

2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group

Appendices Only

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
(510) 844-2800

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

Table of Contents

Appendix A: Sampling Approaches and Implementation

Appendix B: On-Site Data Collection Form

Appendix C: NTG Methods, Survey Instruments and Site Results

C.1 Nonresidential NTG Methods

C.2 Large Nonresidential NTG Survey Instruments

C.3 Detailed Site-Specific NTG Results

Appendix D: Site Reports – IN A SEPARATE ATTACHMENT

Appendix E: Supplementary POC Methods and Results

Appendix F: Responses to Comments Received on Draft Report

Appendix A

Additional Sampling Information

A.1 Sampling Methodology and Description

This appendix describes the proposed approach to developing and implementing sampling to support the overall evaluation for the PG&E Fabrication, Process, and Manufacturing contract group. This appendix provides additional information on sampling steps to that presented in Section 3 of the main report. There are several important questions that must be addressed in order to develop any sample design. These include the following:

- What are the sampling design variables and sampling domains?
- Which statistical method will be used to estimate sampling size and precision levels? What type of stratification, if any, will be utilized?
- What is the desired level of statistical confidence and precision? Or conversely, what statistical precision will be achieved for a desired (or affordable) number of sample points?
- What level of variance is expected for the sampling design variables?
- Each of these questions is addressed in the remainder of this section.

A.1.1 Gross Impact Sampling Design Variables and Domains

In any sample design, the first question to ask is which sampling variables and research objectives are important enough to be used in defining separate domains of study for the evaluation. As discussed here, a domain of study is a sub-population for which sufficient sample will be allocated to achieve estimates of savings with a pre-assigned precision goal, e.g., 90/10. In general, the total sample size of a study is directly proportional to the number of sampling domains, e.g., doubling the number of domains doubles the overall sample size.

The CPUC ED identified three energy metrics as being at the core of the 2006-2008 impact evaluations. These are: energy savings associated with electric energy (kWh/year), electric demand (peak kW), and natural gas energy (therms).

There are two key regulatory reasons for these design variables, both of which are tied to the utilities' risk/reward mechanism for PY2006-2008 as described in D.05-04-051. First, the reward or penalty will be based on the "performance earnings basis" or "PEB", which is simply net benefits to ratepayers (calculated as resource benefits minus costs) associated with program impacts. The PEB is calculated separately for each IOU. Second, the Decision prescribes that IOUs must achieve a minimum of 85 percent of the Commission-adopted savings goals, based on a simple average of the percentage of each individual GWh, MW, and, as applicable, million therm goal they achieve. In addition, they also must meet a minimum of 80 percent of the goal for *each* individual savings metric.

It is useful to observe that although one could theoretically develop a sample plan using the PEB itself as the design variable, which would be the most direct approach to translating evaluation results back into the earnings mechanism, there are a number of separate parameters and data types that go into the PEB calculations such as incremental costs, program costs, energy savings, load shapes, effective useful lives, and avoided costs. Because these data come from a variety of different sources and are not all available within utility tracking systems, sampling directly on PEB may not yet be practical but is something to consider for the future.

The first sampling domain that we considered was the IOU. Since the PEB is to be carried out for each IOU separately, it stands to reason that each IOU should be a separate sampling domain. The only thing of interest to note with respect to this sampling domain is that, all else being equal, if a utility is a sampling domain and the targeted precision is the same across utilities, then smaller utilities (e.g., SDG&E) would have higher sample-to-population ratios and thus higher evaluation costs to program impact ratios than larger utilities (this is not an issue for this evaluation contract group, since this evaluation addresses only PG&E; however, this is an issue for contract groups that have multiple utilities in scope).

We now discuss the sampling variables and consideration of the sampling domains such as energy savings metrics, high-impact program-measure groups, and programs.

With respect to the savings metrics, we first note that if the risk/reward mechanism was based solely on estimating the PEB, one could integrate electric and gas savings into a single design variable that is closely correlated with avoided costs, such as source Btu. However, since the reward component of the mechanism is triggered based on achieving both an average of 85% of the Commission-adopted savings goals across GWh, MW, and million therms, and a minimum of 80% for each individual savings metric, it becomes important to plan for a desired precision level for each savings metric. In the case of energy and peak demand, these metrics are for the same fuel and both occur for every electric project. In addition, there is a correlation between energy and demand savings, though the strength of

this correlation varies significantly across measure types. Overall, though, we believe that the electric energy and demand metrics should be addressed with the same sampling domain of electric energy savings projects. We could have made the design variable either kWh or kW or we could have constructed a variable that reflected the relative contribution of kWh and kW to avoided cost benefits. We utilized kWh as it is generally more robust within utility tracking systems and was consistent with the overall direction provided by the Energy Division.

With respect to therms, although there are some projects that have both electric and gas impacts, there are many projects that are either electric only or gas only. The majority of gas impacts are typically associated with gas-only measures. As a result, we needed to decide whether to make gas projects a separate sampling domain. In general, gas impacts are currently a much smaller contributor to total avoided cost benefits at the IOU level than are electric impacts. If gas only had to be represented as a share of its contribution to avoided cost benefits, it should then be allocated sample points proportional to avoided cost benefits (or source Btu as a proxy). However, due to the structure of the PEB, one could argue that gas impacts should be estimated for a targeted precision level (due to the 80% trigger) and, thus, be considered a separate sampling domain. Doing so, however, could have at times result in a significant over-sampling of gas relative to electric impacts that would produce an evaluation cost to savings impact ratio for gas that would be several fold, if not sometimes an order of magnitude higher, than for electric. Nonetheless, the structure of the PEB necessitated inclusion of gas as a sampling domain for contract groups such as this one with significant portions of a given utility's gas portfolio.

As discussed in Section 2, with respect to high-impact measure groups and individual measures, the evaluation team determined that only one individual PG&E industrial sector electric measure warranted its own sampling domain. This measure was oil well pump-off controller (POCs), which accounted for 34% of PG&E claimed electric savings for the programs in this contract group. POCs were also selected as a sampling and analysis domain because it was preferred that a consistent engineering and measurement approach be developed and implemented for estimating gross savings for this measure. It was determined that the remaining measures were too numerous and the associated savings too low at the measure level to warrant additional measure-level sampling domains. This was also because of the high cost per site of conducting industrial savings analyses.

When we combined the energy savings metrics with the high-impact measure analysis, we arrived at the three sampling domains for this evaluation: POCs, all other non-POC electric measures, and all natural gas measures.

Another possible sampling domain considered was each individual IOU program. Understanding program impacts is of interest in order to capture any significant program-induced variations in gross or net impacts across programs and draw out lessons learned on which programs are more or less effective than others and why. This is certainly of interest for large programs but can become very expensive in the industrial sector if there are many very small programs. There is a great disparity in the contribution of different programs to the total IOU-fuel impacts in many contract groups (as well as for the IOU portfolio as a whole). Defining IOU programs as domains *with equivalent precision requirements* would likely have resulted in small programs being grossly over-sampled as compared with large programs. For example, our research planning analysis showed that if IOU programs were a sampling domain for this contract group, with a targeted relative precision of 10% at the 90% confidence interval, then the smallest programs in this contract group would end up with several fold more evaluation dollars expended on them per unit of impact as compared with the largest and most important program in the contract group (PGE2004).

Another consideration in this sector was that managing sampling domains over the course of the evaluation was likely to be challenging no matter which domains were selected – because the population of projects was changing over time – but that fewer domains would be much easier to manage given the uncertainty in forecasted program goals and measure mixes.

A.1.2 Sampling and Extrapolation Methods, Confidence and Precision

The PG&E Fabrication Process and Manufacturing (“PG&E Fab”) contract group uses 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*.¹ This approach was also used to develop program realization rates for the 2002, 2003 and 2004-2005 Statewide SPC program impact evaluations.

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).

¹ 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

To more formally investigate the expected precision levels for the 2006-2008 PG&E Fab impact evaluation the precision level achieved for two relevant past evaluations was first reviewed -- for the combined 2002-2003 and 2004-2005 SPC impact evaluation samples. 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.² From these past studies, we calculated error ratios (*er*) of 0.35 to 0.45 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 and 2004-2005 SPC results, we used the case weights to calculate the stratified ratio estimator of B , denoted \hat{B} , as follows:

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

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):

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

² 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.

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

The resulting precision levels for alternative samples are shown in Figure A-1 and Figure A-2 below for the calculated *er* of 0.35 and 0.45 as well as a range of error ratios that might occur in a large (N=5,000) and small (N=100) 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 and 2004-2005 SPC evaluations because the scope of those impact efforts was much smaller than the expected M&V scope of the 2006-2008 evaluation. The more limited 2002-2003 and 2004-2005 impact scopes 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 savings estimation process. Conversely, the error ratio expected for the 2006-2007 verification sample and analysis was hypothesized to be possibly lower than 0.35, since verification rates are usually high and variation low for programs with mandatory verification included in the implementation process (as is the case with some of many of the programs in this contract group).

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

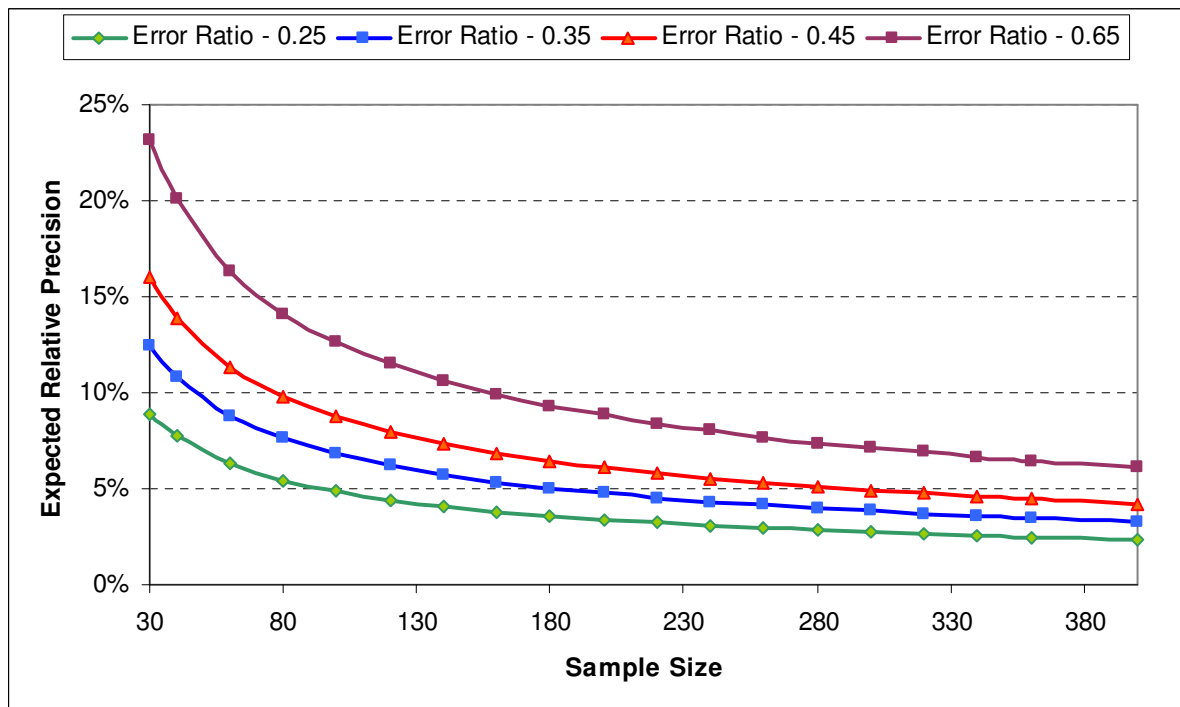
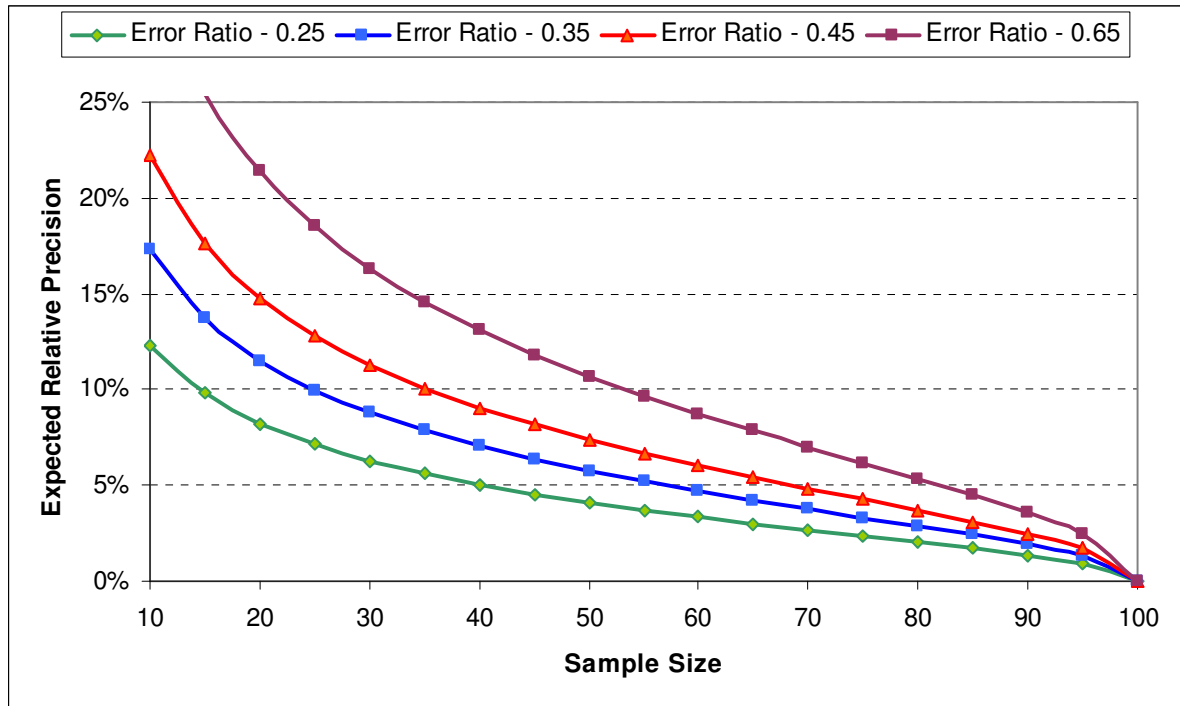


Figure A-2: Expected Relative Sampling Precision (at 90% Confidence Level) Versus Sample Size with Stratified Ratio Estimation for Varying Error Ratios and Small Population (N=100)



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.

A.1.3 2006-2007 Verification Sample Plan

The sample design for the PG&E Fab evaluation started with the verification sampling activities for 2006-2007 conducted in 2008. Because a significant portion of the evaluation work for this contract group was conducted utilizing the M&V protocols and produced site-

specific reports and realization ratios, each site required extensive engineering resources and time to complete. Consequently, we intended to nest the 2006-2007 verification sample within the 2006-2008 impact sample so that the engineering team could conduct some of the impact evaluation analysis in conjunction with completing the verification work. Thus the sampling approach was to first construct a rough sample plan for the overall evaluation, and then to allocate the 2006-2007 sample proportional to the impacts installed through Q2 2007 as compared with the programs' 2006-2008 goals.

The originally proposed sample plan for the PGE2004 impact evaluation is shown in Table A-1 below. Since the Q2, 2007 tracking extract provided by PG&E indicated that program PGE2004 was the only industrial program with significant activity, the verification sampling approach focused only this program. The sample size was drawn to provide 90/10 confidence/relative precision for electric verification and for gas verification. Taking the overall impact evaluation sample as an initial target, points were then allocated to the 2006-2007 verification sample such that the fraction of total sample points was roughly equal to the percent of impacts installed to date as a percent of goal. The 2006-2007 verification sample is shown in Table A-1. The electric 2006-2007 verification sample was composed of 30 projects and the gas sample was composed of 12 projects.

Table A-1: PG&E Fabrication, Process, and Manufacturing Estimated Reported Impacts through Q2 2007 and M&V Sample Points and Relative Precision

Utility	ProgramID	Program	NetkWh	NetkW	NetTherms	Estimated Number of Electric Measure Projects	Estimated Number of Gas Measure Projects	~# of Sample Points - Electric	~# of Sample Points - Gas	Estimated Relative Precision - Electric (w. er=0.35)	Estimated Relative Precision - Gas (er=0.35)
PG&E	PGE2004	Fabrication, Process and Heavy Industrial Manufacturing	164,935,530	26,390	12,310,200	330	31	50	20	90/8	90/9

Sample Stratification

Consistent with the Evaluation Framework Study's recommendations, we stratified our verification sample by size of savings in five strata, where stratum 1 included projects with the largest savings and stratum 5 included projects with the lowest savings. Strata boundaries were drawn so that each stratum would represent 20% of the population. Table A-2 summarizes the population tracking data by stratum using the Q2, 2007 tracking database extract, as well as the initial sample design. Table A-2 indicates the following:

- For the electric projects, the first two size strata include the largest 11 electric records, which represent 3% of total electric records and 39% of electric energy savings.

- For the gas projects, the first four strata include the largest 7 gas records, which represent 25% of total gas records and 77% of gas energy savings.

The proposed electric and gas samples were drawn randomly within each stratum. The electric sample included a census of projects in the first two strata, while the gas sample included a census of projects in the first four strata.

Table A-2: PGE2004 Energy Savings by Stratum and Verification Sample Design as of Q2, 2007

Electric sample

Strata	PGE2004 Program		Verification Sample		Verification Sample Percent	
	N measures	kWh	N measures	kWh	N measures	kWh
1	4	13,418,544	4	13,418,544	100%	100%
2	7	13,728,673	7	13,728,673	100%	100%
3	19	14,478,387	7	5,505,550	37%	38%
4	55	13,772,779	6	1,407,347	11%	10%
5	295	14,053,717	6	243,923	2%	2%
Total	380	69,452,101	30	34,304,038	8%	49%

Gas sample

Strata	PGE2004 Program		Verification Sample		Verification Sample Percent	
	N measures	Therms	N measures	Therms	N measures	Therms
1	1	1,264,741	1	1,264,741	100%	100%
2	1	1,059,000	1	1,059,000	100%	100%
3	2	1,378,722	2	1,378,722	100%	100%
4	3	1,515,022	3	1,515,022	100%	100%
5	21	1,533,766	5	687,462	24%	45%
Total	28	6,751,251	12	5,904,947	43%	87%

Final Verification Sample Design

A Q4, 2007 database extract was received after site visits were either completed or scheduled at 38 out of the 42 sample points in Table A-2. As shown in Table A-3 program activity at the end of Q4, 2007 nearly doubled with respect with electric energy savings and nearly tripled with respect to gas energy savings as compared to Q2, 2007. To capture program activity from Q3-Q4 in the 2006-2007 Verification Report, electric and gas projects were re-stratified using the Q4, 2007 database extract. The remaining untouched sample points from the original sample design (2 electric and 2 gas projects) were then replaced with 3 new electric and 2 gas projects that were installed in Q3-Q4. In an effort to capture the largest possible fraction of energy savings in Q3-Q4, the largest projects installed in Q3-Q4 were selected. The redesigned sample is presented in Table A-3.

Table A-3: PGE2004 Energy Savings by Size Strata, and Final Verification Sample Design as of Q4, 2007

Electric sample

Strata	PGE2004 Program		Verification Sample		Verification Sample Percent	
	N measures	kWh	N measures	kWh	N measures	kWh
1	5	25,532,696	5	25,532,696	100%	100%
2	13	30,780,474	7	16,065,405	54%	52%
3	26	28,696,698	6	8,054,174	23%	28%
4	76	29,409,735	7	3,046,359	9%	10%
5	500	28,620,416	6	371,758	1%	1%
Total	620	143,040,018	31	53,070,393	5%	37%

Gas sample

Strata	PGE2004 Program		Verification Sample		Verification Sample Percent	
	N measures	Therms	N measures	Therms	N measures	Therms
1	1	4,919,708	1	4,919,708	100%	100%
2	1	4,063,495	1	4,063,495	100%	100%
3	1	1,264,741	1	1,264,741	100%	100%
4	4	2,968,672	4	2,968,672	100%	100%
5	35	3,929,202	5	1,125,424	14%	29%
Total	42	17,145,818	12	14,342,040	29%	84%

It was not possible to verify one electric project and one gas project that had been included in the original verification sample. The gas sample point that could not be completed was re-allocated within the same stratum using backup sample. The electric site that could not be completed was cancelled since the redesigned electric sample had one project more than the plan (31 electric projects in Table A-3 vs. 30 electric projects in Table A-2).

During subsequent discussions with ED and its consultants, it was suggested that project stratification from Q2, 2007 be preserved. Table A-4 contains the final disposition for the 2006-2007 verification sample; strata 1-5 include projects completed through the end of Q2, 2007, while the new stratum 6 consists of the projects installed in Q3-Q4, 2007 only. The verification sample captures 37% of electric energy impacts and 84% of gas impacts for the PGE2004 Program through Q4, 2007.

