

Net-to-gross Evaluation of 2013-14 Commercial Quality Maintenance Programs (HVAC3)

California Public Utilities Commission

Date: December 6, 2016

CALMAC Study ID CPU0117.02





LEGAL NOTICE

This report was prepared under the auspices of the California Public Utilities Commission (CPUC). While sponsoring this work, the CPUC does not necessarily represent the views of the Commission or any of its employees except to the extent, if any, that it has formally been approved by the Commission at a public meeting. For information regarding any such action, communicate directly with the Commission at 505 Van Ness Avenue, San Francisco, California 94102. Neither the Commission nor the State of California, nor any officer, employee, or any of its contractors or subcontractors makes any warrant, express or implied, or assumes any legal liability whatsoever for the contents of this document."

Table of contents

1	EXECUTIVE SUMMARY	1
1.1	Evaluation activities	1
1.2	Findings and Conclusions	2
2	INTRODUCTION.....	4
2.1	Objectives	4
2.2	Program logic	5
2.3	Gross savings	7
3	NTG EVALUATION	10
3.1	Contractor survey objectives	10
3.2	Contractor survey implementation	10
3.2.1	Sample design	10
3.2.2	Survey methods	12
3.3	NTG analysis	13
3.3.1	Question Frame3	14
3.3.2	Question Attr3	15
3.3.3	Question Attr2	16
4	FINDINGS	18
4.1	Net savings results	18
4.1.1	NTGRs	18
4.2	Spillover and Other Market Effects	23
4.2.1	Methodology	23
4.2.2	Findings	25
5	CONCLUSIONS AND RECOMMENDATIONS.....	28
	APPENDIX A. INSTALLATION CONTRACTOR SURVEY.....	A-1
	APPENDIX B. RETROSPECTIVE NTGR AND POTENTIAL SPILLOVER CONFIDENCE INTERVALS	B-1
	APPENDIX C. ATTRIBUTION SCORING	C-1
	APPENDIX D. POSSIBLE MARKET EFFECTS.....	D-1
	APPENDIX E. PUBLIC COMMENTS & RESPONSES.....	E-1

List of figures

Figure 1. NTG attribution-analysis flow	13
Figure 2. Estimating spillover for each contractor based on response to question Spill2	24

List of tables

Table 1. Combined program NTGR, PY 2013-14	2
Table 2. Program actions and targeted actors.....	6

Table 3. QM vs. Tune-up.....	6
Table 4: 2013-14 <i>ex ante</i> and <i>ex post</i> gross kWh savings by measure and program.....	7
Table 5. Realization rates of evaluated IOU QM programs and measure groups.....	8
Table 6. Tracking data records with contractor name merged.....	11
Table 7. Number of contractors and total savings by IOU and program.....	11
Table 8. Contractor strata.....	12
Table 9. NTG assignment decision based on response to Frame3 question.....	14
Table 10. NTG assignment decision based on response to Frame3a question.....	14
Table 11. NTG assignment decision based on response to Attr3 question.....	15
Table 12. NTG assignment decision based on response to Attr3a question.....	16
Table 13. NTG assignment decision based on response to Attr3b question.....	16
Table 14. NTG assignment decision based on response to Attr2 question.....	16
Table 15. Program kWh savings with gross realization rate and NTGR applied, PY2013-14.....	18
Table 16. Measure-level kWh NTGRs across programs, PY 2013-14.....	19
Table 17. PG&E Commercial QM NTGRs.....	20
Table 18. PG&E AirCare Plus NTGRs.....	21
Table 19. SCE QM NTGRs.....	21
Table 20. SDG&E Deemed Incentive NTGRs.....	22
Table 21. Spillover assignment decision based on response to Spill1 question.....	23
Table 22. Spillover assignment decision based on response to Spill2 question.....	23
Table 23. Combined program spillover estimates.....	25
Table 24. PG&E Commercial QM spillover estimates.....	25
Table 25. PG&E AirCare Plus spillover estimates.....	25
Table 26. SCE QM spillover estimates.....	26
Table 27. SDG&E Deemed Incentives spillover estimates.....	26
Table 28. Program Type retrospective NTGRs.....	B-1
Table 29. PG&E Commercial QM retrospective NTGRs.....	B-1
Table 30. PG&E AirCare Plus retrospective NTGRs.....	B-1
Table 31. SCE QM retrospective NTGRs.....	B-1
Table 32. SDG&E Deemed Incentive retrospective NTGRs.....	B-2
Table 33. Program type retrospective spillover estimates.....	B-2
Table 34. PG&E Commercial QM retrospective spillover estimates.....	B-2
Table 35. PG&E AirCare Plus retrospective spillover estimates.....	B-2
Table 36. SCE QM retrospective spillover estimates.....	B-3
Table 37. SDG&E Deemed Incentives retrospective spillover estimates.....	B-3
Table 38. PG&E Commercial QM attribution.....	C-1
Table 39. SCE QM attribution.....	C-1
Table 40. SDG&E Deemed Incentives attribution.....	C-1
Table 41. Draft report comments from members of the public and DNV GL team responses.....	E-1

