the net-to-gross ratio for the basic level of rigor and with additional participant-specific documentation for the standard level of rigor.

The Energy Division convened a committee of ED staff, consultants and evaluators to develop a standard framework for the measurement of net-to-gross ratios[25] for residential and small commercial programs in a systematic and consistent manner using the SRA approach. The approach was designed to fully comply with the *Evaluator Protocols*. Energy Division and its consultants developed the *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches* to provide more detailed guidance than was available in the *California Evaluator Protocols*.

For purposes of this evaluation, participants who were involved in the decision-making process at each participating small commercial site were interviewed to measure the

---

[23] Authored by The Residential and Small Commercial Net-to-Gross Ratio Working Group. Please see Appendix A-2 for the full text.

[24] Authored by the Nonresidential Net-to-Gross Ratio Working Group. See Appendix A-3 for the full text.

[25] Currently, California net impacts are specified as net of free-riders and do not include either participant or non-participant spillover.

program's influence on the respondents' decision-making. The survey obtained highly structured responses concerning the probability that the firm would have installed the same measure(s) at the same time in the absence of the program. The survey also included open-ended and closed-ended questions that focused on the firm's motivation for installing the efficiency measure. These questions covered all the requirements provided in the *Guidelines*, such as multiple questions; efficiency level; likelihood of adoption; timing and quantity; and consistency checks.

The NTGR algorithm derived four separate measurements of free-ridership from different inquiry routes. The first measurement consisted of responses to a series of yes/no questions that measured the impact of the program on the quantity, efficiency, and timing of the purchase. The second measurement consisted of a 0-10 scale that asked the likelihood that the respondent would have purchased the same exact high efficiency measure in the absence of the program. The third measurement combined responses to the quantity and timing questions with responses to a 0-10 scale that asked the respondents' agreement with the statement that, in the absence of the program, they would have paid the additional rebate amount to buy the high efficiency equipment on their own. The final measurement combined responses to the quantity and timing questions with responses to a 0-10 scale that asked respondents' agreement with the statement that the program was a critical factor in their decision to purchase the high efficiency equipment. In cases where responses were inconsistent among the four measurements, an analyst reviewed responses to open-ended questions that asked for clarification of the inconsistency, and recoded the four measurements as needed.

These four measurements were averaged to derive the final free-ridership estimate at the measure level. Prior to finalizing the NTGR algorithm, the committee conducted iterative testing with a partial dataset. This testing contributed to the reliability of the algorithm and its computer coding.

Large Nonresidential Net-to-Gross Methodology

As part of the evaluation of the 2006-08 energy efficiency programs designed and implemented by the four investor-owned utilities and third parties, the Energy Division of the California Public Utilities Commission (CPUC) formed a nonresidential net-to-gross ratio working group that was composed of experienced evaluation professionals. The main purpose of this group was to develop a standard methodological framework, including decision rules, for integrating in a systematic and consistent manner the findings from both quantitative and qualitative information in estimating net-to-gross ratios for large nonresidential customers.

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 domain-level Net-to-Gross Ratios (NTGRs), since other available methods and research designs are not feasible for the types of Large Nonresidential Custom programs that were the subject of this evaluation. For example, in the industrial sector, three barriers are immediately apparent. First, there is an expected very small signal to noise ratio (low statistical power) in a participant/nonparticipant billing analysis i.e., the expected difference in monthly energy use between participants and nonparticipants is too small to detect reliably compared to other sources of variation in kWh that vary greatly across individual industrial sites. In addition, large industrial customers targeted by the program have been contaminated by participation in energy efficiency programs in prior years making it very difficult to find true nonparticipants. Finally, even if the first two problems were absent, the large industrial customers targeted by the program are each unique making it unlikely that one could find a group of nonparticipants that could be matched with participants on critical variables.

This SRA 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 Appendix A-3[26].

The method uses a 0 to 10 scoring system for key questions used to estimate the NTGR rather than using fixed categories that were assigned weights. It 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 helps to ensure that all non-program influences are taken into account in assessing the unique contribution of the program as reflected in the NTGR.

There are three levels of free-ridership analysis. The most detailed level of analysis, the **Standard – Very Large Project** 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.[27] 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. Evaluators exercise their own discretion as to what the appropriate thresholds should be for each of these three levels.

Data Sources. 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.

---

[26] Appendix A-3 contains the detailed Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers, which includes a demonstration of how this methodology complies 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).

[27] 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.

Table 3-4 below shows the data sources that are 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 3-4: 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.

NTGR Questions and Scoring Algorithm. 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. Program influence through vendor recommendations is also incorporated in this score.

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 most important 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 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 both the Timing and Selection and No-Program scores, the maximum score is always used. The rationale for using the maximum value is to capture the most important program element in the participant's decision making. Thus, each 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.

When there are missing data or 'don't knows' to critical elements of each score, one of two options is used. The missing element may be backfilled with a value that represents the average of the lowest and highest extreme values. Alternatively, if it is one of several other elements that are considered in the algorithm, the missing element may simply be excluded from consideration.

The self-reported core NTGR is simply the average of the Program Influence, Timing and Selection, and No-Program Scores, divided by 10.

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, all of this information is integrated into an internally consistent and coherent story that supports a specific NTGR.

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[28].

Once the individual analysts have completed their review, they discuss their respective findings and present their respective rationales for any recommended changes to the Calculator-derived NTGR. The outcome of this discussion is the final NTGR for a specific project.

28 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.

Application of Net-to-Gross Methodologies to the Pipe Insulation HIM Evaluation

As described above, two methods were applied to estimate the pipe insulation net-to-gross ratio: the Joint Simple Self Report Net-to-Gross Ratio Approach, and the Large Nonresidential Net-to-Gross Methodology.

Pipe insulation was installed through the 2006-2008 IOU rebate programs in both small commercial, large commercial, commercial chains and industrial facilities. Sites were assigned to these two net-to-gross evaluation methodologies as a function of several key attributes. Specifically, sites were allocated to the "Large Nonresidential Net-to-Gross Methodology" where the following criteria held:

- Receipt of more than $50,000 in rebates[29]
- A corporate or chain account with 10 or more member sites participating
- Installation of "industrial steam traps"

The first two criteria are associated with large facilities and more sophisticated decision-making processes. The last of these criteria arises in part from the tremendous overlap in participation between pipe insulation and steam trap measures. Nearly 90 percent of the pipe insulation participant population also installed incented steam traps through the 2006-2008 IOU programs. The installation of "industrial steam traps"- as opposed to commercial steam traps - involves a more sophisticated approach to program delivery[30] than commercial steam traps. These differences are substantial enough to consider them indicative of a more sophisticated customer and decision making process. Thus, these sites are allocated to the "Large Nonresidential NTG Methodology."

Sites not meeting the above criteria were subject to the Joint Simple SR approach, established for residential and small commercial customers. This method is designed for the following situations:

- The program or measure addressed involves a fairly uniform technology and application across end-use participants.
- Free ridership can reasonably be addressed at the end-user level. That is, the major influence of the program has been through means such as rebates or promotions that the participating decision-maker is likely to have been aware of at the time of the purchase decision. In these cases, influence of the program on vendors or other intermediaries is assumed not to be a major factor in ultimate purchase decisions.

[29] Given the significant overlap between steam trap and pipe insulation, the rebate level to determine NTG methodology were applied to the combination of the rebate from these two measures over the three-year program cycle.

[30] See Section 4 (Steam Traps) for a more complete description of the indicated differences and implications on customer decision-making processes.

The first criterion holds for pipe insulation installed by commercial sites, as the variability in these applications is limited. The second criterion also holds true for commercial sites, where rebates were granted to the customer and used as a primary sales and promotional tool by the vendor.

There is some degree of pipe insulation customization for the net-to-gross survey instruments. The core net-to-gross questions are maintained. However, lines of questioning regarding standard efficiency equipment are removed. Partial free ridership questions instead center on quantities and insulation thickness.

Extra questions were added to the survey to determine the circumstances of the pipe insulation installations. These lines of questioning revolved around the presence of pipe insulation prior to the retrofit, as well as the age of the pipe insulated through the program. Both insulated pipe and new pipe are disallowed by program guidelines. Sites with these factors require a customized weighting when aggregating net-to-gross scores into a program or measure-level result. The weighting technique is described in greater detail below and is designed to ensure the aggregate net-to-gross result represents the most appropriate adjustment to ex-post gross therm impact.

Weighting

This section describes the development of weights applied to each sample point when aggregating results.

Gross Impact Weights

The weighting method applied to the gross impact sample is consistent with the design methodology. Recall the sample is pulled from the SCG participant population, and the population of participating sites is divided into five strata. Sites are ranked from the greatest ex-ante therm impact to the smallest, and then size cutoffs are designed such that the total savings in each strata represents an equivalent portion of total ex-ante gross impact. The weight assigned to each point in the gross impact sample is a combination of the following statistics:

Site_thm(ij): Ex-ante therm impact of site (i) belonging to strata (j)

Samp_strat_thm(j): Total ex-ante therm impact of the sampled sites from strata (j)

Pop_strat_thm (j): Total population ex-ante therm impact from strata (j)

Pop_thm: Total population ex-ante therm impact

Weight(ij) = [Site_thm(ij)/Samp_strat_thm(j)] * [Pop_strat_thm(j)/Pop_thm]

Where Weight(ij) is the weight applied to site (i) from strata(j)

<u>*Net-to-Gross Ratio Weights*</u>

SCG Participants
The net-to-gross sampling methodology was random – based on a census approach to the population. Gross impact assignments have a significant effect on the relevance of the net-to-gross result. More specifically, sites with non-program qualifying installations have net-to-gross ratios that should not inform the net-to-gross adjustment applied to realized gross impacts. For this reason we apply a weighting scheme reflective of the ex-post gross impact.

Realization rates are assigned to each site in the SCG net-to-gross sample as follows:

For sites included in the gross impact M&V sample, the site specific realization rate is invoked.

Sites not in the gross impact M&V sample are divided into three categories based on telephone survey data:

a) Site was 100% new construction or a new pipe retrofit
b) There was pre-existing insulation present at time of retrofit
c) Neither a) nor b) hold

Realization rates resulting from the gross impact M&V work are calculated for each of these three categories using weighting methods described above.

The assigned realization rate is interacted with the ex-ante therm impact to yield the weight applied to the net-to-gross score.

The need for this weighting approach is also supported by the underlying relationship between free ridership rates and the attributes that result in program disqualification -- more specifically, the program rule disallowing insulation on pipe installed in new construction or addition applications. Insulation of new pipes is considered industry "standard practice" and thus would also have a higher expected free ridership rate. Similarly, a customer that has pre-existing insulation is likely to repair or replace that insulation naturally and without program intervention. To apply the net-to-gross reports from these sites to a gross impact that already sets impacts under these conditions to zero would be redundant and inaccurate.

PG&E/SDG&E Net-to-Gross Ratio Weight
The PG&E and SDG&E net-to-gross survey was fielded as a census. Given this random sampling, a gross impact weight was applied to the resulting net-to-gross scores. Similar to the gross impact methodology, two approaches were considered for assignment of weights to PG&E/SDG&E net-to-gross ratio survey respondents.

In the first method, the *ex-post* gross impact is invoked as the weighting metric. This ex-post impact is the result of adjustments made based on self-reported data regarding program qualifying status. This approach has the benefit of emphasizing those respondents that installed program qualifying measures. The second approach is to simply invoke an ex-ante gross impact weight. This has the advantage of circumventing the uncertainty around the accuracy of the self-reported data regarding program qualifying status.

In selecting one of these alternative approaches to weighting, we consider the method selected for gross impact adjustments. For example, if gross impact adjustments were based on self-reported telephone survey data, it might be preferable to use an ex-post gross impact weighting since it will emphasize decision making surrounding program qualifying installations. On the other hand, and as happens to be the case, the on-site work in SCG territory is leveraged to provide gross impact adjustments, then the ex-post impact estimate for each point in the telephone survey sample is a uniform percent of ex-ante. This yields an equivalent weighting scheme to using ex-ante gross impact. Reasons behind the selection of the latter method are discussed in more detail below in the Section 3.4.

3.3 Validity and Reliability

This evaluation took steps to increase both the validity and reliability of measurement for each of the parameters estimated. The evaluation worked to minimize response bias for survey based results and recruitment and undertook uncertainty analyses of gross impact calculations.

3.3.1 Minimizing Response Bias for Survey Based Results and Recruitment

The evaluation conducted telephone surveys to support the estimation of both ex-post gross impact and net-to-gross ratios. Surveys were conducted through Itron’s Computer Aided Telephone Interview (CATI) Center, with the exception of one large industrial site interview which was conducted by an energy consultant. A key step to ensuring the validity and reliability of the pipe insulation analysis was to minimize non-response bias. One specific step taken was to attempt to contact a respondent multiple times at different times of the day, different days of the week, and different weeks of a month. For commercial and industrial customers, the research team called during normal business hours. Callbacks were also scheduled for respondents at a time that was most convenient for them. For larger industrial customers, the team also spoke with their utility representatives to ensure that the correct contact name and phone number for the site was being used. In cases where an industrial applicant was hesitant to speak with Itron, the utility representative was asked to intercede on the evaluation team’s behalf. Finally, incentive gift cards were offered to respondents to encourage those very reticent to participate.

3.3.2 Net-to-Gross Ratio Estimation

Net-to-gross ratio estimation was based on self-report analyses utilizing telephone surveys for both the commercial and industrial applications. The net-to-gross telephone survey instruments are designed to produce valid and reliable net-to-gross ratio results. During the pre-test of the NTGR survey instrument, reliability tests are conducted using the CATI software. Any problem areas were detected and corrected at this time.

Telephone surveys sought out the key decision-maker to respond to the net-to-gross batteries. These batteries included multiple questions to develop the scores that were designed to increase the reliability of the results. In cases where responses were inconsistent across the multiple questions, the respondent was asked to clarify their response, and given a chance to change a previous response. If the response patterns remained inconsistent, an experienced energy consultant reviewed the case. The consultant reviewed responses to open-ended questions that asked for clarification of the inconsistency and recoded the individual responses to resolve the inconsistency.

3.3.3 Gross Therm Impact Measurement Validity and Reliability

Formal uncertainty analysis was completed on a selection of representative sites to quantify the validity and reliability of the site specific gross impact evaluation results presented in this report. The analysis is performed to understand the major contributors, and range of uncertainty in the site-specific measurement of therm impact from the installation of pipe insulation. The data collection and impact calculation methods are the same for every site, and so not all sites needed to be analyzed to determine which variables have the largest contribution to the overall measure level error.

The uncertainty analysis was performed using Crystal Ball software to run a Monte Carlo simulation to propagate the uncertainty through the energy savings calculation. Monte Carlo is the most appropriate method to use because the calculation is multi-faceted and the mathematics that would be required otherwise would be unduly complex.

The variables with the largest measurement uncertainty are the pre-retrofit pipe surface temperatures because they are based on estimates as opposed to data that can be measured directly. The higher temperature pipes have a larger effect on the calculation than the lower temperature pipes because of the dependence of heat transfer on temperature difference.

Table 3-5 shows a summary of uncertainty results for a pipe insulation site analyzed with Monte Carlo analysis. The final therm impact value has a relative margin of error of 9 percent at a 90 percent confidence interval. This value meets the minimum uncertainty requirements for the evaluation process.

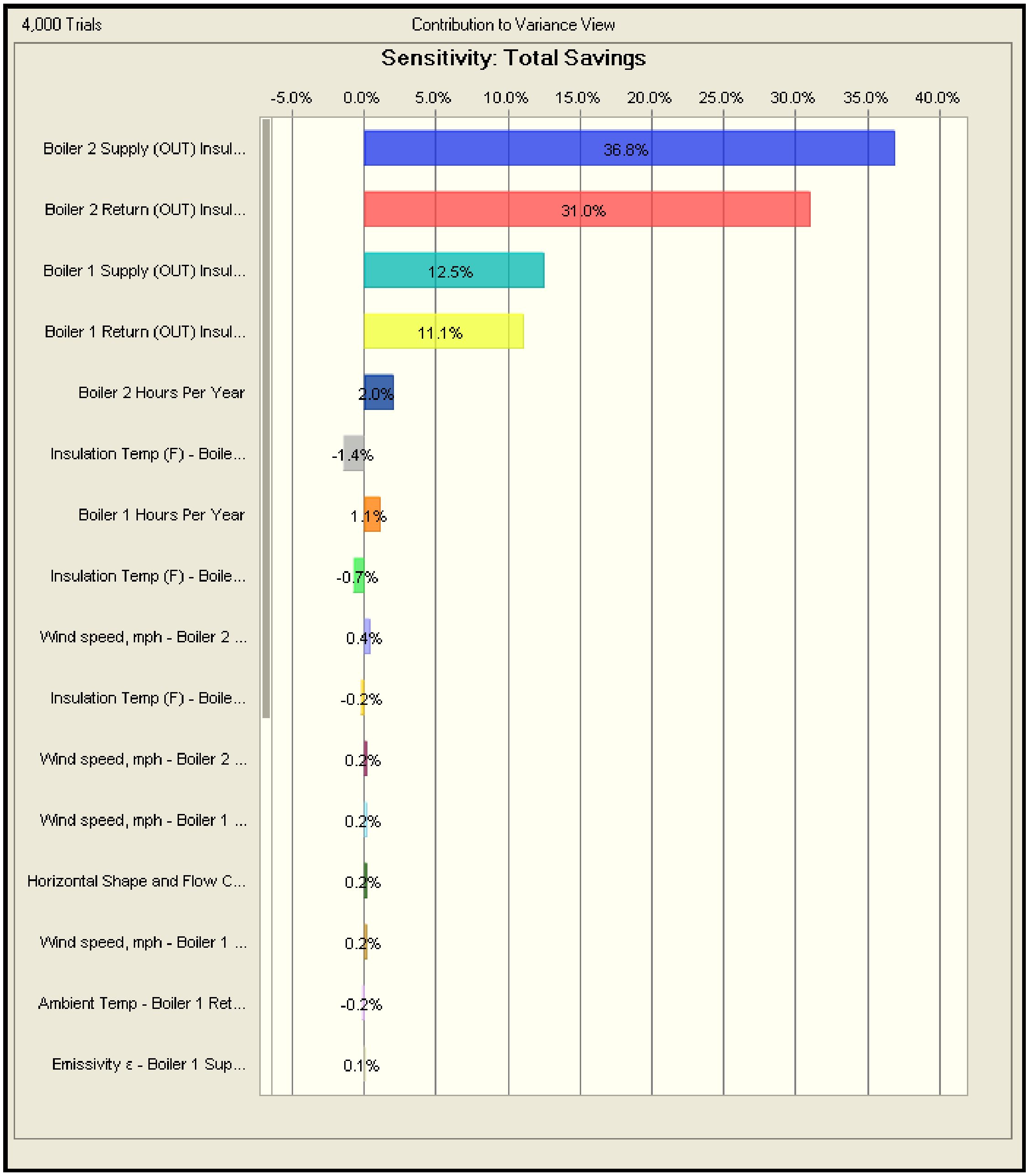
Table 3-5: Summary of Uncertainty Results for a Pipe Insulation M&V Site

Ex Post Annual Energy Savings Summary	Therms/yr	Standard deviation	Margin of Error at 90% CI	Relative Margin of Error at 90% CI
Pre retrofit therm consumption for piping AHL	34,231.44	1,525.22	2,508.99	7%
Post retrofit therm consumption for piping AHL	7,070.24	309.21	508.66	7%
Total Savings	26,161.01	1,438.90	2,366.98	9%

To understand how to reduce the relative margin of error around the estimated therm impact, it is useful to look at an error sensitivity chart of the total savings measurement. Figure 3-1 is a sensitivity chart for the site summarized in Table 3-5 above.

Figure 3-1 below shows the effect of the relevant parameters on the precision of the impact calculation based on the uncertainty assumptions and after 5,000 trials of the Monte Carlo simulation. The larger the percentage shown, the more significant is the contribution of the variable's uncertainty to the overall variance in the result. The values in the vertical axis represent cells in a Microsoft Excel spreadsheet used to perform the simulation.

Figure 3-1: Sensitivity to Variance of Total Savings



The first four bars represent pre-retrofit pipe temperatures for different sections of pipe. These variables had the largest measurement uncertainty and therefore are also the largest contributors to the variance in the savings calculation results. Because the sensitivity of the rest of the variables is at least one order of magnitude less, we can conclude that the pre-retrofit pipe temperature is the most significant contribution to error in the savings calculation.

3.4 Findings and Results

This section presents the important results and findings from the study of pipe insulation installed through the California IOU programs through the 2006 - 2008 Program Years. The section begins with a review of the tracking system contents. The review is intended to provide some context for key study results. The tracking system summary is followed by a review of findings from the on-site M&V work completed in SCG service territory. Gross impact findings for PG&E and SDG&E are discussed next. Net impact results are then discussed for each service territory, and finally, the gross impact findings are combined with net-to-gross ratio results to produce final results for each service territory.

Tracking System Summaries

Table 3-6 below shows the SCG tracking system summary of pipe insulation participation through the 2006 – 2008 program years. There are 10 unique measure descriptions related to pipe insulation in the Express Efficiency tracking database[31]. The per unit or per linear foot therm impact claims vary widely by measure. The first two listed are expressed in gross therm impact per square foot instead of per linear foot, and so are not readily comparable to the others. The remaining eight measure descriptions are expressed in linear feet and vary from a low of 2.8 therms per linear foot for *"Pipe Insulation -Hot Water Application < 1" pipe"* to a high of 63 therms per linear foot for *"Pipe Insulation - Medium Pressure Steam >15 psi >= 1" pipe"*.

Nearly three-fourths of SCG gross impact claims for pipe insulation measures are within the *"Pipe Insulation - Medium Pressure Steam >15 psi >= 1" pipe"* measure. This statistic is driven by the higher per unit therm impact claim (63 therms per linear foot). The greatest number of linear feet are found in the measure, *"Pipe Insulation - Medium Pressure Steam <=15 psi < 1" pipe"* which has an associated ex-ante gross therm impact claim of 9.6 therms per linear foot. This measure accounts for 16 percent of total gross impact claims.

In all, the SCG tracking system shows over half a million linear feet of pipe insulation installed, and an average per unit gross impact claim of 28 therms.

[31] SCG delivered pipe insulation through the Express Efficiency Program only, during the 2006 through 2008 program cycles.

Table 3-6: SCG Tracking Database Summary, Pipe Insulation

Tracking Database Measure Name	Percent of Gross Impact Claim	Units Installed	Ex-Ante Therms Per Unit
Pipe Insulation - Hot Water Applic. (sq ft) 1 inch	0%	9,685	2.6
Pipe Insulation - Hot Water Applic. (sq ft) 2 inch	0%	9,440	2.9
Pipe Insulation - Low Pressure Steam Applic. (LF) 1 inch	1%	8,941	13.4
Pipe Insulation - Low Pressure Steam Applic. (LF) 2 inch	1%	6,034	14.3
Pipe Insulation - Low Pressure Steam <=15 psi < 1" pipe (LF)	1%	19,184	6.1
Pipe Insulation - Low Pressure Steam >15 psi >= 1" pipe (LF)	4%	17,377	40.0
Pipe Insulation - Medium Pressure Steam <=15 psi < 1" pipe (LF)	16%	274,882	9.6
Pipe Insulation - Medium Pressure Steam >15 psi >= 1" pipe (LF)	72%	188,122	63.0
Pipe Insulation -Hot Water Applic. < 1" pipe (LF)	0%	1,092	2.8
Pipe Insulation -Hot Water Applic. >= 1" pipe (LF)	5%	46,432	18.0
Total	**100%**	**581,188**	**28.2**

The PG&E tracking database houses five unique measure descriptions associated with pipe insulation, as shown in Table 3-7 below. Per unit claims range from a low of 2.6 therms per linear foot for "*Pipe Insulation – Hot Water – One Inch*" to a high of 14.3 therms per linear foot for "*Pipe Insulation-Low Pressure Steam – Two Inch*". Nearly 90 percent of ex-ante gross therm impact is housed in the *"Pipe Insulation- Low Pressure Steam – One inch"* measure, which has a per linear foot gross impact claim of 13.4 therms.

Overall, PG&E has a gross impact claim of 934 thousand therms, associated with nearly 90 thousand linear feet of pipe insulation.

Table 3-7: PG&E Tracking Database Summary, Pipe Insulation

Tracking Database Measure Name	Percent of Gross Impact Claim	Linear Feet Installed	Ex-Ante Therms Per LF
Pipe Insulation	0%	20	6.74
Pipe Insulation – Hot Water – One Inch	5%	19,727	2.6
Pipe Insulation – Hot Water – Two Inch	2%	5,125	2.9
Pipe Insulation-Low Pressure Steam – One Inch	88%	61,332	13.4
Pipe Insulation-Low Pressure Steam – Two Inch	5%	3,198	14.3
Total	**100%**	**89,402**	**10.4**

As shown in Table 3-8 below, the SDG&E tracking database houses two measure descriptions associated with pipe insulation: *"Pipe Insulation - Hot Water Applic. (sq ft) 1 in"* and *"Pipe Insulation - Low Pressure Steam Applic. (LF) 1 in."* The first represents 6 percent of total ex-ante gross impact claims, at a per square foot impact claim of 2.6 therms. The former measure houses 94 percent of pipe insulation gross claims, with a per linear foot impact claim of 13.4 therms and total installed linear feet of 8,929.

Table 3-8: SDG&E Tracking Summary

Tracking Database Measure Name	Percent of Gross Impact Claim	Units Installed	Ex-Ante Therms Per Unit
Pipe Insulation - Hot Water Applic. (sq ft) 1 in	6%	2,685	2.6
Pipe Insulation - Low Pressure Steam Applic. (LF) 1 in	94%	8,929	13.4
Total	**126,630**	**11,614**	**10.9**

Measure Concentration in the Dry Cleaner Business Segment

All three IOUs achieved a large portion of total pipe insulation ex-ante gross therm savings through the dry cleaning business segment, as shown in Table 3-9 below. For both SCG and PG&E, dry cleaners account for about 80 percent of participating sites and between two-thirds and three-fourths of total ex-ante gross impact claims. Concentration within the SDG&E participant population is smaller, with about half the sites and one-fifth of total ex-ante claims.

Table 3-9: Pipe Insulation Ex-Ante Impact Summary: Dry Cleaning Business Type

Dry Cleaners	SCG	PG&E	SDG&E	Total Statewide
Percent of Sites that Are Dry Cleaners	78%	83%	52%	78%
Percent of Ex-Ante Therm Impact from Dry Cleaners	64%	74%	22%	64%
Total Sites (All Business Types)	1725	205	21	1951
Total Ex-Ante Gross Therm Impact (All Business Types)	16,400,122	934,033	126,630	17,460,784

3.4.2 Gross Impact Findings

Engineering analysis was performed on a total of 66 sites where pipe insulation was installed within the SCG service territory. For each site, therm impact was calculated against a bare pipe baseline. Then, for cases where insulation installed through the program was determined to violate program qualifying criteria, gross impact was reduced accordingly. Impact for pipes that did not qualify for program incentives are set to zero. Remaining pipe is assigned impact versus a bare pipe baseline.

As shown in Table 3-10 below, on-site data collected from 66 sites, and representing 54,414 linear feet of pipe insulation, produced a realization rate of 7.9 percent, and a corresponding 90 percent confidence interval that extends from 7.4 to 8.4 percent.

Table 3-10: Summary of SCG Gross Impact Realization Rate Findings by Program Qualifying Status

Result	Total	New Construction	Partially New Pipe	Pre-Existing Insulation	Program Qualifying
Realization Rate	7.9%	0.0	12.1%	1.7%	25%
Sites	66	3	2	43	18
Linear Feet	54,414	21,650	3,759	17,394	11,610
Percent of Linear Feet	100%	40%	7%	32%	21%
90 Percent Confidence Interval Lower Bound	7.4%	-	2.0%	1.5%	22.4%
90 Percent Confidence Interval Upper Bound	8.4%	-	22.3%	1.8%	27.9%
Margin of Error	0.5%	-	10.0%	.01%	2.8%
Relative Margin of Error	6%	-	84%	8.7%	11%
Error Ratio	0.29	-	0.72	0.35	0.28

A significant portion of sites (43 of 66) had insulation prior to the program incented retrofit. This was a more frequent finding at the smaller retrofits – the sites with pre-existing insulation were about 65 percent of sites, but just about 32 percent of the total linear feet in the sample. Three of the largest installations were identified as new construction, representing less than 5 percent of sites, but 40 percent of the linear feet inspected. These two segments produced very low realization rates, at 1.7 and 0.0 percent, respectively.

Twenty-seven percent of the 66 sites analyzed were determined to have program qualifying installations. These 18 sites represent 21 percent of the total linear feet and have an estimated realization rate of 25 percent. Two sites were partially new construction or new pipe. These sites have a realization rate of a little more than 12 percent, and represent 7 percent of the total linear feet studied.

The moderate size of the realization rate for program qualifying sites is driven largely by finding lower than assumed operating hours. SCG pipe insulation work papers[32] state an assumed annual operating time of 7,752 hours per year based on the assumption that steam

32 Prepared for Southern California Gas Company in July 2006 by Energy and Environmental Analysis, Inc. titled Pipe Insulation (Non-space Conditioning) Work paper for PY2006-2008, B-REP-06-599-03B Revision B

systems operate under pressure for 24/7, except for six weeks per year for plant maintenance. Dry cleaners often have much more moderate operating hours. The average over the 47 dry cleaners in the M&V sample was about 2,400 hours per year. The average among the other business types also fell short of this mark, at 4,964 hours per year.

Another area where the work paper assumptions are not supported by site investigations is in the assumed ambient or environmental conditions surrounding the pipe. The assumed environmental conditions were taken from ASHRAE literature as stated on page 3:

"ASHRAE uses an ambient temperature of 65 °F and 7.5 mph wind speed for their tables of recommended thicknesses for pipe insulation."

These values would seem appropriate to assume if the pipes are operating in outdoor conditions. If the pipes were outside, then the temperature and wind speed estimates would be fairly accurate, if not even slightly conservative in terms of predicted heat loss and energy savings. However, the vast majority of the sites that were surveyed did not have any outside pipes and had much higher ambient temperatures and no wind speed in the area around the piping system. The higher ambient temperature and lack of wind causes less heat loss from the pipes and results in less energy saved by insulating them.

Sites studied in this evaluation had high ambient temperatures, and little outdoor pipe. The high ambient temperatures resulted from enclosed work spaces with machinery and equipment that produces ambient heat. Dry cleaners typically operate with very high ambient temperatures, with an average measured ambient temperature around insulated pipes of about 90 degrees. Other business types studied also had high ambient temperatures, with a mean ambient temperature value for non-dry cleaners of 81 degrees.

<u>Discussion of Dry Cleaner Results</u>

As noted previously, the dry cleaners segment accounts for 78 percent of pipe insulation sites in the SCG tracking system and about 64 percent of total ex-ante gross impact therm claims. The dry cleaner segment has a lower realization rate than other business type segments (4.6 versus 15.3 percent). This is due to a combination of many factors, including finding lower-than-assumed operating hours, and higher-than-assumed ambient air temperatures. In addition, there was a high likelihood of pre-existing insulation at these sites.

Setting aside the issue of program qualifying status, the laundries achieve just 12 percent of the ex-ante therm impact claim. That is, the realization rate would be 12 percent if we calculated the gross impact relative to bare pipe on all dry cleaner installations, not just program qualifying. In addition, there is a substantial portion of laundry sites with non-program qualifying installations. Thirty-four of the 47 laundries in the on-site M&V sample had pre-existing insulation before the retrofit. The final gross impact realization rate for dry cleaners in SCG service territory is 4.6 percent. The final gross impact realization rate for other business types is 15.3 percent, as shown in Table 3-11 below.

Table 3-11: SCG Gross Impact Realization Rate Findings: Dry Cleaners Versus Other Business Types

Segment Description	Sites	Realization Rate	Lower Bound	Upper Bound	Margin of Error	Relative Margin of Error	Error Ratio
Dry Cleaners	47	4.6%	4.4%	4.8%	0.22%	5%	15%
Other Business Type	19	15.3%	13.8%	16.8%	1.5%	10%	26%
Total	**66**	**7.9%**	**7.4%**	**8.4%**	**0.5%**	**6%**	29%

Gross Impact Findings – PG&E Service Territory

In some ways the SCG and PG&E Express Efficiency pipe insulation programs appear similar. The business type distribution within PG&E is similar to SCG. Both have high proportion of dry cleaners, which account for 86% of PG&E participating sites and 78% of SCG participating sites. However, neither SDG&E nor PG&E allow pipe with less than one-inch diameter to be insulated through the program, while SCG does allow pipe less than one inch diameter[33].

Thirty-eight surveys were completed with PG&E pipe insulation participants. Telephone surveys collected basic information regarding the insulated pipes, including the age of the pipes insulated through the program and whether there was insulation present on the pipes prior to retrofit. Self-reported data, along with the percent of surveyed sites and the associated gross therm impact is shown in Table 3-12 below. Table 3-12 also compares the PG&E distribution with the on-site sample from the M&V work completed in SCG territory.

Among the 38 PG&E participant respondents, 21 reported there was insulation present on the pipes prior to program-incented retrofit; 12 reported the insulation was installed on new pipes only; and 2 stated that part of the insulation was installed on pipes that were new. Of the remaining 3 respondents, one could not provide information regarding the age or insulated condition of the pipes prior to retrofit, and 2 confirmed program qualifying status.

Relative to on-site finding for the 66 SCG sites, PG&E reported a higher incidence of non-program qualifying installs, both on a site basis (87 versus 70 percent) and ex-ante gross therm claim basis (83 versus 68 percent).

[33] These smaller pipe diameters were commonly found at the dry cleaning sites, where the vertical pipe "drop downs" typically had a one-half inch diameter.

Table 3-12: PG&E Telephone Survey Sample Summary

Segment Description	PG&E Participant Respondents	Gross Therm Impact	Percent of Sites	Percent of Gross Impact	Percent of Sites (SCG)	Percent of Gross Impact (SCG)
Program Qualifying	2	13,389	5%	8%	27%	27%
New Construction	12	76,668	32%	45%	5%	31%
Pre-Existing Insulation	21	64,069	55%	38%	65%	37%
Partially New Construction	2	13,440	5%	8%	3%	4%
Don't know/Refused	1	1,514	3%	1%	0%	0%

If we examine only the dry cleaners in the PG&E telephone survey sample, the distribution of non-program qualifying installations is similar to SCG. As shown in Table 3-13 below, among the PG&E respondents 75 percent reported pre-existing insulation, while 72 percent of the SCG on-site sample had pre-existing insulation. These sites represent similar portions of total ex-ante impact claims in each sample, 72 percent of PG&E and 74 percent of SCG. A difference is apparent regarding insulation on new pipe or new construction, which is reported by 3 of the dry cleaners in PG&E service territory, but none in SCG territory.

Table 3-13: PG&E Telephone Survey Sample Summary: Dry Cleaners

Segment Description	PG&E Participant Respondents	Gross Therm Impact	Percent of Sites	Percent of Ex-Ante Gross Impact	Percent of Sites (SCG On-Site Sample)	Percent of Ex-Ante Gross Impact (SCG On-site Sample)
Program Qualifying	1	6,512	4%	8%	28%	26%
New Construction	3	11,846	13%	15%	0%	0%
Pre-Existing Insulation	18	55,563	75%	72%	72%	74%
Partially New Construction	1	1,688	4%	2%	0%	0%
Don't know	1	1,514	4%	2%	0%	0%

As discussed in the Methodology Section above, two approaches were considered for adjusting PG&E gross impact pipe insulation claims. The first utilizes the self-reported data regarding program qualifying status.

- The 12 respondents that report insulation was installed on new pipes are assigned an ex-post impact of zero.

- The two with partially new construction are assigned a proportional reduction corresponding to the percent of pipes reported to be new[34].

- The 21 respondents indicating that there was insulation present are assigned a reduced gross impact. Telephone survey data does not support a self-report based estimate of the percent of pipe insulated at the site that had pre-existing insulation. However, the site work performed in SCG territory confirms that for cases where pre-existing insulation was present, it was present on nearly all the pipes. Among the sites with insulation prior to retrofit, an average of 89 percent of the pipes insulated through the program had pre-existing insulation. In accordance, a downward adjustment of 89 percent is applied.