Table A-4: PGE2004 Energy Savings by Size Strata, and Final Verification Sample as of Q4, 2007

Electric sample

Strata	PGE2004 Program		Verification Sample		Verification Sample Percent	
	N measures	kWh	N measures	kWh	N measures	kWh
1	4	13,418,544	4	13,418,544	100%	100%
2	7	13,728,673	7	13,728,673	100%	100%
3	19	14,478,387	7	6,503,371	37%	45%
4	55	13,772,779	4	885,130	7%	6%
5	295	14,053,717	5	219,750	2%	2%
6	240	73,587,918	3	17,810,354	1%	24%
Total	620	143,040,018	30	52,565,823	5%	37%

Gas sample

Strata	PGE2004 Program		Verification Sample		Verification Sample Percent	
	N measures	Therms	N measures	Therms	N measures	Therms
1	1	1,264,741	1	1,264,741	100%	100%
2	1	1,059,000	1	1,059,000	100%	100%
3	2	1,378,722	2	1,378,722	100%	100%
4	3	1,515,022	2	1,008,242	67%	67%
5	21	1,533,766	4	624,682	19%	41%
6	14	10,394,567	2	8,983,203	14%	86%
Total	42	17,145,818	12	14,318,590	29%	84%

A.1.4 Gross Impact Sample

Evaluation Sample Drawn in Q2, 2008

In August 2008 a High Impact (HIM) Memo Response for the PG&E Fabrication, Process and Manufacturing Evaluation Contract Group was submitted to ED. Based on an analysis of program activity through Q1, 2008, and in response to ED’s shift of emphasis from program evaluation to HIM evaluation, the memo proposed a redistribution of EM&V sampling points across the PG&E Fab programs. Using the PG&E Fab record disposition as of Q2, 2008 and the program goals filed by PG&E for the 2006-2008 program cycle, the memo proposed assigning 117 sample points (86 electric and 31 gas) to the PGE2004 program, and 33 sample points (electric) to the third party programs in the PG&E Fab Group. The following sections discuss our strategy of allocating the 117 sample points for program PGE2004.

Electric Sample

The EEGA data extract through Q2, 2007, the verification sample format, and the goals filed by PG&E for program PGE2004 were used to allocate the proposed 86 electric sample points for PGE2004 by time period and sample stratum. In particular, measure installation date was used to group the tracking data into three time periods: (1) all of 2006 and Q1-Q2, 2007, (2)

Q3-Q4, 2007, and (3) Q1-Q2, 2008. The Q2, 2008 extract indicated that program PGE2004 was on track for reaching goals, so sample provisions were made for a fourth and final time period: Q3-Q4, 2008. Similar to the verification sample design, the gross kWh savings reported in each time period were used to proportionally assign sample points to each time period. Table A-5 shows the disposition of the kWh impacts in the tracking extract, by time period. The number of sample points proposed for time period Q3-Q4, 2008 assumed that goals for Program PGE2004 would be reached by the end of calendar year 2008.

Table A-5: PGE2004 Sample Disposition by Time Period Using Program Achievements as of Q2, 2008

Time Period	PGE2004 Program			Ver. Sample	M&V Sample
	N measures	Gross kWh	Net kWh	N measures	N measures
Thru Q2, 2007	381	69,452,101	53,487,290	27	29
Q3-Q4, 2007	241	73,587,918	60,655,172	3	31
Q1-Q2, 2008	163	23,037,656	18,810,007	0	10
Program Goal			164,935,530		86
Expected Q3-Q4, 2008		39,951,664	31,983,061	0	16

The EM&V sample design then followed the methodology described above for the verification sample. For consistency with the verification work, the number of strata and the strata boundaries as defined in the verification sample design were preserved. Table A-6 shows the resulting disposition of kWh impacts by stratum and time period.

Table A-6: PGE2004 Electric Achievements as of Q2, 2008 by Time Period and Stratum

Strata	Gross kWh Strata bounds	N measures			Gross kWh		
		Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008	Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008
1	2,369,046	4	5	2	13,418,544	23,984,080	5,577,744
2	1,618,033	7	5	0	13,728,673	10,248,766	0
3	445,992	19	28	8	14,478,387	22,433,278	6,457,501
4	146,167	55	37	14	13,772,779	9,502,618	3,704,611
5	0	296	166	139	14,053,717	7,419,175	7,297,799
All	-	381	241	163	69,452,101	73,587,918	23,037,656

Table A-7 presents the EM&V sample disposition, which targets equal sample allocation across strata. If the population in a given stratum was smaller than the target sample size, the number of measures in the next stratum was increased to make up for the difference.

Table A-7: Proposed PGE2004 Electric Sample Disposition by Time Period and Stratum Using Achievements as of Q2, 2008

Strata	Gross kWh Strata bounds	N measures			
		Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008	Q3-Q4, 2007
1	2,369,046	4	5	2	
2	1,618,033	7	5	0	
3	445,992	9	7	4	
4	146,167	4	7	2	
5	0	5	7	2	
All	-	29	31	10	16

The 30 electric measures included in the verification sample were then nested into the proposed EM&V sample. Incremental EM&V sample points were drawn to supplement the 30 sample points already selected for the Verification study. To allow room for integration of pre-M&V sites, and to avoid over- or under-sampling by time period in case program PGE2004 came short of- or exceeded goals, only 25 sample points from Q3-Q4, 2007 and 5 sample points from Q1-Q2, 2008 were drawn at the time. In order to capture a maximum of kWh impacts, this limited sample included a census of measures from the two high-impact strata and a random selection of measures from the lower-impact strata, as shown in Table A-8 below.

Table A-8: Limited Incremental Electric Sample for Program PGE2004 - Disposition by Time Period and Stratum as of Q2, 2008

Strata	Gross kWh Strata bounds	N measures			
		Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008	Q3-Q4, 2007
1	2,369,046	0	2	2	
2	1,618,033	0	5	0	
3	445,992	0	6	3	
4	146,167	0	6	0	
5	0	0	6	0	
All	-	0	25	5	0

Gas Sample

The same approach was used to allocate a total of 31 gas sample points by time period and sample stratum. Measure installation date was used to group the tracking data into the same three time periods, and sample provisions were made for the Q3-Q4, 2008 time period. The gross Therms savings reported in each time period were used to assign sample points to each time period.

Table A-9 shows the disposition of the Therm impacts by time period. Since program PGE2004 had already exceeded its goals by 2.9 million Therms as of Q2, 2008, only three points were assigned to time period Q3-Q4.

Table A-9: PGE2004 Incremental Sample Disposition by Time Period Using Program Achievements as of Q2, 2008

Time Period	PGE2004 Program			Ver. Sample	M&V Sample
	N measures	Gross Therms	Net Therms	N measures	N measures
Thru Q2, 2007	33	6,751,251	5,256,571	10	10
Q3-Q4, 2007	15	10,394,567	7,438,154	2	14
Q1-Q2, 2008	6	3,242,434	2,518,570	0	4
Program Goal			12,310,200		31
Expected Q3-Q4, 2008			?	0	3

For consistency with the verification work, the number of strata and the strata boundaries defined in the verification sample design were preserved. Table A-10 shows the disposition of Therm impacts by stratum and time period.

Table A-10: PGE2004 Therm Achievements as of Q2, 2008 by Time Period and Stratum

Strata	Gross Therms Strata bounds	N measures			Gross Therms		
		Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008	Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008
1	1,264,741	1	2	1	1,264,741	8,983,203	2,179,147
2	1,059,000	1	0	0	1,059,000	0	0
3	624,204	2	0	1	1,378,722	0	1,037,385
4	477,292	3	0	0	1,515,022	0	0
5	0	26	13	4	1,533,766	1,411,364	25,902
All	-	33	15	6	6,751,251	10,394,567	3,242,434

Table A-11 presents the EM&V gas sample disposition, which targets equal allocation across strata. If the population in a given stratum is smaller than the targeted number of measures, the number of measures in the next stratum is increased.

Table A-11: Proposed PGE2004 Gas Sample Disposition by Time Period and Stratum using Achievements as of Q2, 2008

Strata	Gross Therms Strata bounds	N measures			
		Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008	Q3-Q4, 2007
1	1,264,741	1	2	1	
2	1,059,000	1	0	0	
3	624,204	2	0	1	
4	477,292	2	0	0	
5	0	4	12	2	
All	-	10	14	4	3

Similar to the electric sample, the 12 verification sample points were nested into the EM&V sample. To avoid over-sampling in the low-impact strata, only the two high-impact sample points from Q1-Q2, 2008 were drawn at the time. This allowed room for integration of any pre-M&V sites, as well as for any other adjustments that might occur before the end of calendar year 2008.

Table A-12: Limited Additional Gas Sample for program PGE2004 - Disposition by Time Period and Stratum using Program Achievements as of Q2, 2008

Strata	Gross Therms Strata bounds	N measures			
		Thru Q2, 2007	Q3-Q4, 2007	Q1-Q2, 2008	Q3-Q4, 2007
1	1,264,741	0	0	1	
2	1,059,000	0	0	0	
3	624,204	0	0	1	
4	477,292	0	0	0	
5	0	0	0	0	
All	-	0	0	2	0

Final Evaluation Sample

In March 2009, a HIM Plan Addendum for the PG&E Fabrication, Process and Manufacturing Evaluation Contract Group was submitted to ED. Based on an analysis of program data through Q3, 2008, the Addendum proposed a distribution of M&V sampling points across the PG&E Fab Programs. In particular, the memo suggested assigning 103 electric sample points and 27 gas points to projects across all PG&E Fab programs.

It is important to note that in Q4, 2008, ED made the decision that 58 new construction-related project records from program PGE2004, representing 1.15 million kWh and 97 thousand Therms gross ex ante savings, would be included in the New Construction Codes and Standards (NCCS) evaluation contract group. Thus, in what follows, the PG&E Fab

“population” refers to the PG&E Fab extract as of Q4, 2008 minus the 58 records included in the NCCS evaluation.

Electric Sample

A size-stratified sample of 30 electric points were drawn from the tracking extract for PGE2004 through Q4, 2007, to support the 2006-2007 Verification Report. Following the release of the Q2, 2008 tracking extract, an additional size-stratified sample of 30 M&V electric points was drawn from program PGE2004. The PG&E Fab third party programs had much lower activity levels at the time, and were therefore not included in the sampling effort. Thus, following the Q2, 2008 tracking extract, a size-stratified sample of 60 points had been pulled.

In addition to the verification and M&V sample designs, a list of 19 pre-install projects with a reasonable chance of being installed by the end of Q4, 2008 was established for the PG&E Fab programs. These 19 projects spanned six of the PG&E Fab programs (PGE2004, PGE2042, PGE2046, PGE2058, PGE2081 and PGE2087) and pre-installation visits were made at these sites to collect baseline data for the measures to be installed. When the Q4, 2008 program extract became available, we verified that fourteen of the 19 pre-install projects, corresponding to 16 records in the tracking extract, were marked as “installed.” These 16 records were included in the final EM&V sample for the PG&E Fab programs.

Table A-13 shows the disposition of the electric projects from the PG&E Fab tracking extract, as well as the number of projects sampled earlier in the evaluation and those proposed as incremental sample, by stratum. For consistency with the verification and EM&V work completed earlier in the evaluation, the number of strata and the strata boundaries as defined in the verification sample design were retained. We continued to use the gross kWh savings to proportionally assign sample points to each stratum (“target sample” in Table A-13). After accounting for the sample drawn previously, and for the pre-install sample, the “balance” column in Table A-13 shows the number of points that were still needed for the M&V Evaluation. The “realistic” column then re-distributed sample points so that oversampling in certain strata was incorporated into the final sample disposition.

Table A-13: Electric Sample Disposition by Stratum Using Program Achievements as of Q4, 2008

Strata	Population			Sample				
	N	gross kWh	% gross kWh	Target N	Through Q208	Pre-Install	Balance N	Realistic N
1	34	142,447,707	26%	27	11	3	13	12
2	28	55,226,868	12%	12	13	0	-1	0
3	201	169,110,120	37%	38	15	7	16	15
4	260	70,534,548	16%	16	10	5	1	0
5	889	45,255,421	10%	10	11	1	-2	0
All	1,412	482,574,664		103	60	16	27	27

The 27 projects required in the last column of Table A-13 were drawn by random sampling in strata 1 and 3, respectively. This random sample was then combined with the 16 pre-install projects to form an overall “incremental” M&V sample. Table A-14 shows a comparison between the sample already drawn through Q2, 2008 and this “incremental” M&V sample, by stratum.

Table A-14: Comparison of Existing and Proposed Incremental Electric Sample Using Program Achievements as of Q4, 2008

Strata	Sampling Through Q2, 2008				Proposed Incremental Sample			
	N pop	kWh pop	N sample	kWh sample	N pop	kWh pop	N sample	kWh sample
1	11	42,980,369	11	42,980,369	23	99,467,338	15	66,738,243
2	13	25,841,755	13	25,841,755	15	29,385,113	0	0
3	54	42,428,365	15	11,995,637	147	126,681,756	22	19,059,822
4	105	26,833,841	10	2,632,968	155	43,700,706	5	1,812,302
5	602	28,916,859	11	429,847	287	16,338,562	1	42,731
All	785	167,001,189	60	83,880,576	627	315,573,474	43	87,653,097

Table A-15 compares the program achievements through Q4, 2008 with the overall sample disposition, by stratum. The overall sample is close to the target sample shown in Table A-13, and includes nearly a census of stratum 1 (large) projects installed in the PG&E Fab programs.

Table A-15: Distribution of Electric Projects by Stratum in the PG&E Fab Population and Final Sample Using Program Achievements as of Q4, 2008

Strata	Population				Total Sample			
	N	% N	gross kWh	% gross kWh	N	% N	gross kWh	% gross kWh
1	34	2%	142,447,707	30%	26	25%	109,718,612	64%
2	28	2%	55,226,868	11%	13	13%	25,841,755	15%
3	201	14%	169,110,120	35%	37	36%	31,055,459	18%
4	260	18%	70,534,548	15%	15	15%	4,445,270	3%
5	889	63%	45,255,421	9%	12	12%	472,578	0%
All	1,412	100%	482,574,664	100%	103	100%	171,533,673	100%

Table A-16 shows a comparison between the end use disposition in the PG&E Fab population and the end use disposition in the final sample. Since the sample was randomly selected by stratum, the distribution of sample projects by end use is very similar to that in the population. The sample has a slightly higher percentage of POC projects and a slightly lower percentage of lighting and HVAC projects than the program population of projects.

Table A-16: Distribution of Electric Projects by End Use in the PG&E Fab Population and Final Sample Using Program Achievements as of Q4, 2008

End Use	Population				Sample			
	N	% N	gross kWh	% gross kWh	N	% N	gross kWh	% gross kWh
POC	656	46%	163,650,384	34%	41	40%	67,622,421	39%
Lighting	116	8%	45,832,432	9%	9	9%	7,803,370	5%
HVAC	82	6%	16,585,308	3%	3	3%	4,877,928	3%
Other Electric	558	40%	256,506,539	53%	50	49%	91,229,954	53%
All	1,412	100%	482,574,664	100%	103	100%	171,533,673	100%

As shown in Table A-17, the estimated confidence and precision for this sample design, at the time of the sample design, was 90/6 if the error ratio was 0.35 and 90/8 if the error ratio was 0.5. If the error ratio was 0.35, the estimated confidence and precision for the 41 POC projects would be 90/9, and for the remaining electric projects 90/7. Assuming the error ratio was no higher than 0.40, the estimated confidence/precision for both POCs and non-POCs would be equal or better than 90/10. The actual error ratios that resulted from the evaluation sample are provided in Section 4 of this report.

Table A-17: Confidence and Precision Estimates, Under Alternate Error Ratio Estimates, for PG&E Fab Electric Sample Using Program Achievements as of Q4, 2008

End Use	N		Precision	
	Population	Sample	er=0.35	er=0.5
POC	656	41	90/9	90/12
Other Electric	756	62	90/7	90/10
All Electric	1,412	103	90/6	90/8

Gas Sample

The approach used for electric sampling was employed to isolate the 14 gas sample points drawn through Q2, 2008 and to allocate the balance of 13 gas sample points by sample stratum. For consistency with the verification and EM&V work already completed, the number of strata and the strata boundaries defined in the verification sample design were preserved, and the gross Therms savings were used to proportionally assign sample points to

each stratum. Gas sampling did not have to incorporate pre-install projects, because the only gas pre-install project that was targeted for the PG&E Fab group had not been installed as of Q4, 2008.

After accounting for the sample drawn through Q2, 2008, the “balance” column in Table A-18 shows the distribution of 13 sample points that were still needed for the M&V Evaluation. The “realistic” column in Table A-18 then redistributes the sample points so that the total number of sample points by stratum does not exceed the number of projects in the population.

Table A-18: Gas Sample Disposition using Program Achievements as of Q4, 2008

Strata	Population			Sample				
	N	gross kWh	% gross kWh	Target N	Through Q208	Pre-Install	Balance N	Realistic N
1	8	24,058,701	53%	14	4	0	10	3
2	1	1,059,000	3%	1	1	0	0	0
3	6	4,878,873	19%	5	3	0	2	6
4	5	2,720,285	5%	2	2	0	0	1
5	132	7,427,520	20%	5	4	0	1	3
All	152	40,144,380		27	14	0	13	13

Thirteen projects distributed as shown in the last column of Table A-18 were selected by random sampling. Table A-19 shows a comparison between the sample already drawn through Q2, 2008 and the new incremental EM&V gas sample, by stratum.

Table A-19: Comparison of Existing and Proposed Incremental Gas Sample Using Program Achievements as of Q4, 2008

Strata	Sample Through Q2, 2008				Proposed Incremental Sample			
	N pop	Therm pop	N sample	Therm sample	N pop	Therm pop	N sample	Therm sample
1	4	12,427,091	4	12,427,091	4	11,631,610	3	8,877,121
2	1	1,059,000	1	1,059,000	0	0	0	0
3	3	2,416,107	3	2,416,107	3	2,462,766	6	5,217,255
4	3	1,515,022	2	1,008,242	2	1,205,263	1	581,275
5	42	2,811,094	4	624,682	90	4,616,427	3	382,360
All	53	20,228,314	14	17,535,123	99	19,916,066	13	15,058,011

Table A-20 compares the program achievements through Q4, 2008 with the final sample disposition, by stratum. The sample is close to the target sample shown in Table A-18, and includes a census of projects from strata 1, 2 and 3 (largest three strata). Note the power of the stratification in that the sample captures 80 percent of the claimed savings with only 27 sample points (representing roughly 20 percent of the number of projects in the population).

Table A-20: Distribution of Gas Projects by Stratum in the PG&E Fab Population and Final Sample using Program Achievements as of Q4, 2008

Strata	Population				Sample			
	N	% N	gross Therm	% gross Therm	N	% N	gross Therm	% gross Therm
1	8	5%	24,058,701	60%	7	26%	21,304,212	65%
2	1	1%	1,059,000	3%	1	4%	1,059,000	3%
3	6	4%	4,878,873	12%	9	33%	7,633,362	23%
4	5	3%	2,720,285	7%	3	11%	1,589,517	5%
5	132	87%	7,427,520	19%	7	26%	1,007,042	3%
All	152	100%	40,144,380	100%	27	100%	32,593,134	100%

Table A-21 shows a comparison between the gas end use disposition in the PG&E Fab population and the end use disposition in the final gas sample. Since the sample selection was random by stratum, the distribution of sample projects by end use is very similar with the end use distribution in the population. The final sample has a slightly higher percentage of heating boilers and a slightly lower percentage of process boilers than the program population of projects.

Table A-21: Distribution of Projects by End Use in the PG&E Fab Population and Final Gas Sample using Program Achievements as of Q4, 2008

End Use	Population				Sample			
	N	% N	gross Therm	% gross Therm	N	% N	gross Therm	% gross Therm
Process Boiler	50	33%	8,569,002	21%	4	15%	4,854,154	15%
Heating Boiler	27	18%	17,237,876	43%	9	33%	17,151,040	52%
Boiler Controls	8	5%	99,355	0%	0	0%	0	0%
Other Gas	67	44%	14,238,147	35%	14	52%	10,857,940	33%
All	152	100%	40,144,380	100%	27	100%	32,863,134	100%

As shown in Table A-21, the estimated confidence/precision levels for the final sample are 90/10, assuming an error ratio of 0.35. The actual error ratios and confidence and precision levels that resulted from the evaluation are presented in Section 4.

Table A-22: Confidence and Precision Estimates, Under Alternate Error Ratio Estimates, for PG&E Fab Gas Sample using Program Achievements as of Q4, 2008

End Use	N		Precision	
	Population	Sample	er=0.35	er=0.5
Boiler	85	13	90/15	90/21
Other Gas	67	14	90/14	90/20
All Gas	152	27	90/10	90/15

A.1.5 Net-to-Gross Sample Design

The original research plan submitted to the CPUC for the PG&E Fab contract group included net-to-gross evaluations by program based on “300 sample points or one-half of the program size, whichever is lowest.”

Based on a program tracking database extract for Q3, 2008, a net-to-gross sample of 350 points was drawn for program PGE2004. Similarly, a net-to-gross sample of 164 points was drawn for third-party programs in the PGE Fab contract group. The number of points drawn was higher than 300 or one-half of program size as of Q3, 2008 to allow for backups in case of survey non-response. The M&V sample was nested in the net-to-gross sample.

After the Q4, 2008 tracking database was received and the final M&V sample was finalized, the sample points that were added to the M&V sample in Q4, 2008 were also added to the net-to-gross sample. Since our evaluation approach had shifted focus from program-level to contract group-level, the resulting net-to-gross sample size already exceeded the “300 or one-half of program size” rule. The sample was not further expanded with additional measures installed in Q4, 2008.