1 EXECUTIVE SUMMARY

This report presents the methodology and findings from DNV GL's net-to-gross (NTG) evaluation of the 2013-14 California IOU commercial HVAC programs that focus on quality maintenance (QM) and related tune-up programs. The goal of the HVAC tune-up and QM programs is to derive energy savings from increased and improved maintenance of HVAC systems. This evaluation answers the question "to what extent did the programs cause an increase in maintenance actions deemed to save energy?" In short, net savings are about program attribution for taking maintenance actions and gross savings are about how much energy the actions saved, regardless of why they were taken.

1.1 Evaluation activities

To achieve the primary evaluation objective of determining net-to-gross ratios (NTGRs), DNV GL performed the following evaluation activities:

- Conducted phone surveys with participating HVAC contractors
- Calculated NTGRs using the responses from these HVAC contractor surveys

We considered various approaches to quantifying the NTG ratio for the QM and tune-up programs. The midstream design of the QM and tune-up programs means that participating customers do not have a complete picture of program effects. The 2015 HVAC QI/QM decision-making study found that customers had little awareness of or interest in QM and mostly just cared whether or not the equipment worked. Customers would not know whether the program had encouraged their local HVAC contractor to adopt higher quality maintenance service than they had offered before. Customers also wouldn't be able to tell that the maintenance service they had received through the program was a higher quality than the maintenance service that they would have received if the program did not exist.

Conducting quasi-experimental study between participating and non-participating contractors to compare practices presents challenges to measuring program attribution. With such long-standing programs there is always the risk of self-selection effects. There may be factors – e.g., smaller size, less sophistication, targeting the low-cost market – which make the HVAC companies which do not participate in the programs very different than those that do. If these differences are significant, then this would raise questions as to whether the nonparticipating HVAC companies were a valid comparison group for the participants. Since the QI/QM programs do have paperwork requirements which might be burdensome for smaller companies, this barrier to program entry might lead to self-selection effects based on company size and sophistication.

There is also the possibility that some of the non-participating contractors may have participated in previous versions of the QM and tune-up programs. This would make it more difficult to net out program effects. Finally in our previous evaluations of HVAC programs we have often found it difficult to complete interviews with non-participating contractors. Because they tend to be smaller, they are more difficult to reach in the office. The fact that they are not receiving any program benefits also makes them less willing to cooperate with research designed to improve program performance.

In conclusion, we have determined that surveys with participating contractors are the most promising means of collecting the information needed to understand the mechanisms for program attribution and calculating net-to-gross ratios. The proposed method asks concrete questions to get at quantifiable, identifiable aspects of program effect on each of the steps necessary to get energy savings from HVAC quality maintenance and tune-up programs.

The participating HVAC contractor survey guided respondents through a series of framing questions designed to remind individuals about the function of the program, when it started, and the type of HVAC maintenance work they were performing before participating in the program to provide full context for informed responses. Rather than use a non-participant survey we determined the baseline by asking the contractor about their practices before they participated in the program. This avoids the risk of self-selection effects if there are significant differences between the participant and non-participant groups. The survey instrument then focused on a series of NTG-related questions designed to assess program influence and attribution. This NTG question battery focused primarily on three topics:

- the HVAC maintenance services offered by the contractor before working in the program
- the percent of HVAC maintenance work done through the program that would not have occurred without the program and
- whether there were particular HVAC measures which were more likely to not have occurred without the program

The second bullet includes the effect of sales efforts on the contractor side and the decision making process on the customer side leading to maintenance work installed. DNV GL scored the responses to the question battery and combined them into program attribution scores that ranged from 0%-100% for each respondent measure. A more detailed explanation of how we scored the survey responses appears in section 3.3.

We expanded the respondent scores (for each measure) to the population using the statistical technique of ratio estimation, which takes into account both individual strata case weights and the savings amounts associated with the measures each contractor implemented providing program-level NTGRs.

DNV GL applied the *ex post* NTG ratio to the *ex post* gross energy and demand savings to arrive at *ex post* net program energy and demand savings. All values can be found in section 4.1.1.