This approach results in a gross impact claims reduction of 80 percent. That is, this approach results in a gross realization rate of 20 percent.

The second approach uses the more precise data collected on-site in SCG service territory. Further, it accepts the ex-ante gross impact estimates of PG&E, and only applies a proportional adjustment equal to the "non-program qualifying" penalty measures in SCG service territory. The adjustment is the ratio of SCG program qualifying ex-post measured impact to all SCG ex-post measured impact.

The second approach yields a gross impact downward adjustment of 65 percent. The gross impact realization rate is 35 percent, with a 90 percent confidence interval of 1.4 percent, extending from 33.5 to 36.4 percent.

This method is somewhat conservative when applied to PG&E, as PG&E reports higher levels of non-program qualifying installations. However, data collected in SCG service territory shows that while the telephone data was generally confirmed by on-site data, there were some cases where the two sources did not line up. Among the 66 on-sites, 43 were noted to have had some insulation prior to retrofit. Comparing these results to the phone survey data, 5 sites reported having insulation prior to retrofit that was not confirmed through on-site and telephone follow up work. In addition, there were 20 cases where pipe was found to have been insulated prior to retrofit, but this was not reported over the phone. The new construction data was relatively consistent across the phone survey and on-site sources for all but one case, where a large portion of the pipe insulated was new but not reported as such over the phone. The three sites found to be 100 percent new construction were consistent in the telephone survey reports.

Both adjustment factors were calculated and considered. Due to a high propensity among PG&E respondents to report non-program qualifying installations, the latter approach was found to yield the higher gross-impact realization rate and was used in the final adjustment to PG&E gross impact claims. This approach has the benefit of being founded on site specific engineering findings. Also, the evidence provided by the PG&E survey data indicates that

[34] One of these sites reported ten percent was installed on new pipe, and one site reported twenty percent.

non-program qualifying installations were at least as frequent within PG&E service territory as they were in SCG service territory.

Gross Impact Findings – SDG&E Service Territory

Only two survey completes were attained in SDG&E territory. One of these respondents reported there was insulation present on the pipes prior to the program incented retrofit. The second respondent reported that all of the pipes insulated at the site were new. Despite these findings, two surveys are simply not enough data to support an adjustment to ex-ante claims.

3.4.3 Unit Energy Savings Findings

Data collected from the 66 on-sites was analyzed based on a bare-pipe baseline to produce estimates of therm savings for each pipe run, regardless of program qualifying status. Results are summarized in Table 3-14 below.

Separate tabulations are done for dry cleaners versus other business types, due to the similarity in operating hours and conditions within that business types. Average therm impact per linear foot is shown for dry cleaners and non-dry-cleaners, for various temperature and pipe diameter ranges. Ninety percent confidence intervals surrounding these results are also shown.

Dry cleaning establishments have measured impact between 1.8 and 8.0 therms per linear foot. None of the dry cleaning facilities in the sample had pipe diameters in excess of 3 inches. Other establishments were more likely to have larger diameter pipes. The highest measured impact is 88 therms per linear foot, for the highest temperatures range (300.1 to 400 degrees Fahrenheit) and greatest diameter category (greater than 3 inches).

Table 3-14: Unit Energy Savings Results from SCG Service Territory On-Site M&V Work

Dry Cleaners	Pipe Size Category (diameter)	Temp. Category (Fahrenheit)	Linear Feet Sampled	Sites Sampled	Therms per Linear Foot	90% CI Lower Bound	90% CI Upper Bound	Precision
No	< 1 inch	<=200	1,672	3	1.0	0.6	1.5	43%
No	< 1 inch	200.1-300	1,553	11	3.5	3.3	3.8	6%
No	< 1 inch	300.1-400	210	2	10.1	9.5	10.8	7%
No	1-3 inches	<=200	14,354	5	3.9	3.6	4.2	8%
No	1-3 inches	200.1-300	5,422	11	16.0	15.2	16.8	5%
No	1-3 inches	300.1-400	4,589	12	31.2	29.0	33.4	7%
No	> 3 inches	<=200	5,310	4	13.6	11.2	16.0	18%
No	> 3 inches	200.1-300	4,202	6	34.2	29.4	39.0	14%
No	> 3 inches	300.1-400	1,041	5	88.0	75.6	100.5	14%
Yes	< 1 inch	<=200	1,720	13	1.8	1.8	1.9	3%
Yes	< 1 inch	200.1-300	7,003	47	2.6	2.5	2.6	1%
Yes	< 1 inch	300.1-400	1,017	20	4.7	4.6	4.8	2%
Yes	1-3 inches	<=200	910	7	4.1	3.8	4.4	8%
Yes	1-3 inches	200.1-300	2,019	19	4.7	4.6	4.9	3%
Yes	1-3 inches	300.1-400	3,393	34	8.0	7.8	8.1	2%

3.4.4 Net-to-Gross Results

This section summarizes the net-to-gross ratio estimation results for the pipe insulation participant populations within the SCG and PG&E service territories.

The number of achieved net-to-gross survey completes was 248 for SCG, 39 for PG&E and 2 for SDG&E. The participating population was largely dry cleaning establishments (78 percent of the SCG and 83 percent of PG&E population). In many cases, language barriers between business staff and surveyors created difficult circumstances for successfully completing telephone surveys. Gift card incentives were offered to increase response rates within the PG&E and SDG&E service territories, with some moderate success (13 additional completes).

SCG Net-to-Gross Ratio Estimation Results

Large Nonresidential NTG Methodology Results

Among the 1,725 SCG pipe insulation participants, 108 were assigned to the Large Nonresidential NTG methodology. Among these, surveys were completed with 30 customers. All of these customers qualified for the Basic level of rigor. Among these 30 customers net-to-gross ratio scores were successfully generated for 29 customers. (One survey was deemed inconclusive.) The ex-post gross therms were zero for 11 of these sites. The remaining 18 sites contributed to the final net-to-gross ratio estimate for the large nonresidential segment, of 71.7 percent.

Consistency checks were triggered for 7 of these 29 respondents. All seven of the respondents that triggered consistency checks were installers of both steam traps and pipe insulation. Summaries of the analysis and final net-to-gross results regarding these consistency checks are presented in Appendix B-6.

Small Commercial NTG Methodology Results

Two hundred and nineteen net-to-gross surveys were conducted with small commercial pipe insulation installers. Among those, 211 interviews resulted in valid net-to-gross scores. Fifty-two valid scores were assigned a weight of zero based on a zero ex-post gross impact. The remaining 159 sites contributed to a mean net-to-gross score of 72.8 percent.

As summarized in Table 3-15 below, the overall SCG participant population net-to-gross ratio is 72.2 percent, with a 90 percent confidence bound ranging from 72.0 to 72.4 percent. Analysis of the stability of the free ridership scores is presented in Appendix A-6.

Partial Free Ridership

No partial free ridership adjustments are made in the calculation of net-to-gross ratios.

A "Decision Maker Survey[35]" was administered by the on-site engineer to capture partial free ridership and the possibility of alternative baseline equipment with which net-to-gross ratios might be informed. The survey yielded only one respondent that claimed to have considered a different alternative to the installed insulation. This respondent claimed to have considered a greater thickness of insulation than what was installed through the program. More specifically, the customer was considering 2 inch insulation but installed 1 inch insulation through the program. The respondent claimed that since the program did not provide extra incentives for 2 inch insulation, the 1 inch insulation was more desirable. However, it is not logical to deduce that in the absence of the program that this customer would have installed 2 inch insulation. The same rebate was available to alleviate costs for both 1 inch and 2 inch insulation.

[35] See Appendix A-4 for the full text of the Decision Maker Survey.

Table 3-15: Summary of SCG Net-to-Gross Ratio Findings

Statistic Description	Total	Large Nonresidential	Small Commercial
Respondents	248	29	219
Valid NTG Scores	240	29	211
Respondents with zero Ex-Post Gross Impact	63	11	52
Net-to-gross Ratio	**72.2%**	**71.7%**	**72.8%**
90 CI Lower Bound	72.0%	71.1%	72.4%
90 CI Upper Bound	72.4%	72.3%	73.0%
Relative Margin of Error	0.3%	0.8%	0.4%

PG&E Net-to-Gross Ratio Estimation Results

A gross impact weight was applied to PG&E net-to-gross ratio estimates for combining each score into a program level result. The weights are based on ex-ante gross therm impact. Recall the approach selected for adjusting PG&E gross impact claims did not yield site-specific estimates of ex-post gross impact. Please see Section 3.2.3 for a more detailed discussion of this approach.[36] The resulting PG&E net-to-gross ratio estimates are shown in Table 3-16 below.

Table 3-16: Summary of PG&E Net-to-Gross Ratio Findings

Statistic Description	Total	Large Nonresidential	Small Commercial
Respondents	39	4	35
Valid NTG Ratios	35	4	31
Net-to-gross Ratio	**49.2%**	**48.8%**	**49.4%**
90 CI Lower Bound	47.7%	44.6%	47.7%
90 CI Upper Bound	50.7%	53.1%	51.0%
Relative Margin of Error	3.1%	8.7%	3.4%

SDG&E Net-to-Gross Ratio Estimation Results

Due to the completion of only two net-to-gross surveys, estimation of a net-to-gross ratio for the SDG&E service territory based on these primary data is not possible.

3.4.5 Final Results

Final results for the Pipe Insulation HIM Evaluation are shown in Table 3-17 below. The final realization rate estimate for SCG territory is 5.7 percent, with a relative margin of error at the 90 percent confidence level of 6 percent.

[36] A PG&E net-to-gross ratio was calculated using an approach that invoked ex-post gross impact weights based on self-reported program qualifying status, and yielded an estimated net-to-gross-ratio of 23.0 percent.

Final realization rate results for the PG&E service territory are 17.2 percent, with a relative margin of error at the 90 percent confidence level of 8.7 percent.

Table 3-17: Summary of Final Pipe Insulation Realization Rates for SCG and PG&E Service Territories

Final Realization Rate Results	SCG	PG&E
Gross Impact Realization Rate	**7.9%**	**35%**
Sample Size	66	38
90% Confidence Bounds	7.4 - 8.4%	33.5 - 36.4%
Relative Margin of Error	6%	4%
Net-to-gross Ratio	**72.2%**	**49.2%**
Sample Size	248	38
90% Confidence Bounds	72.0-72.4%	47.7-50.7%
Relative Margin of Error	0.3%	3.1%
Final Realization Rate	**5.7%**	**17.2%**
90% Confidence Bounds	5.4-6.1%	16.3-18.1%
Relative Margin of Error	5.9%	8.7%

3.5 Recommendations

Prospective Recommendations for Program Standards and Delivery

Controls should be instituted to ensure compliance with program guidelines.

Program guidelines are designed to support delivery of insulation to segments with lower free ridership and where expected impact is higher. For these reasons, sites with pre-existing insulation and sites installing new pipe are excluded from program qualification.

As discussed throughout this report, the majority of participating sites were found to be in violation of one or more of these program rules. An improvement to the expected outcome of similar programs going forward would be related to the institution of effective enforcement of these program guidelines.

Controls should be instituted to ensure that incented insulation is not installed on pipe with pre-existing insulation. In addition, controls should ensure that incented insulation is not installed in new construction applications, new pipe additions, and/or pipe replacements.

At minimum, verification of these characteristics should be provided by the installation contractor and the customer prior to distribution of incentive money. Another more stringent alternative would be to require IOU representatives to inspect sites prior to approving incentive applications. This approach is highly recommended for large installations of insulation, and installations on industrial sites.

Recommended Revisions to Pipe Insulation Work Paper Assumptions and Ex-Ante Impact Claims

SCG Service Territory

The work papers that formed the basis for the ex-ante therm impact claim for pipe insulation installations in SCG territory have evolved over the 2006 – 2008 period. There are at least five submissions of the work paper, with dates ranging from January 2006 through September 2008. Some of the key parameters remain consistent over all the papers, including the 65 degree ambient temperature and a boiler efficiency of 80 percent. Other assumptions and analytical approaches have evolved over this period.

- B-REP-06-599-03, dated January 2006
- B-REP-06-599-03A, dated June 2006
- B-REP-06-599-03B, dated July 2006
- B-REP-06-599-03C, dated January 2007
- ICF Report #20807E, dated September 2008

The work paper that is most consistent with the ex-ante impact claims is that dated January 2007. In this version, the pipes are divided into diameter categories for ex-ante therm assignment. Nearly 80 percent of claimed linear feet of pipe insulation fall into two categories consistent with these work papers. These two measures and their associated linear feet and ex-ante impact claims are shown in Table 3-18 below.

Table 3-18: Most Common Measures and Associated Ex-Ante Gross Therm Impact Claim, for SCG Service Territory

Tracking Database Measure Name	Percent of Gross Impact Claim	Linear Feet Installed	% of Units Installed	Ex-Ante Therms Per Unit
Pipe Insulation - Medium pressure steam < 1" pipe (LF)	16%	274,882	47%	9.6
Pipe Insulation - Medium pressure steam >= 1" pipe (LF)	72%	188,122	32%	63.0

The work paper dated January 2007, document B-REP-06-599-03C, has several assumptions that led to problems in the ex-ante claims. These assumptions caused the estimates of therm impact to be too high. In particular, the assumed operating hours and the ambient air temperatures and wind-speeds are found to be inconsistent with conditions observed in participating sites. These differences contributed to disparity between the ex-ante therm impact and impact calculated upon inspection of participating sites. Table 3-19 below summarizes the measure descriptions, the impact per linear foot and several other key assumptions.

Table 3-19: SCG Pipe Insulation Work Paper B-REP-06-599-03C, January 2007, Key Statistics

Work Paper Summary -SCG	Therms per Linear Foot	Operating Hours	Ambient Temperature	Wind Speed	Pipe Diameter Assumed for Calculations	Pressure Assumed for Calculations
Hot water <1 " diameter, 1 " insulation	2.8	2,425	65	7.5	0.75	-
Hot water >= 1 " diameter, 1" insulation	18	7,752	65	7.5	2	-
Low Pressure steam < 1" diameter, 1 " insulation (<15 psi)	6.1	2,425	65	7.5	0.75	11
Low Pressure steam > = 1" diameter, 1 " insulation (0-15 psi)	40	7,752	65	7.5	2	11
Medium Pressure steam < 1 inch diameter, 1 " insulation	9.6	2,425	65	7.5	0.75	86
Medium Pressure steam >= 1 inch diameter, 1 " insulation	63	7,752	65	7.5	2	86

One of the weaknesses of this work paper is the pairing of pipe with greater than one inch diameter with operating hours of 7,752 per year. These operating hours are consistent with industrial operations, but much of the 1 inch diameter pipe retrofit through the program was located in small commercial facilities with operating hours closer to 2,425 per year.

Among the sites inspected, the average operating hours were 3,164 per year. The dry cleaners that make up a large portion of program participation and ex-ante gross impact (78% and 64%, respectively) had an even lower measured mean operating hour value, 2,432. Non dry-cleaners have average operating hours of 4,974 hours per year.

Assumptions about the environmental conditions in which the piping systems operate are also generally inconsistent with data collected from participating sites. The assumed environmental conditions include an ambient temperature of 65 °F and 7.5 mph wind speed. If the retrofit pipes were outside, then these temperature and wind speed estimates would be fairly accurate, if not even slightly conservative in terms calculating energy impact. However, the majority of the sites inspected in this evaluation effort did not have any outside pipes, and had much higher ambient temperatures and no wind speed in the area around the piping system.

Eighty five percent of the inspected linear feet of pipe insulation were found installed on indoor pipes with no exposure to wind. With no air flow around the pipes, the convective heat transfer is in a state of free convection, which transfers less heat from the pipe.

Mean recorded ambient temperatures around pipe insulated through the program were quite high, at 87 degrees Fahrenheit[37], with some sites having ambient temperatures in excess of 100 °F. Thus, the difference in temperature between the pipe surface and the surrounding air was often much lower than what was implied by work paper assumptions.

Another assumption contributing to an over-estimation of ex-ante therm impact was regarding the diameter of insulated piping systems. Insulation installed on pipe with greater than or equal to one inch diameter was assumed to be well represented by an assumption of a 2 inch diameter pipe. Among the dry cleaning sites inspected for this evaluation, the mean diameter of the pipes insulated through the program was 1.1 inches. However, this includes a portion of pipe that is less than 1 inch in diameter. If the mean diameter calculation is performed on pipes with more than 1 inch diameter, the mean pipe diameter becomes 1.3 inches, still well below the 2 inch assumption.

In contrast, among the sites inspected that were not dry cleaning sites, the mean pipe diameter of insulated pipes is 2.6 inches, and the mean diameter of pipe with greater than 1 inch diameter is 2.8. Overall, the average pipe diameter found on inspected sites was 1.5 inches. The average among pipe with diameter equal or greater than one inch is 1.7 inches, still falling short of the 2 inch mark.

<u>Revisions and Current Status of SCG Pipe Insulation Work Paper</u>

Substantial revisions to ex-ante impact claims and methods are found in a work paper titled "ICF Report #208807E" dated September 2008. The organization of impact claims presented in this September 2008 document is quite different from previous work papers. Key statistics from this new work paper are shown Table 3-20 below.

[37] Ambient temperatures are not typically seasonally adjusted unless the pipes are located outside. It was found that heat originating from indoor equipment modulated temperature. For example, boiler room temperatures are modulated by the boiler itself. In the dry cleaner "pressing area" the temperature is modulated by the steam presses, or in interstitial spaces by the hot pipes themselves. Each site is unique, however, and treated as such. Where appropriate, the ambient temperatures were adjusted for year round averages.

Table 3-20: SCG Pipe Insulation Work Paper, ICF Report #20807E, September 2008, Key Statistics

Work Paper Summary ICF Report #20807E	Therms per Linear Foot	Operating Hours	Ambient Temp	Pipe Diam. Assumed for Calc	Pressure Assumed for Calc
Small Commercial					
Hot Water <1 inch diameter pipe	2.9	2,425	65	0.75	-
Hot Water >=1 inch diameter pipe	5.7	2,425	65	2	-
Low Pressure steam, <=15 psi, <1 " diameter	6.2	2,425	65	0.75	11
Low Pressure steam, <=15 psi, >=1 " diameter	12.6	2,425	65	2	11
Medium Pressure steam, >15 psi, <1 " diameter	9.7	2,425	65	0.75	86
Medium Pressure steam, >15 psi, >=1 " diameter	19.9	2,425	65	2	86
Large Commercial					
Hot Water <1 inch diameter	5.2	4,380	65	0.75	-
Hot Water <1 inch diameter	10.4	4,380	65	2	-
Low Pressure steam, <=15 psi, <1 " diameter	11.1	4,380	65	0.75	11
Low Pressure steam, <=15 psi, >=1 " diameter	22.8	4,380	65	2	11
Medium Pressure steam, >15 psi, <1 " diameter	17.5	4,380	65	0.75	86
Medium Pressure steam, >15 psi, >=1 " diameter	35.9	4,380	65	2	86
Industrial					
Hot Water <1 inch diameter	9.2	7,752	65	0.75	-
Hot Water <1 inch diameter	18.4	7,752	65	2	-
Low Pressure steam, <=15 psi, <1 " diameter	19.7	7,752	65	0.75	11
Low Pressure steam, <=15 psi, >=1 " diameter	40.3	7,752	65	2	11
Medium Pressure steam, >15 psi, <1 " diameter	31	7,752	65	0.75	86
Medium Pressure steam, >15 psi, >=1 " diameter	63.5	7,752	65	2	86

In the most recent (September 2008) work paper, pipe insulation measures continue to be grouped by hot water, low pressure steam, and medium pressure steam. They also continue to be sub-categorized by the pipe diameter—divided into greater than or equal to 1 inch, and less than 1 inch pipe diameter categories. The major difference is that each of these categories assumes operating hours consistent with the business type where the insulation is installed. That is, operating hours are assumed to be 2,425 for small commercial, 4,380 for large commercial and 7,752 for industrial. This new grouping is a significant improvement, and is expected to increase the accuracy of new claims.

The ambient temperature assumption remains at 65 degrees, which should be considered for further revision, for reasons discussed in detail above. Wind speeds should be set to zero to reflect the predominance of indoor applications. Further, the pipe diameter assumptions for the small commercial category should be revised to reflect the concentration of smaller pipe diameters in small commercial facilities. A value of 1.3 inch diameter for the small commercial, greater than one inch pipe diameter category would reflect the on-site findings of this evaluation.

Table 3-21 below shows the therms per linear foot that result from the site inspections and engineering analysis presented in this evaluation. The results *exclude* all consideration of non-program qualifying status, reflecting only the operating hours, ambient temperatures, and other key conditions present at the inspected sites. Results are shown for dry cleaners separately from other business types. Under the most recent work paper, the range of values claimed by dry cleaners is between 2.9 to 19.9 therms per linear, with all steam pipe installations exceeding 6.2 therms per linear foot.

Table 3-21: Unit Energy Savings and Other Key Impact Assumptions, Dry Cleaners versus Other Business Types

Results Description*	Total	Dry Cleaners	Other Business Types
Therms per Linear Foot	**6.9**	**3.7**	**14.7**
Sample Size	66	47	19
90% Confidence Bounds (therm/LF)	6.7-7.1	3.66 – 3.74	13.3 – 14.7
Relative Margin of Error	3%	1%	5%
Other Key Statistics			
Operating Hours	3,381	2,416	5,536
Pipe Diameter	1.5	1.1	2.3
Ambient Temperatures	86.7	89.1	81.3

* Based on 66 on-site inspections and engineering analysis completed in SCG service territory with 2006-2008 Express Efficiency participants installing pipe insulation.

Further research is needed to recommend a set of ex-ante claimed therm impact values to use going forward. It is recommended that statewide consistency be maintained and that the findings of this report and other pertinent research contribute to new DEER values that can be sourced by utility and third party program implementers and planners.

PG&E and SDG&E Service Territory

PG&E tracking system data for the 2006-2008 program cycle indicates that the pipe insulation offering has been largely subscribed to by dry cleaning establishments, with 83 percent of participating sites and 74 percent of the total pipe insulation ex-ante gross impact claim.

The measure installed in participating dry cleaners is nearly all *"Pipe Insulation low pressure steam – 1 inch"*. This measure makes up 173 of 178 measure line items of pipe insulation installed in dry cleaners in the PG&E tracking system over the 3 year cycle. The associated impact for this measure is 13.4 therms per linear foot.

As presented in this report, the on-site M&V work completed in SCG service territory indicates that pipe insulation installed in dry cleaning establishment achieves a lower impact, at 4.7 therms per linear foot (this includes impact from insulation installed on pipe with diameter greater than or equal to one inch[38]).

There are some assumptions identified in work papers that might be considered for review in order to improve ex-ante claims in future program cycles. The PG&E document entitled "Work Paper PGECOHVC104, Pipe Insulation, Revision #1" references several assumptions that may warrant attention, as follows:

- Base case: 2 inch diameter horizontal pipe, bare surface
  This is an average pipe diameter for commercial and industrial systems.

- Operating hours: 8,760 hours
  This is assuming an average of 12 hours of operation per day, 5 business days per week, 52 weeks per year.

- Ambient temperature: 75ºF (the installation is assumed to be indoors)

The 2 inch diameter assumption for pipe insulated within the dry cleaning segment is not confirmed by the site visits to similar commercial buildings in SCG service territory. PG&E does not allow pipe less than 1 inch in diameter to be insulated through the program. The mean diameter for similar diameter pipe among SCG dry cleaners is 1.3 inches.
It is difficult to know with certainty that the piping systems of dry cleaning establishments within PG&E service territory carry the same typical diameters, but nonetheless the possibility should be considered carefully.

The next assumption that bears mention is the assumed operating hours of 8,760. First, it is noted that this number is not consistent with the text that follows claiming 12 hours per day, 5 days per week and 52 weeks per year. The product of these figures is 3,120, which is very close to the number found in SCG territory through the on-site work, which is 3,164 hours

[38] PG&E and SDG&E require a minimum of one inch pipe diameter to qualify for program incentives.

per year. Note, however, that dry cleaners have lower confirmed operating hours, at 2,416 hours per year.

The assumption regarding the placement of insulation on indoor pipe is reasonable, and confirmed by on-site inspection work in SCG territory.

The PG&E work paper assumption of an ambient temperature of 75 degrees is an area of some concern. This assumption is lower than what was observed during on-site inspections in SCG service territory. Though SCG has different climate characteristics, most of the pipe inspected was indoors, and it was found that indoor piping has ambient temperatures predominantly modulated by equipment operations, not climate-related factors.

The SDG&E pipe insulation claim is made up primarily[39] of one measure "Pipe Insulation – Low Pressure Steam Applic. (LF) 1 in". The associated ex-ante therm impact for this measure is 13.4 therms per linear foot. PG&E also has a measure with similar description that also is associated with an ex-ante claim of 13.4 therms per linear foot. Since the claims are similar per unit between PG&E and SDG&E, it is recommended that SDG&E program implementers review work papers for the same issues as discussed above in regards to PG&E. It is to be noted that the SDG&E population of pipe insulation participants is small (21 pipe insulation sites) and less concentrated in the dry cleaning segment, with just 52 percent of sites and 22 percent of ex-ante therm impact claims.

[39] Ninth-four percent of ex-ante impact claims

4

Steam Traps

4.1 Evaluation Objectives for Steam Traps

This section describes and outlines the evaluation objectives for the High Impact Measure (HIM) Steam Traps. The evaluation provides ex-post estimates of gross savings for commercial, industrial low pressure, and industrial high pressure steam traps. It also provides separate estimates of the net-to-gross ratio (NTGR) for commercial and industrial applications of steam traps. The evaluation of steam traps was conducted under the rules described in the CPUC EM&V Protocols,[40] which specify minimum sample sizes, the required precision, data collection techniques, certain minimum analysis techniques, and formats for documenting and report results to the CPUC.

Though a majority of the claimed energy impacts from steam traps do not originate from programs evaluated under this contract group,[41] ED sought a more even division of HIM evaluation responsibility. The evaluation of steam traps was therefore assigned to the Southern California Industrial and Agricultural Evaluation Contract.

Therm savings claims for steam traps are concentrated in the SCG Express Efficiency Program, the SDG&E Small Business Super Saver Program, and the PG&E Commercial Mass Market Program. Steam traps qualify as an HIM based on their ex ante estimated annual therm savings contributions to the overall statewide portfolio and to the IOU portfolios of SCG, SDG&E, and PG&E.[42] The ex ante therm savings from steam trap replacement are equal to approximately 25 million therms annually, or 18% of the total annual statewide portfolio therm savings for the 2006 – 2008 programs. As Table 4-1 shows, steam traps contribute approximately 14%, 24%, and 7% of the annual therm savings of the portfolios of PG&E, SCG, and SDG&E, respectively.

[40] California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professionals. Prepared by TecMarket Works for the California Public Utilities Commission. April 2006. p.19-62.

[41] The statewide claimed savings from steam traps are generated mostly from programs being evaluated under the Small Commercial Contract Group.

[42] A high impact measure, or HIM, is defined as one that contributed more than 1% of portfolio energy savings by IOU and fuel type for the 2006-08 program cycle.

Table 4-1: Annual Therm Savings of Steam Traps by IOU

Utility	Annual Therm Savings from Steam Traps	Annual Total Portfolio Therm Savings	Percentage of Total Therm Savings
PG&E	9,197,461	67,928,927	14%
SCG	15,252,338	62,290,003	24%
SDG&E	537,187	7,667,793	7%
TOTAL	**24,986,986**	**137,886,723**	**18%**

There is considerable uncertainty surrounding the energy savings attributable to steam traps. The current SCG Express Efficiency Program, SDG&E Small Business Super Saver Program, and the PG&E Commercial Mass Market Program have established deemed savings values for steam traps installed in commercial, low pressure industrial, and high pressure industrial applications. The savings calculations used to determine the deemed values of steam trap savings depend on engineering algorithms whose individual components require on-site observation and/or are hard to verify. The existing deemed savings values are largely derived from a small sample of commercial dry cleaning installations[43] and evaluations of steam traps within a large Canadian utility.[44] The analysis undertaken for this evaluation helps to clarify uncertainties and determine appropriate savings levels for steam trap installations in commercial and industrial applications within California.

The deemed savings values for commercial steam traps have been recently evaluated within the PG&E and SCG service territories using a billing analysis of commercial dry cleaning steam traps.[45] SDG&E did not evaluate its savings values prior to this analysis. The PG&E billing analysis led to a significant downward adjustment to PG&E commercial steam trap ex ante per unit claims. The SCG billing analysis found no evidence that commercial dry cleaning steam traps save energy. SCG and SDG&E chose to not adjust their ex ante per unit claims following the SCG billing analysis.

This HIM evaluation of commercial steam trap applications incorporates site-specific data collection using a phone survey. The site specific data collection helps to inform a commercial billing analysis and to determine a self reported NTGR. The phone survey data was analyzed separately for PG&E and the Sempra (SDG&E and SCG) utilities.[46] The additional site specific information and the longer time period available for the evaluation

---

[43] kW Engineering. Steam Trap Survey and Billing Analysis Report. Prepared for Southern California Gas Company. December 2006.

[44] Enbridge Gas Distribution, Inc. Enbridge Steam Saver Program, 2005 and Enbridge "Steam Saver Program" Steam Boiler Plant Efficiency Update to Year End, 2005. March 2006.

[45] KEMA, Inc. Steam Trap Impact Assessment. Prepared for Pacific Gas & Electric Company. October 2007 .Business Analysis Economic Research, Steam Trap Billing Analysis. Prepared for Southern California Gas Company. April 2008.

[46] The billing analysis for SCG and SDG&E were joined into a single analysis due to common per unit energy savings estimates. In addition, the number of SDG&E phone respondents are too low to validate a single billing analysis.

helps to clarify the gross savings achieved from commercial dry cleaning steam trap applications, while the NTG analysis allows for the calculation of net ex-post savings achieved from these applications.

The evaluation approach for industrial steam traps uses a phone survey to collect site-specific self-report information to calculate a net-to-gross ratio, and on-site data collection to evaluate gross savings. The industrial evaluation is complicated by the variety of types of traps installed, the multiple functions of the traps, and the multiple segments using industrial steam traps. Given the uncertainty associated with the quantification of gross savings from industrial applications, the evaluation methodology chosen for this study first collects site-specific information to help clarify usage parameters associated with the operation of the steam traps. Site specific data includes hours of operation, pressure at boilers and individual steam traps, boiler efficiency, orifice size, steam trap makes and models, total number of traps, number of traps replaced, and method of steam trap failure. The site-specific data are then used in engineering algorithms to calculate ex-post estimates of industrial steam trap savings.

The commercial and industrial approaches chosen for this evaluation help to clarify the observable gross savings attributable to the measure and the sector specific NTGRs associated with steam trap applications.

4.2 Methods Used for Steam Traps

This section describes the methodology used to determine the ex-post gross therm savings and the NTG ratios from steam traps in both commercial and industrial applications. The section begins with a description of the sample methodology used for phone and on-site survey data collection. Additional data to support the estimation of savings for the industrial steam trap retrofits was retrieved from participant rebate applications and steam trap manufacturers. The section concludes with a description of the weighting methodology.

4.2.1 Tracking Data

Table 4-2 lists the distribution of tracking system therm savings claims by IOU and measure category. The SCG and SDG&E per site therm impacts are much greater on average than those claimed by PG&E. This is likely due to the application of a greater savings claim by SCG and SDG&E on a per steam trap basis. For example, the per unit claimed savings for a SCG or SDG&E commercial steam trap was 139 therms, while PG&E only claimed 45.87 therms or 33% of the per unit claims made by the Sempra Utilities.[47]

[47] PG&E chose to reduce their per steam trap ex ante claims following an evaluation completed in 2007 (KEMA, Inc. Steam Trap Impact Assessment. Prepared for Pacific Gas & Electric Company. October 2007).

Table 4-2: Tracking System Summary

IOU	Measure Name	Sites*	Percent of Sites	Therms	Percent of Therms
PG&E	Steam Trap - Commercial - Any Pressure	1,054	29%	1,009,287	4%
	Steam Trap - Industrial High Pressure Steam (>15 Psig)	41	1%	8,006,268	32%
	Steam Trap - Industrial Low Pressure Steam (<15 Psig)	15	0%	181,907	1%
	Total	**1,108**	**31%**	**9,197,461**	**37%**
SCG	Steam Trap Replacement - Commercial <12hr/Day (Dry Cleaners)	505	14%	1,112,623	4%
	Steam Trap Replacement - Commercial/Other	1553	43%	3,348,010	13%
	Steam Trap Replacement - Industrial<15 Psig/Other Commercial 12-24 Hr/Day	55	2%	789,487	3%
	Steam Trap Replacement - Industrial<15 Psig	8	0%	129,846	1%
	Steam Trap Replacement - Industrial>15 Psig	165	5%	9,872,373	40%
	Total	**2244**	**63%**	**15,252,338**	**61%**
SDG&E	Steam Trap Replacement- Commercial - Other	205	6%	493,461	2%
	Steam Trap Replacement - Industrial<=15 Psig	2	0%	12,250	0%
	Steam Trap Replacement - Industrial>15 Psig	1	0%	31,476	0%
	Total	**208**	**6%**	**537,187**	**2%**
All IOUs	**All Measures**	**3560**	**100%**	**24,986,986**	**100%**
All IOUs	**Total Commercial**	**3295**	**92%**	**5,963,381**	**24%**
All IOUs	**Total Industrial**	**273**	**8%**	**19,023,606**	**76%**

*The number of sites listed above may be larger than the number of sites listed in later tables because a site is counted for each type of measure it installs. Energy savings are calculated at the measure level, ensuring that there is no double counting.

4.2.2 Surveys and Sampling Methodology

Telephone surveys were conducted using Computer Assisted Telephone Interviewing (CATI), and for larger industrial applications (with relatively large therms savings) data collection was completed by an energy consultant working in collaboration with an engineer. CATI surveys were applied to the commercial steam trap applications and industrial sites with lower therm savings. One commercial site was surveyed by an energy consultant due to the site's relatively large therm savings.

The first round of the Small Commercial Verification work included telephone surveys of 210 steam trap participants.[48] Of the 210 steam trap participants surveyed, 185 sites installed commercial steam traps and 25 sites installed industrial steam traps. These surveys included the small commercial NTG battery and a sizable battery of steam trap questions; however data from these surveys were not sufficient to support the HIM evaluation, particularly from the 25 industrial steam trap sites.[49] For instance, the industrial sites had not been asked questions from the industrial net-to-gross battery.[50] The research team wanted to collect supplemental data from all of the 210 sites that were surveyed previously to aid in the evaluation of gross therm savings from steam trap retrofits. In addition to re-contacting the 210 steam trap participants using callback surveys designed by the steam trap research team, additional surveys were conducted with commercial and industrial steam trap participants in an attempt to meet quotas for the commercial and industrial phone survey sample designs presented in Table 4-4 and Table 4-6 below. The callback surveys, as well as the commercial and industrial survey instruments, are presented in Appendix B-1. Frequency tables for each question asked in the telephone survey instruments are presented in Appendix B-2.

Commercial Telephone Surveys

The sample design for the commercial phone survey sample was developed to provide the necessary information for the ex post gross savings and the NTG analysis of commercial steam traps by IOU. The survey stratification for the commercial evaluation was limited to IOU, given the relatively homogeneous application of commercial steam traps and the similarity in the per customer size of the applications. Table 4-3 presents the IOU survey stratification for the commercial steam trap evaluation.

---

[48] The commercial and industrial steam trap retrofits evaluated in this section were financed by the utility Small Commercial programs. The first round of the Small Commercial Verification study implemented telephone and on-site verification work in support of the evaluation being conducted by the CPUC Small Commercial Contract Group. Following the development of the High Impact Measure Groupings the steam trap HIM analysis was placed within the Southern California Industrial and Agricultural Contract Group.

[49] The small commercial NTG battery was taken from the Residential and Small Commercial NTG Ratio Working Group's Guidelines for Estimating Net-to-Gross Ratios Using the Self-Report Approaches and the Algorithm for the Residential and Small Commercial Consistent Free Ridership Method, which was prepared for the Energy Division of the CPUC, October 2007.