Table A-23 below summarizes the NTG sample design for electric measures installed in the PG&E Fab programs, and Table A-24 shows the same for the gas measures installed in the PG&E Fab programs.

Overall, extremely high percentages of the claimed electric (69 percent) and gas (87 percent) savings were captured in the net-to-gross samples.

Table A-23: PG&E Fab Net-to-Gross Sample Design – Electric Measures as of Q4, 2008

Strata	PGE Fab Programs		NTG Sample		NTG Sample Percent	
	N records	Gross kWh	N records	Gross kWh	N records	Gross kWh
1	34	142,447,707	27	116,620,304	79%	82%
2	28	55,226,868	20	39,485,533	71%	71%
3	201	169,110,120	147	122,784,283	73%	73%
4	260	70,534,548	171	44,892,447	66%	64%
5	889	45,255,421	188	8,802,084	21%	19%
Total Electric Measures	1,412	482,574,664	553	332,584,649	39%	69%

Table A-24: PG&E Fab Net-to-Gross Sample Design – Gas Measures as of Q4, 2008

Strata	PGE Fab Programs		NTG Sample		NTG Sample Present	
	N records	Gross Therms	N records	Gross Therms	N records	Gross Therms
1	8	24,058,701	8	24,058,701	100%	100%
2	1	1,059,000	1	1,059,000	100%	100%
3	6	4,878,873	6	4,878,873	100%	100%
4	5	2,720,285	5	2,720,285	100%	100%
5	132	7,427,520	34	2,038,612	26%	27%
Total Gas Measures	152	40,144,380	54	34,755,471	36%	87%

Southern California Industrial and Agricultural Program Evaluation ON-SITE Data Collection Form

1.1 INTERVIEW INFORMATION

Company Name / App. No. : _____

Street Address: _____

Facility Representative(s): _____

Phone / Email: _____

SIC Code (if blank see SIC codes
in Lookup Tables) _____

Reported Building Type _____

Electric and Gas Account Information

Verify that all accounts at the site are listed in this table.

Account Type	Account Number	Baseline Annual Energy	Post-Retrofit Annual Energy	Notes

Projects Evaluated

Evaluator	Date of Site Visit	IOU Application Number	Itron Assigned Project No.	Measure(s) Evaluated

1.2 DESCRIPTION OF FACILITY

Primary Services or Products	
Total floor space of this facility	ft ²
Conditioned floor space (this facility)	ft ²
Year business established at site	
Obtain project invoices	Obtained / Not obtained
Customer requested copy of report	Yes / No
Customer requested copy of raw data	Yes / No

(Reports and raw data can be provided to the customer after the project is completed in 2009.)

Site Characteristics

Business Hours

Day Type	Pre-Retrofit Operating Hours	Closed All Day?	Open 24 hours?	Partial Occupancy %	Average # of Occupants?
Weekdays	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		
Saturday	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		
Sunday	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		
Other	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		

Day Type	Post-Retrofit Operating Hours	Closed All Day?	Open 24 hours?	Partial Occupancy %	Average # of Occupants?
Weekdays	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		
Saturday	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		
Sunday	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		
Other	From _____ to _____	<input type="checkbox"/>	<input type="checkbox"/>		

Seasonal variations in the level of occupancy or use:
Does evaluated measure(s) operate when facility is closed?
Are there any regularly scheduled plant shut downs when the measure does not operate? If so when does this occur, how many hours and how many days

Closed Holidays: *Check all that apply below or => N/A*

Number of Closed Holidays per year	_____
------------------------------------	-------

Enter "0" above if they never close. Do not read through the list below, just check the holidays that the site contact mentions or ask a general question about which holidays are closed days, and check that the number above is consistent.

New Year's Day	<input type="checkbox"/>	Labor Day	<input type="checkbox"/>
Martin Luther King Day	<input type="checkbox"/>	Columbus Day	<input type="checkbox"/>
Presidents Day	<input type="checkbox"/>	Veterans Day	<input type="checkbox"/>
Memorial Day	<input type="checkbox"/>	Thanksgiving Day	<input type="checkbox"/>
July 4 th	<input type="checkbox"/>	Christmas Day	<input type="checkbox"/>

1.3 Interview Facility Representative

1) Early retirement under the SPC 04-05 Evaluation requires calculation of energy savings using the existing equipment as the baseline for energy use (verses the current standards), but only for the remaining useful life of the equipment. This can apply to all measures, particularly lighting and equipment replacement. If the measure is an early retirement measure:

- a) At the time the equipment was replaced, how many years were left in its useful life (without major repairs which may have led to replacement)? _____
- b) How old was the equipment that was removed and replaced? _____
- c) Was the existing equipment fully functional, fully functioning but with significant problems, or non-functional? _____
- d) How often was major non-scheduled maintenance required and of what type? _____
- e) How often had the equipment failed recently, and over what time period?

- f) How satisfactory was the performance of the old equipment? _____
- g) How long would the old equipment have met the technical and performance needs of the facility? _____

2) Determination of baseline condition:

a) Did you consider any alternatives to the [DESCRIBE MEASURE] installed/through the PROGRAM that you would have implemented in the same time frame if the program had not been available? By the same time frame I mean within 6 months of the time when you participated in the program. Which of the following describes the alternatives you considered? (check all that apply):

- i) I did not consider any alternatives (SKIP TO Q#3)
- ii) I considered fewer units of the measure
- iii) I considered a different model or efficiency level
- iv) I considered both fewer units and a different model
- v) Other (specify)

b) Did you evaluate any of these alternatives at the same time as you evaluated the MEASURE that you eventually installed through the PROGRAM?

NO: (IF NO skip to Q#2c)

YES: Which of the following best describes the most likely alternative that you evaluated?

- i) Fewer high efficiency units (e.g., controls, VFDs, lights). How many units would you have installed? _____
- ii) A standard efficiency version of the same equipment (or one that meets code or other regulatory requirements). What criteria, code or other requirement

would you have used to determine the efficiency of this equipment?

iii) Equipment more efficient than code, but less efficient than we installed through the program. Do you know the efficiency rating or model number of the equipment that you would have installed? If yes, record: _____ If not, ask: In percentage terms, about how much less efficient would this equipment have been compared to the program qualifying equipment you installed? _____

iv) Repair/rewind/refurbish the existing equipment. How long do you think the repaired/rewound/refurbished equipment would have lasted before requiring replacement? _____

v) Something else (specify)

c) In the absence of the rebate from the PROGRAM, is it more likely that you would have done nothing or is it more likely that you would have installed the alternative that you just described? (IF ALTERNATIVE MORE LIKELY: Can you provide any notes or other documentation regarding your exploration?)

3) Does the customer have any reason to believe that there will be any changes in the operation of the primary measure?

a) Changes in hours _____

b) Changes in load _____

c) Impact on annual kWh savings _____

d) Impact on kW savings _____

4) Any perceived non-energy benefits, e.g., increased production, increased comfort, new equipment, environmental branding, etc.? _____

5) Were there any drawbacks to the energy efficiency measure? _____

6) Was there a production increase when the new measure was installed? _____ If answer YES, then:

a) Was the production increase enabled by the new equipment? _____

b) Would you have increased your production if you had not installed the new equipment? _____

7) Record all measure specific contextual data. (see Measure Specific list in Lookup Tables)

1.4 MONITORING

IOU Application Number:

Itron Project ID:

Site Characteristics to be Verified (that could affect the measure impact or approach)

Data Collection Method Description

*

The following types of measurement equipment will be used in this evaluation including metering interval and duration for each instrument:

Num.	Measurement Type	Equipment	Duration (weeks)	Interval (minutes)

Sensor Calibration and Quality Assurance

Questions to Ask on the Phone or On-Site

1.5 Lookup Tables

1.5.1 Two-Digit Agricultural & Manufacturing 1987 SICs

- 01 Agricultural production- crops
- 02 Agricultural production- livestock
- 07 Agricultural services
- 08 Forestry
- 09 Fishing, hunting, and trapping
- 20 Food and kindred products
- 21 Tobacco manufactures
- 22 Textile mill products
- 23 Apparel and other textile products
- 24 Lumber and wood products
- 25 Furniture and fixtures
- 26 Paper and allied products
- 27 Printing and publishing
- 28 Chemicals and allied products
- 29 Petroleum and coal products
- 30 Rubber and miscellaneous plastics products
- 31 Leather and leather products
- 32 Stone, clay, glass, and concrete products
- 33 Primary metal industries
- 34 Fabricated metal products
- 35 Industrial machinery and equipment
- 36 Electrical and electronic equipment
- 37 Transportation equipment
- 38 Instruments and related products
- 39 Miscellaneous manufacturing industries

1.5.2 Measure Specific Contextual Data

Heating System

- Winter occupied setpoint (F)
- Monitored heating system type (furnace, air/water/ground source heat pump, boiler)
- Monitored heating system year of installation

All Non-Residential Comfort Cooling Measures

- Summer occupied setpoint (F)
- Total non-backup capacity in tons associated with measure
- Monitored system type—type of coils in supply air fan (refrigerant, chilled water)
- Monitored system supply air flow control strategy (constant, variable volume, or cycling)
- Monitored system outside air strategy (none, fixed %, fixed cfm, economizer)
- Monitored compressor type (reciprocating, screw, centrifugal, scroll, other)
- Monitored packaged unit or chiller make & model number

Water-Side Measure on Chilled Water-Based Cooling System

- Chilled water temperature control strategy (constant, reset based on OAT, reset based on load, other)
- Condenser water temperature control strategy (constant, OATdb reset, OATwb reset, load reset, other)

Supply Air Fans

- Predominant summer supply air temperature setpoint for areas affected by measure (F)
- Supply air temperature control scheme for system affected by measure (constant, reset, manually adjusted, other)
- Supply air pressure reset control scheme for system affected by measure (constant, reset, manually adjusted, other)
- Monitored fan type (forward curved, back inclined, airfoil, vane axial, other)
- Monitored fan flow control (constant volume, cycle, VSD, inlet vane, outlet damper, variable pitch, other)
- Monitored motor nameplate hp, volts, amps, efficiency, and power factor

Pumps (Chilled Water and Condenser Water)

- Monitored pump flow control (constant volume, cycle, VSD, throttle, other)
- Monitored motor nameplate hp, volts, amps, efficiency, and power factor

Cooling Towers

- Condenser water temperature control strategy (constant, OATdb reset, OATwb reset, load reset, manual reset, other)
- Fan control strategy (single speed, two-speed, variable speed, multiple motors, combination)

Process Refrigeration - Heat Rejection Side Measures

- Condenser approach temperature (F)
- Minimum head pressure setpoint (psi)

Process Refrigeration - Evaporator Side Measures

- Defrost type (hot gas, resistance, timer, etc.)
- Load type (refrigerated storage, frozen storage, chilling product, freezing product)

Agricultural Pumping

- Acres under irrigation

**Methodological Framework for Using the Self-
Report Approach to Estimating Net-to-Gross
Ratios for Nonresidential Customers**

**Prepared for the Energy Division, California Public Utilities
Commission**

By

The Nonresidential Net-To-Gross Ratio Working Group

Final Version

TABLE OF CONTENTS

1. OVERVIEW OF THE LARGE NONRESIDENTIAL FREE RIDERSHIP APPROACH	1
2. BASIS FOR SRA IN SOCIAL SCIENCE LITERATURE	2
3. FREE RIDERSHIP ANALYSIS BY PROJECT TYPE	2
4. SOURCES OF INFORMATION ON FREE RIDERSHIP	2
5. NTGR FRAMEWORK	5
5.1. NTGR Questions and Scoring Algorithm	6
5.1.1. Timing and Selection Score	7
5.1.2. Program Influence Score	8
5.1.3. No-Program Score	8
5.1.4. The Core NTGR	9
5.2. Data Analysis and Integration	9
5.3. Accounting for Partial Free Ridership	13
6. NTGR INTERVIEW PROCESS	15
7. COMPLIANCE WITH SELF-REPORT GUIDELINES	16

Appendix A: References

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

Appendix C: NTGR Scoring Algorithm and Example

Appendix D: Demonstration of Compliance with the CPUC/ED Guidelines for Estimating Net-to-Gross Ratios Using the Self-Report Approach

Acknowledgments

As part of the evaluation of the 2006-08 energy efficiency programs designed and implemented by the four investor-owned utilities (Pacific Gas & Electric Company, Southern California Edison Company, Southern California Gas Company, and San Diego Gas and Electric Company) 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. The working group, listed alphabetically, was composed of the following evaluation professionals:

- Michael Baker, SBW Consulting
- Fred Coito, KEMA
- Kevin Cooney, Summit Blue Consulting
- Tim Drew, Energy Division, CPUC
- Jennifer Fagan, Itron, Inc.
- Miriam Goldberg, KEMA
- Nick Hall, TecMarket Works
- Kay Hardy, Energy Division, CPUC
- Ken Keating
- John Reed, Innovologie LLC
- Richard Ridge, Ridge & Associates
- Mike Rufo, Itron, Inc.
- Eric Swan, KEMA (formerly of RLW Analytics, Inc.)
- Christina Torok, Itron, Inc.
- Philippus Willems, PWP, Inc.

A public webinar was conducted to obtain feedback from the four investor-owned utilities and other interested stakeholders. The questionnaire was then pre-tested and, based on the pre-test results, finalized in November 2008.

1. OVERVIEW OF THE LARGE NONRESIDENTIAL FREE RIDERSHIP APPROACH

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

This approach preserves the most important elements of the approaches previously used to estimate the NTGRs in large nonresidential customer programs¹. However, it also incorporates several enhancements that are designed to improve upon that approach, for example:

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

It is important to note that the NTGR approach described in this document is a general framework, designed to address all large nonresidential programs. In order to implement this approach on a program-specific basis, it might need to be somewhat customized to reflect the unique nature of the individual programs.

¹ Such as, for example, the NTGR method used to evaluate NTGRs for the California Standard Performance Contracting Program.

2. BASIS FOR SRA IN SOCIAL SCIENCE LITERATURE

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

More specifically, the SRA is a mixed method approach that involves asking one or more key participant decision-makers a series of structured and open-ended questions about whether they would have installed the same EE equipment in the absence of the program as well as questions that attempt to rule out rival explanations for the installation (Weiss, 1972; Scriven, 1976; Shadish, 1991; Wholey et al., 1994; Yin, 1994; Mohr, 1995). In the simplest case (e.g., residential customers), the SRA is based primarily on quantitative data while in more complex cases the SRA is strengthened by the inclusion of additional quantitative and qualitative data which can include, among others, in-depth, open-ended interviews, direct observation, and review of program records. Many evaluators believe that additional qualitative data regarding the economics of the customer's decision and the decision process itself can be very useful in supporting or modifying quantitatively-based results (Britan, 1978; Weiss and Rein, 1972; Patton, 1987; Tashakkori and Teddlie, 1998).²

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

3. FREE RIDERSHIP ANALYSIS BY PROJECT TYPE

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 savings³ 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 must exercise their own discretion as to what the appropriate thresholds should be for each of these three levels.

4. SOURCES OF INFORMATION ON FREE RIDERSHIP

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

² *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches*, October 15, 2007, pg. 3.

³ 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.

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

2. **Decision-Maker Surveys.** When a site is recruited, one must also determine who was involved in the decision-making process which led to the implementation of measures under the program. They are asked to complete a Decision Maker survey. This survey obtains highly structured responses concerning the probability that the customer would have implemented the same measure in the absence of the program. First, participants are asked about the timing of their program awareness relative to their decision to purchase or implement the energy efficiency measure. Next, they are asked to rate the importance of the program versus non-program influences in their decision making. Third, they are asked to rate the significance of various factors and events that may have led to their decision to implement the energy efficiency measure at the time that they did. These include:
 - the age or condition of the equipment,
 - information from a feasibility study or facility audit
 - the availability of an incentive or endorsement through the program
 - a recommendation from an equipment supplier, auditor or consulting engineer
 - their previous experience with the program or measure,
 - information from a program-sponsored training course or marketing materials provided by the program
 - the measure being included as part of a major remodeling project
 - a recommendation from program staff, a program vendor, or a utility representative
 - a standard business practice
 - an internal business procedure or policy
 - stated concerns about global warming or the environment
 - a stated desire to achieve energy independence.

In addition, the survey obtains a description of what the customer would have done in the absence of the program, beginning with whether the implementation was an early replacement action. If it was not, the decision maker is asked to provide a description of what equipment would have been implemented in the absence of the program, including both the efficiency level and quantities of these alternative measures. This is used to adjust the gross engineering savings estimate for partial free ridership, as discussed in Section 5.2.

This survey contains a core set of questions for **Basic** NTGR sites, and several supplemental questions for both **Standard** and **Standard – Very Large** NTGR

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

3. **Vendor Surveys.** A Vendor Survey is completed for all **Standard** and **Standard-Very Large** NTGR sites that utilized vendors, and for **Basic** NTGR sites that indicate a high level of vendor influence in the decision to implement the energy efficient measure. For those sites that indicate the vendor was very influential in decision making, the vendor survey results enter directly into the NTGR scoring. The vendor survey findings are also be used to corroborate Decision Maker findings, particularly with respect to the vendor’s specific role and degree of influence on the decision to implement the energy efficient measure. Vendors are queried on the program’s significance in their decision to recommend the energy efficient measures, and on their likelihood to have recommended the same measure in the absence of the program. Generally, the vendors contacted as part of this study are contractors, design engineers, distributors, and installers.
4. **Utility and Program Staff Interviews.** For the Standard and Standard-Very Large NTGR analyses, interviews with utility staff and program staff are also conducted. These interviews are designed to gather information on the historical background of the customer’s decision to install the efficient equipment, the role of the utility and program staff in this decision, and the name and contact information of vendors who were involved in the specification and installation of the equipment.
5. **Other information.** For **Standard – Very Large Project** NTGR sites, secondary research of other pertinent data sources is performed. For example, this could include a review of standard and best practices through industry associations, industry experts, and information from secondary sources (such as the U.S. Department of Energy's Industrial Technologies Program, Best Practices website URL, <http://www1.eere.energy.gov/industry/bestpractices/>). In addition, the Standard- Very Large NTGR analysis calls for interviews with other employees at the participant’s firm, sometimes in other states, and equipment vendor experts

from other states where the rebated equipment is being installed (some without rebates), to provide further input on standard practice within each company.

Table 1 below shows the data sources used in each of the three levels of free-ridership analysis. Although more than one level of analysis may share the same source, the amount of information that is utilized in the analysis may vary. For example, all three levels of analysis obtain core question data from the Decision Maker survey.

Table 1: 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	√	√	√

¹Only performed for sites that indicate a vendor influence score (N3d) greater than maximum of the other program element scores (N3b, N3c, N3g, N3h, N3i).

²Only performed for sites that have a utility account representative

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

Appendix B provides the full battery of Decision Maker and Vendor survey questions along with notes, for each NTGR level, regarding which questions are asked (denoted by an “X”), and the intended uses of the information in the NTGR analysis. In the case of Basic sites, “TRIGGER” means that a vendor influence score greater than the maximum of other program element scores (N3b, N3c, N3g, N3h, N3i) triggers a vendor survey. In the case of Standard and Standard-Very Large NTGR sites, “TRIGGER” means that a score of 6 or greater triggers a further investigation. A copy of the complete survey forms (with lead-in text and skip patterns) are contained in *Final Large Nonresidential NTGR Survey Instruments.XLS* that is available upon request.

5. NTGR FRAMEWORK

The Self-Report-based Net-to-Gross analysis relies on responses to a series of survey questions that are designed to measure the influence of the program on the participant’s decision to implement program-eligible energy efficiency measure(s). Based on these

responses, a NTGR is derived based on responses to a set of “core” NTGR questions. The NTGR includes the effects of deferred free ridership (i.e., accelerated adoption).

5.1. NTGR Questions and Scoring Algorithm

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

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

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

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

5.1.1. Timing and Selection Score

For the Decision Maker, the questions asked are:

I'm going to ask you to rate the importance of the program as well as other factors that might influence your decision to implement [MEASURE.] 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 implement this specific [MEASURE] at this time.

- Availability of the PROGRAM rebate
- Information provided through a recent feasibility study, energy audit or other types of technical assistance provided through PROGRAM
- Information from PROGRAM training course
- Information from other PROGRAM marketing materials
- Recommendation from a vendor/supplier (If a score of greater than 5 is given, a vendor interview is triggered)

For the Vendor, the questions asked (if the interview is triggered) are:

I'm going to ask you to rate the importance of the [PROGRAM] in influencing your decision to recommend [MEASURE] to [CUSTOMER] and other customers. 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.

1. Using this 0 to 10 scale where 0 is “Not at all important” and 10 is “Very Important,” how important was the PROGRAM, including incentives as well as program services and information, in influencing your decision to recommend that CUSTOMER install the energy efficiency MEASURE at this time?
2. And using a 0 to 10 likelihood scale, where 0 denotes “not at all likely” and 10 denotes “very likely,” if the PROGRAM, including incentives as well as program services and information, had not been available, what is the likelihood that you would have recommended this specific energy efficiency MEASURE to CUSTOMER?
3. Now, using a 0 to 100 percent scale, in what percent of sales situations did you recommend MEASURE before you learned about the [PROGRAM]?
4. And using the same 0 to 100 percent scale, in what percent of sales situations do you recommend MEASURE now that you have worked with the [PROGRAM]?