DNV GL used the NTG methodology described in this report for four of the five California IOU HVAC programs, which accounted for 83 percent of total cross-program savings. We did not calculate NTGRs for SDG&E's Direct Install program because our approved NTG methodology relied on measuring program attribution from the perspectives of independent HVAC contractors. The SDG&E Direct Install program did not use this type of contractor but instead used program-dedicated direct install contractors, for whom our NTG methodology was not suited. Evaluating the NTG for the SDG&E program will require a different approach such as self-report participant surveys.

1.2 Findings and Conclusions

The error bounds for the program-level NTGRs were high due to large variation in contractor responses and a low number of responses in some programs. Due to this, we decided to produce NTG results where the program-level NTGRs were combined for all programs, shown in **Table 1**.

Table 1. Combined program NTGR, PY 2013-14

Program Type	n	Electric (kWh)		Electric Demand (kW)		Gas (therms)	
		NTGR	90% CI	NTGR	90% CI	NTGR	90% CI
All Programs	48	0.36	±0.17	0.53	±0.38	0.25	±0.31



Because the error bounds, even for the blended NTGRs in Table 1, are relatively high, we are cautious about drawing conclusions or making recommendations for program redesign. However, the fact that the kWh NTGRs are less than 50% for most of the measures, three of the four programs, and for the blended program designs indicates that these HVAC programs are having only modest impacts on the maintenance practices of these HVAC contractors. Other evidence from the contractor surveys – such as the small percentage of HVAC contractors reporting changes in their maintenance service offerings due to program participation – support the conclusion that these effects are modest.

The HVAC contractor surveys also revealed little evidence for spillover savings or any long-term market effects. The electric spillover savings ranged from zero to two percent. The vast majority of the contractors are reporting that they offered these maintenance services before joining the program. So, with a few exceptions, the program is not teaching them to offer new services which they did not offer before. There was some limited evidence that the program had encouraged a few of these contractors to improve the quality of their maintenance services. However, there was also evidence that the higher quality of these maintenance services were dependent on the program incentives. Finally, the generally low spillover savings casts further doubt on the prospect for market effects. If the contractors viewed offering these maintenance services as a good value proposition even without the program incentives, one would expect the program spillover results to be a lot higher.

There were a few encouraging results. As noted, the NTGRs for the RCA measure were much higher than those for other maintenance measures. These NTGRs were also higher across three different programs. These results suggest that the programs are influencing the adoption of this measure group. However, the gross savings report also found low gross realization rates for this measure. Therefore this measure is unlikely to contribute significant savings despite the higher NTGRs.¹

¹ In addition, experts from the CPUC HVAC Laboratory Project Coordination Group now suggest that refrigerant charge should only be adjusted when the system is significantly over or under-charged, HVAC Lab Testing Meeting. PG&E San Ramon Advanced Technologies Center, Thursday, November 12, 2015.

2 INTRODUCTION

This report presents DNV GL's NTG evaluation of the CPUC HVAC3 QM and tune-up programs.

The suite of programs making up QM and related tune-up programs includes two residential programs: PG&E Residential QM and SDG&E Residential QM and five commercial programs: PG&E AirCare Plus, PG&E Commercial QM, SCE QM, SDG&E Deemed Incentives, and SDG&E Direct Install.

The evaluation budget was not sufficient to thoroughly study all aspects of all programs, so DNV GL allocated the budget to evaluate the programs with the most claimed savings. The residential QM programs were excluded from the gross impact evaluation and also from the NTG study because they comprised only 5% of overall program savings. A gross impact evaluation strategy was developed that studied measure groups across all programs, and was able to capture 95% of the commercial program savings.

The NTG analysis included budget for one survey, and DNV GL developed a methodology that could be applied to four programs, comprising 83% of overall commercial program savings. SDG&E's Direct Install program was excluded from the analysis because it was implemented directly by dedicated program implementation contractors who did not perform work outside the program. The *ex ante* NTGR of 0.73 will remain for this program, and though it is higher than the NTGRs found for the evaluated programs, we think that it is representative of the NTGR expected for a direct install program.

The QM programs have historically reported unit-level savings made up of multiple measures installed on one HVAC unit. Several programs moved away from that structure in the 2013-15 cycle, and since we evaluated both QM and tune-up programs, we opted for a measure level evaluation structure consistent with the majority of the QM energy savings claims. For the impact portion of this study, DNV GL evaluated five key measure groups implemented through the IOU's QM and tune-up programs: condenser and evaporator coil cleaning, RCA, economizer repair, thermostat replacement or reprogramming, and supply fan control adjustment. The NTG study maintained consistency with the impact study by investigating the same measure groups.