[50] The industrial NTG battery was taken from the Nonresidential NTG Ratio Working Group's Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers, which was prepared for the Energy Division of the CPUC. April 2009.

Table 4-3: Commercial Steam Trap Tracking Data Summary

Sector	Sites	Total Gross Therm Savings	Average Gross Therm Savings per Site
Commercial PG&E	1,054	1,051,340	997[51]
Commercial SCG	2,036	4,646,492	2,282
Commercial SDG&E	205	514,022	2,507

The NTG sample specifications require either 300 sample responses or half of the population for each HIM.[52] Therefore, the commercial gross savings and NTG survey quota called for 300 sites for PG&E, 300 for SCG, and 103 sites for SDG&E as shown in Table 4-4. These sites were distributed between newly conducted surveys and Small Commercial Verification telephone surveys supplemented by additional steam trap questions. Given the size of the quota relative to the steam trap population, the CATI Center used a census of program participants in an attempt to reach the survey quota.

Table 4-4: Commercial Telephone Survey Sample Design

IOU	Sites	Completed Phone Surveys/Planned Re-Contact	Additional Phone Surveys	Total Phone Surveys Planned	Total Phone Surveys Completed
PG&E	1,054	62	238	300	176
SCG	2,036	96	204	300	325
SDG&E	205	27	76	103	41
Total	**3,295**	**185**	**518**	**703**	**542**

Of the 185 commercial steam trap sites previously surveyed during the Small Commercial Verification effort, 103 were successfully contacted and the additional survey questions completed.[53] Of the 703 phone surveys planned, 542 were completed. The census of the steam trap population did not enable the team to meet quota for PG&E and SDG&E. Given the relatively homogeneous application of steam traps in the commercial population, largely representing dry cleaning and laundry applications, the research team believes that the sample from which data were collected is sufficient for both the ex-post gross savings and NTG analyses.

[51] The per-site steam trap savings are lower for PG&E than the Sempra utility claims because PG&E chose to reduce their per steam trap ex ante claims following an evaluation completed in 2007 (KEMA, Inc. Steam Trap Impact Assessment. Prepared for Pacific Gas & Electric Company. October 2007).

[52] California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professionals. Prepared by TecMarket Works for the California Public Utilities Commission. April 2006. Page 38.

[53] All of the 185 Small Commercial surveys are usable for the commercial net-to-gross analysis because these sites answered the commercial NTG battery during the original survey. Given the final design of the steam trap billing analysis, the 185 Small Commercial survey were used in the commercial gross analysis.

Industrial Telephone Surveys

The industrial telephone survey sample design was stratified by ex ante therm savings, using savings to proportionally allocate the sites into three strata with approximately equal savings. The number of sites per strata and the associated savings per site are presented in Table 4-5. Installations with larger ex ante therm savings are expected to be qualitatively different not only from each other, but also from installations with smaller therm savings. Large industrial installations may have larger pipes, high or low pressure steam distribution networks over a wide area, sophisticated decision makers and a well articulated maintenance policy. The strata were not utility specific due to the small number of sites in the population.

Table 4-5: Industrial Steam Trap Tracking Data Summary

Sector	Strata	Sites	Total Gross Therm Savings	Average Gross Therm Savings per Site
Industrial, all IOUs	1	5	6,581,020	1,316,204
	2	20*	6,644,921	332,246
	3	248	6,590,316	26,574
Total		**273**	**19,816,256**	**72,587**

* According to the Steam Trap HIM Research Plan, there were 21 strata 2 sites and a total of 274 steam trap sites across all strata. One of the strata 2 sites was merged with a strata 1 site as it became clear during the evaluation that these two records comprised a single facility with common decision making. Hence, the total number of steam trap sites was reduced to 273 from the original 274.

The industrial phone surveys were designed to collect information about the site, the site's steam trap installations, the site's self-reported NTG, and to recruit for the industrial steam trap on-site field analysis. The quota for the industrial phase of the phone survey consisted of 137 sites as shown in Table 4-6. The sample design included 25 sites that were previously interviewed during the Small Commercial Verification effort. To supplement their data, these 25 industrial steam trap participants from the Small Commercial Verification effort were contacted again by an energy consultant and/or engineer.

Table 4-6: Industrial Telephone Survey Sample Design

Customer Size	Strata	Sites	Completed Phone Surveys/ Planned Re-Contact	Additional Phone Surveys	Total Phone Surveys Planned	Total Phone Surveys Completed
Large	1	5	1	4	5	4
Middle	2	20	0	20	20	15
Smaller	3	248	24	87	111	105
Total	**-**	**273**	**25**	**112**	**137**	**124**

Due to the high quota relative to the number of sites in the population, the telephone survey was a census of the sites. All strata 1 and eight strata 2 sites were surveyed by an energy consultant under the supervision of an engineer. The remaining surveys were conducted by the CATI Center. The completed number of telephone surveys was slightly lower than the planned quota. The number of completed telephone surveys in strata 1 and 2, however, exceeds fifty percent of the sites in each strata and the number of sites completed in strata 3 is nearly 50% of the sites. The completed number of telephone surveys was only 12 sites less than the planned quota and provided the team with a sufficient number of sites to recruit for the on-site fieldwork and to calculate net-to-gross ratios.

Industrial On-Site Surveys

The industrial evaluation of gross energy savings included on-site data collection. The Small Commercial Contract Group's on-site Survey had previously been completed for 18 industrial sites. To use these sites for the steam trap gross savings evaluation, the research team attempted to gather additional data using a callback phone survey designed by an engineer. This survey instrument is also presented in Appendix B-1. The original plan called for the 18 Small Commercial on-sites to be supplemented with an additional 22 industrial on-site surveys to reach a quota of 40 completed on-site evaluations (see Table 4-7). When the team was unable to contact an original Small Commercial site, the quota for the new sites was increased to replace the Small Commercial site, ensuring that the quota of 40 on-sites was completed.

The research team successfully completed all planned on-site surveys. Data from these field visits were gathered based upon the steam trap on-site visit protocols and data collection forms presented in Appendix B-3. The high quota of strata 1 and 2 sites relative to the population was warranted given the concentration of total program savings in these strata.

Table 4-7: Industrial On-Site Survey Sample Design

Sector	Strata	Sites	Completed On-Sites / Planned Phone Re-Contact	Additional On-Sites Planned	Total On-Sites Planned	Total On-Sites Conducted
Large	1	5	1	3	4	4
Middle	2	20*	0	13	13	13
Smaller	3	248	17	6	23	23
Total	**-**	**273**	**18**	**22**	**40**	**40**

*According to the Steam Trap HIM Research Plan, there were 21 strata 2 sites and a total of 274 steam trap sites across all strata. One of the strata 2 sites was merged with a strata 1 site as it became clear during the evaluation that these two records comprised a single facility with common decision making. Hence, the total number of steam trap sites was reduced to 273 from the original 274.

Vendor Interviews

Interviews with steam trap vendors were conducted to support the NTG analysis for those industrial applications where the vendor was an influential factor in the decision to install steam traps. The HIM research plan stated that these interviews would also be used to support the industrial ex-post gross savings analysis to fill in information concerning the failure types for replaced steam traps as well as steam trap make and model specifications, such as orifice size. However, site representatives were found to often possess the results from steam trap surveys indicating failure type while vendors did not retain this information. The vendor interviews were, therefore, primarily used to gather data for the estimation of NTG ratios for participants heavily influenced by vendor recommendations. Data collected from seven vendor interviews were gathered about their interactions with eight steam trap program participants. The vendor interview telephone survey for the steam trap analysis is presented in Appendix B-1.

4.2.3 Gross Energy Savings

Two approaches are used to estimate the ex-post savings associated with steam trap retrofits, depending on whether the traps being studied were in commercial or industrial settings. The observable savings from the retrofit of steam traps in commercial applications are examined using a statistically adjusted engineering (SAE) billing analysis. Estimates of ex-post savings from industrial applications of steam traps are based on an engineering analysis that relies on the collection of data through telephone surveys and on-site visits. These data are used as inputs to an algorithm used to estimate gross savings obtained from steam trap replacement.

Commercial Applications of Steam Traps

The circumstances under which commercial steam traps are installed are relatively homogeneous and, therefore, the commercial steam traps installed in dry cleaning applications were analyzed using a SAE billing analysis. The billing analysis was estimated separately for PG&E and the Sempra Utilities. The per unit ex ante claimed savings differ between PG&E and the Sempra Utilities necessitating separate evaluations. Since there were only 41 commercial phone survey respondents from SDG&E, SCG and SDG&E were combined into a Sempra Utilities SAE analysis and result segment.[54]

The SAE approach using a prescriptive ex ante savings value per trap replaced carries with it some risk. The ex ante savings estimates are insensitive to the site-specific relationship of replaced failed traps to the sites total number of traps replaced. In addition, the ex ante estimate of savings does not vary with the length of time that the failed traps were in poor condition. Over 70% of the phone survey sites report replacing all of their steam traps. The high removal share may be justified by the low incremental cost of the measure, the short expected useful life, and the difficulty associated with quickly and inexpensively determining

[54] The per unit ex ante claims are the 139 therms per rebated steam trap for both SCG and SDG&E. The per unit ex ante claims for PG&E were 45.87 therms or 33% of the SCG and SDG&E level.

that a trap is leaking.[55] However, the high removal share is likely to result in the removal of both working and failed traps.[56] The work paper's ex-ante savings calculations assumed that 100% of the traps were replaced for commercial applications and that only 27% of the replaced traps were failed open or partially blocked.[57] If the actual failed share is lower or other parameters used to estimate savings are incorrect, the observed savings may be substantially less than the ex ante claimed savings for the site. Furthermore, the length of time a removed steam trap was leaking is unknown. If a trap was only leaking for a short period, the observed savings in the participants billing records will be small relative to a year of ex-ante claimed savings. For example, if a steam trap was only leaking for five months prior to being replaced, a billing analysis that relies on comparing current consumption with consumption 12 months previous will only observe savings for 5/12 months, significantly reducing the estimated realization rate.

The SAE billing analysis combined information for dry cleaning commercial applications from the utility billing frames, the tracking data, phone survey data, weather, and information on the county level unemployment rate.[58,59] The data collected in the phone surveys provided additional information concerning the share of replaced traps that were working poorly or failed, the length of time between failure and replacement, and possible changes in the sites' production.[60]

---

55 The utility workpapers (Energy and Environmental Analysis, Inc. Steam Trap Workpaper for PY 2006-2008. Prepared for Southern California Gas Company. December 2006) lists the expected useful life of steam traps as 6 years and the expected cost of a commercial steam trap as $77.

56 Approximately 25% of the phone survey sites report that they believe that traps in good working condition were replaced.

57 Energy and Environmental Analysis, Inc. Steam Trap Workpaper for PY 2006-2008. Prepared for Southern California Gas Company. December 2006.

58 Focusing on dry cleaning sites will provide a realization rate estimated using a more homogeneous set of sites. Dry cleaning sites are also the most common commercial application and represent the segment modeled in the workpapers to calculate commercial site savings.

59 The county level yearly change in the unemployment level was used as an independent variable in the model to help control for changes in the economy. During the billing period of the analysis, the economy experienced both an expanding economy and a significant recession. Dry cleaning operations may be very sensitive to changes in the economy.

60 The work papers assume that the owners or site managers at dry cleaners have a limited knowledge of the operational status of their steam traps. If they truly do not know the status of the steam traps, the phone survey trap information may be of little use.

The SAE model specification to assist in determining the impact of steam traps on commercial facility gas usage is the following model for both PG&E and the Sempra Utilities:[61]

$$\Delta Therms_{it-it-12} = \beta_0 + \beta_1 THMSTSAV_{it} + \beta_2 THMOTSAV_{it} + \sum_{n=4}^{s} \beta_n \chi_{i,t,n} + \varepsilon_{it}$$

where:

$\Delta Therms_{it-it-12}$ = the change in average daily gas usage between month *t* and average daily gas usage during month *t*-12 at commercial site *i*

$THMSTSAV_{it}$ = the average daily engineering estimates of savings for installed steam traps at site *i* in month *t*[62]

$THMOTSAV_{it}$ = the average daily engineering estimates of savings for other gas measures installed under the program at site *i* in month *t*[63]

$\chi_{n,i,t}$ = a set of other independent variables for site *i* in period *t*, such as changes in unemployment at the county level. This variable also includes a set of time fixed effects to further control for changes in the economy that impacted all sites in the analysis.

ε_{it} = a random error term.

The model is dependent on extensive billing data. The first step aggregates the account or meter level billing data to the site level. This is necessary because the installed steam traps may result in savings to gas provided through multiple meters. Once the billing data is aggregated and calendarized, the bills are used to develop the dependent variable as the difference between the site's average daily consumption in a given monthly period and the same period 12 months prior. The participant's per-unit change in average daily gas usage between billing period *t* and *t*-12 is modeled as a function of the engineering estimate of average daily savings from steam traps at the site, the engineering estimate of average daily savings from the installation of other rebated gas measures at the site, a time series fixed effect and other available relevant independent variables.[64]

---

[61] Sites from SCG and SDG&E were combined in the billing analysis due to the small number of SDG&E sites. The ex ante per unit steam trap claims are the same for SDG&E and SCG, 139 therms.

[62] The average daily engineering estimates of savings were derived from the yearly claimed savings provided by the utilities. The yearly values were divided into 12-month shares using the fraction of gas used at dry cleaning sites that state in the phone survey that they do not have gas heat. The 12 monthly shares were then divided by the number of days in the month to derived the average daily estimate.

[63] In the population of rebated steam trap sites approximately 50% of sites installed steam traps and some other rebated gas measure during the program period. The large majority of the other gas rebated measures were pipe insulation installations.

[64] The team also estimated models including the change in electricity usage as an independent variable. The use of electricity as an independent variable significantly impacted the number of sites included in the analysis due to the difficulty in matching a site's gas consumption with it electricity consumption. It was found that the reduction is sample size was driving the change in the realization rate not the inclusion of the change in electricity consumption.

For the first 12 months after the new steam traps or other gas measures are in place, the engineering estimate of savings will be non-zero. In all other months, the engineering estimate of savings will be zero. The coefficient on this variable will represent the portion of the predicted impacts of the steam traps actually detected in the bills. Gas usage from October 2004 to July 2009 was provided by PG&E and data from October 2005 to June 2006 was used in the PG&E billing analysis. For the Sempra Utility analysis, SCG provided bills from January 2004 to June 2009 for a billing analysis data set from January 2005 to June 2009. SDG&E provided bills from August 2004 to June 2009 and the data used for the billing analysis was from August 2005 to June 2009.[65]

During the development of the steam trap analysis, preliminary estimates of savings from the pipe insulation analysis became available. These results indicated that the commercial pipe insulation therm savings were significantly less than the ex ante claims. The steam trap analysis incorporated the reduced pipe insulation savings prior by multiplying the site level pipe insulation savings by 0.045 or 4.5% of the ex ante claim. The estimated realization rate for pipe insulation will be relative to the new engineering estimate.

The time series fixed effect variables are a series of binary (0,1) variables for each year and month. The inclusion of a time series fixed effect is designed to help control for the average effects of observed and unobserved independent variables that vary across time. The estimated coefficient is the average influence of the time period on the change in usage. The fixed effect was included in the model to help control for changes in the economy that influenced all sites but changed over time. During the beginning of the billing period, the economy was growing while during the end of the billing period the economy was in recession. The fixed effect calculates an average effect of each period on the change in usage, allowing the steam trap savings variable to capture the time invariant effect of retrofitting steam traps.

Multiple specifications of the model were estimated using different site specific variables, including variables to control for remodeling, increasing and decreasing numbers of employees, and square footage changes. The results section provides the estimates of the billing analysis and a discussion of the influence of other independent variables.

Industrial Applications of Steam Traps

This section describes the key parameters used to estimate gross energy savings for industrial steam traps, along with measurement techniques, associated measurement uncertainties, and the engineering algorithm used to calculate savings estimates for industrial steam trap applications.

To arrive at the methodology used in this evaluation, a secondary literature search was conducted, which included an assessment of previous steam trap measure evaluations,

[65] The billing analysis model needs bills from t-12 to compare to current consumption, leading to the one year difference between the billing data received and the bills used in the SAE model.

program workpaper-based methods, and phone and email conversations with Department of Energy (DOE) and Enbridge steam trap experts. [66] The review clarified our approach and the specific data that needed to be collected on-site. While the utilities rely upon different prescriptive ex-ante savings for commercial, industrial low pressure, and industrial high pressure trap retrofits, the assumed parameter values used to estimate the industrial savings values do not account for the multiple functions of low and high pressure steam traps in industrial settings. Some of the key parameters that affect the estimation of ex-post energy savings include orifice size, annual steam trap operating hours, and average steam inlet pressures at the trap, boiler efficiency, and steam trap failure type.

Key Parameters

Table 4-8 below presents the key engineering parameters used to complete gross impact calculations and the data sources from which the information was collected. The measurement approach used for parameters is a combination of self-reported data, steam trap audits,[67] and field observations. Descriptions of these parameters and how they were collected are presented below. The observed on-site values of the key parameters are provided later in this section.

---

[66] Energy and Environmental Analysis, Inc. Steam Traps Workpaper. Prepared for Southern California Gas Company. PY2006-2008. December 2006.

Griffin, B. and D. Johnson. The Enbridge "Steam Saver" Program: Steam Boiler Plant Efficiency Update to Year-End 2005. Prepared for Enbridge Gas Distribution, Inc. March 2006.

Manczyk, H. Estimating the Cost of Steam Loss Through the Orifice of a Steam Trap. Manczyk Energy Consulting.

Pacific Gas and Electric Company. Nonresidential Demand Response Program Procedures Manual. 2008.

U.S. Department of Energy, Federal Energy Management Program. Pacific Northwest National Laboratory. Federal Technology Alert: Steam Trap Performance Assessment. 1999.

U.S. Department of Energy, Office of Industrial Technologies. Alliance to Save Energy. Steam Digest. 2001.

U.S. Department of Energy, Office of Industrial Technologies. Improving Steam System Performance: A Sourcebook for Industry. 2002.

U.S. Department of Energy, Office of Industrial Technologies. Energy Matters: Steam Challenge Special Supplement. January 2006.

kW Engineering. Steam Trap Survey and Billing Analysis Report. Prepared for Southern California Gas Company. December 2006.

Yarway Corporation. Simple Techniques for Sampling Steam Traps – Steam Trap Application Data STA-9. 1988.

Yarway Corporation. Estimating Steam Loss – Steam Trap Application Data STA-7. 1980

[67] Steam trap audits document the results of detailed examinations of the steam traps within a facility. They can be conducted by trained in-house maintenance staff or by third party inspectors. The audits contain information about whether steam traps are functioning properly, leaking, or have failed. Audits contain important information that allows facilities to determine whether they are losing steam, and hence, wasting energy.

Table 4-8: Key Parameters Used in Gross Impact Calculations for Industrial Applications

Engineering Parameters	Utility Tracking Data	Industrial Phone Survey	Application Data	Small Commercial On-Site and Phone Survey	Industrial On-Site Survey
Pressure at Individual Traps	X*	X*		X	X
Boiler Efficiency				X†	X
Steam Trap Manufacturer and Model			X	X	X
Steam Trap Orifice			X	X	X‡
Failure Type				X	X
Annual Steam Trap Operating Hours				X	X
Number of Traps Replaced	X	X	X	X	X

* These data sources only specify whether the steam trap pressure is > or < 15 psig.
† Nameplate efficiency is taken from steam trap manufacturers' websites.
‡ Orifice size is found from the make and model of steam traps, which are collected on-site.

Pressure at Individual Traps

Pressure data were collect on-site wherever possible. If the steam pressure was not available during the on-site data collection effort, the data were collected over the telephone from the steam system manager.

Boiler Efficiency

The efficiency of the steam generation boiler plays an important role in the calculation of energy savings. This information was collected during the on-site surveys. In the majority of the cases, boiler efficiency was obtained from the boiler operator. Boilers over 2 million Btu/hr are subject to air quality restrictions that require the boiler operator to test the boiler efficiency at regular intervals. Boiler efficiency was also available for sites where the boiler had been recently serviced for regular maintenance. If the boiler efficiency was not available from the operator, then flue gas analyses were performed by the on-site engineer whenever possible. In cases where the boiler efficiency was not obtained by the above methods, it was estimated based on the make, model number, and age of the boiler.

Steam Trap Orifice Diameter

Orifice diameter is one of the key parameters in the calculation of energy savings. Steam trap manufacturer, model number, and pipe size were collected for each rebated steam trap from steam trap audits and invoices, when available. The orifice diameter can be obtained from the manufacturer specification sheets in most cases, once the trap has been properly identified. When orifice size was not available from specification sheets, phone inquiries were made to the steam trap manufacturers.

Steam Trap Failure Type

Baseline steam trap failure type is also a critical parameter in the calculation of energy savings for replaced steam traps. The failure type was collected from on-site surveys, phone surveys, or the steam trap audit provided by the facility manager/maintenance staff. An adjustment factor was calculated to account for the fact that actual steam loss varies from the theoretical steam loss based on the failure type. This adjustment was termed the leak factor in the savings calculation equation; this factor varies from 0 to 1 depending on the failure type. Failure type and leak factor are classified into the following categories:

Failing or Blowing Through. A trap is considered failed open or blowing through if it is blowing live steam into the condensate return line. For these cases, the actual leak rate is the maximum theoretical leak rate and hence the leak factor is 1.

Failed Closed or Blocked. A trap is considered failed closed if it is not passing condensate and is backing up into the steam line. Traps that are failed closed will have an indirect impact on plant efficiency, but do not lead to actual steam loss. Since the steam loss from blocked steam traps is zero, the leak factor of zero is used for these traps in energy savings equations.

Leaking. A trap is referred to as leaking if the trap is passing live steam through the condensate return line. In some survey audits, this is further classified into small, medium, and large leaks. Based on the findings from on-site surveys, Itron decided to use a leak factor of 0.75 for all the traps that are failed leaking since the onsite evaluation found that the majority of the traps in this category have large or medium leaks. The leak factor will account for the fact that the actual leak rate is less than the theoretical leak rate.

Unknown. For some of the traps, the failure type was unavailable from the survey audit or from the site contact. In such cases, a leak factor of 0.5 was used to account for the fact that the actual leak rate is less than the theoretical leak rate. In the cases where failure type is unknown, the uncertainty in the savings calculation will be higher because of the lack of direct information. Note that the standard assumption in the literature on energy savings from steam trap retrofits, including the SCG workpaper upon which the IOUs base their ex-ante energy savings estimates, is an average leak rate of 0.5 since failure type is unknown.

Annual Steam Trap Operating Hours

Annual steam trap operating hours were collected from the on-site data collection, additional phone surveys for Small Commercial Verification on-site surveys, and steam trap audits. In cases where a steam trap's operating hours are zero, the ex-post energy savings from a retrofit of that steam trap is assumed to equal zero.

Verification

For each on-site impact assessment completed by a research engineer, he/she verified essential tracking data including the locations, sizes, configurations, makes, and models of the replacement traps. In addition to verifying the above mentioned parameters, Itron also verified that the steam trap was a retrofit and not a new addition to their system and that the trap is currently installed and operational. Engineers also collected enough information about steam generation systems to make sure that utility gas was used to produce the steam associated with the steam traps. Traps that qualify for the rebate programs offered by the IOUs are for steam trap replacements that are currently in operation and rely upon natural gas supplied by the California IOUs. New traps, traps that are not installed and in operation, and/or traps that do not use natural gas supplied by the IOU do not qualify for the rebate and therefore, have ex-post gross savings set equal to zero.

Industrial Steam Trap Gross Savings Algorithm

Annual therm savings from the replacement of faulty steam traps can be calculated using the following equation:

$$\text{Annual Therm Savings Per Trap} = \frac{M \times LF \times CF \times OH \times HV}{100 \times BE}$$

where:

M = Steam flow rate (lb/hr)

LF = Leak Factor[68]

CF = Condensation Factor

OH = Annual operating hours of steam trap[69]

HV = Heat of vaporization of steam (kBtu/lb) [70]

BE = Boiler Efficiency[71]

[68]These parameters are explained in detail in the previous subsection. (Key Parameters)

[69]Ibid.

[70]Heat of evaporation of steam produced for any given pressure can be obtained from the steam tables.

[71]Ibid.

Steam Flow Rate (M)

There are many equations to calculate the steam flow rate. Itron undertook a literature search and determined that the modified Napier's equation (also known as Spirax Sarco equation) is widely used in the industry to calculate the steam flow rate.[72] The literature search bibliography is provided in Appendix B-4. As per the Napier's equation

$$Steam\ Flow\ Rate\left(\frac{lb}{hr}\right)=\frac{Orifice\ Area\,(sq.in.)\times Steam\ Pressure\,(psia)\times Discharge\ Coefficient}{70}$$

PSIA steam pressure = Steam gage pressure + 14.7 PSI

Discharge Coefficient = 0.62 [73]

Condensate Factor (CF)

Condensate, when passing through the orifice, prevents steam leakage. A condensate factor is used to account for this. The Spirax Sarco literature indicates that for a fully loaded and properly sized trap, the condensate blocks steam approximately 25% of the time. For all the traps, a condensate factor of 0.75 was used.

4.2.4 Net-to-Gross Evaluation

This section presents the evaluation methodologies used to estimate the NTG ratios for the retrofit of steam traps in commercial and industrial applications. Methodological frameworks were developed by working groups for programs evaluated under the CPUC Small Commercial and Industrial Contract Groups and summaries of these guidelines are presented in the NTG Methodology subsection of the Pipe Insulation HIM chapter of this report.[74] The guidelines were applied for each evaluation with the understanding that certain adjustments could be made to account for program- and/or measure-specific characteristics. The remainder of this section briefly describes the methodologies used to estimate NTG ratios for commercial and industrial sites that replaced their steam traps.

Small Commercial Applications

The self-report approach established for residential and small commercial customers is applicable to the commercial market segment, and hence the commercial application of steam trap retrofits. This method is designed for the following situations:

---

[72]The Spirax Sarco equation is used by Embridge Gas Distribution, Inc., as well by SCG in its steam trap workpaper.

[73]The discharge coefficient takes into account the fact that steam traps have a geometry that is much more complex than a simple orifice. This coefficient is applied to the flow of steam through a simple orifice and depends on the shape of the orifice. For calculation purposes, discharge coefficient (0.6) of sharp-edged orifice is used.

[74] The small commercial and industrial NTG methodological frameworks developed by the NTG Working Groups for the CPUC evaluations are presented in full in Appendix A-2 and Appendix A-3.

- The program or measure addressed involves a fairly uniform technology and application across end-use participants.
- Free ridership can reasonably be addressed at the end-user level. That is, the major influence of the program has been through means such as rebates or promotions that the participating decision-maker is likely to have been aware of at the time of the purchase decision. In these cases, influence of the program on vendors or other intermediaries is assumed not to be a major factor in ultimate purchase decisions.

The first criterion holds for steam traps installed by commercial sites, as the variability in these applications is limited. The data for the steam trap NTG analysis was gathered using CATI administered phone surveys, including the Small Commercial Verification Survey and the Commercial Steam Trap Survey. Phone survey instruments used for the commercial steam trap sites are included in Appendix B-1.

Large Commercial and Industrial Applications

The Large Nonresidential NTG methodology is applicable for industrial, commercial chains, and large commercial applications. Despite the label, this method is adaptable to small and medium-sized commercial and industrial customers. Separate batteries of questions and interviewing procedures were developed for three rigor levels: Basic, Standard, and Standard-XL. The Basic rigor represents the simplest set of questions and interviewing procedures. Standard rigor questions are of medium complexity. Standard-XL rigor questions are the most complex and comprehensive. Customers are assigned to one of these three levels based on the size and complexity of their projects. For large nonresidential applications, the requirement is to use the Basic level of rigor for all projects with rebates of less than $50,000, the Standard level of rigor for all projects with rebates between $50,000 and $200,000, and the Standard-XL for applications with rebates over $200,000. The steam trap NTG evaluation included Basic and Standard level rigor projects. The Basic and Standard NTG batteries used in the steam trap analysis are included in the survey instruments presented in Appendix B-1.

The NTG analysis for large commercial and industrial sites that installed steam traps relied upon 128 sites that completed the Basic or Standard battery of NTG questions. Ninety-four NTG surveys were conducted by the CATI center and 34 were conducted by an energy consultant supervised by an engineer. Of the 128 sites, 125 of them are industrial sites while the other three are commercial sites that had enough ex-ante therm savings to go through the industrial NTG battery. One went through the Standard NTG battery and the other two went through the Basic NTG battery.

NTG Questions and Scoring Algorithm

The core net-to-gross questions were maintained in the phone surveys; however, the research team added extra questions in an attempt to capture information unique to steam trap replacement practices and program influence versus other factors. These questions were added to both the Residential/Small Commercial surveys and Large Nonresidential surveys.

For this technology, however, there are additional issues that may suggest free ridership. One such area relates to ongoing steam trap maintenance practices. Steam trap maintenance programs were common among industrial applicants and may indicate that the measure would have been installed in the absence of the program. Evidence of an ongoing steam trap maintenance program was used to help clarify the NTG ratio in those situations where the industrial site gave conflicting responses during the net-to-gross battery of questions.

4.2.5 Weighting Methodology

This section describes the development of weights applied to each sample point when aggregating results. This methodology was used for both industrial and commercial analysis.

Gross Impact Weights

As described in the section above, the participant population was divided into 6 strata based on therm savings, rebate amount, and measure type. The site weight for each site within the sample population was calculated by dividing the ex-ante therm savings for that site by the total ex-ante therm savings of the sampled sites from those strata. The weight for site *i* in strata *j* is:

$$W_{SSSi} = \frac{ExAnteTherms_i}{\sum_i ExAnteTherms_i}$$

where the denominator is summed from *i* to the number of sampled sites in the strata. Using that weight, the mean weighted realization rate is calculated for each stratum.

$$Weighted\ Mean\ \mathrm{Re}alization\ Rate_j = \frac{\sum_i W_{SSSi} * RR_i}{\sum_i W_{SSSi}}$$

The mean weighted realization rate is calculated for each stratum across the gross verification sample sites and is applied to the ex-ante gross therm savings across the entire population for the stratum to derive the weighted gross ex-post therm savings for the population.

Net Impact Weights

The net-to-gross sampling and the gross telephone sample are nearly identical. The methodology used to develop the net-to-gross weight is very similar to the methodology to develop the gross impact weights. The primary difference between the gross weights and the net weights is that the net weights depend on ex post therms, not ex ante. The site weight for each site within the telephone survey sample was calculated by dividing the gross ex-post therm savings for that site divided by the total gross ex-post therm savings of the sampled sites from those strata. The formula to calculate strata weights is the same as the gross impact formula above. The sites weights are used to calculate the weighted mean by net-to-gross ratio stratum. The formula for the weighted mean net-to-gross ratio for strata j is listed below.

$$Weighted\ NTGR_j = \frac{\sum_i W_{SSSi} * NTG_i}{\sum_i W_{SSSi}}$$

Mean weighted NTGR is calculated for each stratum across the gross verification sample sites and is applied to the ex-ante gross therm savings across the entire population for the stratum to derive the weighted net ex-post therm savings for the population.

4.3 Validity and Reliability

This evaluation took steps to increase both the validity and reliability of measurement for each of the parameters being estimated. The evaluation worked to minimize response bias for survey based results and recruitment and undertook uncertainty analyses both before and after on-site visits were conducted.

4.3.1 Minimizing Response Bias for Survey Based Results and Recruitment

The evaluation conducted telephone surveys to support the development of the impact parameters. Several phone surveys were undertaken, including free ridership surveys for both the commercial and industrial applications, gross evaluation phone surveys for both the commercial and industrial sectors, and callback surveys to collect additional information from sites previously interviewed under the Small Commercial Verification effort. The phone surveys were designed differently for the commercial and industrial applications due to the differences in information required for the analyses. All phone surveys were implemented as a census of the given population of participants.

All the commercial surveys were conducted using Computer Aided Telephone Interviewing (CATI), while the industrial surveys were conducted by either the CATI Center or an energy consultant working in conjunction with an energy engineer. A key step to ensuring the validity and reliability of the steam trap analysis was to minimize non-response bias. One specific step taken was to attempt to contact a respondent multiple times at different times of the day, different days of the week, and different weeks of a month. For commercial and industrial customers, the research team called during normal business hours. Call backs were also scheduled for respondents at a time that was most convenient for them. For larger industrial customers, the team also spoke with their utility representatives to ensure that the correct contact name and phone number for the site was being used. In cases where an industrial applicant was hesitant to speak with the evaluators, the utility representative was asked to intercede on the evaluation team’s behalf.

4.3.2 Free Ridership and Net-to-Gross

Free ridership and NTG were based on self report analyses utilizing telephone surveys for both the commercial and industrial applications. The questions and algorithms used to estimate free ridership and net-to-gross ratios were pretested prior to full scale implementation. Other steps taken to ensure the validity and reliability of the NTG analyses include a stability analysis of the commercial steam trap NTG estimation and an in-depth examination of inconsistent answers to NTG questions asked of industrial steam trap sites. The commercial steam trap NTG stability analysis is presented in Appendix B-5 and the consistency checks completed for the applicable industrial steam trap sites are provided in Appendix B-6. Free ridership and NTG telephone surveys attempted to identify the key decision-maker and then ask the battery of questions. These batteries included multiple questions to develop the scores to increase the reliability of the results. In cases where responses were inconsistent among the multiple questions, an experienced energy consultant reviewed responses to open-ended questions that asked for clarification of the inconsistency and recoded the individual responses to resolve the inconsistency.

4.3.3 Commercial Billing Analysis

A billing analysis was conducted for commercial dry cleaning steam trap applications. The utility-specific per unit ex ante energy savings estimates were used as the steam traps savings inputs to the analysis. As many commercial dry cleaning sites installed both steam traps and pipe insulation, the gross energy savings estimates developed as part of the pipe insulation evaluation were used as inputs for the pipe insulation savings in the steam trap model. Including the pipe insulation savings term in the regression equation helps to ensure that the steam traps savings realization rate does not include the pipe insulation savings. The billing analysis also included variables to control for changes in economic activity, including a county-level change in the unemployment rate, and a time series fixed effect. Both of these variables help to ensure that changes in the macro economy are not attributed to the steam traps savings variable. The billing analysis statistically adjusts the ex-ante and engineering impacts to help correct and/or improve the reliability of the savings numbers.

4.3.4 Industrial Uncertainty Analysis

Uncertainty analyses were performed for the industrial steam trap engineering analysis both before and after on-site visits were conducted. The uncertainty analysis conducted during the planning phase of the on-site engineering evaluation helped the research team determine which parameters contribute most to the overall uncertainty of therm savings from steam trap retrofits. Armed with this information, the on-site engineer had a clearer understanding of the parameters that were most essential in reducing the uncertainty of the savings analysis. The essential parameters became a central focus of the on-site data collection effort.

The uncertainty analysis was performed again after the on-site data was collected from the industrial steam trap sites. The data collected during the on-sites was used to help refine the team’s understanding of the degree of uncertainty surrounding the parameters. Six scenarios were investigated to span the range of uncertainties that were encountered in the field due to different data collection techniques. For example, in some cases the team was able to obtain boiler efficiency information from an efficiency test whereas in other cases the boiler efficiency was estimated from nameplate data. Three scenarios were completed for high pressure traps and three for low pressure traps. For each pressure type, one minimum uncertainty scenario was run, in which all parameters used in the estimation of therm savings were assumed to be measured with the minimum level of uncertainty. The team also ran two maximum uncertainty scenarios for each pressure type; in one the steam leak factor is known and in the other it remains unknown. Two maximum uncertainty scenarios were run for each pressure type in order to show the significance of the steam loss parameter.

The uncertainty analysis was performed using Crystal Ball software to run a Monte Carlo simulation to propagate the uncertainty through the engineering savings calculation. The results of the uncertainty analysis after the on-site visits were completed are presented in the table below. The mean and standard deviation vary across the three low pressure or high pressure scenarios because the values of the input variables did not remain consistent in the various scenarios. For instance, there is a higher uncertainty in number of traps when there are more steam traps since a small number of traps can be counted more accurately than a large number of traps. So, in the minimum uncertainty case a small number of traps was used, and in the maximum cases a large number of traps was used. Similarly, there is less uncertainty in the hours of operation when the facility operates 24 hours a day, seven days a week then when the facility operates five days a week. Hence, the hours of operation were different in the high and low uncertainty scenarios. The variable with the largest contribution to the error in the maximum uncertainty cases is the steam leak factor, which is determined by trap failure type.