5. And, using the same 0 to 10 scale where 0 is “Not at all important” and 10 is “Very important”, how important in your recommendation were:
 - a. Training seminars provided by UTILITY?
 - b. Information provided by the UTILITY website?
 - c. Your firm’s past participation in a rebate or audit program sponsored by UTILITY?

If the Vendor interview is triggered, a score is calculated that captures the highest degree of program influence on the vendor’s recommendation. This score (VMAX) is calculated as the MAXIMUM value of the following:

1. The response to question 1
2. 10 minus the response to question 2
3. The response to question 4 minus the response to question 3, divided by 10
4. The response to question 5a.
5. The response to question 5b.
6. The response to question 5c.

Note that vendors are asked an additional question regarding other ways that their recommendations regarding the measure might have been influenced. Their responses are not used in the direct calculation of the NTGR but are potentially useful in making adjustments to the core NTGR.

The Timing and Selection Score is calculated as:

The highest of the responses to the first four decision maker questions and, if the vendor interview has been triggered, the VMAX score multiplied by the score the decision makers assigned to the vendor recommendation.

5.1.2. Program Influence Score

The questions asked are:

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

The Program Influence score is calculated as:

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

5.1.3. No-Program Score

The questions asked are:

1. Regarding the installation of this equipment, if the PROGRAM had not been available, using a likelihood scale from 0 to 10, where 0 is “Not at all likely” and 10 is “Extremely likely” how likely is it that you would have installed exactly the same item/equipment, using a 0 to 10 scale, where 0 is not at all likely and 10 is extremely likely?

2. IF 1>0. You indicated that there was an “X” in 10 likelihood that you would have installed the same equipment if the PROGRAM had not been available. When do you think you would have installed this equipment? Please express your answer in months
 - a. _____ within 6 months? (Deferred NTG Value=0)
 - b. _____ 7 to 47 months later (Deferred NTG Value=(months-6)*.024)
 - c. _____ 48 or more months later (Deferred NTG Value =1)
 - d. _____ Never (Deferred NTG Value=1)

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

The No-Program Score is calculated as:

10 minus (the likelihood of installing the same equipment multiplied by one minus the *deferred net-to-gross value* associated with the timing of that installation).

5.1.4. The Core NTGR

The self-reported core NTGR in most cases is simply the average of the Program Influence, Timing and Selection, and No-Program Scores, divided by 10. The one exception to this is 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 is based on the average of the Program Influence and No-Program scores only.

5.2. Data Analysis and Integration

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

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

- Account Representative Interview
- Utility Program Manager/Staff Interview
- Utility Technical Contractor Interview
- Third party Program Manager Interview
- Evaluation Engineer Interview
- Gross Impact Site Plan/Analysis Review
- Corporate Green/Environmental Policy Review (if mentioned as important)
- Corporate Standard Practice Review (if mentioned as important)
- Industry Standard Practice Review (if mentioned as important)
- Corporate payback review (if mentioned as important)
- Review relevant codes and standards, including regulatory requirements
- Review industry publications, websites, reports such as the Commercial Energy Use Survey, historical purchase data of specific measures etc.

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

Establishing Rules for Data Integration

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

decision-maker interview, when there is no appropriate decision-maker interview, or when there is critical missing data on the questionnaire, and how to incorporate qualitative information on deferred free-ridership.

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

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

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

Sometimes, *all* the quantitative and qualitative data will clearly point in the same direction while, in others, the *preponderance* of the data will point in the same direction. Other cases will be more ambiguous. In all cases, in order to maximize reliability, it is

essential that more than one person be involved in analyzing the data. Each person must analyze the data separately and then compare and discuss the results. Important insights can emerge from the different ways in which two analysts look at the same set of data. Ultimately, differences must be resolved and a case made for a particular NTGR. Careful training of analysts in the systematic use of rules is essential to insure inter-rater reliability⁴.

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

The outcome of the reconciliation by two analysts determines the final NTGR for a specific project. Again, the reasoning behind the “negotiated” final value must be thoroughly documented in a workpaper, while a more concise summary description of the rationale can be included in the NTGR Calculator workbook (e.g., Analyst 1 and Analyst 2 agreed that the NTGR score should have been higher than the calculated value of 0.45 because of extensive interaction between program technical staff and the customer, but they disagreed on whether this meant the NTGR should be .6 or .7. After discussion, they agreed on a NTGR of .65 as reflecting the extent of program influence on the decision).

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

- Vendor interview data will be used at times in the direct calculation of the NTGR. It will also be used to provide context and confirming/contradictory information for Standard-Very Large decision maker interviews.
- Qualitative and quantitative information from other sources (e.g., industry data, vendor estimates of sales in no-program areas, and other data as described above) may be used to alter core inputs only if contradictions are found with the core survey responses. Since judgments will have to be made in deciding which information is more compelling when there are contradictions, supplemental data are reviewed independently by two senior analysts, who then summarize their findings and recommendations and together reach a final NTGR value.

⁴ 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.

- Responses will also be used to construct a NTGR “story” around the project; that is they will help to provide the context and rationale for the project. This is particularly valuable in helping to provide guidance to program design for future years. It may be, for example, that responses to the core questions yield a high NTGR for a project, but additional information sources strongly suggest that the program qualifying technology has since become standard practice for the firm or industry, so that free ridership rates in future years are likely to be higher if program rules are not changed.
- Findings from other non-core NTGR questions (e.g., Payback Battery, Corporate Policy Battery) are also be used to **cross-check the consistency** of responses to core NTGR questions. When an inconsistency is found, it is presented to the Decision Maker respondent who is then be asked to explain and resolve it if they can. If they are not able to do so, their responses to the core NTGR question with the inconsistency may be overridden by the findings from these supplemental probes. These situations are handled on a case-by-case basis; however consistency checks are programmed into the CATI survey instrument used for the Basic and Standard cases.

Finally, some analysis of additional information beyond the close-ended questions that are used to calculate the Core NTGR could be done for the **Standard NTGR**. For example information regarding the financial criteria used to make capital investments, corporate policy regarding the purchase of energy efficiency equipment or the influence of standard practice in the same industry as the participant could be taken into account and used to make adjustments to the Core NTGR in a manner similar what is done for the Standard – Very Large NTGR.

5.3. Accounting for Partial Free Ridership

Partial free-ridership can occur when, in the absence of the program, the participant would have installed something more efficient than the program-assumed baseline efficiency but not as efficient as the item actually installed as a result of the program.

In situations where there is partial free ridership, the assumed baseline condition is affected. Absent partial free ridership, the assumed baseline would normally be based on existing equipment (in early replacement cases), on code requirements (in normal replace on burnout cases), or on a level above current code (e.g., this could be a market average or value purposefully set above code minimum but below market average; in this case, the definition and requirement would typically be defined by a specific program’s baseline rules). In some cases, there may be a “dual” baseline (more specifically, a baseline that changes over the measure’s EUL) if the project involves early replacement plus partial free ridership. In such cases, the baseline basis for estimating savings is the existing equipment over the remaining useful life (RUL) of the equipment, and then a baseline of likely intermediate efficiency equipment (e.g., code or above) for the remainder of the analysis period (i.e., the period equal to the EUL-RUL). When there is partial free ridership, the baseline equipment that would have been installed absent the program is of an intermediate efficiency level (resulting in lower energy savings than that assumed by the program if the program took in situ equipment efficiency as the basis for

savings over the entire EUL). A related issue with respect to determination of the appropriate baseline is whether the adjustment made, if any, from the in situ or otherwise claimed baseline in the ex ante calculation, is whether the adjustment applies to the gross or net savings calculation.

Assignment of Partial Free Ridership Effects to Gross versus Net. In past evaluations, partial free ridership impacts have principally been incorporated into the net-to-gross ratio. This is because most partial free ridership is induced by market conditions, rather than by non-market factors. Market conditions refer primarily to standard adoption of a technology by a particular market segment or end user as a result of competitive market forces or other end user-specific factors. The key determining principle with respect to application of the adjustment to the net-to-gross ratio is whether there is a level of efficiency, below the efficiency of the measure for which savings are paid and claimed, but above what is required by code or minimum program baseline requirements that the end user would have implemented anyway without the program. Conditions that cause this adjustment to be made to gross savings rather than the net-to-gross ratio may include factors such as

- changing baseline equipment to meet changed business circumstances (such as increased production/throughput, changes in occupancy, etc.);
- compliance with environmental regulations, indoor air quality requirements, safety requirements; or
- the need to address an operational problem.

Each project should be examined separately for partial free ridership and a determination should be made based on the unique circumstances of each installation of whether an adjustment to gross savings or the net-to-gross ratio is warranted.

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

1. Now I would like you to think one last time about what action you would have taken if the program had not been available. Supposing that you had not installed the program qualifying equipment, which of the following alternatives would you have been MOST likely to do?
 - a. Install fewer units
 - b. Install standard efficiency equipment or whatever required by code
 - c. Install equipment more efficient than code but less efficient than what you installed through the program
 - d. repair/rewind or overhaul the existing equipment
 - e. do nothing (keep the existing equipment as is)
 - f. something else (specify what _____)
2. (IF FEWER UNITS) How many fewer units would you have installed? (It is okay to take an answer such as ...HALF...or 10 percent fewer ... etc.)

3. (IF MORE EFFICIENT THAN CODE) Can you tell me what model or efficiency level you were considering as an alternative? (It is okay to take an answer such as ... 10 percent more efficient than code or 10 percent less efficient than the program equipment)
4. (IF REPAIR/REWIND/OVERHAUL) How long do you think the repaired/rewound/refurbished equipment would have lasted before requiring replacement?

In addition, these same partial free ridership questions should be asked during the on-site audit for a given project. This latter interview will be conducted by the project engineers. The collected information helps the gross impact and NTG analysis teams gain a more complete understanding of the true project baseline and equipment selection decision. These decision maker questions are included in the Excel version of the CATI-based Standard and Basic decision maker survey instrument as well as in the Standard-Very Large instrument.

Data Analysis and Integration Procedures. In cases where partial free ridership is found and it is determined that the adjustment should be made to the net-to-gross ratio, the following procedure should be used:

On the net side, the adjustment is based on the intermediate baseline indicated by the decision maker for the time period in which the intermediate equipment would have been installed. The calculation of energy saved under this intermediate baseline is done, and then divided by the savings calculated under the in situ baseline. The resulting ratio is then multiplied by the initial NTGR which was previously calculated using only the 'core' scoring inputs. The effect of this adjustment is to reduce the NTGR further to reflect the effects of the revealed partial free ridership.

In all cases, the Gross Impacts and NTG analysis teams will need to carefully coordinate their calculations to ensure that they are not inadvertently adjusting the savings twice for the same partial free ridership, i.e., through adjustments both to the gross savings calculation and to the NTG ratio.

6. NTGR INTERVIEW PROCESS

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

All but the Standard -Very Large interviews should be conducted using computer-aided telephone interview (CATI) software. Use of a CATI approach has several advantages: (1) the surveys can be customized to reflect the unique characteristics of each program, and associated program descriptions, response categories, and skip patterns; (2) it drastically reduces inaccuracies associated with the more traditional paper and pencil method; and (3) the process of checking for inconsistent answers can be automated, with follow up prompts triggered when inconsistencies are found.

7. COMPLIANCE WITH SELF-REPORT GUIDELINES

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

Appendix B

References

- Blalock, H. (1970). Estimating measurement error using multiple indicators and several points in time," *American Sociological Review*, 35, pp. 101-111.
- Bogdan, Robert and Steven J. Taylor. (1975). *Introduction to qualitative research methods*. New York: John Wiley & Sons.
- Britan, G. M. (1978). Experimental and contextual models of program evaluation. *Evaluation and Program Planning*, 1: 229-234.
- Cochran, William G. (1977). *Sampling techniques*. New York: John Wiley & Sons.
- Crocker, L. and J. Algina. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.
- Cronbach L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- DeVellis, R.F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage Publications, Inc.
- Duncan, O.D. (1984). *Notes on social measurement: Historical and critical*. New York: Russell Sage.
- Guba, E. G. (1978). Toward a methodology of naturalistic inquiry in educational evaluation. *CSE Monographic Series in Evaluation No. 8*. Los Angeles: Center for the Study of Evaluation.
- Hall, Nick, Johna Roth, Carmen Best, Sharyn Barata, Pete Jacobs, Ken Keating, Ph.D., Steve Kromer, Lori Megdal, Ph.D., Jane Peters, Ph.D., Richard Ridge, Ph.D., Francis Trottier, and Ed Vine, Ph.D. (2007). *California Energy Efficiency Evaluation: Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*. Prepared for the California Public Utilities Commission.
- Lyberg, Lars, Paul Biemer, Martin Collins, Edith De Leeuw, Cathryn Dippo, Norbert Schwarz, and Dennis Trewin. (1997). *Survey measurement and process quality*. New York, NY: John Wiley & Sons.
- Madow, William G., Harold Nisselson, Ingram Olkin. (1983). *Incomplete data in sample surveys*. New York: Academic Press.
- Maxwell, Joseph A. (2004). Using Qualitative Methods for Causal Explanations. *Field Methods*, Vol. 16, No. 3, 243-264.

- Megdal, Lori, Yogesh Patil, Cherie Gregoire, Jennifer Meissner, and Kathryn Parlin (2009). Feasting at the Ultimate Enhanced Free-Ridership Salad Bar. *Proceedings of the International Energy Program Evaluation Conference*.
- Mohr, Lawrence B. (1995). *Impact analysis for program evaluation*. Thousand Oaks, CA: Sage Publications, Inc.
- Netemeyer, Richard G., William O. Bearden, and Subhash Sharma. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: SAGE Publications.
- Patton, Michael Quinn. (1987). *How to use qualitative methods in evaluation*. Newbury Park, California: SAGE Publications.
- Ridge, Richard, Philippus Willems, and Jennifer Fagan. (2009). Self-Report Methods for Estimating Net-to-Gross Ratios in California: Honest! *Proceedings from the 19th National Energy Services Conference*.
- Ridge, Richard, Philippus Willems, Jennifer Fagan and Katherine Randazzo. (2009). The Origins of the Misunderstood and Occasionally Maligned Self-Report Approach to Estimating the Net-To-Gross Ratio. *Proceedings of the International Energy Program Evaluation Conference*.
- Rogers, Patricia J., Timothy A. Hacsí, Anthony Petrosino, and Tracy A. Huebner (Eds.) (2000). *Program theory in evaluation: Challenges and opportunities*. San Francisco, CA: Jossey-Bass Publishers.
- Rossi, Peter and Howard E. Freeman. (1989). *Evaluation: A systematic approach*. Newbury Park, California: SAGE Publications.
- Sayer, Andrew. (1992). *Method in social science: A Realist Approach*. New York: Routledge.
- Sax, Gilbert. (1974). *Principles of educational measurement and evaluation*. Belmont, CA: Wadsworth Publishing Company, Inc.
- Schumacker, Randall E. and Richard G. Lomax. (1996). *A beginner's guide to structural equation modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Scriven, Michael. (1976). Maximizing the power of causal explanations: The modus operandi method. In G.V. Glass (Ed.), *Evaluation Studies Review Annual, Vol. 1*, pp.101-118). Beverly Hills, CA: Sage Publications.
- Shadish, Jr., William R. and Thomas D. Cook, and Laura C. Leviton. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage Publications, Inc.
- Stone, Arthur A., Jaylan S. Turkkan, Christine A. Bachrach, Jared B. Jobe, Howard S. Kurtzman, and Virginia S. Cain. (2000). *The science of the self-report: Implications for research and practice*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

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

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

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

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

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

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

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

Appendix B

Net-to-Gross Questions and Uses of Data by Level of NTGR Analysis

Note: A more detailed version of this survey, with skip patterns and complete response categories, is available in Excel format from the NTG Working Group or at <http://www.energydataweb.com/cpuc/default.aspx>

DECISION MAKER SURVEY

	Question Text	Basic	Standard and Standard – Very Large
	<p>Introduction</p> <p>Hello, my name is _____ from COMPANY NAME and I am calling about your recent participation in PROGRAM NAME. Are you the person who was most involved with the decision to participate in the PROGRAM NAME? [IF YES, CONTINUE]. We are interviewing firms that participated in the PROGRAM NAME in 2006 and 2007 to discuss the factors that may have influenced your decision to participate in the program. The interview will take about 20 minutes. The questions on this survey pertain to work completed by your company at this current address, excluding other locations.</p>		
	<p>WARM-UP QUESTIONS</p>		
A1	<p>First, according to our records, you participated in PROGRAM NAME on (approximate date). [READ: Program Description. PROGRAM NAME promotes energy efficiency improvements in commercial/industrial facilities. The program offers (choose all that apply): energy audits to help identify applicable measures, feasibility studies to analyze the energy and cost savings of recommended measures, incentives to help cover a portion of the cost of implementing energy efficient measures, etc. Is that correct?</p>	X	X
	<p>Yes, No, DK, Refused</p>		
A2	<p>Next, I'd like to confirm the following information regarding the measures you implemented through the program: (READ: PROJECT DETAILS INCLUDING SERVICES RECEIVED, MEASURES INSTALLED, KEY DATES, PARTICIPATING VENDORS, ETC.) Does that sound right?</p>	X	X
	<p>Yes, No, DK, Refused</p>		
A3	<p>Why did you decide to implement MEASURE NAME? Were there any other reasons?</p>	X	X
	<p>a. Record VERBATIM</p>		
	<p>b. DK/Refused</p>		
	<p>NET-TO-GROSS BATTERY</p>		
N1	<p>When did you first learn about PROGRAM? Was it BEFORE or AFTER you first began to think about implementing MEASURE?</p>	X	X
	<p>a. Before (Skip to N3)</p>		
	<p>b. After</p>		
	<p>c. DK/Refused</p>		

N2	Did you learn about PROGRAM BEFORE or AFTER you decided to implement the specific MEASURE that was eventually adopted or installed?	X	X
	a. Before		
	b. After		
	c. DK/Refused		
	<i>READ: Program Description: As I mentioned earlier, [PROGRAM NAME] promotes energy efficiency improvements in commercial/industrial facilities. The program offers (choose all that apply): energy audits to help identify applicable measures, feasibility studies to analyze the energy and cost savings of recommended measures, incentives to help cover a portion of the cost of implementing energy efficient measures, etc. I'm going to ask you to rate the importance of the program as well as other factors that might influence your decision to implement [MEASURE.] 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.</i>		
N3	Now, using this 0 to 10 rating scale, where 0 means "Not at all important" and 10 means "Very important," please rate the importance of each of the following in your decision to implement this specific [MEASURE] at this time. [CUSTOMIZE LIST OF FACTORS FOR PROGRAM BEFORE ASKING THEM TO SCORE THE FULL LIST. ROTATE PRESENTATION OF ITEMS. FOLLOW UP WITH "And is there anything else that I may have missed?" RECORD AS p. Other (SPECIFY)]		
	a. The age or condition of the old equipment	X	X
	b. Availability of the PROGRAM rebate	X	X
	c. Information provided through a recent feasibility study, energy audit or other types of technical assistance provided through the PROGRAM (probe on when and by whom?)	X	X
	d. Recommendation from a vendor/supplier (If >5, Vendor interview may be triggered)	TRIGGER	TRIGGER
	e. Previous experience with PROGRAM?	X	X
	f. Previous experience with this MEASURE?	X	X
	g. Information from PROGRAM training course?	X	X
	h. Information from other PROGRAM marketing materials?	X	X
	i. A recommendation from an auditor or consulting engineer	X	X
	j. Standard practice in our business/industry (IF >5, ask standard practice battery)	X	TRIGGER
	k. Endorsement or recommendation by PROGRAM staff, PROGRAM vendor, or UTILITY representative	X	X
	l. Corporate policy or guidelines (If >5 ask Policy questions)	X	TRIGGER
	m. Payback on the investment (If >5 ask payback battery)	X	TRIGGER
	n. General concerns about the environment	X	X
	o. Specific concerns about global warming	X	X
	p. Specific concerns about achieving energy independence	X	X
	q. Other (SPECIFY)	X	X
N4	Now I'd like to ask you a last question about the importance of the program to your decision. Again using the 0 to 10 rating scale we used earlier, where 0 means "Not at all important" and 10 means "Very important," please rate	X	X

	the overall importance of PROGRAM versus the other factors we just discussed in your decision to implement the specific MEASURE. I'd like you to give me a 0 to 10 score for the 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.		
	a. _____ rating of the importance of PROGRAM NAME	X	X
	b. _____ rating of the importance of Other Factors	X	X
	<i>Now I would like you to think about the action you would have taken with regard to the installation of this equipment PROGRAM had not been available.</i>		
N5	Regarding the installation of this equipment if the PROGRAM had not been available, how likely is it that you would have installed exactly the same item/equipment, using a 0 to 10 likelihood scale, where 0 is not at all likely and 10 is extremely likely?	X	X
N6	<i>IF N5>0.</i> You indicated in your previous responses that there was a X in 10 likelihood that you would have installed the same equipment if the PROGRAM had not been available.	X	X
	When do you think you would have installed this equipment? (Please answer in months)_____		
	a. _____ ..within 6 months? NTGR = 0		
	b. _____ .. 6 – 47 months later (NTGR=(months-6)*.024)		
	c. _____ ..4 or more years later (NTGR=1)		
	g. _____ ..Never (NTGR=1)		
	PARTIAL FREE RIDERSHIP BATTERY	GROSS IMPACT	GROSS IMPACT
P1	Now I would like you to think one last time about what action you would have taken if the program had not been available. Supposing that you had not installed the program qualifying equipment, which of the following alternatives would you have been MOST likely to do?: <ul style="list-style-type: none"> a. Install fewer high efficiency units (e.g., controls, VFDs, lights) b. Install standard efficiency equipment or whatever required by code c. Install equipment more efficient than code, but less efficient than we installed through the program d. Repair/rewind/refurbish the existing equipment e. do nothing (keep the existing equipment as is) f. Something else (specify) 		
P4	If P1=a: How many units would you have installed? Record number of units or percentage of units actually installed		
P5			
P6	If P1=c: Can you tell me what model or efficiency level you were considering as an alternative? (It is okay to take an answer such as ... 10 percent more efficient than code or 10 percent less efficient than the program equipment)		
P7	If P1=d: How long do you think the repaired/rewound/refurbished equipment would have lasted before requiring replacement?		
P8			
P9			
	Additional Decision Maker Questions		