The goal of the HVAC3 programs is to derive energy savings through increased and improved maintenance of HVAC systems. Gross savings is an estimate of energy saved by the specific measures promoted by the maintenance programs; DNV GL previously produced a report on the gross findings for these programs. With the method described here we want to answer the question "to what extent did the program influence an increase in maintenance actions deemed to save energy?" In short, "net" is about program attribution for taking maintenance actions and "gross" is about how much energy the actions saved, regardless of why they were taken.

2.1 Objectives

The goal of the HVAC tune-up and QM programs is to derive energy savings from increased and improved maintenance of HVAC systems. This evaluation answers the question "to what extent did the programs cause an increase in maintenance actions deemed to save energy?" This is formalized as the NTGR, the ratio of the net savings to the gross savings. The net savings are those attributable to a particular energy efficiency program and the gross savings are those actually achieved through the program. The primary goal of this evaluation is to determine the NTGR for the QM and related HVAC tune-up programs in California.

The program name "quality maintenance" implies that measures installed through the program are of higher quality than those installed outside the program. Neither this study, nor the gross impact study has

attempted to assess the quality of measures installed through this program compared to those installed outside the program. In the NTG evaluation we posit that when a contractor installs a measure, they install it in the same way regardless if it is installed through the program. The gross impact study compared the installed measure to a baseline of the pre-maintenance or “not installed” condition. In order to objectively answer the question of whether adoption of a given measure through the QM program delivers higher savings than un-incentivized maintenance measures, we would need to perform a baseline study of the energy impact of non-program HVAC maintenance activities. This study attempts to make some assessment of whether the program changed contractor maintenance practices, but cannot fully assess the quality of services performed either within or outside the program.

2.2 Program logic

One way that energy efficiency programs achieve net savings is by addressing barriers in the market that prevent energy savings from occurring naturally. The statewide QM and local tune-up programs’ logic documentation identify three barriers that the programs seek to overcome:

- Barrier 1: Lack of value proposition awareness
- Barrier 2: Reluctance to search for qualified QM contractors
- Barrier 3: Contractors often do not have enough time to evaluate unit efficiency during maintenance checks

A 2015 California HVAC quality installation (QI)/QM decision-making study confirmed these barriers.² The study found that customers generally had low awareness of QM and how to judge quality. Customers tended to stick with the contractors they knew rather than searching for and comparing contractor offerings. Study respondents cited cost and reliability as the two primary decision-making factors when making decisions about HVAC maintenance.

In determining the NTG estimation methodology, we started from a general logic model of the programs that covered the actions taken by the programs, the target groups of the actions, and the expected outcomes. We divided the actions taken by the QM and tune-up programs into two broader categories: first, informative activities addressing awareness and search costs (Barriers 1 and 2), and second, incentives, which address Barrier 3 by reducing the incremental cost of quality maintenance. Table 2 shows the types of actions taken by QM and tune-up programs and the target groups of the actions.

² Steiner, Ellen and Malinick, Todd, EMI. *California HVAC Quality Installation/Quality Maintenance Customer Decision-Making Study*. Southern California Edison, April 15, 2015. Accessed August 24, 2015. http://www.calmac.org/publications/CDM_Report_2015-04-15_FINAL.pdf

Table 2. Program actions and targeted actors

Categories	Information	Incentive
Contractor	Training Marketing Checklists Guidelines	For HVAC assessments For specific services
Customer	Marketing Program Endorsement/ Validation of: - contractors - specific services	For HVAC assessments For specific services

The HVAC QM and tune-up programs attempt to generate savings by providing participating contractors with training, marketing, checklists, and contact information, and by lowering the price point to entice customers to implement a measure. The programs only generate net savings if they motivate an increase in actions that save energy.

While the QM and tune-up programs use many of the same work papers and measures to claim gross savings, they have different designs and implementation strategies that should be accounted for in the NTG evaluation. Table 3 compares the program approaches.

Table 3. QM vs. Tune-up

Characteristic	QM	Tune-up
Gross Savings	Based on HVAC components cleaned, maintained, or replaced	Based on HVAC components cleaned, maintained, or replaced
Persistence	Multi-year contracts with regular maintenance visits increase the likelihood of consistent performance over time	Single visit to identify and maintain HVAC components that need maintenance
Selling point	Intended to compel both contractor and customer to keep the HVAC system(s) at a high level of performance throughout the contract period	Customer pays for a one-time package of maintenance that ensures a level of performance; can elect for extras if measures are identified and recommended
Program Type	<i>Market Transformation:</i> The program is designed to move the market to a long-term service contract model that ensures well-maintained HVAC equipment	<i>Resource Acquisition:</i> The program is designed to get more of the energy saving repairs done today

Despite the differences in implementation approach, in each of the programs the key actors who have the most knowledge of the program's influence are the contractors and technicians who provide the services for which the programs claim energy savings.