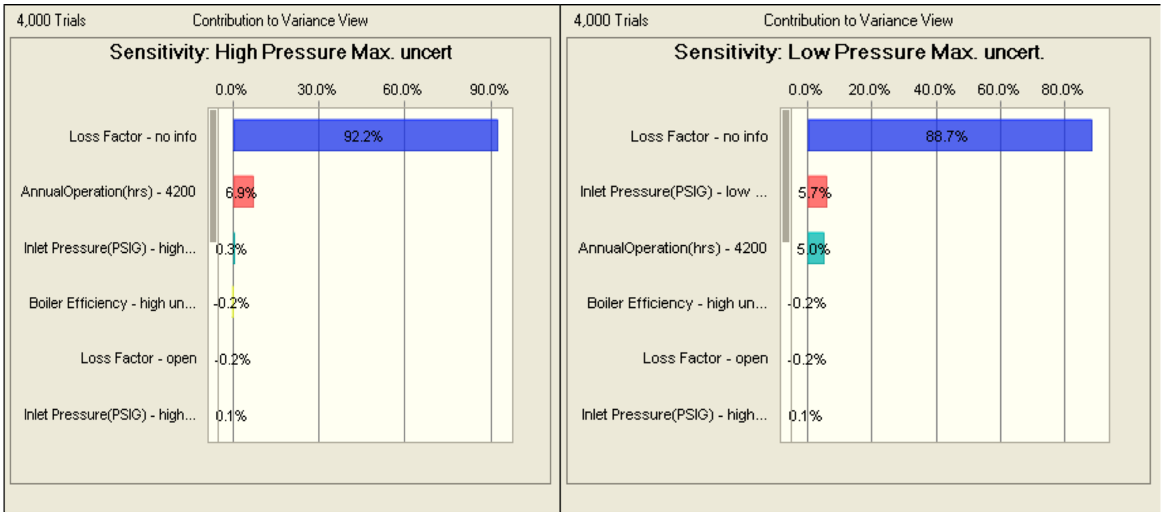
As Table 4-9 shows, knowing the steam leak factor reduces the maximum uncertainty from 149% to 47% for high pressure steam traps, and reduces the uncertainty from 159% to 63% for the low pressure traps.

Table 4-9: Relative Precision of Uncertainty Analysis Estimating Average Therm Savings from Steam Trap Retrofits

Uncertainty Analysis Scenarios	Mean Therm Savings	Standard Deviation	Relative Precision at 68% CI	Relative Precision at 90% CI
High Pressure, Minimum Uncertainty of all parameters	169,832	6,213	4%	6%
High Pressure, Maximum Uncertainty with known Leak Factor	525,124	151,807	29%	47%
High Pressure, Maximum Uncertainty, with unknown Leak Factor	521,005	472,250	91%	149%
Low Pressure, Minimum Uncertainty of all parameters	16,966	996	6%	10%
Low Pressure, Maximum Uncertainty with known Leak Factor	52,742	20,236	38%	63%
Low Pressure, Maximum Uncertainty, with unknown Leak Factor	52,290	50,550	97%	159%

The importance of the steam leak factor can also be seen in the following two charts that show each variable's contribution to the variance in the estimation of gross therm savings from steam trap retrofits. For the high pressure traps with maximum uncertainty, the leak factor is assumed to be unknown. In this case, the steam leak factor contributes over 92% of the uncertainty in the calculation of gross therm savings. While the leak factor's contribution to overall uncertainty is slightly lower for the low pressure trap maximum uncertainty scenario, it is still close to 89%.

Figure 4-1: Maximum Uncertainty Scenarios with Unknown Steam Leak Factors for High and Low Pressure Traps



In cases where the steam leak factor is known, the operating hours and average inlet pressure at the steam trap are the parameters that are significant in the uncertainty analysis. This can clearly be observed in the additional sensitivity scenario charts presented in Appendix B-7.

Based on the uncertainty analysis, it is clear that knowledge of the steam leak factor is essential to improving the validity and precision of the gross therm savings estimates from steam trap retrofits. One reliable source of steam leak factor information for industrial steam traps is steam trap audits that have been conducted prior to the retrofit of faulty and/or failed steam traps. Making results of steam trap audits a requirement would considerably reduce uncertainty in the therm savings estimation.

4.4 Findings and Results

Detailed findings and results from the analyses of steam traps in commercial and industrial applications are presented in this section. Gross ex-post savings for steam traps in commercial and industrial applications are first analyzed and results of the therm savings per trap are presented. In addition to the gross ex-post savings per trap, the net-to-gross analysis results are shown. The results of the NTG analyses are used to calculate net ex-post savings per trap in both commercial and industrial applications.

4.4.1 Commercial Applications of Steam Traps

Commercial steam traps represented 92% of the sites installing rebated steam traps but only 24% of the savings. The typical commercial steam trap applicant was a small dry cleaner replacing all or nearly all of their traps. The relatively homogeneous nature of the applicants and the size of the ex ante per site claims led the team to select an SAE billing analysis as the evaluation method. The objective of the SAE billing analysis is to determine the first-year energy impacts for steam traps installed in commercial applications. The billing analysis is estimated separately for steam traps installed in PG&E and for those installed in the Sempra Utilities.[75] The billing analysis is specified using customer billing data, independent variables gathered during the telephone survey, customer-tracking data that indicated the timing of the installation of the rebated measures and the energy impacts, weather data, and information on monthly county level unemployment rates.[76]

Data Aggregation

The billing analysis was performed at the site level, necessitating the aggregation of the account level billing data to a unique site level. Once the billing data were aggregated to the site level, the billing data, tracking data, and the phone survey data were merged. During the aggregation process, care was taken to ensure that the impacts for multiple gas measures rebated and installed at a site over multiple periods were aggregated to the site level. The merging of the survey, tracking, and billing data led to the development of the analysis database by Site ID. Table 4-10 lists the number of sites by utility from the tracking data and then the number of sites following the merge of the tracking, billing, and survey data sources.

Table 4-10: Number of Commercial Sites by Utility

Utility	Tracking Data Site Count	Tracking + Phone Survey + Billing Data	Remove Non-Dry Cleaners	Final Billing Analysis Data
PG&E	1054	172	157	150
SCG	2,036	330	323	314
SDG&E	205	37	34	33
Total	**3295**	**539**	**514**	**497**

Once the data were merged into a site level dataset, they were censored to only include dry cleaning sites.[77] After the commercial, non-dry cleaning sites were removed, the data were further reviewed prior to undertaking the billing analysis. The screens applied to the aggregated dataset included screens for sites with incomplete billing data, billing data with very low usage, and billing data displaying highly unusual patterns that may indicate that

[75] The billing analysis combined the sites from SDG&E and SCG because there were only 33 SDG&E sites and the ex ante per unit claims were the same for SDG&E and SCG.

[76] The information on the installation date differed by utility. PG&E provided information in their tracking data on the application date, the verification date, and the paid date. For PG&E we used the application date. SDG&E and SCG only provided one date, the paid date. It is likely that PG&E's date is closer to the actual time of installation than the SCG or SDG&E dates.

[77] Restricting the sample to dry cleaning sites has a small impact on the number of sites and results in a more homogeneous data set better suited to a billing analysis. Some of the non–dry cleaning sites have substantially larger consumption than the dry cleaning site. In addition, the workpaper determination of savings was modeled on a dry cleaning site.

portions of the bills were missing, that the site had changed ownership, or that other significant changes were occurring at the site. For PG&E, seven sites were removed due to billing issue, while nine sites were removed from SCG and one site was removed for SDG&E. The right-most column in Table 4-10 lists the number of sites in the final billing analysis dataset.

SAE Model Results

Table 4-11 presents the estimates for the PG&E and the Sempra steam trap models. The model includes independent variables that control for the ex ante savings estimates for steam traps, pipe insulation, and other rebated gas measures. In the PG&E model, the savings from pipe insulation and other gas measures are combined into one variable given the small number of sites that installed steam traps and any other gas measures under the program. In the Sempra model, the savings from pipe insulation and other measures are represented separately in the analysis. Prior to implementing the model, the team incorporated the engineering adjustments to the pipe insulation savings that were estimated in the pipe insulation evaluation.[78] In particular, the ex ante savings for pipe insulation were reduced to 4.5% of their original level prior to estimating the model.

The model performed well in terms of estimating the effects of the rebated steam trap installations on the change in consumption. The estimated coefficients for steam trap savings is shown to be significantly different from zero with a better than 95% probability. However, the estimated coefficients for steam trap savings show that the savings are substantially less than the ex ante claims for both PG&E and the Sempra Utilities.

Table 4-11: Commercial Steam Trap Models with 12 Months Prior Billing Data– PG&E and Sempra[79]

Regressor	PG&E Coefficient	PG&E T-Statistic	Sempra Coefficient	Sempra T-Statistic
Daily Steam Trap Savings	-0.30338	-2.31	-0.11761	-5.97
Daily PI and Other Savings	-0.75905	-2.53		
Daily PI Savings			-0.80276	-7.57
Daily Other Gas Savings			-1.54744	-10.31
Change in HDD*Have Gas Heat	0.03137	0.29	0.00373	0.07
Change in unemployment	-0.17287	-1.17	0.02731	0.25
Increase in Site	0.77814	2.23	0.73544	4.20
Decrease in Site	-0.58643	-1.93	-0.17355	-1.43
Site Remodel			-0.98293	-2.97
R^2	0.0219		0.0588	

[78] See Chapter 3 of this report.

[79] The coefficient values for the time fixed effects are not reported in the table but are included in the model.

The model results presented in Table 4-11 show that the realization rate for commercial steam traps installed in PG&E's territory was 30.34% and statistically significant while the realization rate for the Sempra Utilities was 11.76% and statistically significant. The realization rate for the PG&E model is relative to their ex ante claims of 45.87 therms per trap while the Sempra Utilities claimed 139 therms per trap.[80]

In the Sempra model it was also possible to estimate the realization rate for pipe insulation and other energy efficiency gas measures that were rebated through the program and installed at dry cleaners that also installed steam traps. The estimated realization rate for pipe insulation was 80.27% while the realization rate for other gas measures was 154.74%. The pipe insulation realization rate is calculated after taking into account the ex-post pipe insulation engineering evaluation findings for dry cleaners. The pipe insulation engineering evaluation reduced the ex post pipe insulation savings to 4.5% of their ex ante value. Using a 95% confidence interval, the estimated realization rate for pipe insulation in the Sempra billing analysis model is not statistically different from the ex-post engineering estimate of pipe insulation savings. These results further support the low realization rate found in the engineering analysis of pipe insulation in dry cleaning applications.[81] The realization rate on other gas measures is relatively large and statistically different from zero. The primary measure in the other gas measure group is boiler cleaning.

The other independent variables in the model include a weather variable interacted with the site having gas heat, a 12 month change in the county level unemployment rate, a binary variable indicating the site did something likely to increase their usage (increased their hours, employees, or square footage), something likely to decrease their usage, a binary indicator that the site remodeled, and time series fixed effects.[82] The model shows that an increase or decrease in the sites hours, employees, or square footage led to a change in usage of the

[80] The per-site steam trap savings are lower for PG&E than the Sempra utility claims because PG&E chose to reduce their per steam trap ex ante claims following an evaluation completed in 2007 (KEMA, Inc. Steam Trap Impact Assessment. Prepared for Pacific Gas & Electric Company. October 2007). It is not surprising that the estimated realization rate is further reduced by this evaluation. The previous evaluation results were based on an analysis that restricted the sites to those that did not experience an increase in gas usage following the installation of steam traps.

[81] Given that the billing analysis coefficient for pipe insulation is 80% and the 95% confidence interval of this coefficient includes 100% of the 4.5% engineering realization rate, the steam trap billing analysis findings supports the use of the updated engineering estimates for pipe insulation. The pipe insulation ex-post engineering savings estimate was evaluated using a sample drawn from pipe insulation sites. The steam trap analysis was focused on steam trap sites and the billing realization rate of pipe insulation within this analysis is for sites that installed both pipe insulation and steam traps. Given the similarity of the results and the specificity of the pipe insulation analysis, the billing analysis results should be used to support the pipe insulation findings.

[82] The coefficients for the time series fixed effects were not presented in Table 4-11 due to space limitations. The time series effects tend to be positive during the early periods of the analysis and negative during the later periods.

expected sign. In the Sempra Utilities, sites that remodeled also decreased their usage.[83] The 12 month change in unemployment was not statistically significant due to the inclusion of the time series fixed effects.[84]

Confidence Intervals, Statistical Precision, and SAE Savings Estimates

Table 4-12 includes steam trap realization rates and confidence intervals calculated using a 90% level of precision for both the PG&E and Sempra participant samples. The realization rate for both territories is statistically different from zero, though significantly below a 100% realization rate.

Table 4-12: Small Commercial Steam Trap Realization Rates and Precision Bounds

Service Area	Coefficient Estimates	Standard Error	Lower Bound 90% Precision	Upper Bound 90% Precision
PG&E	-0.30338	0.13130	-0.0874	-0.5194
Sempra Utilities	-0.11761	0.01970	-0.0852	-0.1500

The SAE realization rates are applied to the ex ante per unit term estimates in Table 4-13 to calculate ex post estimates of savings. The application of realization rates to the ex ante therm estimate leads to ex post estimates that are substantially smaller. The ex post estimates for SCG are slightly larger than those for PG&E, but within the 90% confidence interval of the PG&E estimates.

Table 4-13: Small Commercial Steam Trap Gross Ex Ante and Gross Ex Post Estimates of Savings Per Trap

Service Area	Ex Ante Therms Per Trap	Coefficient Estimates	Ex Post Therm Savings Per Steam Trap	Lower Bound 90% Precision	Upper Bound 90% Precision
PG&E	45.87	-0.30338	13.92	4.01	23.82
Sempra Utilities	139	-0.11761	16.35	11.84	20.85

Net-to-Gross Ratio and Net Ex-Post Commercial Results

The commercial self report net-to-gross methodology was implemented to estimate the NTG ratio for commercial steam trap applications. Table 4-14 lists the number of sites used in the

---

[83] The remodel variable was tried in the PG&E model. Too few sites remodeled, however, to allow the model to estimate the coefficient. The remodel variable is a binary variable that equals one for the 12 months following remodeling.

[84] If the model is specified without the time series fixed effects the 12 month change in unemployment has a negative and statistically significant effect on the change in usage. The team chose to keep the unemployment change variable because it allows the effect of the macro economy to vary by geographic location while the time series fixed effects vary over time but are fixed geographically.

self report methodology, the average NTG ratio by strata, the upper and lower bounds and the relative precision. PG&E was found to have the lowest NTG ratio at 0.62 while the Sempra Utilities had similar NTG ratios of 0.70 and 0.72. The confidence interval is tightest on the SCG NTG estimate, consistent with the larger sample size while the interval is largest for SDG&E, the utility with the smallest number of survey respondents.

Table 4-14: Small Commercial Net-to-Gross Ratio

Strata	Site Count of Net-to-Gross Phone Surveys	Workpaper NTG Ratio	Realized NTG Ratio	90% Lower Bound	90% Upper Bound	Relative Precision
PG&E	174	0.96	0.62	0.58	0.66	0.06
SCG	309	0.96	0.70	0.67	0.73	0.05
SDG&E	38	0.96	0.72	0.64	0.80	0.11

Table 4-15 combines the results from the NTG and the gross analysis to produce the ex-post net savings per steam trap. The gross ex ante savings per trap was 45.87 therms for PG&E and 139 therms for the Sempra Utilities. After applying the estimated realization rate and the NTG ratio, the ex post net therm savings per trap were 8.63 therms for PG&E, 11.43 therms for SCG, and 11.78 therms for SDG&E.

Table 4-15: Small Commercial Net Ex-Post Therm Savings

Strata	Total Gross Ex-Ante Therms	Number of Traps	Gross Ex-Ante Therms per Trap	Realization Rata	NTG Ratio	Ex-Post Net Therms per Trap
PG&E	1,051,340	22,920	45.87	0.30	0.62	8.63
SCG	4,646,492	33,428	139	0.12	0.70	11.43
SDG&E	514,022	3,698	139	0.12	0.72	11.78

4.4.2 Industrial Applications of Steam Traps

Industrial steam traps represented only 8% of the sites installing rebated steam traps but accounted for 76% of the ex-ante savings claims. Industrial steam traps were installed in a broad array of different industrial settings, ranging from large refineries and defense manufacturing sites to light industrial sites producing antique jewelry and chair coasters. The different types of applications were associated with many different types of traps, hours of operation, and steam pressure. Given the very heterogeneous nature of steam trap applications, the team chose to analyze industrial steam traps using an on-site data collection approach. The collection of on-site data focused on the key parameters used to calculate the steam and energy savings per trap.

The following section presents the engineering estimates of the gross ex-post therm savings of industrial traps. The section begins with a comparison of the work paper-based gross ex-ante therm savings per trap to the calculated gross ex-post therm savings per trap. This is followed by a summary of the mean values of the key parameters used to estimate gross ex-post therm savings by industrial low pressure and high pressure traps. The section concludes with the presentation of the self-report net-to-gross ratio and the net ex-post therm calculations.

Gross Ex-Ante and Ex-Post Therm Savings

Prior to conducting on-site visits in support of the engineering analysis, the research team developed a summary of the tracking data associated with participants who received rebates for steam traps in industrial applications. Table 4-16 presents the tracking data summary by high and low pressure traps and strata.[85] The strata for the industrial sites were defined to equalize the total ex-ante savings per strata. Measure type was not a determinant of strata since the stratification occurred at the site level and individual sites can install both high and low pressure traps. All but two sites in the industrial steam trap population exclusively had either high or low pressure traps.

Table 4-16: Summary of Industrial Steam Traps Tracking Data

Measure Type	Strata	Number of Sites	Gross Ex-Ante Therm Savings	Number of Traps	Gross Ex-Ante Therm Savings per Trap
High Pressure	1	5	6,581,020	2810	2,342
	2	19	6,295,296	2,688	2,342
	3	171	5,780,056	2,468	2,342
Low Pressure	1	0	0	0	0
	2	1	349,624	548	638
	3	79	810,260	1,270	638
Total		**273***	**19,816,256**	**9,784**	**N/A**

* The number of sites by measure type sum up to 275 since two sites fall into high pressure and low pressure trap categories. In actuality, there are 273 unique sites.

A majority of the industrial sites installed high pressure steam traps. In addition, total gross ex-ante savings and the number of traps installed are much higher at high pressure sites (18.65 million therms vs. 1.2 million therms and 7,959 high pressure vs. 1,818 low pressure traps).

A summary of the industrial steam trap on-site engineering analysis results is presented in Table 4-17. While only 41 of the 273 industrial steam trap sites were visited during the on-site data collection effort, the on-site's focus on the largest sites led to over 50% of the traps being reviewed in the on-site gross ex post savings analysis.

[85] The low pressure trap psig is assumed to be less than or equal to 15 psig while the high pressure trap is above 15 psig. This distinction between high and low pressure traps is based upon the rebate and ex-ante savings structure established by the IOUs for steam trap retrofits.

Table 4-17: Industrial Steam Traps On-Site Data and Results

Measure Type	Strata	Number of On-sites	Number of On-Site Traps*	Gross Ex-Ante Therm Savings per Trap	Gross Ex-Post Therm Savings per Trap	Strata RR
High Pressure	1	4	2,400	2,342	9,660	4.12
	2	13	1,954	2,342	3,395	1.45
	3	18	296	2,342	1,382	0.59
Low Pressure	1	0	0	0	0	0
	2	1	548	638	128	0.20
	3	5	62	638	1,946	3.05
Total		**41**	**5,260**	**N/A**	**N/A**	**N/A**

* The number of on-site traps is taken from the tracking data.

Based on the collection of key parameters for the on-site sample, gross ex-post therm savings per trap were calculated. A notable result from this analysis is that for the sample of sites included in the engineering analysis, gross ex post savings exceed gross ex ante savings for strata 1 and 2 high pressure steam trap sites as well as for strata 3 low pressure trap sites. The high ex post savings value calculated for strata 1 high pressure traps stems from the fact that these sites are all refineries that operate their steam traps for more annual hours and at higher pressures than assumed in the SCG work paper. This is discussed in detail in the key parameters section. The team also found a relatively high ex post savings for strata 3 low pressure steam trap sites. This can be explained by the high weighted average steam trap inlet pressure found on-site for the sampled strata 3 low pressure sites relative to the assumed pressure used in the work papers (61.9 vs. 10.9 assumed in the work papers). The on-site data collection revealed that many of the sites rebated under the low pressure specification actually operated as high pressure traps.

Table 4-18 presents the weighted mean realization rate, the lower and upper bounds and the relatively precision by measure type and strata. From this table, it is clear that the average realization rates are greater than 1 and that the average is very imprecisely estimated. The relative precision using a 90% confidence bound is consistently above 30%, with a maximum of 161% for strata 3 low pressure traps. The site level realization rate for strata 3 low pressure traps varies from 0 to 1,400%, leading to a very imprecise estimate of the mean realization rate. The lack of precision in the realization rate estimates is due to the highly variable nature of industrial steam trap applications and site level operations.

The HIM Plan specified a 90/14 confidence and precision level. Unfortunately the ex ante estimates of variability were substantially lower than what the team found on site. Therefore, while the team successfully completed all of the on-site data collection efforts, meeting quota for all of the strata, the precision of our final estimates fell far short of the expected level.

Table 4-18: Weighted Mean Realization Rates by Measure Type and Strata for Industrial Steam Traps

Measure Type	Strata	Weighted Mean Realization Rate	Lower Bound 90% Precision	Upper Bound 90% Precision	Relative Precision
High Pressure	1	4.12	2.83	5.42	0.31
	2	1.45	0.66	2.23	0.54
	3	0.59	0.25	0.93	0.58
Low Pressure	1	0.00	0.00	0.00	0.00
	2	0.20	N/A*	N/A*	N/A*
	3	3.05	-1.88	7.99	1.61

* Because there is only one low pressure strata 2 site, no standard deviation or lower and upper bounds are applicable.

Key Parameters

The values of the key parameters determine the gross therm savings from the retrofit of steam traps. The following discussion and tables present the average values of some of the salient parameters used to estimate gross ex-post savings from industrial steam trap replacement. Mean parameter values are presented by measure and facility type, where facilities are categorized as refineries or non-refineries. Refineries were found to differ significantly from other industrial sites that primarily rely upon high pressure steam traps; therefore results for these categorizations seem to be most meaningful in understanding the energy savings from replacing steam traps.

Pressure

Table 4-19 provides the weighted mean pressure for high and low pressure steam traps by facility type. The weighted average pressure for refineries is 165 psig, which is much higher than the weighted mean pressure from the SCG work paper (85.9 psig). For non-refineries, the weighted mean pressure for high pressure traps is close to the work paper assumption. For the low pressure traps, mean weighted pressure is slightly greater than 15 psig even though these traps were rebated by the IOUs as traps that operate at less than or equal to 15 psig. This slightly higher average pressure for non-refinery low pressure traps can be explained by the small sample size for this facility type and measure category (n = 6) and due to the inclusion of two sites that are operating their steam traps at 53 psig and 110 psig even though these traps were rebated as ones that operate at less than or equal to 15 psig.

Table 4-19: Summary Statistics of Pressure for Industrial On-site Analysis

Facility Type	Measure Type	Work Paper Pressure	Weighted Mean Pressure	Standard Deviation	Minimum Pressure	Maximum Pressure
Refinery	High Pressure	85.9	165.36	11.41	0	790
Non-Refinery	High Pressure	85.9	86.57	2.78	0	600
Non-Refinery	Low Pressure	10.9	17.76	1.79	0	110

Operating Hours

As Table 4-20 shows, the average weighted annual operating hours for high pressure steam traps in refineries is 8,011. Though the refineries run their traps for 8,760 hours throughout the year, the weighted mean annual operating hours are 8,011 after taking into consideration those traps at refineries that were not operational and/or not installed. For the non-refinery facility type, the annual operating hours are obtained from the facility's steam trap audit or from the site contact. Average annual operating hours for high and low pressure traps are considerably lower than the work paper assumption of 7,752 hours for both measure types.

Table 4-20: Summary Statistics of Operating Hours for Industrial On-site Analysis

Facility Type	Measure Type	Work Paper Hours	Weighted Mean Operating Hours	Standard Deviation	Minimum Operating Hours	Maximum Operating Hours
Refinery	High Pressure	7,752	8,011	215.36	0	8,760
Non-Refinery	High Pressure	7,752	4,860	142.95	0	8,760
Non-Refinery	Low Pressure	7,752	3,951	173.21	0	8,760

Failure Type

Table 4-21 provides the types of failure for high pressure and low pressure steam traps. For the high pressure traps, the most common trap failure is leakage. Forty-one percent of the high pressure steam traps failed due to leakage, 16% failed closed, and 16% were blowing through. The failure mode is unknown for approximately 17% of high pressure traps. For low pressure traps, the most common failure type is blowing through (67%), followed by traps that have failed closed (27%). No low pressure traps included in the on-site field analysis were found to be leaking and a smaller fraction of the low pressure traps had failed for an unknown reason (5%). Rebated traps that were new installations or not being used as replacements, as well as traps that were purchased but not installed are not included in this table.

Table 4-21: Distribution of Types of Failure of Industrial Steam Traps

Type of Failure	High Pressure		Low Pressure	
	Number of Traps	Percent of Traps	Number of Traps	Percent of Traps
Blowing Through	734	16%	406	67%
Closed	751	16%	163	27%
Leaking	1,866	41%	0	0%
Unknown	799	17%	32	5%
Total	**4,150**	**90%**	**601**	**99%**

* Total number of traps is less than the number included in the on-site survey because failure type is not applicable for new traps and traps that have not been installed, but for which rebates were paid to participants. This is also the reason that the percent of traps total is less than 100%.

Boiler Efficiency

The team collected boiler efficiency information for 27 sites. For a majority of these sites, boiler efficiency was provided the site contact. For a handful of sites however, the on-site engineer conducted a flue gas analysis to determine boiler efficiency. For the remaining 13 sites, boiler efficiency was unavailable either because:

- The site contact did not have information regarding boiler efficiency,
- The on-site engineer was unable to perform flue gas analysis, and
- The facility did not have a natural gas boiler.

Table 4-22: Boiler Efficiency Distribution

No of Sites	Boiler Efficiency
27	Collected from On-Site, Phone Survey, or through Flue Gas Analysis
8*	Estimated
5**	No Natural Gas Boiler

* Boiler efficiency was estimated from nameplate if the make and model number of the boiler was available. Efficiency was estimated as 80% where boiler information was unavailable.
** These sites either use waste heat from cogeneration to generate steam or use boiler fueled by digester gas.

For the 27 sites that Itron was able to collect boiler efficiency information, the weighted mean boiler efficiency was calculated and is provided in Table 4-23. As shown in this table, the weighted average boiler efficiency is over 83%, which exceeds the assumption of 80% made in the SCG work paper.

Table 4-23: Weighted Boiler Efficiency

Number of Sites	Work Paper Efficiency	Weighted Mean Boiler Efficiency	Standard Deviation
27	80%	83.3%	0.599

Orifice Size

Itron was able to collect orifice size for industrial steam traps either from manufacturer specification sheets or directly from steam trap manufacturers through phone inquiries. This information was essential to the estimation of therm savings for steam trap retrofits. In the savings calculation algorithm, orifice size of the baseline trap was used whenever available. These data were available from steam trap audits in many cases. If baseline information regarding orifice size was not available, then orifice size of the installed trap was used in the therm savings equation. Orifice size has a relatively large impact on the savings calculation as this parameter is squared in the equation used. For some traps, orifice size is proportional to the pipe size and hence pipe size can have a large impact on therm savings.

Net-to-Gross Ratio and Net Ex-Post Industrial Results

The industrial self report net-to-gross methodology was used to estimate NTG ratios for the industrial sites using data gathered from the telephone surveys. Table 4-24 lists the number of sites used in the industrial net-to-gross methodology, the average NTG ratio for high and low pressure traps, the upper and lower bounds and the relative precision. The NTG ratio for high pressure traps was 0.52 with a relative precision of 0.10 while the NTG ratio for low pressure traps was 0.57 with a precision of 0.09. The estimated NTG ratio is substantially lower than the work paper assumption of 0.96.

Table 4-24: Industrial Net-to-Gross Ratio

Measure Type	Site Count of Net-to-Gross Phone Surveys	Workpaper NTG Ratio	Realized NTG Ratio	Upper Bound	Lower Bound	Relative Precision
High Pressure	89*	0.96	0.52	0.57	0.47	0.10
Low Pressure	37*	0.96	0.57	0.62	0.52	0.09

* The number of sites by measure type sum up to 126 since one site falls into high pressure and low pressure trap categories.

Table 4-25 combines the gross ex-post realization rate with the NTG ratio to determine the net ex-post realization rate and savings per trap by pressure level. The average realization rates for low and high pressure industrial steam traps is greater than 200%. The NTG ratio for high and low pressure traps was 0.52 and 0.57, respectively. By combining the exceptionally high realization rate and the reduced NTG ratio, the final net ex post therm estimate for high pressure traps is 2,630 therms and 794 therms for low pressure traps. The estimated net ex-post therms per trap are slightly higher than the work paper gross ex ante values due to the very high gross realization rate.

Table 4-25: Measure Level Realization Rates, NTG Ratios, and Net Ex-Post Therm Savings per Industrial Steam Trap

Measure Type	Total Gross Ex-Ante Therms	Number of Traps	Gross Ex-Ante Therms per Trap	Gross Realization Rate	NTG Ratio	Net Ex-Post Therms per Trap
High Pressure	18,656,372	7,966	2,342	2.15	0.52	2,630
Low Pressure	1,159,884	1,818	638	2.19	0.57	794

4.5 Recommendations

Evaluation findings and recommendations stemming from the evaluation of commercial and industrial steam traps is presented in this section.

4.5.1 Findings & Recommendations for Future Programs

Commercial Steam Trap Recommendations

The commercial steam trap analysis revealed that the per trap realized therm savings were 14 therms for PG&E and 16 therms for Sempra Utilities. These results support the conclusion that the average commercial steam trap retrofit saves substantially less than the work paper assumptions. The billing analysis methodology used to evaluate commercial steam traps, however, does not help to clarify which of the work paper assumptions used to calculate the ex ante savings is inaccurate. [86] Onsite inspections of sites prior to the trap retrofit would be needed to determine the appropriate steam trap parameters.[87]

Over 70% of the commercial steam trap sites in the Sempra Utility service territories installed multiple rebated gas measures over the 2006-2008 program cycle. The high incidence of multiple rebated measures enabled the steam trap billing analysis in the Sempra Utility territory to estimate a realization rate for pipe insulation and other gas measures. The realization rate for pipe insulation, in combination with the engineering adjustments to savings, indicate that the ex post savings for pipe insulation is substantially lower than the ex ante savings. The realization rate for "other" gas measures, however was 150% of the ex ante values and the coefficient value was statistically different from zero at a 99% level.

[86] An evaluation of the work paper key assumptions would require an onsite evaluation using a pre-inspection methodology. This is the approach taken by kW Engineering to develop the work paper numbers (kW Engineering, *Steam Trap Survey and Billing Analysis Report*, Sponsored by The Southern California Gas Company, December 2006.). This paper was provided as an attachment to the SCG work papers upon request.

[87] During the development of the HIM Plan for steam traps, a proposal was put forth that an onsite review of pre-retrofit commercial sites be undertaken. The billing analysis approach was chosen due to the similarity with previous analyses, the homogeneous nature of the measure, and the time associated with an onsite review given the limited time available once HIM measures were designated. In addition, a pre-retrofit review of sites would have necessitated that the onsite analysis be undertaken on sites installed in 2009.

While other gas measures were made up of many measures, the majority of the savings represented a limited number of sites who received a rebate for boiler cleaning.

While the team did not attempt to evaluate the merit of the boiler cleaning measure, it does help to refocus the results from the analysis. Steam traps work on a boiler system with many functioning parts. Focusing a dry cleaner steam savings program on a single component of the system may miss other cost-effective boiler system savings opportunities.

Given the short payback associated with replacing a failing steam trap, cleaners should be willing to replace the measure if they are aware that it has failed. The work paper emphasizes the assumption that most dry cleaner owners do not know when their traps have failed. The findings from the steam trap billing analysis, both the low realization rate on traps and the high realization rate on other measures, supports a more broadly focused steam assessment program. The steam assessment program could be designed to provide a high quality professional audit, supplying the dry cleaner with information on measures that need replacement and rebates for measures whose replacement is costly.

Industrial Steam Trap Recommendations

The on-site evaluation of industrial steam traps focused on gathering site specific information on the key parameters that determine the steam and energy savings from industrial steam traps. The key parameters collected on site include the steam pressure, the hours of operation, and the boiler efficiency. The team also requested data on the type of trap failure from the steam system manager. Many of the sites were able to provide the team with a steam trap survey that clearly identified the failure type of the rebated traps. This information helped in reducing the uncertainty associated with the steam trap savings calculation.[88] The team also collected orifice size information from the manufacturers for all the steam traps reviewed during the onsite analysis.

Savings Variability

The data collection led to the determination that, for the sites surveyed, the mean energy savings from steam traps were substantially higher than the ex ante values. The ex ante gross savings for low pressure traps were 638 therms while the ex post gross savings were 1,398 therms and it resulted in an average realization rate of 219%. The ex ante gross savings for high pressure traps were 2,342 therms while the ex post gross savings were 5,033 therms and it resulted in an average realization rate of 215%.

While the average realization rate for high and low pressure traps was substantially above 100%, the precision of these estimates is poor. The variability in the savings per trap per site is very high. For low pressure traps in the third savings strata, the average realization rate is

[88] The reliability and variability analysis clearly indicates that the failure type is the most uncertain variable in the calculation of steam trap savings. Given the uncertainty surrounding this parameter which can vary from zero to one, the utility work papers assume a value of 0.50 for the failure type.

219%, but the 90% lower bound is -188% while the upper bound is 800%.[89] The extremely high variability results in site level realization rates in this strata ranging from a low of 0 to a high of 1,400%.

Of the 41 on-sites completed during the evaluation, 6 sites had zero savings while 5 sites had realization rates of over 400%. Table 4-26 lists the utility, number of traps, and rationale that was used to determine the site level savings from the rebated traps with were zero savings. For all the sites in the Table 4-26, steam traps were rebated even though the traps did not have an impact on the natural gas consumption of the facility; hence savings for these sites were zero. Sites with zero savings included sites that were non-operational due to bankruptcy, sites with boilers that were not operating due to their failure to pass California Air Resource Board standards, sites where all rebated traps had failed closed, and sites whose boiler fuel was not provided by the utility. Even sites with non-zero realization rates had traps with zero savings. Individual traps could receive zero savings if they were new, were not retrofit traps, were failed closed, or were stored instead of installed. Sites with extremely high realization rates tended to operate at higher pressure than the work paper assumptions, have more traps failed open, and possessed larger orifice sizes.

Table 4-26: Industrial Sites with Zero Steam Trap Savings

Utility and Number of Rebated Traps	Rational for Zero Savings
PG&E, 12 Traps	All the rebated traps failed closed.
PG&E, 5 Traps	The boiler that serves the steam traps has been shut down by EPA because of emissions and the site is not planning to replace it with a new boiler.
PG&E, 145 Traps	The site in non-operational due to bankruptcy.
PG&E, 45 Traps	The site did not have a natural gas boiler. Natural gas is used to make their product (H2SO4) and waste heat from stacks is recovered to make steam.
PG&E, 2 Traps	Cogeneration system waste heat recovery is the source of steam. Site does not use natural gas as fuel for cogeneration system, they use 100% digester gas..
SCG, 69 Traps	All the rebated traps failed closed.

Recommendation due to Variability

The extremely high variability in per trap savings strongly supports the conclusion that industrial steam traps should not be rebated as a prescriptive measure. Prescriptive measures should be limited to measures that are relatively homogeneous in their application and their per unit savings, industrial steam trap are extremely heterogeneous in their application and their savings. The results from this evaluation lead to the recommendation that industrial steam traps be rebated as a custom measure.