	PAYBACK BATTERY (If payback importance >5)		
N10	What financial calculations does your company make before proceeding with installation of a MEASURE like this one?		X
N11	What is the cut-off point your company uses before deciding to proceed with the investment?		X
N12	What was the result of the calculation for MEASURE: a) with the rebate? b) without the rebate?		X
	<i>INVESTIGATE INCONSISTENT RESPONSE</i>		
N13	What competing investments, if any, were considered for the funds that were allocated to the adoption of MEASURE?		X
N14	Why was MEASURE chosen over these other investments		X
	CORPORATE POLICY BATTERY (If corporate policy importance >5)		
N15	Does your organization have a corporate environmental policy to reduce environmental emissions or energy use? Some examples would be to "buy green" or use sustainable approaches to business investments.		X
N16	What specific corporate policy influenced your decision to adopt or install MEASURE?		X
N17	Had that policy caused you to adopt the MEASURE at this facility before participating in this program?		X
N18	Had that policy caused you to adopt the MEASURE at other facilities before participating in this program? When and where?		X
N19	Did you receive an incentive for a previous [MEASURE]? If so, please describe.		X
	STANDARD PRACTICE BATTERY (If standard practice importance >5)		
N20	How long has MEASURE been standard practice in your industry?		X
N21	Does your company ever deviate from the standard practice? If yes, under what conditions?		X
N22	How did this standard practice influence your decision to install the energy efficiency equipment		X
N23	What industry group or trade organization do you look to establish standard practice for your industry?		X
N24	How do you and other firms/facilities receive information on updates in standard practice?		X
	OTHER INFLUENCES BATTERY		
N25	Who provided the most assistance in the design or specification of MEASURE? Designer or Consultant, Equipment Distributor or Mfr Rep, Installer, Utility rep, or Internal staff	X	X
N26	Please describe the type of assistance that they provided.	X	X
N27	Please state, in your own words, any other factors that influenced your decision to go ahead on this energy efficient equipment/project.	X	X

VENDOR SURVEY

	Question Text	Basic	Standard and Standard Very Large
	Warm Up		
A1	The CUSTOMER indicates that you recommended the installation of [EFFICIENT MEASURE] at their facility at [CUSTOMER LOCATION] on [DATE]. Do you recall making this recommendation?	X	X
	a. Yes		
	b. No		
	c. DK (-8)		
	d. Refused (-9)		
	<i>I'm going to ask you to rate the importance of the [PROGRAM] in influencing your decision to recommend [MEASURE] to [CUSTOMER] and other customers. 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.</i>		
V1	Using this 0 to 10 scale where 0 is "Not at all important" and 10 is "Very Important", how important was PROGRAM, including incentives as well as program services and information, in influencing your decision to recommend that CUSTOMER install the energy efficiency MEASURE at this time?	X	X
V2	And using a 0 to 10 likelihood scale, where 0 denotes "not at all likely" and 10 denotes "very likely," if the PROGRAM, including incentives as well as program services and information, had not been available, what is the likelihood that you would have recommended this specific energy efficiency MEASURE to CUSTOMER?	X	X
V3	Now, using a 0 to 100 percent scale, in what percent of sales situations did you recommend MEASURE before you learned about the [PROGRAM]?	X	X
V4	And using the same 0 to 100 percent scale, in what percent of sales situations do you recommend MEASURE now that you have worked with the [PROGRAM]?	X	X
V4a	In what other ways have your recommendations regarding MEASURE been influenced? [For each mention, ask: And using the same 0 to 10 scale, where 0 is "Not at all important" and 10 is "Very important", how important in influencing your recommendations. . . (INSERT FIRST MENTION, INSERT SECOND MENTION ETC.)]	X	X
V5	And, using the same 0 to 10 scale where 0 is "Not at all important" and 10 is "Very important", how important in your recommendation were		
	a. Training seminars provided by UTILITY?	X	X
	b. Information provided by the UTILITY website?	X	X
	c. Your firm's past participation in a rebate or audit program sponsored by UTILITY?	X	X

	Optional:		
V6	Approximately what percentage of your sales of MEASURE in UTILITY'S service territory are energy efficient models that qualify for incentives from the UTILITY program.	X	X
V7	On a 0 percent to 100 percent scale, in what percent of sales situations do you encourage your customers in UTILITY territory to purchase program qualifying [MEASURES]?	X	X
V8.	(IF LESS THAN 100) In what situations do you NOT encourage your customers to purchase energy efficient models if they qualify for a rebate? Why is that?	X	X
V9	Of those installations of EQUIPMENT in UTILITY service territory that qualify for incentives, approximately what percentage do not receive the incentive?	X	X
V10	Why do they not receive the incentive (open end?)	X	X
V11	Do you also sell MEASURE in areas where customers do not have access to incentives for energy efficient models?	X	X
V12	About what percent of your sales of MEASURE are represented by these areas where incentives are not available?	X	X
V12a	IF AT LEAST 10%: And approximately what percentage of your sales of MEASURE in these areas are the energy efficient models that would qualify for incentives in UTILITY'S service territory?	X	X
V13	Have you changed your stocking practices as a result of the UTILITY program? If yes, how?	X	X
V14	Do you promote energy efficient models equally in areas with and without incentives?	X	X

Appendix C

NTGR Scoring Algorithm and Example

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

1. A ***Timing and Selection*** score that captures the influence of the most important of various program and program-related elements in influencing the customer to select the specific program measure at this time. Program influence through vendor recommendations is also captured in this score.
2. An overall ***Program Influence*** score that captures the perceived importance of the program (whether rebate, recommendation, or other information) in the decision to implement the specific measure that that was eventually adopted or installed. The overall program influence score is reduced by half if the respondent says they learned about the program only after they decided to install the program qualifying measure.
3. A ***No-Program*** score that captures the likelihood of various actions the customer might have taken at this time and in the future if the program had not been available. This score accounts for deferred free ridership by capturing the likelihood that the customer would have installed program qualifying measures at a later date if the program had not been available.

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

Timing and Selection Score

For the decision maker, the questions asked are:

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

- Availability of the PROGRAM rebate
- Information provided through a recent feasibility study, energy audit or other types of technical assistance provided through the PROGRAM
- Information from PROGRAM training course
- Information from other PROGRAM marketing materials
- Recommendation from a vendor/supplier (If >5, a vendor interview is triggered)

For the vendor, the questions asked if the interview is triggered are:

1. On a 0 to 10 scale where 0 is “Not at all important” and 10 is “Very important”, how important was PROGRAM, including incentives as well as program services and information, in influencing your decision to recommend that CUSTOMER install the energy efficiency MEASURE at this time?
2. And using a 0 to 10 likelihood scale, where 0 denotes “Not at all likely” and 10 denotes “Extremely Likely,” if the PROGRAM, including incentives as well as program services and information, had not been available, what is the likelihood that you would have recommended this specific energy efficiency MEASURE to CUSTOMER?
3. Now, using a 0 to 100 percent scale, in what percent of sales situations did you recommend this MEASURE before you learned about the PROGRAM?
4. And using the same 0 to 100 percent scale, in what percent of sales situations do you recommend this MEASURE now that you have worked with the PROGRAM?
5. And, using the same 0 to 10 scale where 0 is “Not at all important” and 10 is “Extremely Important”, how important in your recommendation were:
 - a. Training seminars provided by UTILITY?
 - b. Information provided by the UTILITY website?
 - c. Your firm’s past participation in a rebate or audit program sponsored by UTILITY?

If the vendor interview is triggered, a score is calculated that captures the highest degree of program influence on the vendor’s recommendation. This score (VMAX) is calculated as the MAXIMUM value of the following:

1. The response to question 1
2. 10 minus the response to question 2
3. The response to question 4 minus the response to question 3, divided by 10
4. The response to question 5 a.
5. The response to question 5b.
6. The response to question 5c.

The Timing and Selection Score is calculated as:

The highest of the responses to the first four decision maker questions and, if the vendor interview has been triggered, the VMAX score multiplied by the score the decision makers assigned to the vendor recommendation..

Example:

The decision maker provides responses of 5 for the importance of the rebate, 6 for an audit or feasibility study, 3 for training, 2 for other marketing materials, and 7 for the vendor recommendation, which means a vendor interview is triggered.

The vendor responses are 8 for the significance of the program, 5 for the likelihood of recommending the measure in the absence of the program, 40% for how often the measure was recommended before program awareness and 60% for how often it is recommended after program awareness, 3 for the importance of training, 2 for the importance of the website and 5

for the importance of previous participation. The VMAX score is the greatest of 8, (10-5), (60-40)/10, 3, 2 and 5. So VMAX is 8. This score is multiplied by the importance of the vendor recommendation, to which the decision maker assigned a 7, so the vendor score is 5.6.

The timing and selection score is the maximum of the four decision maker responses (5, 6, 3, and 2) and the vendor score (5.6). Even though the vendor interview was triggered, the vendor score is not as high as the 6 assigned to the importance of the audit or feasibility study, so the timing and selection score is 6.

Program Influence Score

The questions asked are:

1. Did you learn about PROGRAM BEFORE or AFTER you decided to implement the specific MEASURE that was eventually adopted or installed?
2. Again using the 0 to 10 rating scale we used earlier, where 0 means "Not at all important" and 10 means "Very important," please rate the overall importance of PROGRAM versus the most important of the other factors we just discussed in your decision to implement the specific MEASURE that was adopted or installed. This time I would like to ask you to have the two importance ratings -- the program importance and the non-program importance -- total 10.

The program influence score is calculated as:

The program importance response, on the 0 to 10 scale, to question 2. This score is reduced by half if the respondent became aware of the program only after having decided to adopt the program qualifying measure.

Example:

The decision maker says they became aware of the program before deciding to implement the measure, and provides a response of 7 to question 2, which becomes the program influence score.

No-Program Score

The questions asked are:

1. Regarding the installation of this equipment if the PROGRAM had not been available, how likely is it that you would have installed exactly the same item/equipment, using a 0 to 10 likelihood scale, where 0 is not at all likely and 10 is extremely likely?
2. IF 1>0. You indicated in your previous responses that there was an "X" in 10 likelihood that you would have installed the same equipment if the PROGRAM had not been available. When do you think you would have installed this equipment? Please express your answer in months
 - a. _____ Within 6 months? (Deferred NTG Value=0)
 - b. _____ 7 to 47 months later (Deferred NTG Value=(months-6)*.024)

- c. _____ 48 or more months later (Deferred NTG Value =1)
- d. _____ Never (Deferred NTG Value=1)

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

The No-Program Score is calculated as:

10 minus (the likelihood of installing the same equipment multiplied by one minus the deferred net-to-gross value associated with the timing of that installation).

Example

The respondent says there is a 4 in 10 likelihood that they would have installed the same equipment. In response to question 5, the decision maker says they would have installed the qualifying equipment 18 months later, which has a NTGR value of $(18-6)*.024$, or .29 associated with it.

The No-Program score is 10 minus $(4*(1-.29))$, which is 10 minus $4*.71$ or 7.16.

Core NTG Ratio

The self-reported core NTGR in most cases is simply the average of the Program Influence, Timing and Selection, and No-Program Scores, divided by 10. The one exception to this is 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 is based on the average of the Program Influence and No-Program scores only.

Example (Core NTGR)

The NTGR is the average of 6, 8 and 7.2, or 7.1 divided by 10 = .71. This figure is then applied to adjusted gross savings to yield net savings.

Appendix D

Demonstration of Compliance with the CPUC/ED and MEC's Guidelines for Estimating Net-to-Gross Ratios Using the Self-Report Approach

1. Timing of the interview

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

2. Identifying the correct respondent

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

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

3. Set-up questions

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

4. Use of multiple questions

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

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

5. Validity and reliability

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

- *"Tried and true" question wording.* Many of the core questions used in NTGR scoring are substantially the same as those that have been used extensively in previous large C&I program evaluations, such as the last several rounds of evaluation for the California Standard Performance Contracting Program. While the question construct is somewhat

different from in the past, the wording used is essentially the same as has been used previously.

- *Information from supplemental questions and multiple data sources to corroborate and triangulate on the NTGR “story”.* In addition to self-reported information, the NTGR findings for Standard and Standard – Very Large NTGR sites include responses to a number of supplemental questions surrounding the project (e.g., corporate policy, standard industry practice and payback), and the results from an interview with the vendor(s) involved in the project. These findings will be used to converge on a plausible estimate of the NTGR and to help tell the “story” behind the project and its context.
- *Multiple reviewers. Standard - Very Large customer projects are reviewed by two experienced analysts.* The two reviewers seek to develop a NTGR consensus on the project, and resolve any differences of opinion.
- *Identification and explicit consideration of alternate hypotheses.* Respondents are asked about the relative influence of a variety of program and non-program factors.

During the pre-test of the NTGR survey instrument, reliability tests should be conducted using the CATI software. Any problem areas detected should be corrected.

6. Consistency checks

Questions within the NTGR battery that are more likely to produce inconsistent responses have been flagged. These include questions regarding the program’s reported importance in the decision to implement the specified measure, alternative actions in the program’s absence, questions reporting the motivations for doing the project, as well as any closely related supplemental questions. The CATI software should be specifically programmed to flag any inconsistencies, and include follow-up prompts when they are found. Interviewers should be instructed how to administer these follow-up questions to resolve these inconsistencies. Interviewers should make every effort to resolve any inconsistencies before concluding the interview. Examples of the procedures for checking consistency of responses are provided in Section 3.

7. Making the Questions Measure-Specific

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

8. Partial free-ridership

Questions P1-P9 are designed to collect the information necessary to adjust for any partial free-ridership. *However, this adjustment is be made to the **gross savings** estimates and not to the NTGR.*

9. Deferred free-ridership

Question N6 addresses deferred free ridership, and provides specific adjustment factors for each response category. The NTGR algorithm (See Section 5 and Appendix C) text fully explains the specifics of this adjustment.

10. Scoring algorithms

The methodology includes a specific algorithm for developing a NTGR based on responses received. The results of the 0 to 10 scoring are used to develop specific values for each question used to score the NTGR. A description of the scoring algorithm is provided in Section 5 and in Appendix C.

11. Handling unit and item non-response

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

12. Weighting the NTGR

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

13. Ruling out rival hypotheses

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

14. Precision of the NTGR

The calculation of the achieved relative precision of the NTGRs (for program-related measures and practices and non-program measures and practices) is expected to be straightforward. However, the inclusion of more complicated situations involving multiple participant and vendor

interviews as well as the inclusion of additional qualitative information means that the NTGR standard errors may underestimate the uncertainty surrounding the NTGR estimate.

15. Pre-testing the questionnaire

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

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

Specific rules have been established for data integration and these are described in Section 3.

17. Qualified interviewers

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

Standard Decision Maker NTG Survey Instrument Modified 06/22/09

Introduction

AA1	<p>This is %n calling on behalf of the CPUC, [California Public Utilities Commission] from ITRON CONSULTING. THIS IS NOT A SALES CALL. May I please speak with <%CONTACT> ... the person most knowledgeable about your firm's involvement in ...<%CUSTOMER>'s... installation of ...<%MEASURE>..on approximately ...<%INSTALL_DATE>?,</p> <p>1 Yes 2 No</p>	AA7 AA2
AA2	<p>Who would be the person most knowledgeable about your firm's involvement with ...<%CUSTOMER>'s...project that involved the installation of ...<%MEASURE>... on approximately... <%INSTALL_DATE>?,</p> <p>1 Record name 88 Refused 99 Don't know</p>	AA3 Thank and Terminate Thank and Terminate
AA3	<p>May I speak with him/her?</p> <p>1 Yes 2 No (not available right now) SCHEDULE APPOINTMENT</p>	AA4 Reschedule appt.
AA4	<p>This is %n calling on behalf of the CPUC, [California Public Utilities Commission] from ITRON CONSULTING. THIS IS NOT A SALES CALL. I was told that you are the person most familiar with your firm's involvement in ...<%CUSTOMER>'s... installation of ...<%MEASURE>..on approximately ...<%INSTALL_DATE>? __Is this correct?</p> <p>1 Yes 2 No, there is someone else (RECORD NAME) 3 No and I don't know who to refer you to 88 Refused 99 Don't know</p>	AA7 AA5 Thank and Terminate Thank and Terminate Thank and Terminate
AA5	<p>This is %n calling on behalf of the CPUC, [California Public Utilities Commission] from ITRON CONSULTING. THIS IS NOT A SALES CALL. Am I speaking with the person most familiar with your firm's involvement in ...<%CUSTOMER>'s... installation of ...<%MEASURE>..on approximately ...<%INSTALL_DATE>? __Is this correct?</p> <p>1 Yes. 2 Yes, but I need to make an appointment 3 No, but I will give you to the correct person 88 Refused 99 Don't know</p>	AA7 Reschedule appt. AA7 Thank and Terminate Thank and Terminate
AA7	<p>We are interviewing firms that participated in <%PROGRAM> during 2006, 2007 and 2008 to discuss the factors that may have influenced their decision to participate in the program. By receiving a rebate of \$ <%INCENTIVE> through this program, your organization agreed to participate in this follow-up study on your experiences with this program.</p> <p>IF VISIT = 1 We <(VISIT == 1)/Have already visited/will also be visiting> your site to get information on the measures installed. One of our engineers has already visited your site to get information on the measures installed.</p> <p>1 .<%ENGINEER>... spoke to ...<%ONSITEREP> ... on ..<%ONSITEDATE>.;</p>	A1

Your input to this research is extremely important. We will not identify or attribute any of your comments or organization

Before we start, I would like to inform you that for quality control purposes, this call may be monitored by my supervisor. For the sake of expediency, we will be recording this interview.

[If INTERVEWEE requests a contact at their local utility, the following are the appropriate representatives for this evaluation, note

PGE Rob Roffrey - (415) 973-1222
 SCE Ron Cobas - 626-633-3088
 SDGE Sandra Williams 858-636-5802
 CPUC Peter Lai 213-576-7087

A1	<p>According to our records your organization participated in .. <%PROGRAM>... on ...<%INSTALL_DATE>... by installing ...<%MEASURE>. Does this sound right?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>A1b A1a A1a A1a</p>
A1a.	<p>What do you remember installing through this program?</p> <p>77 RECORD VERBATIM 88 Refused 99 Don't know</p>	<p>A1b A1b A1b</p>
IF AUDIT == 1; THEN ASK ELSE A1c		
A1b	<p>According to our records, your organization also received an AUDIT from <%UTILITY>. Is this correct?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>A1c A1c A1c A1c</p>
IF TECH_ASSST == 1, THEN ASK, ELSE A1d		
A1c	<p>According to our records, your organization also received TECHNICAL ASSISTANCE from <%UTILITY>. Is this correct?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>A1d A1d A1d A1d</p>
IF FEAS_STUDY == 1, THEN ASK, ELSE A1e		
A1d	<p>According to our records, your organization also received a FEASIBILITY STUDY from <%UTILITY>. Is this correct?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>A1e A1e A1e A1e</p>
IF RCX == 1, THEN ASK, ELSE A1f		
A1e.	<p>According to our records, your organization also received RETROCOMMISSIONING from <%UTILITY>. Is this correct?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>A1f A1f A1f A1f</p>
IF PTRAIN == 1, THEN ASK ELSE A1g		
A1f.	<p>According to our records, your organization also received PROGRAM TRAINING from <%UTILITY>. Is this correct?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>A1g A1g A1g A1g</p>
A1g	<p>Our records show that your organization received \$ <%INCENTIVE> from ...<%PROGRAM>... for the installation of this equipment. Does this sound correct?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>A1h A1gg A1h A1h</p>
A1gg.	<p>What was the incentive amount that your organization received through the program?</p> <p>77 RECORD VERBATIM 88 Refused 99 Don't know</p>	<p>A1h A1h A1h</p>

[READ] For the sake of expediency, during the balance of the interview, we will be referring to the <-%PROGRAM> as the PROGRAM and we will be referring to the installation of ... <-%MEASURE> as the MEASURE. I will repeat this from time to time during the study as your organization may have installed more than one measure through more than one program.

[READ] I would like to get some information on the VENDORS that may have helped you with the implementation of this equipment. As part of this study, we will be conducting a separate interview with the vendors that worked with you on the implementation of this equipment.