[89] The highest realization rate in the onsite survey for low pressure traps was 1,400%. The realization rate at this site was high due to large pipe diameters and orifice size.

As part of the custom rebate program, steam trap surveys would be required to be completed prior to the receipt of a rebate for steam trap retrofits. The steam trap survey would provide the utility and the site with information on the key parameters needed to calculate the savings from the trap retrofit. The survey would collect information on hours of operation, steam pressure, boiler efficiency, make and model number of the steam traps and most importantly the failure type.

Steam traps can fail if they are blowing through, leaking, or if they are blocked closed. Energy savings from retrofitting steam traps comes from replacing traps that are blowing through or leaking. Retrofitting traps that are blocked will help to maintain the integrity of the steam system, but these do not generally save energy. During the on-site survey, it was determined that 751 of the 5,207 on site rebated steam traps were failed closed. While these traps were technically eligible for a rebate, the utility work papers note that the savings are only achieved for traps that are blowing through or leaking. If a steam trap survey was required prior to the receipt of a rebate, the utility could stop rebating traps that do not save energy.

In addition to making steam traps a custom measure requiring a steam trap survey, it is further recommended that the utility closely monitor the installation of the traps. During the onsite data collection effort, it was found that 421 rebated traps were either not installed or not operational. While some of the non-operational traps were out of use due to reductions in operations associated with the recession, other traps were non-operational because they were in storage.

Post-installation inspections could also help to ensure that the utilities were not rebating traps installed in new, not retrofit, applications. The work paper clearly describes the steam trap rebate measure as a retrofit measure. The telephone survey of industrial customers found that 13 sites had installed only new, non-retrofit trap and that 13 sites had installed a combination of new and retrofit traps. A pre-retrofit steam trap survey followed by post site monitoring could reduce the possibility of rebating steam traps used in new applications.

Industrial steam traps, when installed in appropriate locations, save substantial quantities of energy. The per trap savings, however, is highly dependent on the application of the trap. Traps installed in refineries have extremely high per trap therm savings, while traps installed in smaller industrial or light manufacturing are more at risk in time of economic uncertainty. In addition, the utilities need to carefully examine the site level net of free ridership ratio. Sites with pre-existing trap maintenance programs are likely to have lower net-to-gross ratios, leading to more uncertainty in the net realized therms at the site. It is recommended that the utilities continue with the industrial steam trap program, but transform steam traps from a prescriptive measure to a custom measure. The custom approach should include a pre-retrofit survey, an assessment of the site’s likely free ridership, and post installation inspection.

5

Pump Testing

This chapter of the report evaluates the Pump Test component of the Commercial and Industrial (C&I) Pumping high impact measure (HIM). The C&I Pumping measure consists of both a free pump testing service and financial incentives for energy efficiency improvements to commercial, agricultural and industrial pumping systems. Due to significant differences across these measures in how savings are claimed and estimated, pump tests are disaggregated from the financial incentives for this HIM. This report pertains only to the Pump Testing Measure Impact Evaluation. The focus of this evaluation element was on measuring program influence on pumping system upgrades that occur in the absence of incentives. The evaluation approach applied is most similar to that used for other education program elements, such as audits.

The market potential for this service is quite large with over 40,000 pumping accounts. The Pump test procedure analyzes the efficiency of operations of the bowl, impeller and shaft of the pump, and identifies any related efficiency improvements needed. These measures generally apply to farming irrigation and water distribution companies.

Pump testing is part of Southern California Edison's (SCE) Southern California Industrial and Agricultural Program. As of Q4 2008, pump testing gross savings claims total 32.5 million gross kWh, and 10,810 gross kW. Pump tests are currently not associated with claimed savings for any other PY2006-2008 California Energy Efficiency program or Contract Group. Table 5-1 below summarizes total claimed savings for the pump test measure for the Southern California Agricultural Program.

Table 5-1: Summary of Pump Test Measure Claims by IOU, Program and Contract Group

Measure	Pump Test
IOUs Claiming Measure Savings	SCE
Program(s)	SCE2510
Contract Groups	Southern California Industrial and Agricultural Program
Gross kWh	32,499,989
Gross kW	10,810
Net-to-Gross Ratio	0.75

5.1 Evaluation Objectives for Pump Testing

The main objectives of this evaluation are to calculate both the gross and net impact of a pump test. As described in SCE Work Paper WPSCNRWP0001 (Full Service Pump Efficiency Improvement), gross impacts are currently calculated using the following four assumptions;

- Assumption 1: **Only 50 percent of the pumps tested are actually in need of repair**. This percent was estimated based on yearly tracking database assessments from 2002 to 2006. During this time the percent of pumps needing repairs has gradually decreased from 68.2 to 56.5 percent. Fifty percent was taken as a conservative value.

- Assumption 2: **Only 29 percent of pumps needing repairs are repaired**. This value was derived from three reports; 1994 San Diego Gas and Electric (SDG&E) Agricultural Energy Management Services Program Report, 2002 SCE Pump Test and Hydraulic Services, and the 2006 SCE Pump Test and Hydraulic Services Program Report. These three reports found between a 28.7 and 29.4 percent realization rate for this measure. These numbers were then rounded to 29 percent.

- Assumption 3: **50 percent of the pumps repaired will move to a different energy efficiency measure program** and therefore energy savings will be reported under those programs. 50 percent was selected as a conservative number.

- Assumption 4: **The average pump in need of repair has a potential annual energy savings of 34,092 kWh/year and a potential demand reduction of 11.34 kW**. These numbers were derived using the calculations described below;
  - Annual Energy Savings
    As part of the pump test, the existing Overall Plant Efficiency (OPE_1) is calculated as:

    Existing OPE % = (Water Horsepower / Input Horsepower) x 100 (Eq. 1)

    where water horsepower is calculated as:

    Water Horsepower = GPM x total head / 3960 (Eq. 2)

    GPM = Water flow rate in gallons per minute

    Total head units are feet.

    3960 constant = (33,000 Foot Pounds/Hp) / (8.33 lbs of water/gallon)

    Input Horsepower = kW input to motor x (1.341 horsepower per kW)

    If the existing OPE value is below expected performance level, an improved OPE value will be determined based on either historical tests or CPUC Table of Wire to Water Efficiencies.

    The improved annual pump energy consumption value is calculated as:

    Improved Annual kWh = (Existing OPE / improved OPE) x Existing Annual Pump kWh (Eq. 3)

    The annual kWh savings is calculated as:

    kWh Savings = Existing Annual Pump kWh - Improved Annual kWh (Eq. 4)

    The calculated kWh Savings for each pump test is stored in the Pump Test Tracking system. The kWh Savings for those pumps in need of repair between the years 2002 and 2006 was averaged to arrive at the assumed savings value of 34,092 kWh/year.

– Demand Reduction
The existing demand kW input value is measured during a pump test. The improved kW input value is calculated as:
Improved kW Input = (Existing OPE/ Improved OPE) x Existing kW Input (Eq. 5)
The demand reduction is calculated as:
kW Input Demand Reduction = Existing Input kW – Improved Input kW (Eq. 6)
The calculated kW Input Demand Reduction for each pump test is stored in the Pump Test Tracking system. The kW Input Demand Reduction for those pumps in need of repair between the years 2002 and 2006 was averaged to arrive at the assumed demand reduction value of 11.34 kW.

The first three assumptions are multiplied to calculate the percent of pump tests that result in a non-incented pump repair. This is then multiplied to the kWh and kW savings values from the last assumption and results in a per-pump test kWh and kW savings numbers (2,472 kWh and 0.82 kW).

This evaluation uses the pump test tracking system combined with self-report phone survey data to assess these four assumptions as well as to calculate a net-to-gross ratio.

5.2 Methods Used for Pump Testing

A participant self-report survey was implemented to collect data used in this evaluation. This survey was completed using a computer aided telephone instrument (CATI) and was designed to evaluate the gross and net impacts of a pump test through SCE's Ag Program. The survey gathered information such as the likelihood the customer would have undertaken a test without the program, the preponderance of pump repairs and retrofits following a selected sample of tests, the isolation of those actions that were completed in the absence of any program incentive, the likelihood the repair would have occurred without the test, and an assessment of alternative timelines. The full survey instrument is presented in Appendix D-2.

5.2.1 Gross Impacts

The ex-post gross impact analysis approach focuses on the major parameters used to calculate ex-ante gross impacts. According to SCE Work Paper WPSCNRWP0001 (Full Service Pump Efficiency Improvement) SCE uses the following assumptions to calculate savings from pump tests:

- Assumption 1: Only 50 percent of pumps tested need repairs
- Assumption 2: Only 29 percent of pumps needing repairs are repaired

- Assumption 3: Approximately half of the pumps repaired move to other programs, where incentives are paid, and the savings are not reported for those measures.

- Assumption 4: An average non-incented pump repair results in a gross impact of 34,092 kWh, and 11.3 kW

The ex post gross impact analysis approach leveraged the participant tracking database and the participant surveys to true up the values of these four assumptions.

- *Assumption 1: Information from the tracking database* was assessed to determine the percent of pump tests that resulted in a recommendation to repair. As part of the self-report survey, participants were also asked to recall whether the pump test resulted in a recommendation for repair. However only 48 percent of the respondents gave the same answer that was found in the tracking database. The tracking data was determined to be a more reliable source of information then the self report data for this assumption. Additionally, if the self-report data was used, the tracking database impacts used in Assumption 4 would not be populated for those that incorrectly recalled the test results.

- *Assumption 2*: The telephone survey measured the percent of tested pumps in need of repair that went on to repair the pump. The following question was asked for each pump that was found in the Pump Repair tracking database: "Our records show that in [insert participant program year] your company participated in the SCE Pump Repair Rebate Program at this location, is this correct?" and the following question was asked for each pump: "Since the pump test on [insert pump test date], has your company made any [other] changes or repairs to this pump?"

- *Assumption 3*: Similarly, the telephone survey measured the rate at which pumps are tested, subsequently have repairs, and perform these repairs outside of incentive programs. The pump testing sample was merged against the SCE2510 records involving incentive payments to flag the repairs associated with incentives. Again, impacts for those claims are already credited to the program, and are eliminated as part of the pump testing credit to avoid double-counting. For the pumps that did not appear in the SCE2510 records involving incentives, the following question was asked: "Did your company receive a rebate for repairing this pump?"

- *Assumption 4:* The savings values presented in the pump test tracking database, based on individual pump test reports, were applied as the gross impact associated with retrofit and repair activities identified through the customer survey and determined not to be associated with a rebate.

5.2.2 Net-to-Gross

The net impact approach was to collect and analyze free-ridership self-reports from participants using the telephone survey in accordance with the methodology developed by the Large Nonresidential Net-to-Gross Ratio (NTGR) working group. The working group crafted consistent batteries of survey questions, which were assembled based on lessons learned in past evaluations and the experience of the ED and their contractors. The group developed batteries of questions and associated scoring algorithms for calculating net of free-ridership ratios. The NTG methodology, the non-residential questionnaires and the scoring algorithms underwent public review and can be found in Appendix A of the Southern California Industrial and Agricultural Evaluation Plan, submitted February 28, 2008. That said, the pump test survey was customized for this particular application, including investigations of both pump testing and repair activity without the SCE free pump test services.

A self-report NTG ratio (NTGR) is computed using the following approach. 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 test the pump at this time.

2) A **Program Influence** score that captures the perceived importance of the overall program in the decision to repair the pump versus all other influences. This score is adjusted (i.e., divided by 2) if respondents say they had already considered testing the pump 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. This score accounts for deferred free ridership by incorporating the likelihood that the customer would have completed pump testing and repair actions 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 both the **Timing and Selection** and **No-Program** scores, the maximum score is always used to capture the most important element in the participant's decision making. Thus, each score is always based on the strongest influence indicated by the respondent.

When there are missing data or 'don't knows', to critical elements of each score, one of two options is used. The missing element may be backfilled with a value that represents the average of the lowest and highest extreme values. Alternatively, if it is one of several other elements that are considered in the algorithm, the missing element may simply be excluded from consideration.

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.

Timing and Selection Score

The questions asked were:

I'm going to ask you to rate the importance of the program as well as other factors that might influence your decision to test your pump. 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 test your pump on [insert pump test date].

1. That the test was free
2. Information about the Pump Testing Program or SCE marketing materials
3. The endorsement or recommendation by your SCE Account Representative
4. The age or condition of your pumps
5. Previous experience with SCE's Pump Testing Program
6. Previous experience with pump tests outside of SCE's program
7. A recommendation from a design or consulting engineer
8. The standard practice of pump testing in your business or industry
9. Following a regular pump testing schedule
10. Corporate policy of pump testing

The Timing and Selection Score is calculated as the highest of the responses to the first three questions above.

Program Influence Score

The questions asked are:

1. Did your company first consider having your pump(s) tested before or after learning of SCE's pump testing program?
2. Please rate the overall importance of the SCE Pump Test Program versus the most important other factor in your decision to TEST your pumps. I'd like you to give me a 0 to 10 score for the SCE Pump Test Program's influence and a 0 to 10 score for the influence of the most important other factor so that the two scores total 10.

The Program Influence score is calculated as the response, on the 0 to 10 scale, to question 2, reduced by half if the respondent considered having their pumps tested before they learned about the program.

No-Program Score

The questions asked are:

1. If SCE's pump testing program did not exist, how likely is it that you would have had the pump tested?

2. If the SCE pump test program did not exist, how long would you have waited before having the pump tested?
   a. within 6 months? (NTG=0)
   b. 6 to 47 months later (NTG=(months-6)*(1/42))
   c. 48 or more months later (NTG=1)
   d. Never (NTG=1)
3. How likely is it that your company would have completed the repairs to the pump if it had not been tested? Please use the same 0 to 10 likelihood scale, where 0 is not at all likely and 10 is extremely likely.

The No-Program Score is calculated as 10 minus (the average of the likelihood of testing the pump at the same time and the likelihood of completing the repairs without the test multiplied by one minus the deferred net-to-gross value associated with the timing of the pump test).

NTG Ratio

The self reported NTGR (NTGR) is simply the average of the Program Influence, Timing and Selection, and No-Program Scores, divided by 10.

Estimation and Application of Weights

Weights were used to calculate the NTG ratio. The weight was equal to the ex-ante impact claim associated with each pump. In other words, if there were two pump test records in the tracking database associated with one pump, the weight for that pump would be equal to the sum of the ex-ante impact of those two tests.

5.3 Confidence and Precision of Key Findings

The sample used for the telephone survey was grouped into three strata, where each strata represents one-third of program ex-ante impact claims (also one-third of pump tests performed, since each pump test has the same ex-ante impact assigned). Table 5-2 summarizes the number of pumps, addresses and customers found in each of the resulting three strata. As shown in this exhibit, pumps tested through the program are concentrated into a relatively small number of large customers. The sample was designed to capture 75 percent of the largest customers (strata 1), 60 percent of the medium customers (strata 2) and 250 of the smallest customers (strata 3).

Table 5-2: Telephone Survey Sample Design

Measure	Customers	Addresses	Pumps	Tests	Average Pumps Per Address	Average Tests Per Pump	Planned Completes	Maximum Number of Addresses Surveyed Per Customer
Strata 3	1354	3,370	4,122	4,380	1.2	1.1	**250**	1
Strata 2	156	2,378	3,742	4,382	1.6	1.2	**94**	2
Strata 1	35	1,782	3,471	4,387	1.9	1.3	**26**	5
Total	1,545	7,530	11,335	13,149	1.5	1.2	**370**	1.5

Data collection was completed for all of the pumps at up to 5 addresses for each of the survey completes in Strata 1; 2 addresses for each customer in Strata 2 and 1 address for each customer in Strata 3. The final distribution of survey competes is provided in Table 5-3. Interviews were completed with a total of 376 customers, surrounding 842 pumps at 552 addresses.

Table 5-3: Telephone Survey Sample Disposition

Measure	Planned Completes	Resulting Completes with Customers	Addresses	Pumps	Tests	Average Pumps Per Address	Average Tests Per Pump	Average Number of Addresses Surveyed Per Customer
Strata 3	250	258	258	301	350	1.2	1.2	1
Strata 2	94	92	184	338	469	1.8	1.4	2
Strata 1	26	26	110	203	293	1.8	1.4	4.2
Total	370	376	552	842	1,112	1.5	1.3	1.5

The sample was designed to get a margin of error of less 5 percent, using a coefficient of variance of 0.60. However, the achieved coefficient of variance was 2.3 yielding a margin of error of 11 percent and a confidence interval of 1.4 percent.

5.4 Validity and Reliability

The thrust of this evaluation was to verify work paper assumptions used to derive gross impacts and to assess the net of free-rider influence of the program on the gross savings achieved. Formal uncertainty analysis was not completed to quantify the validity and reliability of the evaluation results presented in this report, but a qualitative assessment follows.

The principal data sources contributing to the gross impact evaluation were a selected mixture of utility tracking system-based findings and self-reports obtained from participating customers. This approach sought to verify work paper methods in some cases, but also relied upon tracking system findings with limited independent effort applied to verify the accuracy of those inputs. The tracking system inputs that were relied upon include the frequency with which pumps that are tested require repairs (an attempt was also made to verify this using telephone self-reports) and the resulting potential energy savings for tests that fail a prescribed pump testing program efficiency threshold.

This Evaluation was guided by the *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals90*. This program was assigned a Basic level of rigor for the primary evaluation objectives and follows the Indirect Impact Protocol. The Indirect Impact Evaluation Protocol is intended for those programs where the primary uncertainty lies in the program's ability to obtain behavior change(s) targeted by the program. Indirect impact evaluations are linked wherever possible to previously measured energy or demand savings estimates that yield savings estimates with the same rigor required by the Basic rigor level for impact evaluations.

Additionally, the survey is designed to produce valid and reliable NTGR results, based on the following:

- *"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.

- *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 were conducted using the CATI software. Any problem areas were detected and corrected at this time.

Future evaluation efforts could seek to further validate utility program pump test results using independent pump testing of wells under pre-retrofit and post-retrofit conditions, as well as an independent assessment of the measure savings attributable to repairs completed. Future studies might also seek to verify self-reported work paper parameters using a verification sample of on-sites nested within a given telephone survey sample. This would provide a method for validating the accuracy of self-reports. To the extent it is possible to obtain independent evaluation-based verification of utility data and test results in parallel with program implementation, such an effort should also be considered as it provides a real-time feedback loop on program tracking and measurement accuracy. In the presence of these additional data sources it should be feasible in future evaluations to accurately quantify the validity and reliability of the evaluation-based results.

[90] TecMarket Works, April, 2006

Nonetheless, upper and lower bounds are estimated where possible throughout section 9.5 below. In general, a standard deviation was calculated for each result and divided by the square root of the n to calculate the standard error. The standard error was then multiplied by 1.645 to calculate the confidence interval. Where results are combined, such as gross impacts multiplied by the NTG ratio, the propagation of error method was used. This method takes into account the error of both measurements to get a combined error. The propagation of error method is calculated as:

$$\mu_{n=1-x} \times \sqrt{\sum_{n=1}^{x} \left(\frac{\varepsilon_x}{\mu_x}\right)^2}$$

ε = standard error

μ = mean

5.5 Findings

The detailed findings for this evaluation are grouped by gross impacts, net impacts and other findings and are presented below.

5.5.1 Gross Impacts

Table 5-4 below presents the gross impact assumptions according to SCE Work Paper WPSCNRWP0001 (Full Service Pump Efficiency Improvement) compared to the evaluation results.

Table 5-4: Gross Impact Assumptions

Assumption		Source	SCE Work Paper Assumptions	Evaluation Results			
				Total	Strata 3	Strata 2	Strata 1
Assumption 1:	Percent of pumps tested needing repairs	Tracking Data	50.0%	60.0%	69.4%	63.5%	43.0%
Assumption 2:	Percent of pumps needing repairs that are repaired	Self-Report Survey	29.0%	20.1%	25.1%	15.6%	20.7%
Assumption 3:	Percent of non-incented pump repairs	Self-Report Survey	50.0%	66.4%	71.2%	69.8%	47.8%
Product of Assumptions 1-3:	**Percent of tested pumps that result in non-incented repairs**	**Calculated**	**7.25%**	**8.00%**	**12.41%**	**6.91%**	**4.26%**
Assumption 4:	Average Potential Gross **kW Per Repair**	Tracking Data	11.34	10.52	7.60	16.40	5.62
	Average Potential Gross **kWh Per Repair**	Tracking Data	34,092	23,392	16,223	33,456	23,320
Product of Assumptions 1-4:	Average Potential Gross **kW Per Test**	**Calculated**	**0.82**	**0.84**	**0.94**	**1.13**	**0.24**
	Average Potential Gross **kWh Per Test**	**Calculated**	**2,472**	**1,872**	**2,013**	**2,311**	**994**

A more detailed description of each evaluation result is presented below:

Assumption 1: Only 60.0 percent of pumps tested need repairs. This percent was derived using the tracking data for those participants that completed the telephone survey. The 842 pumps that were discussed through the survey had a total of 1,112 pump tests over the 3 year evaluation period. 667 out of the 1,112 (or 60.0 percent of) pump tests resulted in a recommendation for repair.

Assumption 2: Customers reported that 20.1 percent of pumps needing repairs are repaired. Only 622 of the 667 recommendations (from assumption 1) were given on the pump test that was discussed in the survey. The other 45 were removed from the rest of this analysis. Of the 622 tests with recommendations, there were 125 pump repairs reported by respondents who answered yes to Q10 or Q14 on the telephone survey.

Assumption 3: Customers reported that 66.4 percent of the pumps repaired did not go through a program where incentives are paid. Of the 125 pump repairs (from assumption 2), 83 of these repairs did not receive an incentive from a rebate program following repairs made to those pumps. This was established based on respondents answering no to Q19 and only includes those respondents that also were not asked Q10.

Assumption 4: According to SCE tracking data, the 83 non-incented pump repairs had an average potential gross impact of 23,392 kWh (with a variation around the mean of +/- 6,971), and 10.5 kW (with a variation around the mean of +/- 2.43). This corresponds with the average savings potential presented in the pump test report for the 83 non-incented pump repairs. Note that the kWh and kW potential gross impact values were not verified as part of this evaluation, but are meant to illustrate the maximum potential impact for a pump repair.

The strata level results are also presented in Table 5-4. Note that the percent of tested pumps that result in a non-incented repair varies notably between the strata. This difference is a result of Assumptions 1 and 3 varying between the three groups. Strata 1 participants had a much lower percent of pumps in need of repair (assumption 1) and are much more likely to participate in a rebate program if they do repair (assumption 3). Strata 3 and 2 are more similar, however, the lower percents in assumption 1 and 2 cause strata 2 to have a much lower "percent of tested pumps that result in a non-incented repair." These results demonstrate that participants with fewer pumps are more likely to need their pump repaired, more likely to repair it, and more likely to not go through a rebate program. And therefore, customers accounting for the largest ex-ante impact in the program would have the lowest ex-post impact if it were to be applied by strata.

5.5.2 Net-to-Gross

The net to gross (NTG) ratio was calculated using self reported survey results, as described in the methodology section, for the 83 non-incented pump repairs. The resulting overall NTG ratio is 0.631 with a 90 percent confidence interval of +/- 0.026. As shown in Table 5-5, the NTG ratio does not vary significantly across stratum. However, this NTG ratio does vary significantly from the ex-ante NTG value of 0.75.

Table 5-5: NTG Ratios by Strata

Strata	NTG	90% CI
Strata 3	0.631	0.037
Strata 2	0.641	0.051
Strata 1	0.619	0.041
Total	0.631	0.026

5.5.3 Other Findings

The final tracking database contained 23 percent fewer records then the Q3 2008 tracking database (16,135 vs. 13,149). According to SCE, records were removed from the final pump test database if they showed up more frequently than every two years.

Additionally, the final tracking database should have been populated with savings estimates for pump tests; however no savings values for gross kWh and gross kW were included. These variables were populated in the Q3 2008 tracking database and a request was made to SCE to provide these variables. The revised final tracking database received from SCE was compared to the Q3 2008 database and it was determined that the new savings values were merged only by pump and not by pump test date, and were therefore not accurate. Another request was made to SCE and a complete data set was received.

5.6 Results

The evaluation finds that the percent of pump tests that result in a non-incented pump repair is 8.00 (+/- 1.39). In other words, 8 out of every 100 pump tests result in a repair and some associated level of impact.

The gross per-pump test kWh impact is calculated as 8 percent * 23,392 kWh and kW impact is calculated as 8 percent * 10.5 kW. The resulting gross impact per test is 0.84 kW (+/- 0.24) and 1,872 kWh (+/- 645).

The net impact is calculated as the gross savings * the NTG ratio (0.63). The resulting net impact per test is 0.53 kW (+/- 0.16) and 1,182 kWh (+/- 410).

When these numbers are applied to the population of 13,149 pump tests that were completed in PY 2006-2008, the gross impacts for the program are 11,067 kW and 24,619,025 kWh.

The strata level results are presented in Table 5-6 below, as well as SCE's Work Paper assumptions and the overall result.

Table 5-6: Evaluation Findings vs. Work Paper Assumptions

Savings Measurement	SCE Work Papers	PY '06-'08 Evaluation Findings				Realization Rate
		Total	Strata 3	Strata 2	Strata 1	
Gross kW/unit repaired	11.34	10.52	7.6	16.4	5.62	92.80%
Gross kW/test	0.82	0.84	0.94	1.13	0.24	102.40%
Gross kWh/unit repaired	34,092	23,392	16,223	33,456	23,320	68.60%
Gross kWh/test	2,472	1,872	2,013	2,311	994	75.70%
Percent of tested pumps that result in non-incented repairs	7.25%	8.00%	12.41%	6.91%	4.26%	110.30%
NTG	0.75	0.63	0.63	0.64	0.62	84.00%
Participant Population PY 06-08	13,149	13,149	4,380	4,382	4,387	
Total Program Gross kW	10,810	11,067				102.40%
Total Program Gross kWh	32,499,989	24,619,025				75.80%

5.7 Recommendations

1.) The kW and kWh savings values should be verified.

The kW and kWh savings values reported for the pump repairs in this evaluation have not been verified and should be treated as such. This task was outside the scope of this evaluation. The value that was used in this evaluation was the potential savings value reported at the time of the pump test and taken from the tracking database and it is important to consider its reliability.

One method to evaluate the kW and kWh savings values is to conduct an on-site verification of pump operating efficiency. This verification can be done by either SCE or a third-party and would be done on a subset of the program participants. This work would seek to answer the following questions:

- Are pre-retrofit OPEs reproducible by another party?
- For those pumps that are repaired outside of a program, what are the post-retrofit OPEs?

By verifying pre-installation OPE (where no customer actions are taken) and post-installation OPE (where efficiency actions have been taken) the savings values would be more reliable.

Another method of evaluation would be to conduct a statistically adjusted engineering billing regression within a telephone survey sample. This work would quantify the savings that are evident from energy bills by comparing pre- and post-repair energy use.

2.) Pump test frequency should be optimized.

As stated earlier, the final tracking database contained 23 percent fewer records as compared to the Q3 2008 tracking database (16,135 vs. 13,149). According to SCE, records were removed from the final pump test database if they showed up more frequently than every two years. Apparently this is a program rule. As a result, if this rule persists, going forward pump tests should not be performed sooner than once every two years. Removing those records was appropriate, as the pump test impacts reflect a program rule enforcement of frequency between tests not being less than two years.

As shown in Table 5-7 below, a little over half of the participants surveyed reported that they test their pumps every 1 to 2 years. And about half of those participants stated that if the program did not exist, they would still test their pumps on the same schedule. This implies that participants have a need to test pumps in some instances at a frequency greater than two years.

Consideration should be given to optimizing any program rules surrounding the frequency of testing. Certainly if a participant suspects they have a pump problem and their previous test was just one year earlier, the pump test is warranted to help identify the extent of the problem. At the same time, the utilities should not be encouraged to allow customers to over-test pumps. Unfortunately the evaluation team does not currently have a suggested solution for optimizing the frequency of pump testing allowed, but recognizes that the program might benefit from optimization.

Table 5-7: Self-Reported Pump Testing Schedule

	Q49. How often would you estimate that your company would get your pump(s) tested if the pump test program did not exist?								
Q45. How often does your company have each pump tested?	**Same Time**	**Wait 0.5 Years to 1 Year**	**Wait 1.5 to 2 Years**	**Wait 2.5 to 3 Years**	**Wait More Than 3 Years**	**Less Often**	**Same**	**Don't Know**	**Total**
Not on a schedule						27%	15%	1%	43%
At least once a year	17%	7%	2%		2%			2%	30%
Every 1.5 to 2 years	10%	2%	3%	4%	1%			3%	21%
Every 2.5 to 3 years	2%	0%	1%	0%					3%
More than every 3 years	2%	0%	0%						2%
Don't know	0%							1%	1%
Grand Total	**31%**	**9%**	**6%**	**4%**	**3%**	**27%**	**15%**	**7%**	**376**

3.) Examine the pump testing and repair practices in areas where free water pump efficiency testing is not offered.

The SCE pump test program has been in place since 1911, and some of the participating customers have been working with SCE for decades. This program legacy confounds the ability of participants to report testing and repair practices in the absence of such a program. For this reason, it is recommended that future studies examine the pump testing and repair practices in areas where free water pump efficiency testing is not offered, and has not been recently offered. This would provide a broader perspective on the value of the testing service, and may be a better foundation for estimating baseline conditions and net-to-gross ratios for the program than self-reported data from SCE participants.

4.) Future evaluations should look for ways to get participants to follow through with repairs.

This evaluation found that only 20 percent of pumps in need of repair, go on to be repaired. Participants who reported that their pump was found to be inefficient when tested and did not follow through with a repair were asked a few follow up questions, one of which was the primary reasons that their company has not taken action to repair the pump. As can be seen in Table 5-8, the number one reason stated for lack of follow through with repairs (reported for almost half of pumps that were tested as inefficient and not repaired) is "Lack of Funding/Repairs Too Expensive." This was followed by the response "Benefit of Repair Outweighs the Cost," reported for 27 percent of the pumps in need of repair but were not repaired.

Table 5-8: Self-Reported Reasons for Not Repairing

Q50. What are the primary reasons that your company has not taken action to repair the pump?	Total	Strata 3	Strata 2	Strata 1
Lack Of Funding/Repairs Too Expensive	46%	43%	49%	43%
Benefit Of Repair Outweighs The Cost	27%	24%	21%	49%
Don't Have Time To Repair It	12%	3%	16%	19%
May Repair/Replace In Future/Waiting For Funding	8%	6%	8%	11%
Pump Is Not Being Used/Selling The Land	5%	7%	6%	0%
Used Only As Backup Pump/Used Lightly	5%	6%	6%	3%
It Still Works Fine	4%	10%	2%	0%
Other Priorities	4%	4%	5%	3%
Other Reason	4%	9%	1%	3%
It Is Efficient Enough	3%	4%	0%	8%
We Lease The Land	2%	7%	0%	0%
N	206	70	99	37

In previous program cycles it has been noted that IOU Non-Residential Audit Program staff sometimes conduct follow-up calls to encourage participants to follow through with the recommendations given on the audit. The purpose of these calls is to obtain program feedback and encourage customers to install recommended measures. These follow-up calls appear to have made a significant impact on the likelihood that the customer will implement recommendations from the Audit.

When a repair is recommended through an audit, SCE could dispatch a repair technician to provide participants with an estimate for the needed repair. It is unclear from this evaluation whether those customers who did not make a repair actually obtained an estimate for the repairs or that they just perceive that the repair would be too costly.

5.) Spillover

Forty-seven of the surveyed customers said that they also test pumps at facilities outside of SCE territory. Those customers were then asked if the non-SCE pump tests were also free, and as shown in Table 5-9, 30 percent of the customers who had pump tests outside SCE territory stated that they were free and for 23 percent of the customers none of those tests are free.

Table 5-9: Pump Tests Outside SCE's Territory

Q59. Are those pump tests free also?	Total	Strata 3	Strata 2	Strata 1
Yes All Are Free	30%	31%	30%	25%
None Are Free	23%	28%	20%	13%
Most Are Free	13%	10%	30%	.
Some Are Free	21%	28%	.	25%
Don't Know	13%	3%	20%	38%
N	47	29	10	8

As illustrated in Table 5-10, the majority of these customers (62 percent) said that their experience with the SCE pump testing program was very important in their decision to have their pumps outside of SCE territory tested. This indicates that SCE's pump testing program is helpful in educating customers on the benefits of pump testing.

Table 5-10: Influence of the Program on Pump Tests Outside SCE's Territory

Q61. How important was your experience with Edison pump testing program in your decision to have these pumps tested? (0 to 10 importance scale)	Total	Strata 3	Strata 2	Strata 1
0 Not At All Important	13%	14%	10%	13%
3	2%	3%	.	.
4	2%	3%	.	.
5	4%	3%	.	13%
6	4%	7%	.	.
7	9%	10%	.	13%
8	28%	14%	60%	38%
9	13%	17%	.	13%
10 Extremely Important	21%	24%	30%	.
Don't Know	4%	3%	.	13%
Average	7.0	6.9	7.8	6.4
N	47	29	10	8

6.) Real-time tracking of customer water pumping volume and electricity usage.

It may be worth assessing a real-time benchmarking program design in which the utility continuously tracks customer pump OPE using interval data of pumped water volume and pump electric usage. Rather than periodic spot testing of pump performance, develop a database of remotely collected interval data and regularly query the database by pump to track performance over time. The service would include regular performance benchmarking updates that examine changes in pump OPE and provide warnings when performance drops below an expected threshold.

Alternatively the program might simply encourage, through information dissemination and other assistance, customer collection and trending of OPE data. The advantage would be that these real time assessments may result in more timely repairs. It would also allow customers to measure the value of a repair by comparing the OPE and pumped volume before and after a repair.

6

SCE Industrial Measures and Agricultural Measures

6.1 Evaluation Objectives

This section describes and outlines the evaluation objectives for the industrial measures in the SCE Industrial Energy Efficiency Program (SCE2509, Industrial Program) and agricultural measures in the SCE Agricultural Energy Efficiency Program (SCE2510, Agricultural Program). The industrial measures consist of: pump-off controllers, refrigeration equipment, controls, variable speed drives, compressors, motors, custom process and lighting improvements. The agricultural measures consist of pumping, motors, irrigation and other agricultural measures, but also some lighting and control measures installed in agricultural facilities. The evaluation provides ex post estimates of gross savings for these measures. The evaluation also provides separate estimates of the net-to-gross ratio for these industrial and agricultural measures.

The savings claims through Q4 2008 for SCE's Industrial Energy Efficiency Program (SCE2509) can be found in Table 6-1 below. These claims total over 129 million kWh and nearly 15,000 kW.

Table 6-1: Gross Savings Impact Summary for SCE2509

		Unadjusted		RR-Adjusted+	
EEGA Program	N records	Gross kW	Gross kWh	Gross kW	Gross kWh
SCE2509	264	16,776	145,467,578	14,931	129,466,144
SCE2510 Ag	1,133	14,646	70,916,981	13,501	65,009,964

+ Savings for some (but not all) measures in program SCE2509 were adjusted by a realization rate of 0.89 starting in Q407. As of Q408, savings for all SCE2509 measures were adjusted by the 0.89 realization rate.

Savings for some (but not all) measures in program SCE2510 were adjusted by a realization rate of 0.89 in Q408.

The parameters examined in the EM&V effort for SCE2509 measures are the gross Realization Rate (RR) and the Net-to-Gross Ratio (NTGR). Estimates for these parameters are applicable to the Southern California Edison Industrial Program. The parameter examined for SCE2510 measures is the Net-to-Gross Ratio (NTGR) only. Estimates for this parameter are applicable to the Southern California Edison Agricultural Program.

The objectives of the Gross Impact analysis for SCE2509 measures are to a) develop ex post estimates of the energy and demand savings for each project in the sample, and b) apply those findings to the full participant population to obtain a complete estimate of program group-level impacts.

The objective of the Net-to-Gross (NTG) analysis for industrial SCE2509 and agricultural SCE2510 measures is to understand the role of non-program factors versus the utilities or third party's energy efficiency program and rebate in influencing the customer's decision to install these measures. The NTG analysis estimates free-ridership and net savings from the measures installed in the Southern California Industrial and Agricultural Programs.

6.2 Impact Methods

6.2.1 Overview and Description of Methods

The methods developed and implemented for this evaluation were designed to be compliant with CPUC evaluation protocols and guidance and to produce the most accurate and defensible results possible given project resources and timelines. The approaches followed in this evaluation were guided by the CPUC Evaluation Protocols, however, the CPUC allowed for slight deviations. These variances were consistent with the intent of the Protocols, were expected to produce reliable and robust results, and consisted of:

- A sampling approach that varies program-specific precision targets in order to maximize the ratio of evaluation spending-to-program spending.
- A flexible application of gross energy and demand rigor at the site and measure level, rather than a pre-determined level of rigor by program.
- A small deviation from the minimum net-to-gross sample size by program (setting sample sizes at half of the program population or 300, whichever was less).
- Use of site-specific M&V impact methods that were a hybrid of IPMVP[91] Option A and B to maximize precision in a cost-effective manner. This hybrid method fell between the Basic and Enhanced rigor levels.

6.2.2 Protocols and Rigor Levels

Protocols

This Evaluation was guided by the California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals[92]. The following protocols were used:

[91] IPMVP refers to the *International Performance Measurement and Verification Protocol*, which specifies alternative measurement and analysis methods that can be used to estimate gross energy and demand savings from a measure installed under a program being evaluated. See www.evo-world.org.

- Impact Evaluation Protocol
  - Gross Energy Impact Protocol
  - Gross Demand Impact Protocol
  - Participant Net Impact Protocol
- Measurement and Verification Protocol
- Sampling and Uncertainty Protocol
- Evaluation Reporting Protocol

Rigor Levels

As mentioned above, the rigor levels for this evaluation deviate slightly from the Protocols in that specific participant sites and their measure mix were used to drive the evaluation type and rigor assignments.

The evaluation type for all sample points used the Protocol Guided Direct Impact for this evaluation. Additionally, the assignment of gross energy and demand rigor was set to a hybrid of Basic and Enhanced for all programs in this contract group. That is, the rigor assignment was included in the site-level analysis planning stage of each sampled project, and those assignments were applied by measure. Based on past experience and the level of certainty required by this evaluation, Itron strongly believed that a blended rigor level between Basic and Enhanced was most appropriate, and that the flexible application of rigor by site/measure ensured that evaluation resources were applied to produce the most reliable results possible given the budgeted resources. Application of rigor at the site level also offered a more robust and flexible approach given the level of uncertainty in the final relative size of each program's accomplishments. The same reasons applied to the NTG rigor assignments which were made at the sampled project and not the program level.

6.2.3 Sampling Methodology and Description

The Southern California Industrial and Agricultural (SCIA) contract group used the ratio-estimation approach for sample design described in Chapter 13 of the *Evaluation Framework Study* and referenced in the *California Energy Efficiency Evaluation Protocols*.[93] This approach was used to develop program realization rates for the 2002, 2003 and 2004-2005 Statewide SPC program evaluations.

---

[92] TecMarket Works, April, 2006, available at http://www.calmac.org/events/EvaluatorsProtocols_Final_AdoptedviaRuling_06-19-2006.pdf

[93] Chapter 13 – Sampling, page 358, of the TecMarket Works, 2004. *2002 Evaluation Framework Study*, prepared by TecMarket Works for Southern California Edison Company, June. http://www.calmac.org/publications/ California_Evaluation_Framework_June_2004.pdf

A key input to the ratio-estimation sample planning methodology is the error ratio (*er*) that is expected to result, given the evaluation sample size selected (the *error ratio* is defined below). As with the a priori use of the expected coefficient of variation in other sampling methods, the variance in the parameter of interest is not known prior to completing the evaluation work. Instead, analysts must estimate the *er* from other related studies and work or summarize expected sampling results across a range of possible *er* (as is often done with confidence levels).

To more formally investigate the expected precision levels for the 2006-2008 SCE Industrial and Agricultural impact evaluation, the precision level achieved for a relevant past evaluation was first reviewed -- for the combined 2002-2003 SPC impact evaluation sample the precision estimation process was carried out, as described for ratio estimation-based samples in Chapter 13 of the *Evaluation Framework Study*. Specifically, the error ratio was calculated and the precision expected was estimated, with alternative sample sizes as described on pages 358 and 365, respectively, using the results from the 2002-2003 SPC ratio estimation process.[94] Specifically, we calculated an error ratio (*er*) of 0.35 using the following formula:

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

where

$\gamma = 0.8$

$e_i = y_i - \hat{B} \, x_i$

w_i is the case weight,

x is the tracking estimate of savings for each project, and

y is an estimate of the estimated savings from the ex post evaluation.

Based again on the 2002-2003 SPC sample, we used the case weights to calculate the stratified ratio estimator of B, denoted $\hat{B}$, as follows:

$$\hat{B} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_{i=1}^{n} w_i \, y_i}{\sum_{i=1}^{n} w_i \, x_i}$$

[94] See Chapter 7 of Quantum Consulting, 2005. *2003 Statewide Nonresidential Standard Performance Contract (SPC) Program Measurement and Evaluation Study*, prepared by Quantum Consulting, Inc. for Southern California Edison Company, SCE Study ID: SCE0206.01, December.

We then estimated relative precision of $\hat{B}$, at the 95 and 90 percent confidence levels, for alternative sample sizes using the equation below (which includes finite population correction):

$$rp = 1.96\sqrt{1-\frac{n}{N}}\frac{er}{\sqrt{n}} \quad 95\%\ CL$$

$$rp = 1.645\sqrt{1-\frac{n}{N}}\frac{er}{\sqrt{n}} \quad 90\%\ CL$$

The resulting precision levels for alternative samples are shown in Figure 6-1 and Figure 6-2 below for the calculated *er* of 0.35 as well as a range of error ratios that might occur in a large and small program population. We took under consideration that error ratios might be somewhat higher for the 2006-2008 impact evaluation than they were for the 2002-2003 SPC because the scope of the 2002-2003 M&V effort was much smaller than the expected M&V scope of the 2006-2008 evaluation. The more limited 2002-2003 scope may have resulted in a higher fraction of cases in which evaluation engineers defaulted the realization rate to 1.0 because they were not able to conduct a more rigorous analysis than was conducted as part of the program's savings estimation process.

Figure 6-1: Expected Relative Sampling Precision (at 95% Confidence Level) Versus Sample Size with Stratified Ratio Estimation for Varying Error Ratios and Large Population (N=5,000)

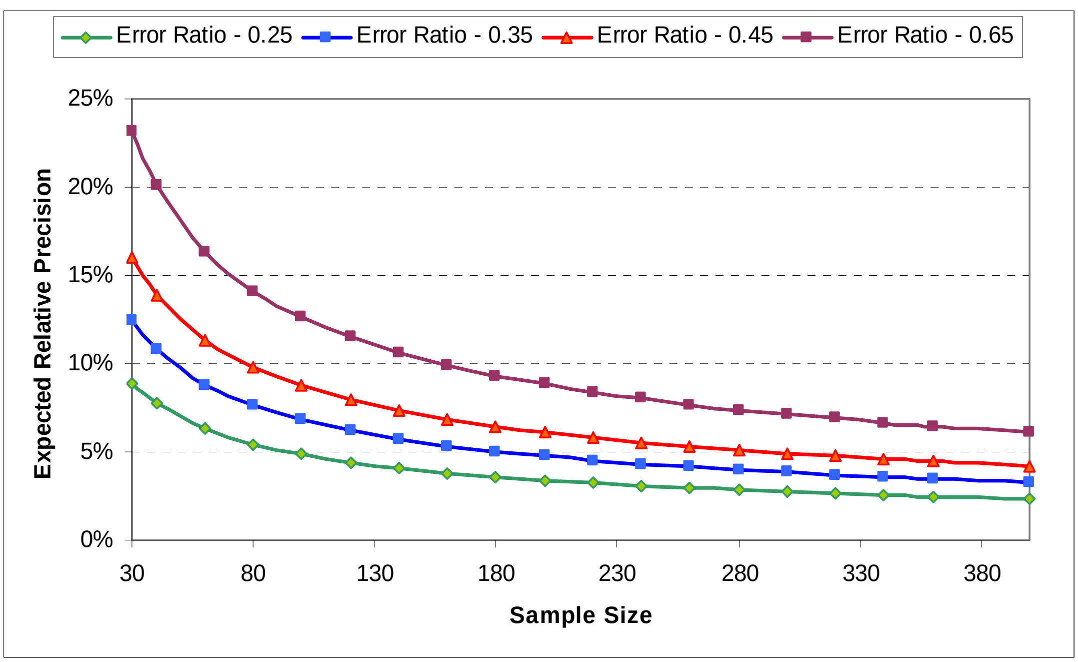
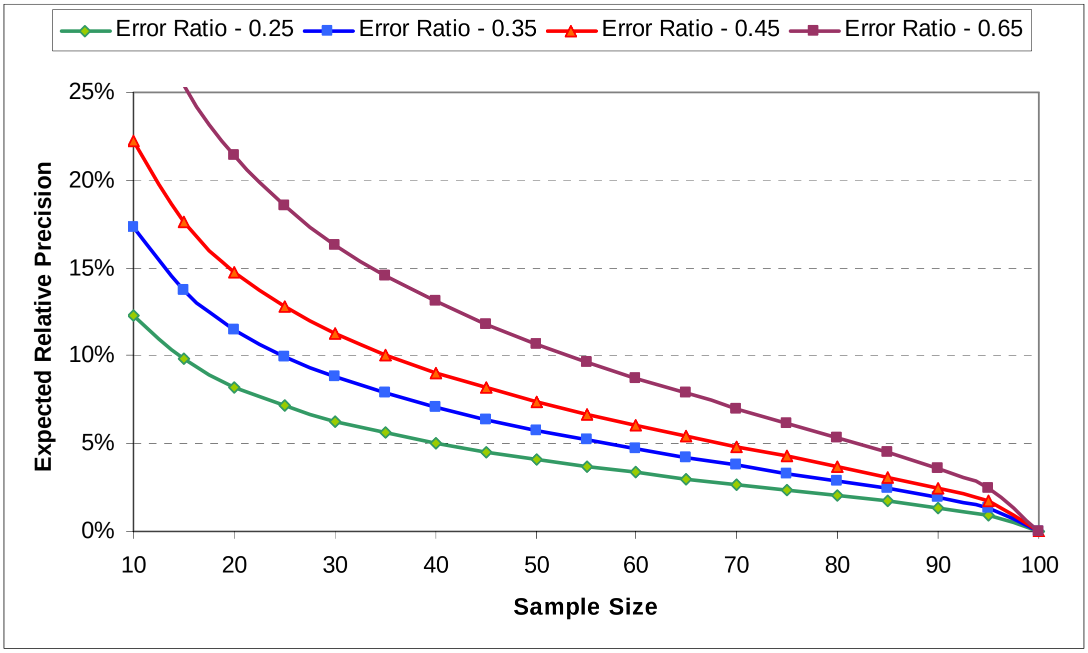


Figure 6-2: Expected Relative Sampling Precision (at 90% Confidence Level) Versus Sample Size with Stratified Ratio Estimation for Varying Error Ratios and Small Population (N=100)



The results in the figures are generally consistent with the example given in the *Evaluation Framework Study* (p. 366) and shows that precision levels as a function of sample size are highly non-linear.

Perhaps the most important aspect of any sample design for programs that address medium and large nonresidential customers is the use of stratification based on the amount of savings associated with each project. In implementing size stratification, typically projects are grouped into 3 to 5 strata from largest to smallest within which total savings are relatively equal for each stratum. It is not uncommon to find a 100-fold difference in average savings between the stratum with the largest and smallest projects (for example, the difference between strata 1 and 5 for the 2004-2005 SPC Evaluation was 75 fold). The improvement in sampling efficiency that can result from size stratification in the nonresidential sector can often be an order of magnitude decrease in sample sizes that would otherwise be required.

6.2.4 Gross Impact Sample

The sample size proposed in the Final Research Plan submitted for the SCIA contract group was 60 for each of the two SCE programs (SCE2509 and SCE2510.) The targeted confidence/ relative precision for both programs were 90/7.

A list of installed Industrial Energy Efficiency and Agricultural projects received from SCE in March 2008 was used to design an initial evaluation sample for the SCIA Contract Group. The sample design stratified the population of measures for each program (SCE2509

Industrial and SCE2510 Agricultural) into 5 strata and randomly selected measures for evaluation.

The target sampling was conducted at the finest level of segmentation supported by the tracking system (measure level) instead of site level. This made the sampling approach more transparent and simplified the measurement and evaluation planning process. Sampling at the site level would have required an additional sampling step, such as sampling measures within sites, and would have increased the difficulty of applying the results back to the population.

Based on the list of installed industrial projects obtained in March 2008, and on the goals submitted by SCE for the 2006-2008 evaluation cycle, a proportional sample of 30 points was selected for program SCE2509. Similarly, a proportional sample of 10 points was selected for program SCE2510. Note that the pump testing portion of program SCE2510 was not included in the project list obtained from SCE in March 2008.

Table 6-2 below summarizes the initial sample design for program SCE2509 and Table 6-3 summarizes the initial sample design for program SCE2510.

Table 6-2: SCIA Sample Design using List of Installed Measures received in March 2008 – Program SCE2509 Industrial

		Population				Initial Sample			
Strata	Gross kWh Strata Bounds	N records	%N	Gross kWh	% Gross kWh	N records	%N	Gross kWh	% Gross kWh
1	2,639,821	3	4%	10,042,290	20%	3	10%	10,042,290	27%
2	1,707,578	5	6%	10,130,537	20%	5	17%	10,130,537	28%
3	1,022,427	8	10%	9,575,854	19%	8	27%	9,575,854	26%
4	487,846	15	19%	10,584,351	21%	8	27%	5,699,650	16%
5	0	49	61%	10,128,220	20%	6	20%	1,140,516	3%
All		80	100%	50,461,252	100%	30	100%	36,588,847	100%

Table 6-3: SCIA Sample Design using List of Installed Measures received in March 2008 – Program SCE2510 Agricultural

		Population				Initial Sample			
Strata	Gross kWh Strata Bounds	N records	%N	Gross kWh	% Gross kWh	N records	%N	Gross kWh	% Gross kWh
1	490,046	2	2%	1,017,311	15%	2	20%	1,017,311	46%
2	320,165	4	4%	1,444,999	21%	2	20%	705,842	32%
3	144,150	8	9%	1,468,064	22%	2	20%	289,302	13%
4	73,690	14	16%	1,433,832	21%	2	20%	165,134	7%
5	0	62	69%	1,386,217	21%	2	20%	29,497	1%
All		90	100%	6,750,422	100%	10	100%	2,207,086	100%

This sample design allowed the SCIA contract group to get started with evaluation activities using the limited information regarding installed projects that was available in March 2008.

Following CPUC’s “Requirements for Evaluating High Impact Measures” directive of July 21, 2008, the SCIA contract group was instructed to finalize impact evaluation work for the samples that had already been drawn. Any remaining funds were re-directed to the evaluation of two High Impact Measures for the Small Commercial contract group – Steam Traps and Pipe and Tank Insulation. Consequently, in addition to one pre-M&V site that was installed in Q4, 2008 in program SCE2509, and that was added to the initial sample of 30, no additional impact sample was drawn for programs SCE2509 and SCE2510.

This report contains impact results for SCE2509 only. Efforts to complete impact evaluation for the 10 sites drawn for program SCE2510 were stopped when the Q4, 2008 extract was received from SCE. That extract, containing 1,133 records and a very diverse set of measures, led to the conclusion that any results based on a sample of 10 points wouldn't be robust enough to support a realization rate result.

6.2.5 Net-to-Gross Sample Design

The original research plan submitted to the ED for SCE's Industrial and Agricultural programs included net-to-gross evaluations by program based on "300 sample points or one-half of the program size, whichever is lowest."

As of Q4, 2008, program SCE2509 (Industrial) had installed 264 measures and program SCE2510 (Agricultural, not including pump testing) had installed 1,133 measures. Five of these 1,133 measures were strip curtains (a HIM) and were included in the Commercial Facilities evaluation. According to the original research plan the net-to-gross evaluations should be based on 132 sample points for SCE2509 and 300 points for SCE2510.

Assuming an error ratio of 0.5 for the net-to-gross analysis, a sample of 132 points for program SCE2509 would provide a NTG result with a precision of 90/5. A sample of 300 points for program SCE2510 would provide a NTG result with a precision of 90/4. Consequently, a net-to-gross evaluation conducted according to the original research plan for this contract group would have produced results with a much higher precision than the 90/10 target that had been set for EM&V analysis.

In a June, 2009 memo to ED, we therefore proposed to scale back the net-to-gross samples to target a 90/10 confidence/ precision level instead of the "300 sample points or one-half of the program size." Assuming an error ratio of 0.5, the estimated 90/10 sample size for program SCE2509 was 55 points. Similarly, the 90/10 estimated sample size for program SCE2510 was 65 points.

As a result of re-directing resources to the analysis of Steam Traps and Pipe and Tank Insulation HIMs, the M&V scope for programs SCE2509 and SCE2510 was limited to the samples drawn in March 2008: 30 points for SCE2509 and 10 points for SCE2510. Scaling back the net-to-gross effort would not interfere with achieving complete overlap between the net-to-gross sample and the M&V sample.

Table 6-4 below summarizes the NTG sample design for program SCE2509 and Table 6-5 summarizes the NTG sample design for program SCE2510. For ease of integration of NTG results and M&V results, the strata boundaries defined for the M&V sample design were used to stratify the population of measures for each program. The NTG sample points were allocated such that the fraction of total sample points in a given stratum was roughly equal to the percent of total program impacts installed in that stratum; the M&V sample was nested within the NTG sample.

Table 6-4: SCIA Net-to-Gross Sample Design – Program SCE2509 Industrial as of Q4, 2008

	SCE2509 Program		NTG Sample		NTG Sample Precent	
Strata	N records	Gross kWh	N records	Gross kWh	N records	Gross kWh
1	8	35,070,679	8	35,070,679	100%	100%
2	10	17,865,601	10	17,865,601	100%	100%
3	13	14,405,448	11	12,321,328	85%	86%
4	46	29,308,941	10	6,543,541	22%	22%
5	187	32,815,475	16	2,543,188	9%	8%
Total	264	129,466,144	55	74,344,337	21%	57%

Table 6-5: SCIA Net-to-Gross Sample Design – Program SCE2510 Agricultural as of Q4, 2008

	SCE2510 Program		NTG Sample		NTG Sample Precent	
Strata	N measures	Gross kWh	N measures	Gross kWh	N measures	Gross kWh
1	17	12,798,362	13	8,919,233	76%	70%
2	21	7,248,408	7	2,395,222	33%	33%
3	87	16,492,260	17	3,818,881	20%	23%
4	136	12,647,153	13	1,468,855	10%	12%
5	867	15,352,880	15	342,742	2%	2%
Total	1,128	64,539,063	65	16,944,932	6%	26%

* Claimed results exclusive of the 5 Strip Curtain records that were included in the Commercial Facilities HIM evaluation.

6.2.6 Approach to Estimating Ex-Post Energy Savings

This evaluation used a similar set of approaches to estimating ex-post gross energy savings as used in previous California IOU industrial program evaluations, relying primarily on site-specific measurement and verification for the impact evaluation. The key steps used to develop an overall savings estimate for the contract group were to:

- independently verify reported measure installation records,
- develop ex-post estimates of the energy savings for each project in the sample, and
- apply these findings to the full participant population to obtain a complete estimate of program impacts.

The approach to the impact analysis consists of a distinct set of steps that are listed below and discussed in the subsections that follow. These steps include:

- Developing and implementing the sample design;
- Obtaining the sample of application files and associated documentation;
- Reviewing the applications and preparing the ex-post analysis plans by site;
- Scheduling and conducting the on-site data collection, conducting project verification, and developing the ex-post impact estimates for each site;
- Preparing detailed, site-specific impact evaluation reports;
- Carrying out a quality control review of the ex-post impact estimates and the associated draft site reports and implementing any necessary revisions;
- Estimating a net-of-free-ridership ratio for each site;
- Extrapolating the final ex-post realization and net-of-free ridership estimates for the sample to the remaining applications; and
- Reviewing each of the steps above with ED and their consultants.

For the sampled participant sites, the engineering analysis methods and degree of monitoring varied from project to project, depending on the complexity of the measure, the size of the associated savings, and the availability and reliability of existing data.

To address the wide range in size and complexity across projects, and to maximize the number of sample points for the evaluation, a multi-tiered level of effort for the site-specific engineering work was planned and implemented, from least to most complex and engineering resource intensive. Table 6-6 below provides an overview of the engineering level of effort tiers. The complexity and heterogeneity of expected projects required that enough effort be spent on each site's M&V to produce an accurate and defensible ex post savings estimate. Increasing the reliability of the ex post estimates is as important as increasing the reliability of the sample (which occurs by increasing the number of sample points). In the table below, we summarize the range of M&V efforts that we incorporated into our sample design and site evaluation resource plans.

Table 6-6: Overview of Engineering Level of Effort Tiers

M&V Tier	Description
Level 1	▪ Largest and most complex projects. Detailed application review, on-site verification, collection of data on key parameters, billing/interval data analysis, calibrated simulation models, spot measurements, long-term post monitoring, pre- verification and short-term measurement. ▪ Approximate hours per site: 200
Level 2	▪ Large, complex projects. Detailed application review, on-site verification, collection of data on key parameters, billing/interval data analysis, engineering models, spot measurements, mid-term post monitoring, pre- verification. ▪ Approximate hours per site: 160
Level 3	▪ Large or relatively complex projects. Detailed application review, on-site verification, collection of data on key parameters, billing/interval data analysis, engineering models, spot measurements, short-term post monitoring, pre-verification. ▪ Approximate hours per site: 100
Level 4	▪ Medium size projects requiring monitoring or metering. Detailed application review, on-site verification, collection of data on key parameters, revised engineering calculations, billing data analysis, and spot measurements, possible short term post monitoring. ▪ Approximate hours per site: 60
Level 5	▪ Smaller, simpler projects. Detailed application review, on-site verification, collection of data on key parameters, revised engineering calculations, billing data analysis, and possible spot measurements. ▪ Approximate hours per site: 30

6.2.7 Obtain Sample Application Records

Details of the sampling approach are covered in the proceeding section. Once a sample of projects was selected, Itron submitted a formal data request to SCE for the application records, including site data, verification records, all savings calculations, and all information transactions. Once those documents were received, the individual engineer assigned to each application conducted an initial interview. This was used to develop the site-specific engineering plan and to assess the need for additional documentation.

6.2.8 Review Applications and Prepare Analysis Plans

For each selected application, the assigned engineer performed an in-depth application review to assess the engineering methods, parameters and assumptions used to generate all adjusted ex-ante impact estimates. Application review served to familiarize the assigned engineer with the gross impact approach applied in the program calculations. This also allowed an assessment of the additional data and monitoring needs that were required to complete each analysis and the likely sources for obtaining those analytic inputs. Data sources included on-site personnel, visual inspection of the systems and equipment, EMS data downloads, spot measurements, short-term monitoring (e.g., less than four weeks), and mid-term monitoring (4 to 8 weeks).

A site specific measurement and verification plan was developed for each site and submitted to ED staff and their consultants for review and approval. This plan outlined the general ex-post impact approach to be used (including monitoring plans), provided an analysis of the current inputs and identified calculations necessary to complete the evaluation. The plan specified what data was required to be collected during the site visit.

The ex-post methods applied varied in complexity from applications that required an entirely new approach, to those that required an independent calculation using the application-based approach, to those that simply required a careful review and verification of the methods and inputs in the ex-ante calculations, and finally to those that required the installation of loggers or other monitoring equipment in support of detailed engineering calculations.

6.2.9 Schedule and Conduct On-Site Data Collection

On-site surveys and data collection were completed for each of the 31 SCE Industrial project applications and for six of the 10 SCE Agricultural project applications in the sample. The On-site data collection form can be found in Appendix D-4. During the site visit, the Itron team engineer met with a customer representative knowledgeable about the equipment and operation, and asked a series of questions regarding operating schedules, location of equipment, and equipment operating practices. During the on-site survey, data identified in the measurement and verification plan was collected, including monitoring records (such as instantaneous spot watt measurements, measured fluid and gas temperatures, data from equipment logs, energy management system (EMS) downloads, Supervisory Control and Data Acquisition (SCADA) system data, equipment nameplate data, system operation sequences and operating schedules, and, a careful description of the baseline condition being modeled. Information was also collected related to net-to-gross analysis and baseline specification, including financial information on project economics, reasons for conducting the project, and remaining useful life of replaced equipment.

6.2.10 Conduct Site-Specific Verification and Impact Calculations

After all of the field data was collected, energy and demand savings were developed based on the on-site data, monitoring data, application information, and third-party implementer records and, in some cases, billing/interval data.

Energy savings calculations were accomplished using methods that include short-term monitoring, simulation modeling, bin models, application of engineering methods and algorithms, analysis of pre- and post-installation billing and interval data, and other specialized algorithms and models. Short-term monitoring was a priority for all sites, and peak demand savings was also estimated for all projects in the sample. In cases where billing/interval data analysis was used to estimate energy savings, peak demand savings were estimated using all data and methods, principally interval data (if available) and engineering calculations based on estimates of operating profiles and coincident peak diversity factors.

6.2.11 Site-Specific Net-to-Gross Analysis

As discussed further below, a detailed net-to-gross analysis was conducted for each project in the net-to-gross sample. All of the gross impact sites were included in the net-to-gross sample and most of the associated net-to-gross interviews were completed. Net-to-gross interviews were conducted by both professional consulting staff and by staff in Itron's computer-aided telephone interviewing (CATI) center depending on the rigor level of the sample point. For the higher rigor level interviews, the net-to-gross interviewing team worked with the engineering team to prepare for both NTG and gross impact surveys, share information collected from each survey, and to discuss issues related to determining the final NTG and gross impact estimates. Care was taken to ensure that the results were internally consistent and did not include any double counting of effects between the two samples.

6.2.12 Site-Specific Analysis Documentation

Detailed documentation is provided in the site report appendix for each site included in the gross impact analysis. The site report documentation includes the following elements:

- Measure Description
- Summary of Program Impact Calculations
- Comments on Program Impact Calculations
- Description of the Impact Evaluation Process
- Impact Evaluation Results
- Supporting Documentation

6.2.13 Quality Control Review and Final Site Reports

The Itron engineering and project management team reviewed and commented on all draft site reports and provided feedback to each assigned engineer for revisions or other improvements. Each assigned engineer then revised the draft reports as necessary to produce the final report as approved by Itron.

The second level of quality control consisted of submitting the draft site reports to the CPUC and its consultants for their review and approval. This review provided an important additional level of quality assurance, thereby enabling the CPUC to make the final decisions on policy-related requirements for project eligibility and baseline specification.

6.2.14 Estimate Verification and Realization Rates, and Impacts for Participant Population

Extrapolation of the site-specific ex-post results to the population was carried out following the statistical procedures for ratio estimation discussed in Section 3.3 and in the *California Evaluation Framework Study*. The population-weighted results are provided in Section 1.3 (Results) of this report.

6.2.15 Approach to Uncertainty

The consideration of uncertainty was an important part in this evaluation. In addition to the uncertainty that is associated with sampling error, there also uncertainty associated with the estimation and measurement of savings. It is well-known that there is uncertainty in both the ex-ante and the ex-post energy savings estimates. Program level savings estimates are affected by both the number of points sampled (sampling error) and the degree to which the measured site level energy savings estimates reflect the "true" savings (measurement error). The purpose of conducting rigorous site-level evaluation of savings for custom efficiency projects in heterogeneous applications is to estimate savings as reliably as possible (i.e., minimize measurement error) given available resources and the need to maintain a large enough sample to provide adequate sampling precision.

If the component of measurement uncertainty could be directly observed then a mathematical trade-off could be made to optimally reduce the uncertainty of both sampling and measurement error, that is, by sampling more points to decrease the standard error of the mean, or by investing more time and money into each individual site to reduce the individual site measurement error. The dispersion of the sample cannot be characterized directly, but it can be inferred, knowing the total uncertainty and the measurement uncertainty if we assume that the two components to the dispersion (measurement error of the "true" value and variation of "true" values within the population) combine in quadrature to yield the combined dispersion of the sample. Sampling error can be estimated a priori using error ratios and standard errors from similar evaluation efforts conducted previously; however, measurement error cannot be easily estimated as the true savings values are not directly observed. Thus, the trade between the number of sample points and level of effort to reduce measurement uncertainty through increased site-specific evaluation efforts requires judgment. This is a value maximization exercise in which a sample size is specified to meet the study requirements and then the available resources are allocated across the sites based on their size and complexity to minimize the measurement error.

The discussion above concerns tradeoffs between sample error and measurement error for the overall evaluation planning effort. In addition, uncertainty was also addressed in developing the M&V plan for each individual site in this evaluation. Site level energy savings uncertainty was addressed in a two-fold effort. In the first step, in the M&V planning stage, uncertainty analysis was used to understand which variables were likely to have the largest contributions to the overall measurement error in order to focus the site M&V effort on reducing error in the most cost-effective and feasible manner. In the second step, after the M&V work and associated analysis was complete, the ex post results were used to revisit the uncertainty analysis to investigate the levels of uncertainty that remained. This information is useful both to better understand the results of the current study and as input to help characterize measurement uncertainty to improve evaluation and M&V planning for similar future studies.

Because the data collection and savings calculation methods are different for varied industrial sites with different types of measures, an uncertainty analysis could not be done in aggregate for the whole program, but instead was needed at the site level. With over thirty sites in the sample, it was not possible to perform detailed uncertainty analysis with statistical methods (e.g., Monte Carlo) on every site.

However, in the M&V planning stage for nearly every site, the individual components driving uncertainty – the variables in an energy savings equation – were analyzed to aid in the design of M&V that would result in reduced uncertainty, even if a formal analysis was not performed.

Measurement uncertainty stems from several sources: instruments to monitor variables such as power, flow, or even on/off periods have potential measurement error. Instrument error can be reduced by choice of more accurate equipment and proper training on equipment use. A mean measured value can also be characterized more precisely by monitoring it for a longer time period, as the uncertainty of the mean is inversely proportional to the square root of the number of measurements made. In most cases, we are interested in the mean value of the variables we measure, so monitoring for a longer period is very valuable in reducing measurement error at a site.

For non-measured values, there was an evaluator-determined uncertainty as to the expected range for that variable, informed by observations on site, conversations with on site staff, and familiarity with possible and expected ranges for the type of variable. As an example, hours can range from 0 to 8,760 hours per year, but for a regularly scheduled application, like factory lighting fixtures on an energy management system, the hours may be known fairly well with only a relatively small range of error.

For those cases for which a Monte Carlo analysis was conducted, each of the key variables were described with an expected range, the type of error distribution (e.g., normal, triangular, lognormal, etc.), and either standard deviations or maximum /minimum values. For some sites, detailed uncertainty analysis was performed using Crystal Ball tm or At Risktm software to run a Monte Carlo simulation and propagate the uncertainty through the energy savings equation or calculation. Monte Carlo is an appropriate method to use on both simple sites and sites where the calculation is complex and there are multiple parameters of import.

Where uncertainty analysis was conducted in the site planning and ex post analysis, it is discussed in the site reports in Appendix D-5.

6.2.16 Approach to Determining Gross Baselines

The evaluation worked diligently in determining an appropriate gross impact baseline for all projects in the sample, always in coordination with free ridership estimation, wherever appropriate. Assessment of both full free ridership and partial free ridership is related to the selection of baselines used for gross savings analysis. Free ridership occurs when the program participant would have installed the program-incented or recommended measure in the absence of the program. Partial free-ridership can occur when, in the absence of the program, the participant would have installed something more efficient than the baseline efficiency specified for the gross savings estimation but not as efficient as the item actually installed as a result of the program. For example, in the absence of the program, a participant states that they would have installed 5 VSDs rather than the 10 installed through the program. Or, as another example, in the absence of the program, a participant might have installed an industrial boiler system with an efficiency of 82% (greater than a program-assumed baseline of the existing efficiency code, e.g., efficiency of 80%) rather than the 85% efficiency that was installed through the program.

A challenge that occurs in a number of industrial projects is how to define the evaluation's baseline for gross savings with respect to program requirements that reference "industry standard practice" as the basis for the baseline. In some cases, the availability of efficiency options above the industry standard practice baseline may leave room for further savings adjustment due to partial free ridership. In other cases, there may be few or no efficiency options above the industry standard practice baseline, the result of which may be low or zero gross savings. Evaluators' choices of baselines may differ from those selected by program administrators for a number of reasons as discussed in the remainder of this subsection.

Differences in baseline choices between evaluators and implementers will lead to differences in savings estimates and evaluation realization rates. Documenting these baseline selection differences and explaining the basis for them is an important part of the industrial evaluation process and final evaluation report findings and recommendations.

Below are several principles that Itron used as guidance for determining the appropriate baseline to be used in calculating the gross savings for a project in the PG&E Fab contract group:

Code or market baselines were used for replace-on-burnout and 'natural turnover'. In situ baselines were only used for the portion of the remaining useful life (RUL) of the pre-existing equipment that was eliminated due to the program. Consideration was given to the specifics of the application with respect to the remaining life, if any, of the pre-existing equipment when selecting the baseline, including:

- In-situ equipment was used as the gross baseline only when the existing equipment was not at the end of its useful life and there was no compelling evidence that the pre-existing equipment had a remaining useful life.
- Code requirements or industry standard practice baseline were used for replace-on-burnout or natural turnover situations.
- Care was taken in the use of industry standard practice baseline with respect to how much, if any, savings adjustments applied to gross versus net savings.

CPUC policy rules and IOU program eligibility rules governed the baseline. Careful review of utility and third-party program and CPUC policy rules were made and adjustments were applied to gross savings in some cases, while in others to net savings. The adjustments were applied to gross when there was clear evidence from program or policy rules that savings claims may not be made nor rebates paid for the case in question. Program rules also came into play with respect to gross baseline requirements, e.g., specifying a given efficiency level or percentage above code. In situations where program or policy rules were in question, the case was reviewed by the Itron project management team, ED's consultants, and ED, with ED making the final judgment on whether rules were violated and whether associated corrections were required in the baseline determination or measure qualification.

Minimum production or energy service requirements govern the baseline. In some situations, a measure for which savings were claimed may be the only acceptable equipment for an application. In such cases, the baseline was set at the minimum needed to meet the requirements. Care was taken to ensure that the production or energy service requirements were not merely preferences, but were fundamentally required. An example would be an industrial process where only a variable-speed drive pumping system could meet the production requirements.

For situations where the baseline conditions are changed (such as production levels), the baseline equipment was defined as the minimum equipment needed to meet the revised conditions. This could result in changes in gross savings if claimed savings were set at pre-installation requirements.