Revision

First let's talk about the EQUIPMENT SUPPLIER/INSTALLER Vendor. We show (READ NAME AND PHONE) ! as the EQUIPMENT VENDOR.[READ NAME AND PHONE NUMBER] Is that correct?
 ! VENDOR NAME... <-%VEND1NAME>
 ! VENDOR PHONE...<-%V1PHONE>
 1 Yes A1h
 2 No A1h1
 88 Refused A1h
 99 Don't know A1h
IF VENDOR1 =2 OR A1h=2, THEN ASK:
 Can we have the VENDOR NAME_____, Their phone number, ___their CONTACT name _____,
 A1h1 Their Cell phone number !___their EMAIL ADDRESS ?
 77 RECORD VENDOR NAME, PHONE NUMBER AND CONTACT INFORMATION A1i
 88 Don't know A1i
 99 Refused A1i
IF VENDOR2 = 1 OR 2, THEN ASK
 Our records show you also used a DESIGN or CONSULTING Engineer. Did you use a DESIGN OR CONSULTING Engineer?
 A1i [READ NAME AND PHONE NUMBER]
 ! VENDOR NAME... <-%VEND2NAME>
 ! VENDOR PHONE...<-%V2PHONE>
 1 Yes A1j
 2 No A1i1
 88 Refused A1j
 99 Don't know A1j
IF VENDOR2 =2 OR A1i=2, THEN ASK:
 Can we have the VENDOR NAME_____, Their phone number, ___their CONTACT name _____,
 A1i1 Their Cell phone number !___their EMAIL ADDRESS ?
 77 RECORD VENDOR NAME, PHONE NUMBER AND CONTACT INFORMATION A1j
 88 Don't know A1j
 99 Refused A1j
IF VENDOR3 == 1 OR 2, THEN ASK
 Our records show you also used a PROGRAM PROVIDED Vendor. Did you use a PROGRAM PROVIDED Vendor? [READ
 A1j. NAME AND PHONE NUMBER]
 ! VENDOR NAME... <-%VEND3NAME>
 ! VENDOR PHONE...<-%V3PHONE>
 1 Yes A2a
 2 No A1j1
 88 Refused A2a
 99 Don't know A2a
IF VENDOR3 ==2, THEN ASK:
 Can we have the VENDOR NAME_____, Their phone number, ___their CONTACT name _____,
 A1j1 Their Cell phone number !___their EMAIL ADDRESS ?
 77 RECORD VENDOR NAME, PHONE NUMBER AND CONTACT INFORMATION A2a
 88 Don't know A2a
 99 Refused A2a

Thanks for helping us with this vendor information. Below, I am going to ask some questions about the implementation of the measure that you installed through the program. Should you remember any vendor information later on, please feel free to volunteer this information at that time, I can record vendor information at any time.

WARM-UP QUESTIONS:

A2a	How did you first become aware of the &MEASURE?		
	1 Bill insert	A2	
	2 Program Literature	A2	
	3 Account representative	A2	
	4 Program provided vendor	A2	
	5 Program representative	A2	
	6 Utility or program website	A2	
	7 Trade publication	A2	
	8 Conference	A2	
	9 Newspaper article	A2	
	10 Word of mouth	A2	
	11 Previous experience with it	A2	
	12 Company used it at other locations	A2	
	13 Contractor	A2	
	14 Other (RECORD VERBATIM)	A2	
	88 Refused	A2	
	99 Don't know	A2	
A2	In your own words, can you tell me why you decided to implement this MEASURE?		Revision
	77 RECORD VERBATIM	N1	
	88 Don't know	N1	
	99 Refused	N1	

NET-TO-GROSS QUESTIONS:

N1	When did you first learn about <%UTILITY>'s PROGRAM? Was it BEFORE or AFTER you first began to THINK about implementing this MEASURE?	
	1 Before	N3
	2 After	N2
	88 Refused	N2
	99 Don't know	N2
N2	Did you learn about <%UTILITY>'s Program BEFORE or AFTER you DECIDED to implement the MEASURE that was installed?	
	1 Before	N3
	2 After	N3
	88 Refused	N3
	99 Don't know	N3

[READ:&PROGRAMDESCR]. Next, I'm going to ask you to rate the importance of the program as well as other factors that might have influenced your decision to implement &MEASURE. 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.

Next, I'm going to ask you to rate the importance of the program as well as other factors that might have influenced your decision to implement this MEASURE. 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 extremely important, so that an importance rating of 8 shows twice as much influence as a rating of 4. Now using this scale please rate the importance of each of the following in your decision to implement the MEASURE at this time.

N3		
N3a.	The age or condition of the old equipment # Record 0 to 10 score (_____) 88 Refused 99 Don't know	N3a. N3b. N3b. N3b.
N3b.	Availability of the PROGRAM rebate # Record 0 to 10 score (_____) 88 Refused 99 Don't know	N3bb N3bb N3bb
	IF N3b > 7, THEN ASK.	
N3bb	Why do you give it this rating? 77 Record VERBATIM 88 Refused 99 Don't know	N3c. N3c. N3c.
	IF &FEAS_STUDY=1, &AUDIT=1, OR &TECH_ASSIST=1, THEN ASK, ELSE N3h Information provided through... !!__<(FEAS_STUDY == 1)/ The Feasibility study/> !__<(AUDIT == 1)/The Facility or System AUDIT/>	
N3c.	!__<(TECH_ASST == 1)/The Technical Assistance # Record 0 to 10 score (_____) 88 Refused 99 Don't know	N3c1. N3c2. N3c2.
	IF N3c > 7, THEN ASK.	
N3c1.	Why do you give it this rating? 77 Record VERBATIM 88 Refused 99 Don't know	N3c2. N3c2. N3c2.
	IF VENDOR1,NE.0,THEN ASK	
N3d.	Recommendation from an equipment vendor that sold you &MEASURE and/or installed it [VENDOR_1] # Record 0 to 10 score (_____) 88 Refused 99 Don't know	IF N3d > N3b, N3c, N3g, N3h, N3I then conduct ve N3e. N3e. N3e.
N3e.	Previous experience with this &MEASURE? # Record 0 to 10 score (_____) 88 Refused 99 Don't know	N3f. N3f. N3f.
N3f.	Previous experience with the utility &PROGRAM or a similar utility program (such as &SIM_PGM?) # Record 0 to 10 score (_____) 88 Don't know 99 Refused	N3g. N3g. N3g.
	IF &PGM_TRAIN=1 OR &UTIL_TRAIN=1 THEN ASK, ELSE N3h	
N3g.	Information from &PROGRAM or &UTILITY training course? # Record 0 to 10 score (_____) 88 Refused 99 Don't know	N3gg N3h N3h
	IF N3g >7, THEN ASK	

Revision

N3gg	Why do you give it this rating?	
	77 Record VERBATIM	N3h.
	88 Refused	N3h.
	99 Don't know	N3h.
N3h.	Information from &PROGRAM or &UTILITY marketing materials?	
	# Record 0 to 10 score (_____)	N3hh.
	88 Refused	N3i
	99 Don't know	N3i
	IF N3h >7, THEN ASK	
N3hh	Why do you give it this rating?	
	77 Record VERBATIM	N3i
	88 Refused	N3i
	99 Don't know	N3i
	IF VENDOR2,NE.0,THEN ASK	
N3i.	A recommendation from a design or consulting engineer [VENDOR_2]	
	# Record 0 to 10 score (_____)	N3j.
	88 Refused	N3j.
	99 Don't know	N3j.
N3j.	Standard practice in your business/industry	
	# Record 0 to 10 score (_____)	N3k.
	88 Refused	N3k.
	99 Don't know	N3k.
	IF VENDOR3,NE.0,THEN ASK	
N3k.	Endorsement or recommendation by [&PGM_VEND] [VENDOR_3]	
	# Record 0 to 10 score (_____)	N3k1
	88 Refused	N3k2
	99 Don't know	N3k2
	IF N3k >7, THEN ASK	
N3k1	Why do you give it this rating?	
	77 Record VERBATIM	N3k2
	88 Refused	N3k2
	99 Don't know	N3k2
N3l.	Endorsement or recommendation by &ACCT_REP	
	# Record 0 to 10 score (_____)	N3ll
	88 Refused	N3m
	99 Don't know	N3m
	IF N3l >7, THEN ASK	
N3ll	Why do you say that?	
	77 Record VERBATIM	N3m
	88 Refused	N3m
	99 Don't know	N3m
N3m.	Corporate policy or guidelines	
	# Record 0 to 10 score (_____)	N3n.
	88 Refused	N3n.
	99 Don't know	N3n.
N3n.	Payback on the investment	
	# Record 0 to 10 score (_____)	N3o.
	88 Refused	N3o.
	99 Don't know	N3o.
N3o.	Were there any other factors we haven't discussed that were influential in your decision to install this MEASURE?	
	1 Nothing else influential	N33
	77 Record verbatim	N3oo
	88 Refused	N33
	99 Don't know	N33
N3oo.	Using the same zero to 10 scale, how would you rate the influence of this factor?	
	# Record 0 to 10 score (_____)	N33
	88 Refused	N33
	99 Don't know	N33

IF N3n.>5, THEN ASK, ELSE CP1

PAYBACK BATTERY (If payback importance >5)

P1	What financial calculations does your company make before proceeding with installation of a MEASURE like this one?	
	77 Record VERBATIM	P2
	88 Don't know	P2
	99 Refused	P2
P2	What is the payback cut-off point your company uses (in months) before deciding to proceed with an investment?	
	1 0 to 6 months	P3a
	2 6 months to 1 year	P3a
	3 1 to 2 years	P3a
	4 2 to 3 years	P3a
	5 3 to 5 years	P3a
	6 Over 5 years	P3a
	88 Don't know	P3a
	99 Refused	P3a
P3a	What was the payback calculation for &MEASURE:(in months) with the rebate from &PROGRAM?	
	# payback in months (___ months) with rebate	P3b.
	88 Don't know	P3b.
	99 Refused	P3b.
P3b	And what was the payback calculation for &MEASURE:(in months) without the rebate from &PROGRAM?	
	# payback in months (___ months) without rebate	P3c
	88 Don't know	CP1
	99 Refused	CP1
	IF P3b<P2, THEN ASK.	
	"Even without the rebate, the &MEASURE project met your company's financial criteria. Would you have gone ahead with it even without the rebate?"	
P3c	77 Record VERBATIM	P3d
	88 Don't know	P3d
	99 Refused	P3d
	IF P3a<P2, AND N3b<5, THEN ASK.	
	"The rebate seemed to make the difference between meeting your financial criteria and not meeting them, but you are saying that the rebate didn't have much effect on your decision, why is that?"	
P3d	77 Record VERBATIM	P3e
	88 Don't know	P3e
	99 Refused	P3e
	IF P3a>P2, AND N3b>7, THEN ASK.	
	"The rebate didn't cause this &MEASURE to meet your company's financial criteria, but you said that the rebate had an impact on the decision to install &MEASURE. Why did it have an impact?"	
P3e.	77 Record VERBATIM	CP1
	88 Don't know	CP1
	99 Refused	CP1
	IF N3m.>5, THEN ASK, ELSE SP1	

CORPORATE POLICY BATTERY (If corporate policy importance >5)

CP1	Does your organization have a corporate environmental policy to reduce environmental emissions or energy use? Some examples would be to "buy green" or use sustainable approaches to business investments. 1 Yes [CAN I OBTAIN A COPY OF THE POLICY?] 2 No 88 Don't know 99 Refused	CP2 SP1 SP1 SP1
CP2	What specific corporate policy influenced your decision to adopt or install the &MEASURE? 1 RECORD VERBATIM [IF NOT ALREADY ASKED IN CP1: CAN I OBTAIN A COPY OF THE POLICY?] 88 Don't know 99 Refused	CP3 CP3 CP3
CP3	Had that policy caused you to adopt the &MEASURE at this facility before participating in the &PROGRAM? 1 Yes 2 No 88 Don't know 99 Refused	CP4 CP4 CP4 CP4
CP4	Had that policy caused you to adopt the &MEASURE at other facilities before participating in the &PROGRAM? 1 Yes [RECORD Locations and Dates] 2 No 88 Don't know 99 Refused	CP5 CP5 CP5 CP5
CP5	Did you receive an incentive for a previous installation of &MEASURE? If so, please describe the amount of incentive received, the approximately timing, and the name of the program that provided it. 77 Record VERBATIM 88 Don't know 99 Refused	CP6 CP6 CP6
CP6	IF CP3=1 OR CP4=1, THEN ASK. If I understand you correctly, you said that your company's corporate policy has caused you to adopt &MEASURE previously at this and/or other facilities. I want to make sure I fully understand how this corporate policy influenced your decision versus the &PROGRAM. Can you please clarify that? 77 Record VERBATIM 88 Don't know 99 Refused	SP1 SP1 SP1

IF N3j.>5, THEN ASK, ELSE O1**STANDARD PRACTICE BATTERY (If standard practice importance >5)**

SP1	Approximately, how long has &MEASURE been standard practice in your industry? # Record Number of Months or Years 88 Don't know 99 Refused	SP2 SP2 SP2 SP2
SP2	Does your company ever deviate from the standard practice? 1. Yes [Under what conditions does your company deviate?] RECORD VERBATIM: _____ 2 No 88 Don't know 99 Refused	SP3 SP3 SP3 SP3
SP3	How did this standard practice influence your decision to install the &MEASURE? 77 Record VERBATIM 88 Don't know 99 Refused	SP3a SP3a SP3a

	Could you please rate the importance of the &PROGRAM, versus this standard industry practice in influencing your decision to install &MEASURE. Would you say the &PROGRAM was much more important, somewhat more important, equally important, somewhat less important, or much less important than the standard practice?	
SP3a	1 Much more important	SP4
	2 Somewhat more important	SP4
	3 Equally important	SP4
	4 Somewhat less important	SP4
	5 Much less important	SP4
	88 Don't know	SP4
	99 Refused	SP4
SP4	What industry group or trade organization do you look to to establish standard practice for your industry?	
	77 Record VERBATIM	SP5
	88 Don't know	SP5
	99 Refused	SP5
SP5	How do you and other firms in your industry receive information on updates in standard practice?	
	77 Record VERBATIM	O11
	88 Don't know	O11
	99 Refused	O11

IF N3o.>5, THEN ASK, ELSE N33.

OTHER INFLUENCES BATTERY (If other influences importance >5)

	Who provided the most assistance in the design or specification of &MEASURE? [DO NOT READ: Was it: the Designer, the Consultant, the Equipment Distributor, the Mfr Rep, the Installer, the Utility rep, or Internal staff?]	
O11	1 Designer	O12
	2 Consultant	O12
	3 Equipment distributor	O12
	4 Installer	O12
	5 &UTILITY account representative	O12
	6 &PROGRAM staff	O12
	77 Other: (Record VERBATIM)	O12
	88 Don't know	O12
	99 Refused	O12
O12	Please describe the type of assistance that they provided.	O13
	77 Record VERBATIM	O13
	88 Don't know	O13
	99 Refused	O13
O13	Please state, in your own words, any other factors that influenced your decision to go ahead on this energy efficiency project?	
	77 Record VERBATIM	N33.
	88 Don't know	N33.
	99 Refused	N33.

NET-TO-GROSS QUESTIONS (CONTINUED)

IF ACCT_REP = 1, ACCTREPNAME:= 0, THEN ASK.

N33 We do not have the name of your ACCOUNT REP at <%UTILITY>. Can you give me his or her name?

!! ___ Do you have his/her email address?

! ___ Do you have a phone number for him/her?

! ___ Do you have a cell phone number for him/her? \,

77 RECORD NAME, Phone, Email ETC

88 Refused

99 Don't know

N41

N41

N41

Revision
Revision
Revision

!!! ___ For the sake of expediency, we are referring to the ... <%PROGRAM> ... as the PROGRAM and we are referring to the installation of ...<%MEASURE>... as the MEASURE.

!! ___ I will repeat this from time to time during the study as your organization may have installed more than one measure through more than one program. \;

Next, I would like you to rate the importance of the PROGRAM in your decision to implement this MEASURE as opposed to other factors that may have influenced your decision such as...(SCAN BELOW AND READ TO THEM THOSE ITEMS WHERE THEY GAVE A RATING OF 8 or higher)

! <%N3A> Age or condition of old equipment,

! <%N3D> Equipment Vendor recommendation

! <%N3E> Previous experience with this measure

! <%N3F> Previous experience with this program

! <%N3I> Recommendation from a design or consulting engineer

! <%N3J> Standard practice in your business/industry

! <%N3M> Corporate policy or guidelines

! <%N3N> Payback on investment.

If you were given 10 points to award in total, how many points would give to the importance of the program and how many points would you give to these other factors? \

N41 How many of the ten points would you give to the importance of the PROGRAM in your decision?

Record 0 to 10 score (_____)

88 Refused

99 Don't know

N42

N42

N42

N42 and how many points would you give to these other factors? \

Record 0 to 10 score (_____)

88 Refused

99 Don't know

N41a

N41a

N41a

___ We want these two sets of numbers to equal 10.

! <%N41> for Program influence and

! <%N42> for Non Program factors

CONSISTENCY CHECK ON PGM IMPORTANCE SCORE

IF N41 &PROGRAM>6 AND N3b, N3c, N3g, N3h, N3k AND N3l ALL<4, THEN ASK N41a. ELSE IF N41 &PROGRAM<4 AND N3b OR N3c OR N3g OR N3h OR N3k OR N3l>6, THEN ASK N41b. OTHERWISE SKIP TO N5.

When you scored the importance of the program as <%N41>, I would interpret that to mean that the program was quite important to your decision to install this equipment. Earlier, when I asked about the importance of individual elements of the program I recorded some answers that would imply that certain elements of the program were not that important to you. Just to make sure I have recorded this properly, may I please take a second to review?

N41a	77 Record VERBATIM	N5
	88 Don't know	N5
	99 Refused	N5
	IF N3b<4, THEN ASK	

When I asked you about THE AVAILABILITY OF THE PROGRAM REBATE, you gave a rating of ...<%N3B> ... out of ten, indicating that the program rebate was not that important to you. Can you tell me why the rebate was not that important?

N41aa	77 Record VERBATIM	N41ab
	88 Don't know	N41ab
	99 Refused	N41ab

IF N3c<4, THEN ASK

When I asked you about THE INFORMATION PROVIDED THROUGH

!! __<(FEAS_STUDY == 1)/ The Feasibility study/>

! __<(AUDIT == 1)/The Facility or System AUDIT/>

! __<(TECH_ASST == 1)/The Technical Assistance/> !

you gave a rating of ...<%N3C> ... out of ten, indicating that the information provided was not that important to you. Can you tell me why the information provided was not that important?