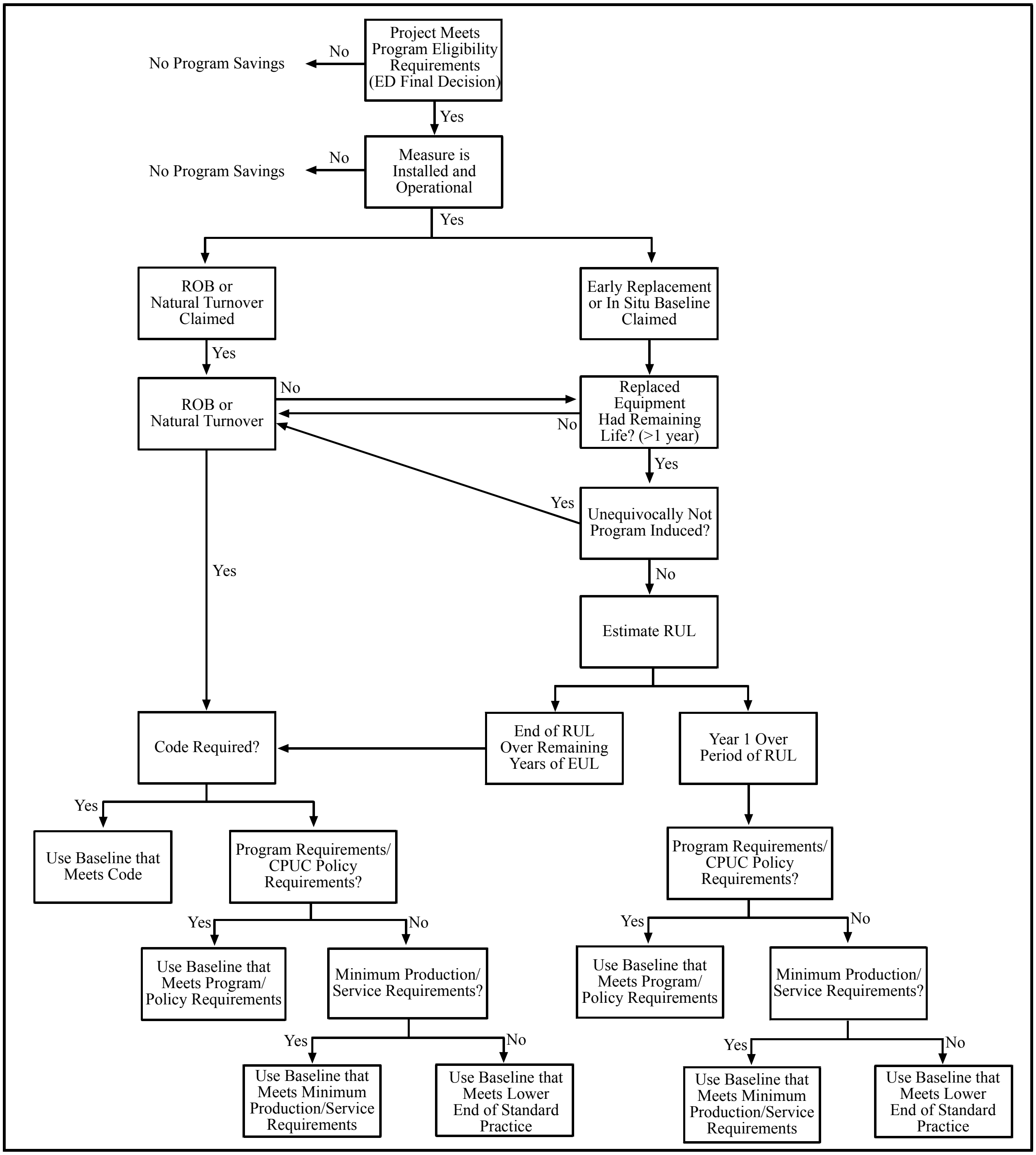
Evaluate early replacement RULs and program inducement. The gross engineering team determined whether there was evidence that early replacement actually occurred, that is, that there was remaining life on the equipment replaced. If so, an estimate was made of the associated RUL. The net team, in consultation with the engineering team, was responsible for determining whether the early replacement was program induced. If the early replacement was not program induced, the gross baseline was set based on the ROB/natural turnover guidelines.

The decision tree which was used as guidance for determining the baseline for gross savings can be found in Figure 6-3 below. The application of site specific baselines, gross and net baseline approaches were reviewed by ED and its consultants.

Figure 6-3: Baseline Guidance

Guidance for Determination of Baseline for Gross Savings

Take Most Efficient of All Applicable Cases



6.2.17 Approach to Estimating Freeridership

As part of the evaluation of the 2006-08 energy efficiency programs designed and implemented by the four investor-owned utilities and third parties, ED formed a nonresidential net-to-gross ratio working group that was composed of experienced evaluation professionals. The main purpose of this group was to develop a standard methodological framework, including decision rules, for integrating in a systematic and consistent manner the findings from both quantitative and qualitative information in estimating net-to-gross ratios.

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 Appendix A-3.

- The method used a 0 to 10 scoring system for key questions used to estimate the NTGR rather than using fixed categories that were assigned weights. It 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 taken into account in assessing the unique contribution of the program as reflected in the NTGR.

- There are three levels of free-ridership analysis. The most detailed level of analysis, the Standard – Very Large Project NTGR, was applied to the largest and most complex projects (representing 10 to 20% of the total) with the greatest expected levels of gross savings.[95] The Standard NTGR, involving a somewhat less detailed level of analysis, was applied to projects with moderately high levels of gross savings. The least detailed analysis, the Basic NTGR, was applied to all remaining projects. Evaluators exercised their own discretion as to what the appropriate thresholds should be for each of these three levels.

Data Sources. 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.

---

[95] 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.

Table 6-7 below shows the data sources that were 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 was utilized in the analysis varied. For example, all three levels of analysis obtained core question data from the Decision Maker survey.

Table 6-7: 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]	√	√	√

Footnotes below reference question numbers provided in Appendix D-1.

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.

NTGR Questions and Scoring Algorithm. The NTGR was calculated as an average of three scores. Each of these scores represented 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 reflected 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. Program influence through vendor recommendations was also incorporated in this score.

2) A **Program Influence** score that captured 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 was determined by asking respondents to assign importance values to both the program and most important non-program influences so that the two total 10. The program influence score was adjusted (i.e., divided by 2) if respondents said they had already made their decision to install the specific program qualifying measure 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 accounted 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 were multiple questions that feed into the scoring algorithm, as was the case for both the **Timing and Selection** and **No-Program** scores, the maximum score was always used. The rationale for using the maximum value was to capture the most important program element in the participant's decision making. Thus, each score was always based on the strongest influence indicated by the respondent. However, high scores that were inconsistent with other previous responses triggered consistency checks and led to follow-up questions to clarify and resolve the discrepancy.

When there were missing data or 'don't knows', to critical elements of each score, one of two options was used. The missing element was sometimes backfilled with a value that represents the average of the lowest and highest extreme values. Alternatively, if it was one of several other elements that were considered in the algorithm, the missing element may simply have been excluded from consideration.

The self-reported core NTGR in most cases was simply the average of the Program Influence, Timing and Selection, and No-Program Scores, divided by 10. The one exception to this was when the respondent indicates a 10 in 10 probability of installing the same equipment at the same time in the absence of the program, in which case the NTGR was based on the average of the Program Influence and No-Program scores only.

Data Analysis and Integration. The calculation of the Core NTGR was generally mechanical and was 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 required 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 were conducted eliciting both quantitative and qualitative data and a variety of program documentation had been collected, all of this information was integrated into an internally consistent and coherent story that supported a specific NTGR.

Sometimes, *all* the quantitative and qualitative data clearly pointed in the same direction while, in others, the *preponderance* of the data pointed in the same direction. Other cases were more ambiguous. In all cases, in order to maximize reliability, it was essential that more than one person was involved in analyzing the data. Each person analyzed the data separately and then compared and discussed the results. Important insights can emerge from the different ways in which two analysts look at the same set of data. Ultimately, differences were resolved and a case made for a particular NTGR. Careful training of analysts in the systematic use of rules was carried out to insure inter-rater reliability.[96]

[96] 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.

Once the individual analysts completed their review, they discussed their respective findings and presented their respective rationales for any recommended changes to the equation-derived NTGR. The outcome of this discussion was the final NTGR for a specific project.

6.3 Results

This section presents the quantitative results of the evaluation. It begins with an analysis of program-specific participation patterns by quarter, and then continues with a detailed reporting of gross and net realization rates for each sampling domain.

Please note that the gross impact findings are limited to the SCE2509 Industrial program only. The limited gross impact sample (10 sample points) for the SCE2510 Agricultural program is not sufficiently robust to support a program population-level level gross realization rate result. However, Net-to-Gross findings are presented for both the SCE2509 Industrial program, and the SCE2510 Agricultural program, since the NTG samples were sufficiently large in both cases.

6.3.1 Quarterly Participation Patterns

Figure 6-4 and Figure 6-5 below summarize the cumulative participation and associated claimed kWh savings by quarter for each SCE program. These data clearly demonstrate the 'hockey stick' effect of participation in these types of Industrial and Agricultural programs, whereby long-lead time projects lead to relatively low savings claims during the early part of the plan period, and dramatically higher savings claims at the end of the plan period. Note that both programs' participation in the 4th quarter of 2008 increased by between 40% and 50% relative to the 3rd quarter of 2008.

Figure 6-4: PY2006-2008 Participation for SCE2509 Industrial

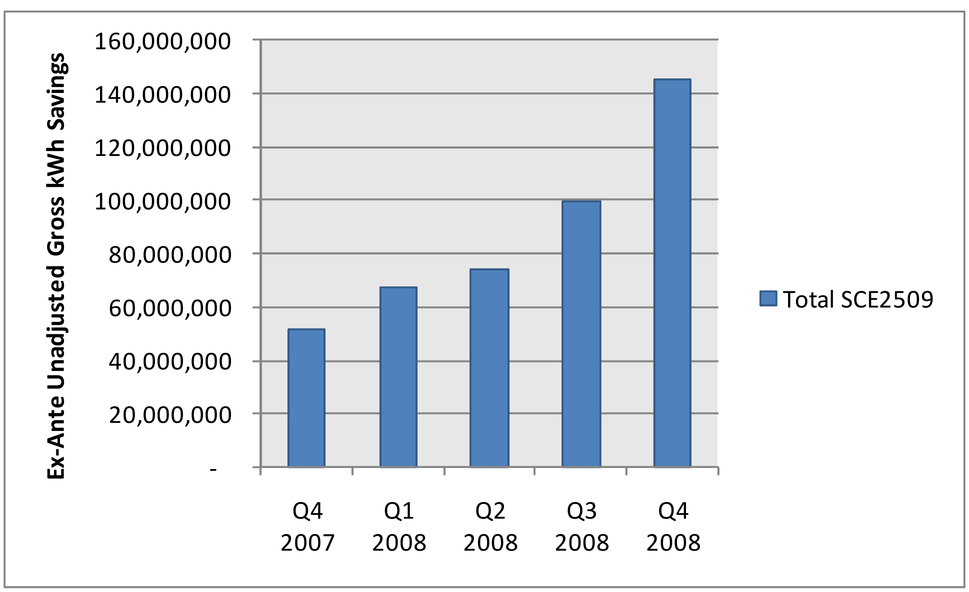
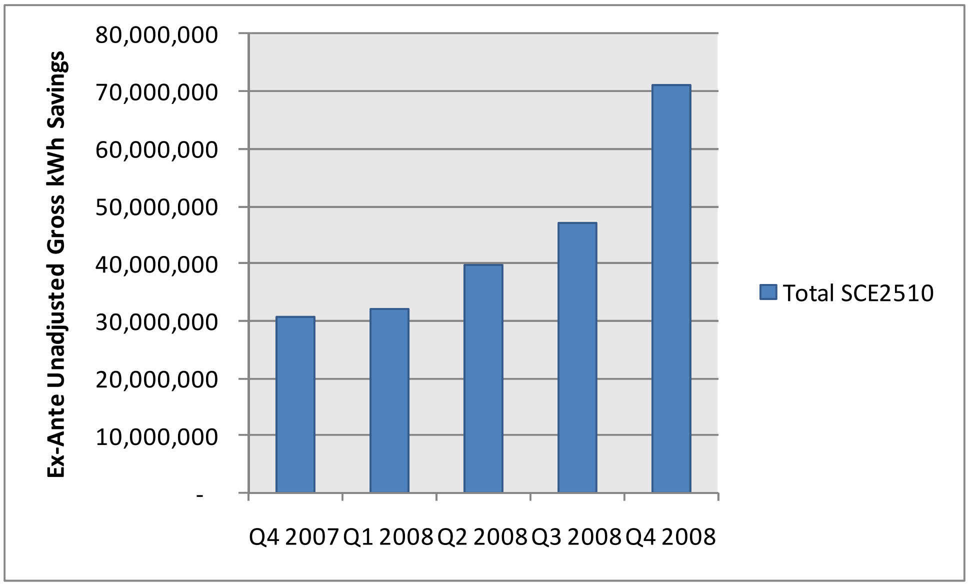


Figure 6-5: PY2006-2008 Participation for SCE2510 Agricultural



6.3.2 Site-Specific Gross Impacts

In this sub-section we present our gross impact results on an unweighted basis by project. Anonymous site-specific results are included in summary tables. The impact evaluation results are based on a combined sample, for all 3 sampling domains, of 30 matched gross and net projects. As described in Section 6.2, a complete M&V plan and an impact evaluation report were developed for each site. The resulting detailed site-specific project descriptions, ex-ante methods, ex-post methods, and ex-post results are provided in the site reports which comprise Appendix D-5.

Ex-ante energy savings from the Installation Report, ex-post savings from this impact evaluation, and associated realization rates are shown in Table 6-8 for each project in the evaluation sample. The last column in the table also includes an explanation of the reasons for the differences found between the ex-post and ex-ante estimates. A description of the retrofit performed at each evaluated site is shown in Table 6-9.

Table 6-8: Summary of Ex-Ante and Ex-Post Savings for All Sampled Projects

Site ID	Ex-Ante Savings			Ex-Post Savings			Gross Realiz. Rate		NTG ratio	Reason for differences
	kW	kWh	Therms	kW	kWh	Therms	kW	kWh		
C001	273.23	2,349,441		66.00	580,090		0.24	0.25	0.23	The ex-ante estimates used an average motor load factor (MLF) of 72%, as compared to the average ex-post measured value of 29%. Also, the ex-post pre-retrofit kW loading was lower, and the ex-post post-retrofit operating hours were higher than those used in the ex-ante calculations.
C002	330.19	2,462,639		476.00	4,169,431		1.44	1.69	0.97	The ex-ante kWh savings was an estimate; the ex-post kWh is calculated using data measured through the SCADA system. By inference, this 10% difference appears to be within expected limits.
C003	1,018.16	2,483,799		406.30	372,918		0.40	0.15	0.93	Reduced hours of operation account for most to the difference between ex-ante and ex-post. Also, ex-post calculations used the demand of the motors including ramp-up and ramp-down periods, while the ex-ante calculations did not.
C004	172.66	1,858,269		53.90	472,477		0.31	0.25	0.77	The ex-post calculated Motor Load Factor (MLF) – the kW drawn to full load kW – was about 35% - 45%, which was considerably less than the ex-ante assumed MLF of 76%. Also, the ex-post pre-retrofit kW loading was lower, and the ex-post post-retrofit operating hours were higher than those used in the ex-ante calculations.
C005	0.00	1,793,421		0.00	1,820,060		-	1.01	Nonresponse	The report cited better sequencing and control strategy of equipment as an explanation of the savings, but by inference, the ex-ante and ex-post results appear sufficiently close not to warrant explanation. Note: No demand savings were submitted with the application.
C006	160.20	1,778,690		134.40	913,517		0.84	0.51	0.63	The ex-ante calculations overestimated the extent to which the motors would unload. The ex-post data shows that the motors operate at constant load when they operate. In addition, operating hours were found to be significantly less than originally predicted
C007	239.41	2,066,079		258.00	2,229,510		1.08	1.08	Nonresponse	The ex-ante and ex-post results appear sufficiently close not to warrant explanation.
C008	180.67	1,519,744		96.80	662,389		0.54	0.44	0.57	Production increases associated with the installed measure may be responsible for increasing whole site energy usage, thereby reducing ex-post savings. Ex-ante and Ex-post results were normalized against pre-install production levels.
C009	129.94	1,116,557		25.30	261,269		0.19	0.23	0.73	The motor load factors measured during the evaluation were much lower than the motor load factors used in the ex-ante analysis. The ex-post baseline pre-retrofit runtimes for the well pumps were shorter than the assumed application baseline.
C010	105.02	919,234		93.10	815,649		0.89	0.89	0.08	The difference between ex-ante and ex-post savings can be attributed to two pumps in which the ex-ante savings projected did not materialize; VFD efficiency of less than 100% for another caused a kWh penalty for one well operating at full speed. Speeds were not reduced appreciably at many of the seven wells in this application.
C011	138.84	1,000,679		116.32	1,238,159		0.84	1.24	1.00	Differences probably due to the usage of assumed operating profiles in the simulation model. Also, the demand reduction estimated from the two demand shifting measures was zero, resulting in the reduced demand savings number.
C012	197.58	1,186,906		0.00	0		0.00	0.00	0.57	Installed equipment is not operating due to environmental issues and there is no date for its operation.
C013	99.68	1,182,321		113.60	1,058,943		1.14	0.90	Nonresponse	The ex-post energy savings values are less than the ex-ante energy savings because the ex-ante calculations overestimated the number of fixtures controlled and the percent of hours saved per controlled fixture (40%).
C014	104.13	909,960		21.90	191,712		0.21	0.21	0.08	Ex-post average annual energy usage was much lower than the ex-ante calculation due to the underestimation of actual operational loading of the new installed equipment and an incorrect use of post-retrofit production data to determine savings.
C015	165.54	1,156,972		101.33	1,007,921		0.61	0.87	Nonresponse	The ex-post demand reduction is less than the ex-ante because the ex-ante baseline assumed that the high pressure system supplied the low pressure air system, which artificially inflated the ex-ante savings impacts.
C016	89.00	1,050,200		84.00	861,339		0.94	0.82	0.78	Ex-post kWh savings were less than ex-ante savings due to lower loading of the baseline pre-melters during non-operational hours. Monitoring data showed that significantly higher loading was required during manufacturing operations and significantly reduced loading was required to maintain liquid slurry during non-operational hours.

Table 6-8: Summary of Ex-Ante and Ex-Post Savings for All Sampled Projects (Continued)

Site ID	Ex-Ante Savings			Ex-Post Savings			Gross Realiz. Rate		NTG ratio	Reason for differences
	kW	kWh	Therms	kW	kWh	Therms	kW	kWh		
C017	95.23	825,991		-6.00	-51,827		-0.06	-0.06	0.50	Both wells produced more water than predicted in the ex-ante case, resulting in comparatively higher ex-post energy demand and use.
C018	80.99	692,913		17.70	156,065		0.22	0.23	0.50	Only one well out of three had a reduced water production rate compared to the geologist's prediction. In the other two wells, ex-post water production exceeded the ex-ante prediction.
C019	69.42	600,853		60.78	532,310		0.88	0.89	Nonresponse	The ex-post water production slightly exceeded the ex-ante prediction, reducing the ex-post savings.
C020	68.53	593,414		228.00	2,000,973		3.33	3.37	Nonresponse	Ex-post results far exceeded the ex-ante prediction. This well had a water-oil ratio of 78% compared to the predicted 99% ratio.
C021	56.96	518,085		28.00	245,578		0.49	0.47	Nonresponse	The measured ex-post motor load factors were lower than ex-ante load factors, which resulted in lower overall kW for both pre-retrofit and post-retrofit cases. In addition, the ex-post baseline pre-retrofit runtimes for the well pumps were shorter than the assumed application baseline due to the observed presence of timers on these wells.
C022	0.00	645,797		0.00	358,477		-	0.56	Nonresponse	The ex post energy savings are lower than the ex ante energy savings because the ex ante calculations overestimated the flow rate of the milk pasteurization and fill lines.
C023	76.54	647,487		39.40	344,815		0.51	0.53	Nonresponse	The installed 800-hp compressor had a lower efficiency than the 700-hp compressor described in the application, and used for ex-ante estimates
C024	67.64	548,399		22.70	406,840		0.34	0.74	Nonresponse	The methodology for calculating the ex-ante energy use was not clear. Compressed air demands were unchanged over the entire period. Economic forces reduced operating times, which may partially account for the reduced energy savings.
C025	41.83	363,383		13.50	116,413		0.32	0.32	0.50	Well D405A was unproductive and excluded from the ex-post results. Well D741 produced a higher proportion of water than was predicted.
C026	27.59	232,562		30.80	269,999		1.12	1.16	0.63	The existing motor was outdated but still had some useful life (early replacement measure.) This older motor operated at low efficiency, while the available motors (current practice) operate at nearly 96% efficiency.
C027	28.48	160,628		1.67	14,311		0.06	0.09	0.70	Only one well of the three had a POC still in use; the other two were converted to injection wells. For the operating well, the ex-ante savings were over estimated due to a higher anticipated load factor (75%) compared to the actual case (28%).
C028	14.24	122,384		14.24	122,384		1.00	1.00	0.23	All three wells were converted to submersible pump systems with variable frequency drives. One of these three, Well VRU 304, operated at least 50 days with a POC prior to its conversion and accounted for the stated annualized savings.
C029	8.90	75,497		14.10	123,122		1.58	1.63	0.10	The ex-ante analysis had assumed that the well runtime would be 70% once the pump off controller was installed. Data obtained during the ex-post site visit shows that the average run time for well # 1 is 44%, and for #2 is 41%, so ex-post savings are higher than ex-ante.
C030	7.12	60,605		10.90	89,980		1.53	1.48	0.67	Two factors contribute to improved ex-post performance. Higher compressor load compared to the ex-ante assumptions resulted in increasing observed efficiency. Also, compressor runtime was slightly shorter than originally estimated.
C045	66.75	788,280		35.40	755,416		0.53	0.96	Nonresponse	The ex-ante estimates did not reflect the demand reduction operating strategy that was in place before the retrofit. The ex-post baseline operating hours were lower than predicted and the ex-ante under estimated the operating hours of the new fixtures.

Table 6-9: Summary of Retrofit Descriptions for All Sampled Projects

Site ID	Measure Description
C001	Installation of 17 pump-off controllers (POCs) for Oil Wells using Standard Rod Beam Pumps
C002	Installation of gamma metric analyzer to provide a reliable and accurate means of achieving consistent raw mix chemistry to improve kiln efficiency
C003	New headers and valves that reduce the water flow rate in an existing descaling system. The measure enabled reduction to 24 nozzles for the narrower sheets.
C004	Install Pump-Off Controller on Thirty (30) Oil Wells
C005	Refrigeration Controls Upgrade: Stage the Compressors; Floating Condenser Discharge Pressure; Control Glycol Chiller; Enhanced Evaporator Temperature Control; Control of new VFDs on Condenser Fan.
C006	Process Water Controls - Installation of controls and VFDs to reduce the energy use of the process cooling system.
C007	SMART well completion: selected perforations of wells 416F and 854 where oil concentrations were modeled to be the highest.
C008	Replace two DVD replicator machines with one higher efficiency DVD replication processing unit
C009	Install Pump-Off Controllers (POCs) on 9 standard rod and beam oil-well pumps
C010	Install Variable Speed Drives (VSDs) on Seven Electric Submersible Pumps (ESPs) for oil wells
C011	Install new VFD compressor, floating head pressure control strategy, and VFDs on evaporator fans (6 measures)
C012	Install a variable speed drive (VSD) on an existing autoclave ventilation fan
C013	Replacement of existing metal halide lighting system with occupancy controlled T5 and T8 fixtures
C014	Carbon absorption/adsorption gasoline vapor recovery system
C015	Install new VFD compressor and a new load/unload compressor
C016	Retrofit a manufacturing facility by replacing three plastic premelters with newer, on-demand premelters and a five-ton chiller with a newer, higher efficiency chiller, and also installing a new high-speed production line
C017	SMART well completion: selected perforations of wells J112 and J186 where oil concentrations were modeled to be the highest.
C018	SMART well completion: selected perforations of wells A561, A551, and A848 where oil concentrations were modeled to be the highest.
C019	SMART well completion for PERL 36 where oil concentrations were modeled to be the highest.
C020	SMART well completion for PERL 35 where oil concentrations were modeled to be the highest.
C021	Install Pump-Off Controllers (POCs) on 14 Oil Wells with Standard Rod Beam Pumps
C022	Installation of Two New High Temperature Short Time (HTST) Plate and Frame Regenerative Heat Exchangers
C023	Replacement of existing air compressors with a single 700 hp centrifugal air compressor
C024	Modify compressed air system to allow reduction in system pressure and replace an oversized air compressor
C025	SMART well completion: selected perforations of wells D405A and D741 where oil concentrations were modeled to be the highest.
C026	Replacement of a 2000 HP water injection pump motor with a 1500 HP pump motor
C027	Install Pump-Off Controllers (POCs) on 3 Oil Wells with Standard Rod Beam Pumps
C028	Install Pump-Off Controllers (POCs) on 3 Oil Wells with Standard Rod Beam Pumps
C029	Install Pump-Off Controllers (POCs) for 2 oil wells; retrofit those two wells and a third well with high efficiency motors
C030	Install a new premium efficiency motor and variable speed drive for an existing gas compressor
C045	Retrofit 400-watt high bay metal halide (MH) and T12 linear fluorescent lighting fixtures with T5HO & T8 linear fluorescent fixtures and extra efficient electronic ballasts

Electric measures frequently installed through the program include pump-off controllers, SMART wells, and variable frequency drives.

6.3.3 Site-Specific Net-to-Gross Results

Table 6-10 provides a summary of the net-to-gross results for all sampled projects. As discussed in the methodology section, the estimated NTGR is an average of three scores: a timing and selection score that reflects the influence of the most important of the program elements in the customer's decision to select the program measure; a program influence score that captures the perceived influence of the program relative to non-program factors in the decision to implement the measure; and a no-program score that captures the likelihood of various actions the customer might have taken in the absence of the program.
As the table shows, NTGRs range from a low of 0.08 to a high of 1.00.

Table 6-10: Summary of Site-Specific Net-to-Gross Results

Itron ID	Type	Timing and Selection Score	Program Influence Score	No-Program Score	NTGR	Adj.	APP ID	MEASURE
C001,C028	Std - VL	7	5	2	0.23	Yes	004-057,1212 018 057	Pump off Controllers
C002	Std - VL	10	9	10	0.97			Low accuracy discrete Analyzer
C003	Std - VL	8	10	10	0.93		2006-00016-B	Modification to Existing Descaling System
C004	Std - VL	9	7	7.2	0.77		GEP 007 012	Pump off Controllers
C006	Standard	9	6	6	0.63		PCTVADT	Controls and VFDs
C007	Standard	8	5	0	0.25		1212 083 015	Smart Wells
C008	Standard	10	4	3	0.57		2007-00063-D	New DVD replicator system
C009	Std - VL	10	6	6	0.73		1212 016 084	Pump off Controllers
C010	Standard	5	1.5	0	0.08	Yes	1212 003 001	VFDs - 7 pumps
C011	Standard	10	10	10	1.00		S-1571	Variable Speed Refrigeration measures
C012	Standard	10	5	2	0.57		3-000-0033-41	Autoclave fan VFD
C014	Standard	4	1.5	0	0.08	Yes	114 007	Refrigeration Vapor Recovery Unit
C015	Standard	10	8	8	0.87		2007-00023-A	Low pressure air compressor
C016	Standard	10	5	8.3	0.78		3-000-1225-95	Pre-melter Project
C017, C018, C025	Std - VL	7	4	4	0.50	Yes	Various	Smart Wells
C026	Standard	9	5	5	0.63		1212 011 079	Motor replacement and downsizing
C027	Std - VL	9	6	5.7	0.69		90 013 066	Pump off Controllers
C029	Standard	10	2	0	0.10	Yes	1212 024 053	Pump off Controllers
C030	Standard	5	5	10	0.67		1212 022 045	Premium efficiency motor and a VSD
C031	Standard	10	10	0	0.50	Yes	2006-009	Pump replacement/refurbishment at Well Site #4
C033	Standard	8	4	6	0.60		2007-021	Pump replacement/refurbishment at Q2 Pump

6.3.4 Overall First-Year Gross Impact Realization Rate for SCE2509 Industrial Program

First, we graphically summarize ex-post versus ex-ante savings estimates for the entire sample across all end-uses.

Figure 6-6 present the ex-ante (tracking system) and ex-post (engineering estimate) savings for the SCE2509 Industrial sample, for kWh and summer demand kW, respectively. The charts also include a unity line, which divides the results into those in which the site-specific realization rates were above one (sites above the line) and below one (sites below the line). Any sites for which the kW impact analysis was inconclusive are excluded from the calculation of the program realization rate (they are not defaulted to realization rates of 1.0).

Figure 6-6: First-Year Ex-Post and Ex-Ante Savings (kWh) for PY2006-2008 SCE2509 Industrial Gross Sample (n =31)

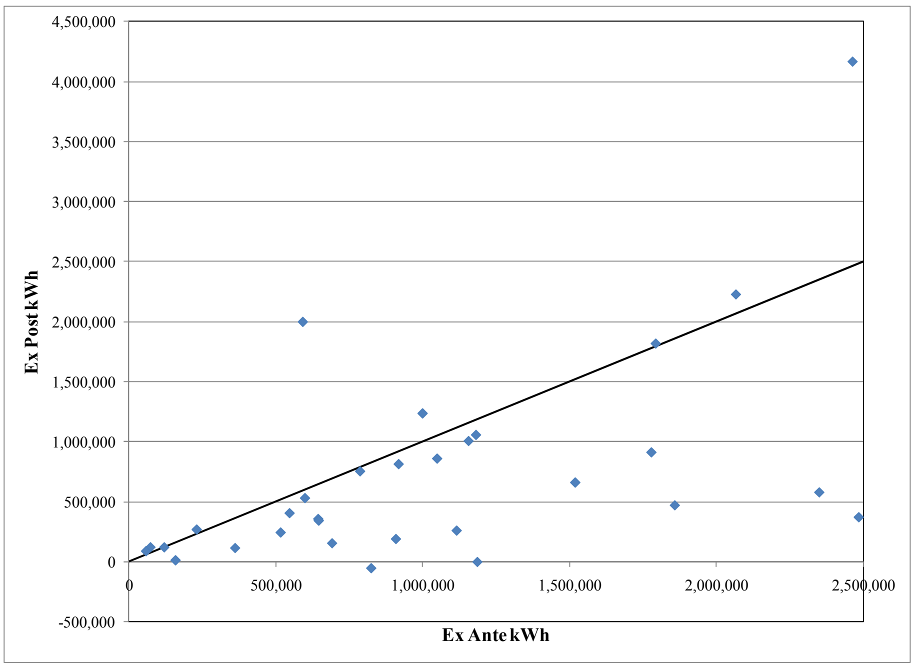
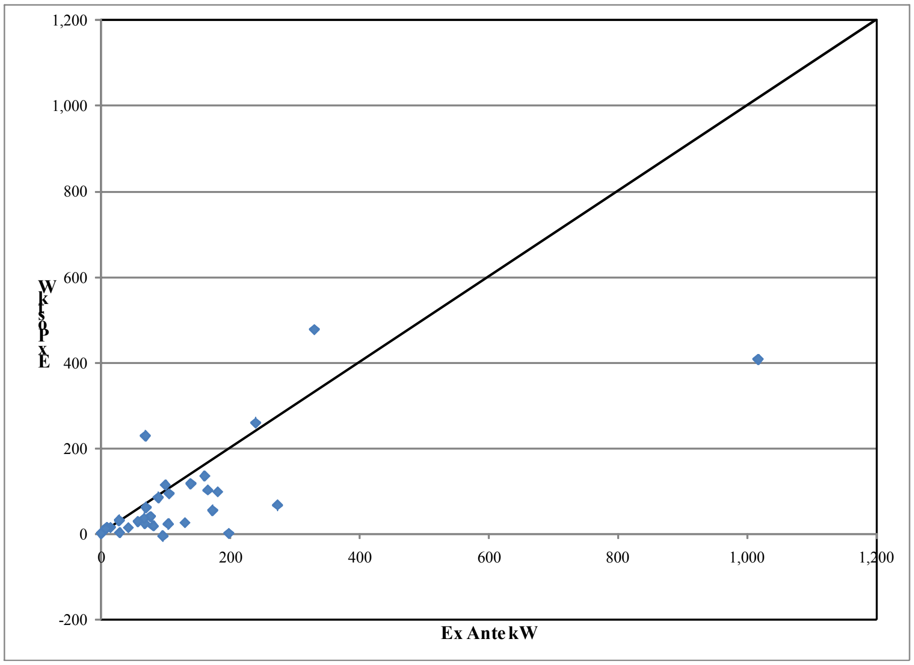


Figure 6-7: Ex-Post and Ex-Ante Savings (kW) for PY2006-2008 SCE2509 Industrial Gross Sample (n = 29)



6.3.5 Weighted Overall Program Gross Realization Rates

To produce the overall realization rate for the SCE2509 sampling domain, the individual realization rates for each of the field sample points were weighted by the size of the energy savings impacts associated with each sample project, and by the proportion of the total program impacts represented by each stratum. The total population impacts for PY2006-2008 are presented in Table 6-12.

Table 6-11 presents statistics for the population and M&V sample completes used to develop the final weighted results for each sampling domain.

Table 6-11: Tracking System and M&V Gross Sample kWh and kW Savings for PY2006-2008 SCE2509 Industrial Gross Sample by Gross Impact Weighting Stratum

	Number of Records		Gross Ex Ante kWh		Gross Ex Ante kW	
Sampling Strata	Population	M&V Completes	Population	M&V Completes	Population	M&V Completes
1	8	3	35,070,679	7,295,878	3,912	1,622
2	10	5	17,865,601	9,016,203	1,853	753
3	13	8	14,405,448	8,522,829	1,745	1,030
4	46	9	29,308,941	5,861,220	3,144	582
5	187	6	32,815,475	1,015,059	4,277	128
All	264	31	129,466,144	31,711,189	14,931	4,114

Across all sampled projects, the gross realization rates by stratum, as well as the overall weighted realization rate and the associated confidence interval are shown in Table 6-12 below. The overall weighted gross realization rate across all sampled projects is 0.72 for kWh and 0.65 for kW. The 90 percent confidence interval for the 0.72 overall kWh gross realization rate is 0.490 to 0.953. The mid-sized Stratum 3 projects were found to have the lowest realization rates.

Table 6-12: PY2006-2008 Gross Impact Realization Rates for SCE2509 Industrial Sample

	RR	
Sampling Strata	kWh	kW
1	0.70	0.58
2	0.68	0.72
3	0.64	0.54
4	0.81	0.73
5	0.73	0.66
Weighted RR	0.72	0.65
90 Percent CI	0.49 to 0.953	0.463 to 0.838
Relative precision	0.321	0.288
N measures in sample	31	29
N measures in population	264	259
ER	1.16	1.00

6.3.6 Net of Free Ridership Ratios

The methodology used to develop the individual, site-specific net-of-free-ridership estimates is summarized in Section 6.2. Here, we present the weighted results both for each sampling domain. To produce an estimate of net-of-free-ridership, the individual net-of-free-ridership ratios for each of the applications in the sample were weighted by the size of the impacts associated with the application and the proportion of the total sampling domain impacts represented by each sampling stratum. Results are presented for both the SCE2509 Industrial program and the SCE2510 Agricultural program, since a sufficient number of interviews was completed for each to provide for robust program-level results.

6.3.7 Overall Results

Table 6-13 and Table 6-14 present statistics for the population and Net-to-Gross sample completes used to develop the final weighted results for each sampling domain. Note that the Net-to-Gross sample is larger than the gross sample; in addition to gross sampled sites, it also includes a number of 'net-only' sites. For both sampling domains, a large number of surveys were completed, representing significant percentages of the total population and providing for robust results.

Table 6-13: PY2006-2008 Net-of-Free-Ridership Evaluation Sample – Tracking System Savings by Gross Impact Weighting Stratum: SCE2509 Industrial Projects

Sampling Strata	Number of Records		Gross Ex Ante kWh		Gross Ex Ante kW	
	Population	Sample	Population	Sample	Population	Sample
1	8	7	35,070,679	32,344,093	3,912	3,601
2	10	9	17,865,601	16,072,180	1,853	1,853
3	13	11	14,405,448	12,207,989	1,745	1,493
4	46	6	29,308,941	3,966,601	3,144	457
5	187	10	32,815,475	1,833,130	4,277	217
All	264	43	129,466,144	66,423,994	14,931	7,621

Table 6-14: PY2006-2008 Net-of-Free-Ridership Evaluation Sample – Tracking System Savings by Gross Impact Weighting Stratum: SCE2510 Agricultural Projects

Sampling Strata	Number of Records		Gross Ex Ante kWh		Gross Ex Ante kW	
	Population	Sample	Population	Sample	Population	Sample
1	17	9	12,798,362	6,128,882	1,669	494
2	21	4	7,248,408	1,227,865	972	79
3	87	14	16,492,260	2,757,189	2,584	335
4	136	6	12,647,153	581,466	3,145	90
5	867	6	15,352,880	134,200	5,101	33
All	1,128	39	64,539,063	10,829,603	13,471	1,030

* Claimed results exclusive of the 5 Strip Curtain records that were included in the Commercial Facilities HIM evaluation.

Applying the same ratio estimation weighting approach referenced in the realization rate discussion in the Methods section, the resulting weighted net-of-free-ridership estimate for kWh savings is 0.63 for the SCE2509 Industrial Program and 0.59 for the SCE2510 Agricultural Program. (Corresponding values for kW for each program are 0.65 and 0.63, respectively.) The kWh and kW values are somewhat higher than the estimate of net-of-free-ridership for the statewide Standard Performance Contracting (SPC) program in the PY2004-2005 evaluation. (This may reflect the more limited scope of this contract group evaluation, which does not include the SPC program.) In addition, they are somewhat higher than the NTGR estimates made in prior SPC evaluations conducted for each program year since the program's inception in 1998. Table 6-15 and Table 6-16 summarize the net-of-free-ridership values by stratum, along with the 90 percent confidence interval, overall, and for each sampling domain.

Table 6-15: PY2006-2008 Net-of-Free-Ridership Ratio: SCE2509 Industrial Projects

Sampling Strata	NTGR*	
	kWh	kW
1	0.77	0.81
2	0.60	0.60
3	0.56	0.57
4	0.57	0.57
5	0.59	0.61
Weighted NTGR	0.63	0.65
90 Percent Confidence Interval	0.595 to 0.671	0.612 to 0.685
Relative Precision	0.059	0.057
N measures in sample	43	43
N measures in population	264	255
Error Ratio	0.26	0.25

* Consistent with current CPUC policy, the Net-to-Gross ratios in this evaluation reflect the effect of free ridership only and exclude any consideration of spillover.

Table 6-16: PY2006-2008 Net-To-Gross Ratio: SCE2510 Agricultural Projects

Sampling Strata	NTGR*	
	kWh	kW
1	0.46	0.56
2	0.49	0.62
3	0.59	0.66
4	0.65	0.58
5	0.71	0.67
Weighted NTGR	0.59	0.63
90 Percent CI	0.521 to 0.667	0.56 to 0.7
Relative Precision	0.123	0.111
N measures in sample	39	36
N measures in population	1,128	1,054
ER	0.47	0.41

* Consistent with current CPUC policy, the Net-to-Gross ratios in this evaluation reflect the effect of free ridership only and exclude any consideration of spillover.

* Claimed results exclusive of the 5 Strip Curtain records that were included in the Commercial Facilities HIM evaluation.

Both sampling domains cover a broad range of custom energy efficiency measures, and results vary widely as a function of the project size, measure type, rationale for the project and economic condition of the company, among other factors. In general, for the SCE2509 Industrial sampling domain, NTGR values are highest for the largest Tier 1 projects, while the small and medium-sized projects exhibit somewhat lower values, similar in magnitude to the values seen over the long-term in prior SPC evaluations. NTGR values for the SCE2510

Agricultural sampling domain exhibit the opposite pattern and steadily increase as the project size declines.

6.3.8 Program-Level NTG Findings

SCE2509 Industrial

The Net-to-Gross surveys of customers in the SCE2509 Industrial Program revealed a moderate level of free ridership (and related moderate program influence) among those customers who installed measures through the program. The average NTG ratios for installed measures were 0.63 (kWh) and 0.65 (kW).

To assess free ridership, these customers were asked to rate the importance, using a 1 to 10 importance rating scale, of a wide range of factors, covering both program and non-program elements.

The findings indicate that both utility/program and non-program factors were important elements in their installation decision. The financial aspects of the project (i.e., payback on the investment and the program rebate) were both rated highly. With respect to utility/program influences, the program rebate and verbal information provided by the SCE account rep were considered the most important. Other factors considered important were their previous experience with SCE's program, and the age and condition of the equipment that was replaced.

Across all customers surveyed, the specific ratings given were as follows:

- **The payback on the investment**, which received an average importance rating of 8.9 out of 10.
- **Previous experience with SCE's program**, which received an average rating of 7.6.
- **The availability of the program rebate,** which was rated a 7.5.
- **Previous experience with installed measure,** which also received a 7.5 rating.
- **Endorsement or recommendation by the SCE Account Rep,** given an average rating of 7.3.
- **The age or condition of the old equipment** given an average rating of 6.7.

Participants were also asked to score the relative importance of the program versus non-program factors in their installation decision, and were given a total of 10 points to split between these two elements. Program factors, were considered somewhat more important than non-program factors as indicated by average importance ratings of 5.7 for the former, and 4.3 for the latter.

Finally, respondents were asked to rate the likelihood of installing the exact same equipment at the same time absent the program. This ‘no program’ rating averaged 4.7 on a 0 to 10 likelihood scale. One interpretation of this rating is that just under half of the projects would have likely gone forward on their own, without the information and technical assistance provided through the audit program.

SCE2510 Agricultural

The Net-to-Gross surveys of customers in the SCE2510 Agricultural Program likewise revealed a moderate level of free ridership (and related moderate program influence) among those customers who installed measures through the program. The average NTG ratios for installed measures were 0.59 (kWh) and 0.63 (kW).

To assess free ridership, these customers were asked to rate the importance, using a 1 to 10 importance rating scale, of a wide range of factors, covering both program and non-program elements.

The findings indicate that both utility/program and non-program factors were important elements in their installation decision. The financial aspects of the project (i.e., payback on the investment and the program rebate) received the highest scores. With respect to utility/program influences, the program rebate and verbal information provided by the SCE account rep were considered the most important. Other factors considered important were their previous experience with SCE’s program, and standard practice in their industry. Across all customers surveyed, the specific ratings given were as follows:

- **The payback on the investment**, which received an average importance rating of 8.2 out of 10.
- **The availability of the program rebate,** which was rated a 7.9.
- **The age or condition of the old equipment** given an average rating of 7.8.
- **Endorsement or recommendation by the SCE Account Rep,** given an average rating of 7.6.
- **Previous experience with SCE’s Agricultural Energy Efficiency program**, which received an average rating of 7.4.
- **Standard practice in your industry,** which received an average rating of 7.1.

Participants were also asked to score the relative importance of the program versus non-program factors in their installation decision, and were given a total of 10 points to split between these two elements. Program factors, were considered more important than non-program factors as indicated by average importance ratings of 6.3 for the former, and 3.7 for the latter.

Finally, respondents were asked to rate the likelihood of installing the exact same equipment at the same time absent the program. This 'no program' rating averaged 6.7 on a 0 to 10 likelihood scale. This suggests that roughly two-thirds of the projects would have likely gone forward on their own, without the information and technical assistance provided through the audit program.

6.3.9 Spillover Results

In accordance with CPUC policy rules, spillover is not considered in the calculation of the NTGR values for PY2006-2008 programs. However, evaluators were directed to research and report on spillover as part of the NTGR research. Research was to be performed on only those projects (1) for which significant program influence was reported, and (2) which had sizable impacts.

Itron routinely administered the standard battery of spillover questions as a part of the NTGR interview process. However, there were no cases of spillover reported that met both of the above conditions. Therefore, there are no spillover results to report. However, Itron will be conducting a separate analysis, outside of this report, to document the percentage of projects in which spillover projects were reported, and will include descriptions of the types of projects undertaken without rebates, and the reasons why rebates were not claimed (if applicable).

6.3.10 Net First-Year Realization Rates

Table 6-17 below presents a comparison, for the SCE2509 Industrial Program, of the evaluation verified net savings with the final program claimed net savings, as obtained from the final PY2006-2008 reports posted on EEGA. These net realization rates are obtained by combining the net-of-free-ridership and gross impact realization rates to produce estimates of net realization rates. The table also includes a calculation of the verified net savings as a percentage of the claimed net savings.

Table 6-17: Comparison of First-Year Evaluation-Based Net Savings with the Final Program-Claimed Net Savings: SCE2509 Industrial Projects

	Electric Savings	
	kWh/year	Avg. peak kW
Tracking		
a. Claimed Gross Savings	145,467,578	16,776
b. Claimed Realization Rate	0.89	0.89
c. Claimed Adjusted Gross Savings (c = a x b)	129,466,144	14,931
d. Claimed NTG Ratio	0.80	0.80
e. Claimed Net Savings (d = c x d)	103,572,915	11,945
Evaluation		
f. Evaluation Gross Realization Rate	0.72	0.65
g. Evaluated Gross Results (g = c x f)	93,438,719	9,710
h. Evaluation NTG Ratio*	0.63	0.65
i. Evaluated Net Results (i = g x h)	59,149,486	6,299
j. Evaluation Net Realization Rate (h = d x f)	0.46	0.42
k. Evaluated Net Savings as a Fraction of Claimed Net Savings (k = i / e)	**0.57**	**0.53**

* Consistent with current CPUC policy, the Net-to-Gross ratios in this evaluation reflect the effect of free ridership only and exclude any consideration of spillover.

As the table shows, the evaluated gross realization rate is 72% for program SCE2509. (An evaluated gross realization rate of 100% would indicate evaluated gross savings which are identical to claimed gross savings.) Another finding is that there are fairly significant differences between claimed and evaluated NTGRs. The primary source of these differences are claimed NTGRs which default to 'old' DEER values of 0.80 (retrofit) and 0.94 (new construction, and are substantially above evaluated values that average 0.63 for the SCE2509 Industrial Program and 0.59 for the SCE2510 Agricultural Program. Finally, evaluated net savings as a percentage of program claimed net savings are in the range from 53% (kW) to 57% (kWh). These values indicate that verified net program savings are on the order of one-half (for SCE2509) of claimed savings, far below program savings estimates. The specific reasons for these low realization rates are discussed in more detail in Section 6.4, Findings.

6.3.11 Project Lifetime Impact Reporting

All of the realization rates presented in the above tables represent first-year results only. The ex-ante and ex-post evaluation kWh and kW projections for the 31 on-site sample points were also used to derive a unit energy consumption (UEC) shape over the lifetime of each evaluated project which takes into account the evaluation team's estimate of savings over the effective useful life (EUL) of the measures.

The gross ex-ante, gross ex-post and net ex-post UEC shapes for the SCE2509 Industrial program are presented in Table 6-18. The net ex-post shape incorporates both the engineering realization rates and the net-to-gross ratios.

Table 6-18: Lifetime Unit Energy Consumption Shapes: SCE2509 Industrial

Year	Calendar Year	Gross ExAnte Claimed kWh UEC	Gross ExPost Verified kWh UEC	Net ExPost Verified kWh UEC	Gross ExAnte Claimed kW UEC	Gross ExPost Verified kW UEC	Net ExPost Verified kW UEC
1	2006	1.000	0.722	0.457	1.000	0.650	0.422
2	2007	1.000	0.690	0.437	1.000	0.617	0.400
3	2008	1.000	0.690	0.437	1.000	0.617	0.400
4	2009	1.000	0.679	0.430	1.000	0.617	0.400
5	2010	1.000	0.623	0.394	1.000	0.548	0.356
6	2011	1.000	0.614	0.388	1.000	0.548	0.355
7	2012	1.000	0.600	0.380	1.000	0.536	0.348
8	2013	1.000	0.589	0.373	1.000	0.536	0.348
9	2014	1.000	0.600	0.380	1.000	0.536	0.348
10	2015	1.000	0.600	0.380	1.000	0.536	0.348
11	2016	0.973	0.600	0.380	0.972	0.536	0.348
12	2017	0.927	0.546	0.346	0.937	0.510	0.331
13	2018	0.927	0.557	0.353	0.937	0.510	0.331
14	2019	0.927	0.557	0.353	0.937	0.510	0.331
15	2020	0.927	0.557	0.353	0.937	0.510	0.331
16	2021	0.000	0.000	0.000	0.000	0.000	0.000
17	2022	0.000	0.000	0.000	0.000	0.000	0.000
18	2023	0.000	0.000	0.000	0.000	0.000	0.000
19	2024	0.000	0.000	0.000	0.000	0.000	0.000
20	2025	0.000	0.000	0.000	0.000	0.000	0.000

6.4 Discussion of Findings and Recommendations

In this section we discuss key findings from this evaluation and provide associated recommendations. We begin with an overall summary and then discuss findings and recommendations related to specific areas such as baseline specification, program influence, ex ante documentation, tracking systems, pump-off controllers, SMART wells, and the evaluation itself.

6.4.1 The Evaluation-based Estimates of Overall Program Savings Realized are Significantly Below Those Estimated by SCE

As shown in the Results section of this report, the overall net realization rate for the SCE2509 industrial program covered in the scope of this CPUC evaluation contract group (see Section 4) is 0.46, with a gross realization rate of 0.72 (kWh) and net-of-free-ridership ratio of 0.63 (kWh). Per-kW realization rates are slightly lower, 0.42 overall with 0.65 for gross and 0.65 for net. These quantitative results indicate that the program is significantly overestimating their savings claims. In addition, the results for the 2006-2008 program cycle show little to no improvement as compared to the historic results for industrial sector programs and may, in fact, be worse.

There are a number of specific findings that help to explain why the ex-post savings estimates are significantly below the ex-ante. A list of these related key findings is provided below along with specific examples of where these particular problems were observed in evaluated projects.

6.4.2 Overall Improvement Needed for the Industrial Program Realization Rates

Previous evaluations have identified many of the same issues that are identified in this evaluation yet these issues and the effects they have on overall program gross and net impact results have not yet been fully addressed. It is recommended that greater efforts are put forth to carefully review evaluation findings and recommendations and to apply corrective actions within the programs that address shortcomings in the accuracy of ex ante methods and results. The question remains regarding how to increase the effectiveness of industrial efficiency programs given the history of the programs and the challenges that the sector presents. We note that, despite these challenges and results, the industrial sector remains an important area for achieving cost effective and significant energy efficiency reductions above those that would otherwise occur due to natural market forces; in additions, programs may remain cost effective even with mediocre realization rates due to the size of the savings opportunities as compared with other sectors.

6.4.3 Problems with Ex Ante Baseline Selection or Modeling of Baseline Parameters

Baseline equipment was incorrectly selected for ex ante analysis in several of the site-specific gross impact (M&V) sample points. In a motor retrofit project it was determined that the existing motor it replaced was near the end of its effective useful life, and so the ex post evaluation selected a standard motor efficiency to represent baseline (after the estimated remaining useful life) rather than the in-situ system, as was the case for the ex ante analysis of impact. These program claims are inconsistent with most of the industrial programs' procedure manual references to "industry standard practice" as the baseline from which savings and incentives are to be estimated. Consequently, in this evaluation we used current industry standard practice to estimate gross savings for applications in which there was strong evidence for use of a replace on burnout or natural turnover baseline (increased and improved use of industry standard practice baselines are discussed further in Section 3). In the motor retrofit noted above, the resulting gross savings were zero after the remaining useful life due to the lack of any alternative to the project implemented by the customer (and the lack of any associated program effect).
Example Site – C026

As in prior SPC evaluations, a number of cases were identified where the assumptions for the program baseline calculations were unverified and undocumented, and ultimately proved to be inappropriate estimates based on ex post measurement and documentation. Increased documentation of input assumptions for savings estimation is needed, particularly, for larger and more complex sites. For example, energy savings calculations for many refrigeration and compressed air projects are based on actual data. These data can be used to calibrate the ex ante baseline models and inputs. However, for several complex projects in the M&V

sample, baseline input assumptions and parameters proved to be relatively poor. The resulting evaluation-based realization rates for these projects were found to be highly variable, suggesting that the underlying ex ante impact estimates are highly uncertain due to the baseline system modeling performed.

It was found that the ex post SMART well technology impact estimates were deeply affected by the use of speculative ex ante baseline conditions -- specifically the expected oil-to-liquid volume ratio. This ex ante baseline assumption in particular affected the ex post estimates because that baseline condition, as specified in the applications themselves, was also applied to derive the ex post savings estimates. There was no other reliable estimate available of the expected baseline well conditions. Geologists provide this service, but the evaluation findings suggest that the observed post-installation well conditions and assumed ex ante baseline conditions are not always aligned, and in fact far apart in some instances.
Example Sites – C017, C018, C019, C020

Similar baseline assumptions were found to affect a VFD air compressor, where the ex ante baseline assumed that the high pressure system also supplied low pressure air requirements. This was found not to be the case.
Example Sites – C015

Recommendation: Put Measures such as SMART wells, with an Inadequate Empirical Basis for Savings Estimates, in the Emerging Technologies Program

We found that SMART wells are a very new technology, lacking empirical data or a compelling engineering theory upon which to base and defend savings estimates. By their nature, these wells and possibly other technologies also present evaluation challenges given the lack of scientific literature and sometimes small samples available. The CPUC and IOUs should develop more explicit criteria for determining whether new measures are included under resource programs or the emerging technologies program. Measures with highly uncertain savings in need of detailed research to establish validity, expected savings, and repeatable algorithms and measurement protocols should be included in emerging technologies.

Recommendation: Improve Baseline Specification.

As discussed above, the selection of appropriate baseline has not been properly addressed by the program (or evaluations for that matter) for many years. Corrective actions that are recommended to address these problems include the following:

a) End the practice of using in situ baselines over the EUL of the measure as the default baseline for estimating savings and paying incentives.
b) Identify projects explicitly in program files as replace-on-burnout, natural turnover,[97] or early replacement.

[97] For the purposes of establishing a baseline for gross savings, natural turnover is replacement of equipment or major renovation that occurs for reasons other than equipment burnout (e.g., environmental compliance, technological obsolescence, management of production risk due to failure of aging equipment, expansion of

c) For the replace-on-burnout and natural turnover cases, which are likely to be the majority if identified properly, baselines should be set based on the efficiency of alternative new equipment, not the existing in situ equipment (see related discussion on defining standard practice below).

d) In the case of early replacement, if a claim is going to made for program-induced early replaced of functioning equipment, and if permitted under CPUC policy, claims should include evidence and documentation of the remaining useful life (RUL) of the equipment replaced, the estimated time at which the equipment would have been replaced in the future, and the effect of the program in accelerating early replacement. In situ baselines should only be claimed in such situations for the period equal to the RUL of the existing equipment; new equipment should then form the basis for the baseline for the remaining portion of the claimed savings.

Recommendation: Clarify and Enforce the Definition of "industry standard practice".

Several savings claims were inconsistent with procedure manual references to "industry standard practice" as the baseline from which savings and incentives are to be estimated. Unfortunately, there is little precision and documentation associated with this term despite its wide use throughout the industry in custom efficiency programs. As a result, different engineers, analysts, program managers, evaluators, and evaluation managers often come to different conclusions on what specific equipment and efficiency levels reflect standard practice. The lack of precision problem is exacerbated in the industrial sector due to the lack of mass market data that can be used to quantify efficiency levels and associated market shares.

It is strongly recommended that, for the next EE program cycle (2010-2012), the CPUC and IOUs should ensure that program and policy references to "industry standard practice" are more precisely defined with respect to program participation, incentive level payments, gross versus net savings attribution, and energy efficiency goal attainment. For example, the CPUC and IOUs could define "industry standard practice" more explicitly with respect to whether it is intended to reflect the mean or typical efficiency, the minimum efficiency available, the efficiency level that a certain percentage of the market exceeds (e.g., 75%) or some other explicit criteria. We recommend that standard practice be defined somewhere in between the market minimum and the mean efficiency of the market. Setting the baseline efficiency higher will generally result in an improvement in program-induced savings and higher net and gross realization rates in future evaluations; however, setting the baseline definition too high could lead to lost opportunities.

6.4.4 Some Measures Were Not Operational

Some of the installed measures had already been disabled at the time the evaluation was conducted. There were a couple examples of this in the M&V sample. In two cases, POC

capacity for current or future needs, etc.) but not for the *primary* purpose of achieving energy efficiency savings.

controlled oil well pumps were converted to either injection wells or submersible pump systems. In another case environmental issues had led to the shutdown of a ventilation fan. Example Sites – C027, C028, C012

Recommendation: Empirically Study the Effective Useful Life of Measures in an Industrial Setting

Due to the uncertain nature of measure persistence associated with equipment operation in an industrial setting, it is recommended that empirically based persistence studies be completed to assess the appropriateness of EUL assumptions used by the programs and evaluators. Measures installed in PY2006-2008 could form the basis for this empirical persistence study.

6.4.5 Ex Post Measurement Uncovers Significant Problems with Ex Ante Calculation-Based Models and Model Inputs

Ex post measurement uncovered an array of ex ante program calculation methods, assumptions and inputs that were not validated. This was the most frequently identified reason for substantial differences between ex post and ex ante estimates of savings. For example, energy savings calculations for many applications incorporated unverified operating hours, motor load factors, other system loads, production levels, operating profiles, efficiency levels and so forth. However, a number of these types of complex projects base such assumed loads, hours of operation or input estimates on limited measurement. Program savings estimates for such projects are thus based on unverified assumptions that can vary widely from site to site. Sometimes in addition to there being no measured data to back up the ex ante models, there also may not be any documentation of how the load varies throughout the year.
Example Sites – C023, C045, C003, C006, C014, C022, C030, C004, C009, C021, C029, C011, C016, C025

Recommendation: Incorporate Greater Levels of Real-Time Measurement and Pre- and Post-Installations Measurement Based Verification

Particularly for projects that are larger and have more uncertain savings, incorporate greater levels of either within program measurement or evaluation-based measurement conducted in parallel with program measure installation. Pre- and post-installation measurement in particular can be useful in not only establishing more robust ex ante savings estimates but can also be incorporated when available by the relevant evaluation team. It is anticipated that these efforts would close the substantial gap that exists between current ex ante models and results and those based on evaluation M&V in this report.

Recommendation: Require a Greater Level of Technical Documentation for the Largest and Most Complex Projects

Consider increases in the level of technical documentation required for the largest, most complex projects. There is a balance between keeping the application process and forms from being overly complex and costly to navigate, while at the same time providing adequate levels of documentation for verification and savings analyses. Application documentation should not be over-simplified given the complexity of measures and range of site-specific characteristics in this program.

Recommendation: Require Better Documentation of Pre-Installation Operating Conditions

Better documentation may also be needed regarding pre-installation or pre-retrofit operating conditions. In particular, large complex projects might be required to submit a greater level of site-specific application data than smaller projects, since (a) they contribute disproportionately to total program savings; (b) the large incentive payments increase the temptation for gaming or fraud; (c) measures implemented are often site-specific or industry-specific, and (d) savings may be very sensitive to baseline conditions.

Recommendation: Aggregate and Approve Fuel Switching and Distributed Generation-Related Projects in One or More Explicit Programs or Clearly Identified Program Elements

If the CPUC approves use of fuel switching, it should require all applications to follow the three-prong test set forth in the CPUC Policy Manual[98] and any other CPUC or other regulatory agency requirements (e.g., those related to GHG reduction goals).

6.4.6 Re-Examine Underperforming Measures

Oil well field measures underperformed relative to the ex ante claim in the majority of the selected M&V sites evaluated. This includes both POCs and SMART wells. For SMART wells it was already suggested that further evaluation of that technology be addressed by the emerging technology program. ***For POCs it is recommended that the substantial evaluation resource from 2006-2008 be used to update ex ante advice filing-based and DEER-based estimates of savings.*** This might yield an enhanced empirical method to derive mean impacts for POCs.

6.4.7 Screening of Measures Included in the Program

It appears that SCE is doing a good job of screening viable custom energy efficiency projects for incorporation in the program. Some common custom program issues were not encountered. The project documentation presented a reasonably clear description of how a given project saves energy and the energy efficiency measures included in the program all appear to have a reasonable basis for claiming energy savings. Ex ante savings estimates were somewhat technically accurate, although, as noted above, some equations and data inputs applied were not well supported or sourced. The baseline condition selected for the impact calculations were also most often reasonable. SMART wells are noted as an

[98] See CPUC Decision 92-10-020, Conclusion of Law 5. The Three Prong Test requires that any fuel switching measures: (1) not increase source-BTU consumption; (2) have a TRC benefit-cost ratio of 1.0 or greater; and (3) not adversely affect the environment. The Three Prong Test does assess total fuel input, in addition to determining if the switch is cost-effective.

exception for the appropriateness of ex ante baseline applied. No apparent project fraud or thoroughly unreasonable impact claims were encountered.

6.4.8 Measures Installed and Operational

With several exceptions noted above, the M&V sample measures were generally verified to be installed and operational.

6.4.9 Program Tracking of Project Status and Progress Towards the Accomplishment of Goals

Since the evaluation undertaken involved real-time sampling of projects under development, in order to support pre-installation M&V objectives, it was necessary to receive regular updates of "pipeline" tracking extracts that identify projects under development, their stage of development and likelihood of completion. The SCE2509 tracking extracts provided for this purpose were found to be of the highest quality for supporting this evaluation objective. The status of each project was immediately clear using this system and proved a valuable tool for real-time evaluation sampling. In addition, tracking of program accomplishments versus goal attainment was clearly and completely evident. ***It is recommended that other programs of a similar type consider and/or adopt this same tracking structure and format.***

6.4.10 Quality of Tracking System Extracts

Measure description information is populated in the tracking system but there is room for improvement in consistently labeling individual measures. Currently applications involving more than one measure appear as a single record, and therefore the measure descriptions tend towards a mixture of rough information concerning the measures installed. ***Tracking system modifications should be considered that would isolate individual records for each measure installed and achieve greater levels of consistency in reporting variables that describe measures and end-uses affected.*** With these improvements in place it would be possible to provide measure-based summary statistics as they relate to program accomplishments. Given current measure labeling practices such evaluation efforts were not deemed reasonable to produce.

To support the impact evaluation, the evaluation team requested application-level information for the M&V sample. This involves a formal request for that information, sometimes extended waiting to receive the request and sometimes multiple iterations to complete each request.

Custom applications often address some fraction of an end-use system at a facility, for example, replacing 30 HID's in a manufacturing facility that has 100 HID's. Such retrofits are often completed in phases. ***It is important that the application paperwork identify what fraction of a particular end-use is retrofit, and program policy should require clear identification of the units involved in each phase.*** Measure installation verification and other evaluation activities are hampered by poor delineation of what particular units are involved in a given project.

The tracking extract for SCE2510 did not incorporate a gross ex ante realization rate adjustment of 0.89 (for some records 1.0 is applied) until Q4 2008. The evaluation was affected by this late incorporation of the realization rate, as the definition of the ex ante claim was then changed, and no longer matched prior extracts nor paperwork. This presented difficulties due to the mid-stream nature of this adjustment and the real-time nature of sampling applied.

Recommendation: Enhance Tracking Systems to Support Measure-Level Tracking

Consideration should be given to enhancing the tracking system to ensure measure-level tracking, with use of common measure descriptions and "reporting" across projects. This might include tracking the relevant size, quantity and efficiency of each item-level measure installation, including the appropriate units. (For example, measure = chiller replacement, number of units = 2, total capacity = 600, units of capacity = rated cooling tons, efficiency = 0.60, efficiency units = kW/ton, and detailed measure type = rotary screw water-source chiller replacement.) Currently the tracking system often lists multiple measures under a single line item, and disaggregation for reporting is either very difficult or not feasible. Working towards a tracking system model that is closer to a prescriptive program model would enhance reporting of measure installations, both within the program and by evaluations.

Recommendation: Make Application-Level Information Readily Accessible to Evaluators

Consideration should be given to making application-level information readily accessible to evaluators through electronic storage of all application files and possibly a retrievable on-line system made available to evaluators. Such a system might provide easy viewing access to the project tracking data plus downloading rights to project documentation in electronic format for each project. This documentation and storage and retrieval system would greatly facilitate the evaluation, while removing a step that commonly impedes evaluation progress: data requests. This level of access and documentation would represent best practice in this area for a custom program.

Recommendation: Tracking System Ex Ante Impacts Adjustments Should be Applied at the Beginning of the Program Cycle

The ex ante impacts stored in the SCE tracking system should not require mid-program cycle adjustments. The application of these adjustment factors were required on the part of the evaluators to match E3 claims, but should have been included in the tracking system format from the start of the program cycle. Extracts are provided to evaluators multiple times in a given program cycle. Such extracts should include neither changes to the impacts achieved during the program cycle, nor structural changes to the database itself.

6.4.11 The Evaluation-based Estimates of Overall Net Program Savings Realized are Below Those Estimated by the Utilities

As shown in Section 3 of this report, the overall net kWh realization rate for program SCE2509 is 0.57. This quantitative result indicates that the Industrial program savings claims are significantly overestimated.

6.4.12 The Programs Need to Improve the Capability of Implementation Staff to Materially Influence Advanced Industrial Efficiency Improvements

Influencing large industrial customers to implement energy efficiency projects that go beyond their normal practices and plans is extremely difficult in practice. To move these customers further along the efficiency spectrum takes time and advanced levels of technical expertise, often requiring expertise in specific production practices and options. In addition, even with the right level of expertise on hand, increasing program influence requires providing advanced energy efficiency options directly to end users at the earliest stages of their equipment or facility modification decision making. There is already significant expertise available at the utility and third-party contractors. This expertise should be built upon and further increased. Development of the depth of technical expertise required to increase the net effects of the programs is a long term endeavor that requires both utility and regulatory support. Industrial technical experts need to know that there will be consistent support over time for efficiency programs if they choose to invest significant portions of their careers in program implementation. End users need to be confident that the suggestions of program staff will work to achieve the targeted savings while also meeting their various production and business requirements.

A related recommendation is to improve the training of Program Staff to enhance their capability to review submitted projects for compliance with program objectives, rules and policies. Training should be provided to address proper baseline specification, enforcement of program and policy rules, reasonableness of claims, and increasing program influence on end user's efficiency-related decisions.

6.4.13 The Programs Need to Enhance their Capability to Get Involved with Projects at the Earliest Possible Stage

Program involvement after the decision to install energy efficient equipment had been made was seen in several projects and is obviously problematic. Program involvement at an early stage to identify large equipment and facility changes helps ensure efficiency opportunities are appropriately considered and maximizes the chance of program influence. Utilization of sales or related tracking systems helps prevent projects from becoming lost opportunities.

Programs should actively work with customers to identify energy efficiency projects (and thus gain customer perceived credit for those efforts) and conversely be cautious of projects that are far along in conception or implementation when the customer learns about available rebates.

6.4.14 The Programs Need to Continue to Build Upon Market-Driven Efficiency

In some cases, high free ridership can be viewed as a positive indicator of strong market driven efficiency. A challenge for the programs is to influence these customers to go even further in their efficiency plans than they would otherwise due to their own internal policies and financial criteria. In one sense, this means setting baselines higher – which can be accomplished by using industry standard practice rather than in situ practice – as the basis for program participation and incentives. It can also mean developing customer specific baselines based on the plans the customer had at the initial point of program interaction.

6.4.15 The Programs Need to Provide Early Project NTG and Baseline Screening for the Largest Projects

The CPUC should strongly consider using early project NTG and baseline screening prior to the incentive being approved for the largest projects and those with significant policy issues such as fuel switching, self generation, and greenhouse gas impacts. Such screening for the largest projects, whereby the baseline claim is reviewed and NTG interviews are conducted just after the implementation decision is made, would help to obtain critical information regarding program influence that may lead to the project being re-defined or dropped.

6.4.16 Other Net-to-Gross Recommendations

The following are overarching free-ridership-related recommendations from previous SPC program evaluations that continue to be relevant:

Recommendation: Consider Limiting or Excluding Incentive Payments to Known Free Riders

When program administrators are incented and permitted to simply exclude known free riders, scarce program funds can instead be utilized on projects that provide net benefits.

Recommendation: Consider Using Incremental Costs to Benchmark and Limit Payments

Limiting payments so that they do not exceed a pre-determined portion of average or customer-specific incremental cost estimates is critical to avoiding grossly overpaying for savings.

Recommendation: Consider Incorporating a Payback Floor

The use of a payback floor (minimum payback level based on energy savings alone) helps to ensure that project generates meaningful and significant energy savings. With a payback floor, the program avoids incenting projects that are primarily being done for reasons other than energy savings (modernization, production efficiency, environmental compliance, etc.)

Recommendation: Set Incentive Levels to Maximize Net, not Gross Program Impacts

Free riders dilute the market impact of program dollars. Payback floors and increasing incentives with increasing payback levels are one approach. Another is to tie incentive levels to individual measures or types of measures that are known to have extremely high or low naturally occurring adoption levels.

Recommendation: Consider Tying Staff Performance to Independently Verified Net Results

Tying performance reviews and bonuses of program staff to verified savings as reported through an independent M&V or impact evaluation process is likely to increase project quality and the accuracy of initial savings estimates. Marketing staff, in particular, should have any financial incentives tied to savings that are independently verified.

6.4.17 Evaluation-Related Findings and Recommendations

The following are recommendations to improve evaluation usefulness and focus:

Recommendation: Involve impact evaluators in large projects and a sample of projects on a real-time basis throughout the program cycle

The timing of evaluation processes should be accelerated. Under the current approach, where evaluation activities commence after or well into the program cycle, many projects are evaluated two to three years after they are implemented. Often, the primary decision maker has left. In a few cases, companies had gone bankrupt or transferred operations to another location. Moving the evaluation process forward in time to occur just after the project is installed would effectively address this problem. This can be accomplished through earlier contracting and implementation of the evaluation, combined with improved utility tracking and early reporting of installations (as well as projects in the pipeline), more frequent sampling and evaluation of projects throughout the program plan period.

Recommendation: Evaluation participation requirements should be strengthened.

In the course of conducting the evaluation, we experienced 'pushback' from many participants who either refused to participate in evaluation surveys and on-sites or declined to provide required data and documentation. This made it difficult to conduct the evaluation efficiently and can lead to systematic bias. Requirements for participating in impact evaluations must be clearly explained, both at the time applicants are paid incentives, and later, when evaluation activities commence. Evaluation participation should be clearly written into program participation and incentive payment agreements.

Recommendation: Conduct a full complement of impact, process, and market evaluations

Due to the relatively small number of customers that participate in these programs, and due to the inherent evaluation advantages in conducting comprehensive impact, process and market evaluation activities in parallel, it is recommended that those activities be combined. Large customer programs and markets are very dynamic and require regular assessment in order for program managers and policy makers to continuously improve them. Most of the effort for the 2006-2008 industrial evaluation focused on impact evaluation, in accordance with the CPUC's evaluation priorities. Future evaluations should consider more integration of process evaluation and market assessment to capture research economies and reduce customer and vendor interview burdens.

Recommendation: Stagger the timing of process and ex post impact tasks so that process evaluations can be conducted and results communicated on a relatively real-time basis

If process and impact evaluations are more integrated in future evaluations, care must be taken to schedule activities and deliverables appropriately. Because of the sometimes long project installation lag after program commitment in these programs, it is important to schedule process evaluation tasks to be conducted during or just after each program year so that results can be utilized to improve program processes for the subsequent program year (rather than producing results only late in the three-year program cycle for use in the next program cycle).

Recommendation: Conduct baseline research to establish standard industry practices for key measures in key industries

Significant research is needed to establish meaningful and defensible data, especially market share, for establishing industry standard practices for measures that are not completely site specific. Improved information on industry standard practices can then inform decisions about which measures to include in the program (and provide incentives for), which could in turn lead to reductions in free ridership.

Recommendation: Conduct Persistence Study of Industrial Sector Savings

Few studies of the persistence of program savings in the industrial sector have been conducted, particularly within the last decade. As noted previously in this section, there were instances of non-operational equipment in the gross impact sample. In addition, in some program years and cycles industrial production levels will be higher or lower depending on economic conditions. Some facilities may also close and stay closed while others may reopen and reutilize efficiency measures. Research is needed to measure the persistence of savings over time under a range of economic conditions. Sufficient time needs to pass in order to maximize the information provided from such persistence studies. We recommend waiting until the recession is completely over and the economy is in full recovery. To accelerate the time at which meaningful results would be obtained, studies can be conducted using earlier program cohorts, for example, going back to the 2002-2003 or 2004-2005 program cycles (or earlier), rather than simple waiting for the 2006-2008 cohort to age.

Recommendation: Conduct Analysis of Incentives by Customer and Industry Type, Further Research on Use of Incentive Caps

Customer incentive caps have been utilized in various forms for many years. During times of low budgets and low goals, caps were set low to spread incentives to a broad pool of participants. More recently, as goals and budgets have significantly increased, caps have increased greatly as well. We are not aware of any systematic study of the effect of the incentives caps. Similarly, research is needed to explore how much total incentive dollars have been distributed across or concentrated within certain customers to determine whether these patterns are aligned and supportive of efficiency policy goals.