N41ab	77 Record VERBATIM	N41ac
	88 Don't know	N41ac
	99 Refused	N41ac

N41ac	IF N3g<4, THEN ASK When I asked you about THE INFORMATION FROM THE PROGRAM or UTILITY TRAINING COURSES, you gave a rating of ...<%N3G> ... out of ten, indicating that the information from the program or utility training course was not that important to you. Can you tell me why this information was not that important?	
	77 Record VERBATIM	N41ad
	88 Don't know	N41ad
	99 Refused	N41ad
	IF N3h<4, THEN ASK When I asked you about THE INFORMATION from the PROGRAM or UTILITY MARKETING MATERIALS, you gave a rating of ...<%N3H> ... out of ten, indicating that this information from the program or utility marketing materials was not that important to you. Can you tell me why this information was not that important?	
N41ad	77 Record VERBATIM	N41ae
	88 Don't know	N41ae
	99 Refused	N41ae
	IF N3k<4, THEN ASK When I asked you about THE ENDORSEMENT or RECOMMENDATION by PROGRAM STAFF or PROGRAM VENDOR, you gave a rating of ...<%N3K> ... out of ten, indicating that this program endorsement was not that important to you. Can you tell me why this program endorsement was not that important?	
N41ae	77 Record VERBATIM	N41af
	88 Don't know	N41af
	99 Refused	N41af
	IF N3l<4, THEN ASK When I asked you about THE ENDORSEMENT or RECOMMENDATION by YOUR ACCOUNT REP ...<%ACCT_REP_NAME>, you gave a rating of ...<%N3L> ... out of ten, indicating that this Account Rep endorsement was not that important to you. Can you tell me why this endorsement was not that important?	
N41af	77 Record VERBATIM	N41b
	88 Don't know	N41b
	99 Refused	N41b
	IF N41 &PROGRAM<4 AND N3b OR N3c OR N3g OR N3h OR N3k OR N3l>6, THEN ASK N41b. OTHERWISE SKIP TO N5. When you scored the importance of the program as <%N41>, I would interpret that to mean that the program was not very important to your decision to install this equipment. Earlier, when I asked about the importance of individual elements of the program I recorded some answers that would imply that certain elements of the program were very important to you. Just to make sure I have recorded this properly, will you please state in your own words why you feel the program was not very important?	
N41b	77 Record VERBATIM	N5
	88 Don't know	N5
	99 Refused	N5
	Now I would like you to think about the action you would have taken with regard to the installation of this equipment if the &PROGRAM had not been available. Using a likelihood scale from 0 to 10, where 0 is "Not at all likely" and 10 is "Extremely likely", if the &PROGRAM had not been available, what is the likelihood that you would have installed exactly the same equipment?	
N5	# Record 0 to 10 score (_____)	N5a.
	88 Refused	N6
	99 Don't know	N6

CONSISTENCY CHECKS

	IF N3b>7 and N5>7, THEN ASK.	
	When you answered ...<%N3B> ... for the question about the influence of the rebate, I would interpret that to mean that the rebate was quite important to your decision to install. Then, when you answered ..<%N5>... for how likely you would be to install the same equipment without the rebate, it sounds like the rebate was not very important in your installation decision. I want to check to see if I am misunderstanding your answers or if the questions may have been unclear. Will you explain in your own words, the role the rebate played in your decision to install this efficient equipment?	
N5a	77 Record VERBATIM	N5aa
	88 Don't know	N5aa
	99 Refused	N5aa
	Would you like for me to change your score on the importance of the rebate that you gave a rating of <%N3B> and/or change your rating on the likelihood you would install the same equipment without the rebate which you gave a rating of <%N5> and/or we can change both if you wish?	
N5aa	77 Record VERBATIM	SP3a
	88 Don't know	SP3a
	99 Refused	SP3a

PROBE ON STANDARD PRACTICE if n3>7, ELSE ASK N9

	In an earlier question, you rated the importance of STANDARD PRACTICE in your industry very highly in your decision making. Could you please rate the importance of the PROGRAM, relative to this standard industry practice, in influencing your decision to install this MEASURE. Would you say the program was much more important, somewhat more important, equally important, somewhat less important, or much less important than the standard practice or policy?	
SP3a	1 Much more important	N9
	2 Somewhat more important	N9
	3 Equally important	N9
	4 Somewhat less important	N9
	5 Much less important	N9
	88 Don't know	N9
	99 Refused	N9
	IF N5>0, THEN ASK.	
	You indicated in your response to a previous question that there was a <%N5> in 10 likelihood that you would have installed the same equipment if THE PROGRAM had not been available. When do you think you would have installed this equipment? Please express your answer in months.	
N9	a. at the same time	TD1
	b. within _____ .months	N9b
	c. Never	N6
	88 Refused	N6
	99 Don't know	N9a.
N9a.	If respondent is having difficulty specifying answer in months...would it have been..	
	a. _____ ..within 6 months?	TD1
	b. _____.. 6 months to 1 year later	TD1
	c. _____.. 1 - 2 years later	TD1
	d. _____ ..2 - 3 years later?	TD1
	e. _____ ..3 - 4 years later?	TD1
	f. _____ ..4 or more years later	N9b
	88 Don't know	N6
	99 Refused	N6
	IF N9>=48 months OR N9a=response f, THEN ASK N9b, ELSE ASK N6.	
N9b.	Why do you think it would have been 4 or more years later?	
	77 Record VERBATIM	TD1
	88 Don't know	TD1
	99 Refused	TD1

DEFERRED FREE RIDERSHIP FOLLOW-UP

INTRO You said that there was an <N5> in 10 likelihood that you would have installed the same equipment about <&N9> months later
 FOR BOTH (OR at the same time) if the PROGRAM had not been available. I'd like to ask a couple of questions to help us estimate at what
 TD1 and point in the future you would definitely have installed new equipment. We understand that you can't know exactly when you
 TD1a would have done this, especially so far into the future. We're just trying to get a sense of how long you think the current
 equipment or process would have kept serving your company's needs before you had to or chose to replace it.

If N9 or N9a ≤ 60 months, ask TD1, ELSE TD1A

TD1 So, again using a 0 to 10 scale, where 0 means not at all likely and 10 means extremely likely, what is the likelihood that you
 would have installed the same equipment within 60 months, or 5 years, later if the program had not been available?

Record 0 to 10 score (_____)

88 Refused

99 Don't know

IF <10 ASK TD2, ELSE GO TO N5a

TD2

TD1A

TD1A

TD2 And what would you say is the likelihood that you would have installed the same equipment within 120 months, or 10 years, later
 if the program had not been available?

Record 0 to 10 score (_____)

88 Refused

99 Don't know

TD1A

TD1A

TD1A

If N9 or N9a > 60 months, ask

TD1A Now, using the same 0 to 10 scale, where 0 means not at all likely and 10 means extremely likely, what is the likelihood that you
 would have installed the same equipment within 120 months, or 10 years, later if the program had not been available?

Record 0 to 10 score (_____)

88 Refused

99 Don't know

N9bb

N9bb

N9bb

CONSISTENCY CHECK ON AGE

IF N3a>6 AND N9>=48 months OR N9a=response f, THEN ASK. ELSE N6.

Earlier when asked about the influence of the age/condition of the old equipment on your decision to install this new equipment,
 you gave me a rating of <%N3A> out of ten. I would interpret this to mean that the age/condition was quite influential in your
 decision to install this new equipment when you did. Perhaps I have either recorded something incorrectly or maybe you could
 explain in your own words the role the age/condition of the existing equipment played in your decision to install this new energy-
 efficient equipment.

N9bb

77 Record VERBATIM

88 Don't know

99 Refused

N6

N6

N6

Revision

PARTIAL FREE RIDERSHIP

	Now I would like you to think one last time about what action you would have taken if the program had not been available. Supposing that you had not installed the program qualifying equipment, which of the following alternatives would you have been MOST likely to do?	
N6	1 Install fewer units	N6a
	2 Install standard efficiency equipment or whatever required by code	SP1
	3 Install equipment more efficient than code but less efficient than what you installed through the program	N6b
	4 repair/rewind or overhaul the existing equipment	N6c
	5 do nothing (keep the existing equipment as is)	SP1
	6 something else (specify what _____)	SP1
	88 Don't know	SP1
	99 Refused	SP1
N6a	How many fewer units would you have installed? (It is okay to take an answer such as ...HALF...or 10 percent fewer ... etc.)	
	77 RECORD VERBATIM	SP1
	88 Refused	SP1
	99 Refused	SP1
N6b	Can you tell me what model or efficiency level you were considering as an alternative? (It is okay to take an answer such as ... 10 percent more efficient than code or 10 percent less efficient than the program equipment)	
	77 RECORD VERBATIM	SP1
	88 Don't know	SP1
	99 Refused	SP1
N6c	How long do you think the repaired/rewound/refurbished equipment would have lasted before requiring replacement?	
	77 RECORD VERBATIM	SP1
	88 Don't know	SP1
	99 Refused	SP1

SPILLOVER QUESTIONS

SP1	Did you implement any additional energy efficiency measures at this facility since your participation in the 2006-2008 Program and before the end of 2008 that did not receive incentives through any utility or government program?		Revision
	1 Yes	SP2	
	2 No	CAFAC1	
	88 Refused	CAFAC1	
	99 Don't know	CAFAC1	
SP2	What was the first Measure that you implemented?		
	77 Record FIRST measure	SP3	
	88 Refused	CAFAC1	
	99 Don't know	CAFAC1	
SP3	What was the second measure?		
	77 Record SECOND measure	SP4	
	88 Refused	SP4	
	99 Don't know	SP4	
SP4	What was the third measure?		
	77 Record THIRD measure	SP5	
	88 Refused	SP5	
	99 Don't know	SP5	
SP5	I have a few questions about the FIRST Measure that you installed. Why are you not expecting a rebate for this measure? Why did you not install this measure through a Utility Program?		
	77 Record VERBATIM	SP5b	
	88 Don't know	SP5b	
	99 Refused	SP5b	
SP5b	Please describe the SIZE, The EFFICIENCY and QUANTITY of this measure.		
	77 Record VERBATIM	SP5c	
	88 Don't know	SP5c	
	99 Refused	SP5c	
SP5c.	Was this measure specifically recommended by a PROGRAM related audit, report or program technical specialist?		
	1 Yes	SP5d	
	2 No	SP5d	
	88 Refused	SP5d	
	99 Don't know	SP5d	
SP5d.	How significant was your experience in the 2006--2008 Program in your decision to implement this Measure, using a scale of 0 to 10, where 0 is not at all significant and 10 is extremely significant?		
	# Record 0 to 10 score (_____)	SP5dd	
	88 Refused	SP5e	
	99 Don't know	SP5e	
SP5dd.	Why do you give it this rating?		
	77 Record VERBATIM	SP5e	
	88 Don't know	SP5e	
	99 Refused	SP5e	
SP5e.	If you had not participated in the 2006-2008 program, how likely is it that your organization would still have implemented this measure, using a 0 to 10 scale where 0 means you definitely WOULD NOT have implemented this measure and 10 means you definitely WOULD have implemented this measure?		
	# Record 0 to 10 likelihood rating (_____)	SP5f	
	88 Refused	SP5f	
	99 Don't know	SP5f	

SP6	<p>I have a few questions about the SECOND Measure that you installed. Why are you not expecting a rebate for this measure? Why did you not install this measure through a Utility Program?</p> <p>77 Record VERBATIM 88 Don't know 99 Refused</p>	<p>SP6b SP6b SP6b</p>
SP6b	<p>Please describe the SIZE, The EFFICIENCY and QUANTITY of this measure.</p> <p>77 Record VERBATIM 88 Don't know 99 Refused</p>	<p>SP6c SP6c SP6c</p>
SP6c.	<p>Was this measure specifically recommended by a PROGRAM related audit, report or program technical specialist?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>SP6d SP6d SP6d SP6d</p>
SP6d.	<p>How significant was your experience in the 2006--2008 Program in your decision to implement this Measure, using a scale of 0 to 10, where 0 is not at all significant and 10 is extremely significant?</p> <p># Record 0 to 10 score (_____)</p> <p>88 Refused 99 Don't know</p>	<p>SP6dd SP6e SP6e</p>
SP6dd.	<p>Why do you give it this rating?</p> <p>77 Record VERBATIM 88 Don't know 99 Refused</p>	<p>SP6e SP6e SP6e</p>
SP6e.	<p>If you had not participated in the 2006-2008 program, how likely is it that your organization would still have implemented this measure, using a 0 to 10 scale where 0 means you definitely WOULD NOT have implemented this measure and 10 means you definitely WOULD have implemented this measure?</p> <p># Record 0 to 10 likelihood rating (_____)</p> <p>88 Refused 99 Don't know</p>	<p>SP7 SP7 SP7</p>
SP7	<p>I have a few questions about the THIRD Measure that you installed. Why are you not expecting a rebate for this measure? Why did you not install this measure through a Utility Program?</p> <p>77 Record VERBATIM 88 Don't know 99 Refused</p>	<p>SP7b SP7b SP7b</p>
SP7b	<p>Please describe the SIZE, The EFFICIENCY and QUANTITY of this measure.</p> <p>77 Record VERBATIM 88 Don't know 99 Refused</p>	<p>SP7c SP7c SP7c</p>
SP7c.	<p>Was this measure specifically recommended by a PROGRAM related audit, report or program technical specialist?</p> <p>1 Yes 2 No 88 Refused 99 Don't know</p>	<p>SP7d SP7d SP7d SP7d</p>
SP7d.	<p>How significant was your experience in the 2006--2008 Program in your decision to implement this Measure, using a scale of 0 to 10, where 0 is not at all significant and 10 is extremely significant?</p> <p># Record 0 to 10 score (_____)</p> <p>88 Refused 99 Don't know</p>	<p>SP7dd SP7e SP7e</p>

SP7dd.	Why do you give it this rating? 77 Record VERBATIM 88 Don't know 99 Refused	SP7e SP7e SP7e
SP7e.	If you had not participated in the 2006-2008 program, how likely is it that your organization would still have implemented this measure, using a 0 to 10 scale where 0 means you definitely WOULD NOT have implemented this measure and 10 means you definitely WOULD have implemented this measure? # Record 0 to 10 likelihood rating (_____) 88 Refused 99 Don't know	CAFAC1 CAFAC1 CAFAC1
CAFAC1	Now, thinking about other facilities operated by your organization in the regions of California that are served by PG&E, SCE, SDG&E or Southern California Gas Company , are you aware of any additional energy efficiency measures implemented at these other facilities since your participation in the 2006-2008 program and before the end of 2008 that did not receive an incentive through a utility or government program? 1 Yes 2 No 88 Refused 99 Don't know	CAFAC2 C1 C1 C1
CAFAC2	What was the first Measure that you implemented? 77 Record FIRST MEASURE 88 Refused 99 Don't know	CAFAC3 CAFAC3 CAFAC3
CAFAC3	What was the second measure? 77 Record SECOND MEASURE 88 Refused 99 Don't know	CAFAC4 CAFAC4 CAFAC4
CAFAC4	What was the third measure? 77 Record THIRD MEASURE 88 Refused 99 Don't know	MEAS1_1 MEAS1_1 MEAS1_1
MEAS1_1	IF CAFAC1=1, THEN ASK, ELSE C1 I have a few questions about .the FIRST MEASURE that you installed. Was this measure part of a <%UTILITY> program or any other utility or government energy efficiency incentive Program? 1 Yes 2 No 88 Refused 99 Don't know	MEAS2_1 MEAS1_2 MEAS2_1 MEAS2_1
MEAS1_2	Why did you not install this measure through a Utility Program? 77 Record VERBATIM 88 Don't know 99 Refused	MEAS1_3 MEAS1_3 MEAS1_3
MEAS1_3	Please describe the SIZE, The EFFICIENCY and QUANTITY of this measure. 77 Record VERBATIM 88 Don't know 99 Refused	MEAS1_4 MEAS1_4 MEAS1_4
MEAS1_4	Was this measure specifically recommended by a PROGRAM related audit, report or program technical specialist? 1 Yes 2 No 88 Refused 99 Don't know	MEAS1_5 MEAS1_5 MEAS1_5 MEAS1_5

Revision

MEAS1_5	How significant was your experience in the 2006--2008 Program in your decision to implement this Measure, using a scale of 0 to 10, where 0 is not at all significant and 10 is extremely significant?	
	# Record 0 to 10 score (_____)	MEAS1_6
	88 Refused	MEAS1_7
	99 Don't know	MEAS1_7
MEAS1_6	Why do you give it this rating?	
	77 Record VERBATIM	MEAS1_7
	88 Don't know	MEAS1_7
	99 Refused	MEAS1_7
MEAS1_7	If you had not participated in the 2006-2008 program, how likely is it that your organization would still have implemented this measure, using a 0 to 10 scale where 0 means you definitely WOULD NOT have implemented this measure and 10 means you definitely WOULD have implemented this measure?	
	# Record 0 to 10 likelihood rating (_____)	MEAS2_1
	88 Refused	MEAS2_1
	99 Don't know	MEAS2_1
	IF CAFAC2=1, THEN ASK, ELSE C1	
MEAS2_1	I have a few questions about .the SECOND MEASURE.that you installed. Was this measure part of a <%UTILITY> program or any other utility or government energy efficiency incentive Program?	
	1 Yes	MEAS3_1
	2 No	MEAS2_2
	88 Refused	MEAS3_1
	99 Don't know	MEAS3_1
MEAS2_2	Why did you not install this measure through a Utility Program?	
	77 Record VERBATIM	MEAS2_3
	88 Don't know	MEAS2_3
	99 Refused	MEAS2_3
MEAS2_3	Please describe the SIZE, The EFFICIENCY and QUANTITY of this measure.	
	77 Record VERBATIM	MEAS2_4
	88 Don't know	MEAS2_4
	99 Refused	MEAS2_4
MEAS2_4	Was this measure specifically recommended by a PROGRAM related audit, report or program technical specialist?	
	1 Yes	MEAS2_5
	2 No	MEAS2_5
	88 Refused	MEAS2_5
	99 Don't know	MEAS2_5
MEAS2_5	How significant was your experience in the 2006--2008 Program in your decision to implement this Measure, using a scale of 0 to 10, where 0 is not at all significant and 10 is extremely significant?	
	# Record 0 to 10 score (_____)	MEAS2_6
	88 Refused	MEAS2_7
	99 Don't know	MEAS2_7
MEAS2_6	Why do you give it this rating?	
	77 Record VERBATIM	MEAS2_7
	88 Don't know	MEAS2_7
	99 Refused	MEAS2_7
MEAS2_7	If you had not participated in the 2006-2008 program, how likely is it that your organization would still have implemented this measure, using a 0 to 10 scale where 0 means you definitely WOULD NOT have implemented this measure and 10 means you definitely WOULD have implemented this measure?	
	# Record 0 to 10 likelihood rating (_____)	MEAS3_1
	88 Refused	MEAS3_1
	99 Don't know	MEAS3_1

IF CAFAC3=1, THEN ASK, ELSE C1

I have a few questions about .the THIRD MEASURE.that you installed. Was this measure part of a <-%UTILITY> program or any other utility or government energy efficiency incentive Program?

MEAS3_1
1 Yes C1
2 No MEAS3_2
88 Refused C1
99 Don't know C1

MEAS3_2 Why did you not install this measure through a Utility Program?
77 Record VERBATIM MEAS3_3
88 Don't know MEAS3_3
99 Refused MEAS3_3

MEAS3_3 Please describe the SIZE, The EFFICIENCY and QUANTITY of this measure.
77 Record VERBATIM MEAS3_4
88 Don't know MEAS3_4
99 Refused MEAS3_4

MEAS3_4 Was this measure specifically recommended by a PROGRAM related audit, report or program technical specialist?
1 Yes MEAS3_5
2 No MEAS3_5
88 Refused MEAS3_5
99 Don't know MEAS3_5

How significant was your experience in the 2006--2008 Program in your decision to implement this Measure, using a scale of 0 to 10, where 0 is not at all significant and 10 is extremely significant?
MEAS3_5
Record 0 to 10 score (_____) MEAS3_6
88 Refused MEAS3_7
99 Don't know MEAS3_7

MEAS3_6 Why do you give it this rating?
77 Record VERBATIM MEAS3_7
88 Don't know MEAS3_7
99 Refused MEAS3_7

If you had not participated in the 2006-2008 program, how likely is it that your organization would still have implemented this measure, using a 0 to 10 scale where 0 means you definitely WOULD NOT have implemented this measure and 10 means you definitely WOULD have implemented this measure?
MEAS3_7
Record 0 to 10 likelihood rating (_____) C1
88 Refused C1
99 Don't know C1

And finally, I have a few questions about the characteristics of your business.

C1. Our records indicate that the primary business code for the facility that installed &MEASURE is &NAICS. Is that correct?
1 Yes C2
2 No C2
88 Don't know C2
99 Refused C2

C2. Please describe the type of work performed at this facility and/or the primary product made or main service provided.
77 Record VERBATIM C3
88 Don't know C3
99 Refused C3

C3. Please describe any changes made to this site since January 2006 that significantly impacted energy usage.
77 Record VERBATIM END
88 Don't know END
99 Refused END

Premise General Information

Please answer the following questions

C4. What kind of premise is this?: P = Part of a bldg B = 1 building, single footprint MF = 1 building w/multiple footprints SM = Small multi-building CM = Campus (multi-bldg) OT = Other	P B MF SM CM OT
C5. What is the total occupied floor area of this premise (excluding enclosed parking garage area)?	_____ ft ²
C5a. If the premise has an enclosed parking garage, approximately what is the floor area?	_____ ft ²
C6. How many buildings are part of this premise?	_____
C7. Is this premise owner-occupied (O) or leased (L)?	O L
C8. What year was this business established at this location?	----
C9. How many full-time equivalent employees work at this premise?	_____

END Those are all the questions I have for you. On behalf of the CPUC, thank you very much for your time.

END OF SURVEY

Business/Building Type Codes

Appendix E

Supplementary POC Gross Impact Methods and Results

Energy and demand impacts resulting from the retrofit of rod beam oil well pumps with pump-off controllers (POCs) are based on site specific measurement activities. Site specific measurement was performed at a number of oil fields operated by a variety of companies throughout California.

There is considerable uncertainty surrounding the accuracy of ex-ante energy and demand impacts attributable to pump-off controllers. Motor power draw and time-of-use are the most important determinants of energy consumption and energy impact. The ex-ante method for calculating energy impacts for POC applications typically incorporates a deemed motor load factor of 20% for both the pre-retrofit (or baseline) case and the post-retrofit operating condition. The utility-run PGE2004 program incorporates a runtime factor based on annual and theoretical production based on well characteristics. Ex-ante demand impacts were calculated by distributing savings over operating hours.

Another important component of impact is the change in motor load factor and kW drawn by the motor in well operation that may occur as a result of the installation of the POC.

Finally, it is important to consider oil well production changes that are unrelated to a POC retrofit. Such changes in production may indicate it is appropriate to normalize production to baseline or post-retrofit levels for the appropriate presentation of measure impact.

The analysis undertaken for this evaluation helps to clarify uncertainties and to determine appropriate impact levels for POC applications. The analysis identifies key parameters for pre and post retrofit data collection activities, and provides details regarding their contribution to the accuracy of impact measurement.

Engineering Approach for POCs

The steps for conducting site-specific gross impact analysis for POCs are described in this appendix. The reader may find additional details on the ex-post engineering approach and rationale in the site report appendix, and projects B001 and B007 in particular.

Measure Description

The measure involves the installation of pump-off controllers (POCs) on rod beam pumps for various oil well fields. Normal operation of the pump motors without pump-off controllers (POCs) is constant (24 hours per day, 7 days per week), except for downtime due to maintenance or repair, and for wells already retrofit with control systems such as older vintage POCs or time clocks. If operating continuously, the constant operation of the pump motors occurs whether oil is efficiently brought to the surface or not, thus wasting energy. A well is said to be ‘pumped off’ when there is little or no fluid at the pump barrel in the reservoir, and therefore low quantities of fluid are brought to the surface. When this occurs, the pump becomes essentially oversized for the current available well production, and is removing fluid faster than the fluid flows to the well bore.

A pump-off controller (POC) will sense when the well is ‘pumped off’ and shut off the pump motor temporarily. While the motor is de-energized, the subsurface fluid recharges (or fills) the well. Sensing is typically performed through a strain gage or ‘load cell’ on the rod beam. This strain corresponds to ‘fillage’ in the well and the amount of fluid being pumped. The off period is manually programmed based on the time the well requires to recharge to a level sufficient for effective pumping (typically initially estimated by field personnel but later often optimized through a reset feature on certain POCs). Electrical usage (kWh) savings are achieved when the POCs shut off the pump motors. Electrical demand (kW) savings generally accrue since the pumps are equally likely to be de-energized at any given time interval. For an individual well, this demand savings is distributed evenly throughout the year, as the well pump motor operation is not affected by daily, weekly, or seasonal patterns. It should be noted that a pump under active POCs, but in an energized mode – during the intervals when the pump is energized and running, the motors will draw *more* power than under the uncontrolled conditions. Nonetheless, on a well field basis, a large scale deployment of POCs will often result in a significant percentage of pumps which will be de-energized at any one time, reducing both the maximum coincident load and any peak demand period of interest.

Measurement & Verification

A primary objective of work at the oil fields is the continuous economical production of oil. Wells are maintained and operated to optimize delivery of the oil/water fluid to the surface by pumping from subsurface deposits. Optimization for maximum oil production for the minimum cost to produce the oil is desirable.

The variability of subsurface characteristics of the oil bearing geology cause unique operation at each well. These unique parameters influence the "on" and "off" periods of the motors which the POCs control, and also affect the pump power required. Measurements of on and off periods are

thus a primary determinant of post-retrofit and pre-retrofit energy use and the resulting energy savings.

POC applications from a single customer and a single well field that use a single ex-ante impact calculation methodology were grouped for M&V purposes. For sampling within each group, wells were tiered (typically into five strata) based on the level of expected energy savings of each well. Sample selection was determined using stratified random sampling and based on a 90 confidence interval with 20% relative precision.

Approach and Formulae

The data collection tools and procedures were designed to accommodate the variations in load from the upstroke and downstroke of the rod beam and regenerative power. Typically, there are 6 to 12 strokes per minute (5 to 10 seconds per complete stroke). The regenerative portion of the stroke, in which power is sent back into the electrical grid, typically accounts for less than 1 second per stroke (10 to 20 percent of total time). Dent Elite Pro_{tm} devices and Wattnode_{tm} pulse loggers were both utilized in data collection. Short term data collection entailed the use of NIST traceable calibrated Fluke_{tm} 1735 and Powersight_{tm} 3000 meters.

Two types of loggers were used to record energy consumption and load characteristics of oil well pumps, the DENT Elite Pro_{tm} loggers and the Hobo Energy Logger Pro_{tm}. The Hobo Energy Logger Pro_{tm} data logging devices use a pulse counter (Wattnode_{tm}) to measure kWh continuously at high sampling intervals. These loggers were set up to record data every 5 minutes, typically for a 14 day period or longer. The pulse counting device captures the waveform continuously and samples at a rate of over 1,000 times per second. The pulse counting devices are accurate for kWh measurement over short periods, but have some notable limitations. Limitations of the pulse counting device include the inability to record power factor, inaccuracy in the measurement of absolute kW.

The DENT Elite Pro_{tm} loggers were set up to collect the pump motor kW, along with voltage, amps and power factor to support the best possible understanding of energy usage characteristics. The DENT_{tm} loggers sample the waveform every three seconds. Typical deployments involved recording the average of these sampling intervals every five minutes. These loggers are accurate to approximately 5% over periods which exceed eight hours (for shorter durations, the capture of the electrical parameters only once every three seconds results in a decrease in accuracy for this type of cyclical and variable load).

Both logging devices measure regenerative power. Regenerative power varies from well to well depending on the balancing of the well. Typically, several wells are connected to one meter, such that the energy introduced into the grid by one of the wells in a group will flow to other nearby wells rather than back to the utility meter.

Baseline (pre-retrofit) simulation and energy consumption measurement was performed at the Kern and Midway oil fields. Baseline conditions were simulated and recorded by metering during a period where the POC devices were disabled. Metering was also used at these fields to measure energy consumption patterns under post-retrofit conditions, with the POC devices enabled. The baseline (pre-retrofit) power draw (kW) was collected by over-riding the POC and allowing the well to simulate the pre-retrofit operation for a period of one week at Kern fields and for a period of two weeks at Midway fields. This test was conducted for eighteen (18) wells in the Kern fields and eighteen (18) wells in the Midway fields. It is reasonable to expect that most wells will stabilize to a baseline or pre-retrofit mode of operation immediately after the well is pumped off. However, longer logging periods allowed for more accurate simulation of performance over the continuous operation of the well pumps under a baseline condition.

At the Kern and Midway oil fields, metering results indicate that the average cycle rate was 8.9 strokes per minute (SPM) for the sample. The average duration for each complete stroke, including the upstroke and downstroke period, was 6.7 seconds, and thus two or three data points were obtained by the DENT loggers for each stroke. Approximately 80,000 to 100,000 strokes for each well were measured to calculate the baseline and the post-retrofit kW. The number of strokes measured for each well provides a representative average kW for the monitored wells. In order to verify DENT metering results, parallel logging was conducted on a subsample of wells using pulse counting devices. Metering results using the pulse counting devices were consistent with the DENT results within approximately 5%.

Post-retrofit energy use data was collected after ensuring the POC devices were re-energized and that controlled well pump operation had fully re-established itself. The latter determination was made by observing several typical on / off cycles. For post-retrofit energy use at the Kern and Midway oil fields, power loggers were installed on the same 36 oil well pumps metered for the baseline. The typical POC control set points for the logged wells are an average pump fillage of 85% and average downtime duration per well of 8 minutes. With these set points, the well pump motor operates the pump until the well fillage is less than 85%. At that point, the well shuts off for 8 minutes. These settings are variable depending on the well and the well field.

The POC bypass baseline simulation test revealed that energy demand during run-time periods is higher under the POC controlled (post-retrofit) conditions than under the uncontrolled, or baseline, conditions. The baseline simulation work indicated that demand in the baseline case was 81.35% of demand under controlled conditions.

Average energy demand was measured to be 81.35% of the well pump under POC operation. The increased demand is due to a lower fillage and fluid pumped in the baseline case with a

continually operating well. Thus, this adjustment factor of 81.35% is used when only post-retrofit kW power draw is available for a pump motor.

The following algorithms were used to calculate energy savings for the wells with simulated pre-retrofit and post-retrofit measurements:

Pre-retrofit (baseline) energy usage (kWh/yr) = measured pre-retrofit kW x pre-retrofit annual operating hours (8760 hours/year x 0.97 availability factor)

Pre-retrofit energy usage (kWh/yr) [used in the absence of the pre retrofit kW] = measured post retrofit kW x adjustment factor (81.35%) x pre retrofit annual operating hours (8760 hours/year x 0.97 availability factor)

Post-retrofit annual operating hours = % on time from SCADA in normal operation x pre-retrofit annual operating hours (8760 hours / year x 0.97 availability factor)

Post-retrofit energy usage (kWh/yr) = measured (operating) post-retrofit kW x post-retrofit annual operating hours (hours/ year)

Energy savings (kWh/yr) = Pre-retrofit energy usage (kWh/yr) - Post-retrofit energy usage (kWh/yr)

Demand savings (kW) = Energy savings (kWh/yr) / 8760 hours/year

The POC demand savings are distributed equally throughout the year. The collected data were analyzed and verified that the load shapes for oil well pumps are flat and random, and do not systematically differ by hour of the day or day of the week. No seasonal effects were expected. Data in this study was collected primarily in the summer and early fall months. The peak period kW demand reduction was calculated using the energy consumption pre-retrofit and post-retrofit, determining the resulting kWh savings, and averaging this savings over the entire year (8,760 hours), as POC controlled and uncontrolled well pump operation is not weather dependent and equally likely to occur at any time of the year.

It was not possible to simulate the pre-retrofit baseline condition in all cases. Some program participants were concerned that over-riding the POC would damage the well and result in a loss of production due to downtime. The post-retrofit kW measurements were obtained easily for most of the participants as there was no disruption to the typical operating conditions. Therefore, it was essential to determine an adjustment factor for a typical oil well to calculate the pre-retrofit kW. This factor would allow better estimation of the pre-retrofit kW based on measurements obtained for post-retrofit kW. The kW pre and kW post were monitored for 18 wells in the Kern field and 18 wells in the Midway field. From this data, an average oilfield baseline kW adjustment factor of 81.35% was determined. As stated above, the baseline kW is lower than the post-retrofit kW due to a smaller quantity of fluid being pumped per stroke.

The ex-ante pre-retrofit annual operating hours for a well were estimated based on the reported schedules available through the California Department of Conservation's Division of Oil, Gas and Geothermal Resources (DOGGR). The DOGGR maintains a database with run-times for the operating oil wells in California. The downtime for the sample of wells selected for the baseline simulation was verified through DOGGR data that reports the annual operating days for a well. The DOGGR confirmed downtime for the pre-retrofit year 2005 for the sample of wells was similar to the estimated ex-ante percentage of downtime, both measured at 2 to 3% for a typical oil well. Therefore, the ex-ante assumed downtime (2 to 3%) for the wells before the installation of the POC was determined to be accurate, and this figure was used as the ex-post baseline (pre-retrofit) hours of operation..

Annual operating schedules and run time hours for all the wells were typically collected from the participant's SCADA (supervisory control and data acquisition) system. For selected wells, time of use current sensing loggers were also deployed and used to verify other collected data. Additionally, run time hours from the POC display panels on certain models recorded the run time for various periods (up to a maximum period of sixty days); this data was also collected and compared to the SCADA data to confirm consistency. The POC calculates the percent on time for a well every day and this calculated percent on time is collected and stored in the SCADA system. Therefore, the well run time records are available only after the installation of POCs.

Oil production was found to be similar between the pre-retrofit baseline and the post-retrofit periods. If production had been found to increase or decrease, energy impact would have been adjusted accordingly. Based on the data collected for the baseline and post-retrofit logged periods, there was no strong basis on which to perform production adjustments.

It should be noted that production data and records are imperfect. Production rates are typically generated at daily or less frequent intervals using a short duration test which may last 30 minutes or less. The data is recorded in the company's SCADA system and also reported by the oil company to DOGGR. The testing setup consists of several wells with a common manifold.

The daily production records for the logged period were collected from the SCADA system. All wells are tested for production levels at random frequencies ranging from once every week to every other day (and in cases more frequently if well operational problems are encountered). The testing frequency is determined by the production engineer based on well production records and other factors. The annual production data for the wells was obtained from the DOGGR database.

The volumetric efficiency factor (VEF) of each well was calculated for the pre-retrofit logged period, the post-retrofit logged period and the annual operating period. The production data for the simulated pre-retrofit logged period of one week and the post-retrofit logged period of three

weeks was obtained from SCADA data. The annual production data for the wells was obtained from the DOGGR database.

Note that according to this data, the average production and VEF slightly decreased as a result of the POC installation. POCs may cause an increase in production over periods longer than the monitored periods, due to fewer maintenance problems and better detection. Certain wells showed an increase in production while other wells showed a decrease in production. However, over longer periods, naturally declining production and other dynamic conditions of the oil fields are complicating factors. There was no production rate adjustment applied to the post-retrofit kW use for either kW or kWh savings, due to the wide variation in production rates and the uncertainty of the figures resulting from the intermittent collection of production data.

Contextual data for rod beam pumps can be obtained from site information and from the CA Department of Conservation's DOGGR (Division of Oil, Gas and Geothermal Resources) database. Typical contextual data for a well is shown below.

- Total well depth: 6,705 ft.
- Total rod length: 6,240 ft.
- kWh/barrel of fluid (oil + water): 1.47 kWh/barrel (annual average)
- Average water cut (percentage of water in the total fluid drawn): 89% annual average
- Average fluid produced per day: 119 barrels/day (based on monthly averages from November 2007 to November 2008)
- Motor horsepower: 75 hp
- Well age: 3 years

Peak demand (kW) study to confirm flat load shapes. A peak kW study was performed to confirm that the peak demand kW savings do not differentiate by hour of the day or day of the week. The load shapes of all the monitored wells from Kern River field (17 wells) and Midway Sunset field (17 wells) were analyzed to determine the peak kW. Note that the data from one (1) well in each field was not used due to logger failure during the monitoring period. Figure 1 and Table 1 present a weekday average, Saturday average, and Sunday average demand profile for each of the wells (total of 34 wells) monitored. Review of the demand profile plots demonstrates that this well pump energy consumption and operation is not sensitive to the time of day and is no different for Saturdays and Sundays. The analysis demonstrated that the load shapes were flat and that peak kW is the same as the average kW at any time during the day. Using statistical tests shown in the table below it was possible to accept the hypothesis that weekday hours of 2 pm - 5 pm are not statistically significantly different than other hours in the week. Therefore, the peak kW demand and average demand reduction throughout the year are the expected average kW reduction in the 2 pm to 5 pm period on summer weekdays.

Figure 1: Load profile of all the monitored wells for baseline and POC-enabled operation

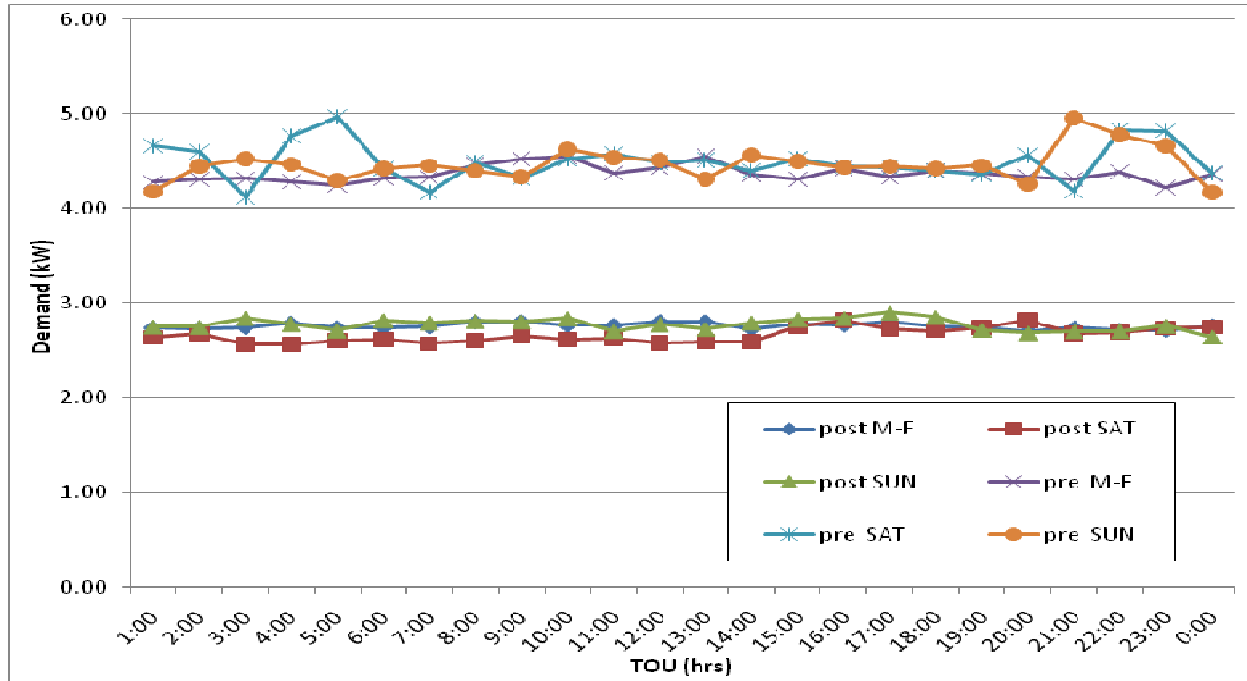


Table 1: Statistical analysis of hourly kW trends for baseline and POC- enabled operation

POCs	Description of Mean	Average Hourly kW	Lower Bound at 90% Confidence Interval	Upper Bound at 90% Confidence Interval
Pre retrofit	All Observations	4.44	4.16	4.73
Pre retrofit	Weekday 2-5	4.36	4.26	4.46
Post retrofit	All Observations	2.73	2.61	2.86
Post retrofit	Weekday 2-5	2.79	2.76	2.81

DENT and Wattnode comparison study. DENT Elite Pro loggers were used for evaluation of many of the sites involving POCs. A study was carried out at several different pump sites to quantify any measurement level uncertainties, including an assessment vs. Wattnode measurements. Measurements with a high-sampling resolution-enabled Wattnode device were compared to the DENT logger-measured values. The study revealed that the average difference in the DENT logger-recorded values and the Wattnode measured values was approximately 5%, which can be characterized as an uncertainty in the measured values for either the Wattnode or the DENT devices. Note that both devices use current transformers (CTs). For most instances of non-invasive, non-permanent metering, split core CTs (verses solid core CTs) are used. The CTs can have inaccuracies of several percent and may be the largest contributor to measurement uncertainty.

Additional tests were performed to compare the 4 Hz Wattnode measurements with the 120 Hz Wattnode measurements (4 Hz and 120 Hz are the maximum data upload rates). The test revealed that the average difference for the 4Hz Wattnode and the 120Hz Wattnode measured values was 0.6% (note that this difference may also be due to CTs used). This study proved that the 120 Hz Wattnode measurements were comparable to the 4 Hz Wattnode measurements. All results are summarized in Table 2 and Table 3 below.

Table 2: Comparison of logged data from two types of loggers for motors with POC operation enabled

DENT vs. Wattnode (4 HZ) Summary		
Test	Test Period	DENT / WattNode 4Hz
1 - OilCo #20	6/19/09 - 6/26/09	105.12%
2-OilCo#2	7/22/09 - 8/3/09	104.99%
3 Oil Co#5	7/23/09 - 8/3/09	105.18%
4 Oil Co#5	8/3/09 - 8/13/09	110.02%
5 Oil Co#5	8/13/09 - 8/21/09	104.16%
DENT/Wattnode 4Hz Error		105.90%
DENT/Wattnode 4Hz Error (without test #4)		104.86%

Table 3: Comparison of logged data from two types of pulse counting devices for motors with POC operation enabled

Wattnode (WN) 4Hz & Wattnode (WN) 120Hz		
Total kWh for 5 days 21 hrs – POC BYPASSED		
304.75	306.57	0.60%
Total kWh for 2 days 3 hrs - POC OPERATION		
19.98	20.21	1.15%
Total kWh for 8 days 0 hrs		
120 Hz WN	4 Hz WN	4Hz WN / 120Hz WN
324.73	326.78	0.63%

Pre retrofit kW study (adjustment factor) summary. The average kW adjustment factor for simulated baseline conditions for pre-retrofit operation without POCs is a factor encompassing results from both the Kern River and the Midway Sunset (MWSS) fields and was calculated as 81.35% (for more details about the study, refer to report B001 or B007). This is applicable when only the post-retrofit operating kW of well pump motors is available and is applied to other well pump motors in other fields. The results are summarized in Table 4. Note that 28 wells are identified in the Table 4. From the total of 36 wells monitored, data from eight (8) wells was not used. Two (2) wells experienced logger failure during the monitoring period, four (4) wells were not bypassed for the entire monitoring period due to POC control initiated by the participating customer due to operational concerns, and two (2) wells were continuously operating during the bypassed and post installation period.

Table 4: Power (kW) adjustment factor results from both the Kern and the MWSS fields

Well Sample Pt.	Well id	Motor size (HP)	Measured pre kW	Measured post kW	Adjustment factor (pre kW/post kW)
1	B007-1	25	12.71	12.96	98.10%
2	B007-2	40	4.86	4.4	110.50%
3	B007-3	30	3.39	2.97	114.10%
4	B007-4	40	3.55	3.67	96.70%
5	B007-5	20	4.92	5.67	86.70%
6	B007-7	5	0.94	1.04	90.40%
7	B007-10	15	2.98	3.08	96.80%
8	B007-12	25	3.32	3.71	89.50%
9	B007-13	20	7.11	8.41	84.50%
10	B007-14	10	1.55	1.66	93.40%
11	B007-15	15	2.35	2.74	85.80%
12	B007-16	5	1.09	1.2	90.80%
13	B007-17	10	2.32	2.35	98.70%
14	B001 -1	30	6.96	13.45	51.70%
15	B001 -3	20	2.43	7.44	32.70%
16	B001 -4	30	4.17	6.13	68.00%
17	B001 -5	30	3.89	5.07	76.70%
18	B001 -6	20	3.16	7.14	44.30%
19	B001 -7	25	4.4	9.39	46.90%
20	B001 -8	30	4.7	6.71	70.00%
21	B001 -9	20	3.27	7.05	46.40%
22	B001 -10	20	2.44	7.38	33.10%
23	B001 -11	20	3.38	2.08	162.50%
24	B001 -12	30	3.88	7.75	50.10%
25	B001 -13	20	3.09	3.16	97.80%
26	B001 -14	10	2.68	119.0%	2.68
27	B001 -16	15	2.4	2.58	93.00%
28	B001 -17	20	2.17	4.37	49.70%
					81.35%