2.3 Gross savings

Previous gross impact evaluation work identified *ex ante* and *ex post* gross savings for programs and measures in the HVAC3 portfolio as shown in Table 4.

Table 4: 2013-14 *ex ante* and *ex post* gross kWh savings by measure and program

Savings	Measure Group	PG&E Commercial QM	PG&E AirCare Plus	SCE QM	SDG&E Deemed Incentives	SDG&E Direct Install
<i>Ex Ante</i> kWh	Coil Cleaning - Condenser		53,539		7,008,173	3,636,362
	Coil Cleaning - Evaporator		23,154		1,852,771	1,925,586
	Economizer	3,068,415	1,460,285		44,000	9,180
	Fan Control	4,367,405	453,308			
	Maintenance			4,439,795		
	RCA	870,455	1,450,631		2,890,034	997,141
	Thermostat	2,712,280	2,906,562			214,693
<i>Ex Post</i> kWh	Coil Cleaning – Condenser	1,354,873	319,956		4,765,519	2,246,860
	Coil Cleaning - Evaporator	107,587	25,229		197,776	128,071
	Economizer	1,319,418	642,525		6,933	5,141
	Fan Control	3,725,681	437,371			
	Maintenance			5,877,152		
	RCA	268,625	424,926		1,034,195	374,057
	Thermostat	2,712,280	2,906,562			214,693

The impact evaluation of 2013-14 Commercial Quality Maintenance Programs estimated the achieved savings and compared them with the expected savings as a ratio called a realization rate. This is inclusive of the *ex post* installation rate and any *ex post* adjustment to the unit energy savings. Table 5 provides a summary of the realization rates for the IOU programs and measure groups evaluated in this study. Empty table cells are those program measure groups that had no *ex ante* claims in a specific program. While the IOU programs address both residential and commercial activities, DNV GL focused on commercial QM measure groups because residential program savings were less than 5% of the overall QM savings. Within the commercial program, the evaluation team addressed high-impact measure groups, those that contributed the largest percentage of *ex ante* savings, across all programs in this evaluation.

The realization rates were generally low; although some were consistent with past evaluation estimates, others were higher than expected. The highest realization rates were for the SCE Commercial QM program. The high realization rate in this program is due to greater than expected frequency of installation of many of the component measures making up the QM measure group.

In general, economizer repair realization rates were not high, but they were noticeably higher than the past evaluation cycle. This is a promising trend considering that previous studies³ have shown high failure rates. Although not reported as a separate measure, the realization rate for economizers in the SCE Commercial QM program was 82%, indicating a significant improvement in the effectiveness of economizer repairs in that program.

Table 5. Realization rates of evaluated IOU QM programs and measure groups

	 Coil Cleaning	 Refrigerant Charge Adjustment	 Economizer Repair	 Thermostat Adjustment	 Supply Fan Control Adjustment
PG&E					
Commercial QM		31%	43%	100%	85%
AirCare Plus	353%	29%	44%	100%	96%
SDG&E					
Deemed Incentives– Commercial HVAC	39%	36%	56%		
Commercial Direct Install	34%	38%	56%	100%	
SCE – Quality Maintenance Measure					
Commercial QM	132%				

The SCE QM program did not report claims on a measure level so a program-wide realization rate is given. The coil cleaning realization rates are an average of condenser and evaporator coil cleaning realization rates.

The primary reason for the low RCA realization rate is a difference between laboratory data collected recently and used in this evaluation and the previous data underlying expected savings calculations.

Similarly, the coil cleaning measure group's low realization rate of 37% (average of evaporator and condenser coil cleaning realization rates) in the SDG&E programs was based on simulations, laboratory data, and field measurement of non-residential package units serviced through these programs compared with *ex ante* claims. The AirCare Plus program had a surprisingly large realization rate for the coil cleaning measure because it reduced claimed savings by approximately 90% as directed by the CPUC Energy Division Disposition of the 2013-14 workpapers issued June 2013.

The thermostat and supply fan measures were not evaluated due to low sample size, high sample variability, and resulting low precision in the *ex post* savings estimate. *Ex ante* savings were used for these measures. We found that a majority of thermostats did not meet program setback requirements during unoccupied periods. In PG&E's Commercial QM program we found the program implementer-supplied data were

³ DNV GL, HVAC Impact Evaluation FINAL Report WO32 HVAC – Volume 1: Report, Jan. 28, 2015. http://www.calmac.org/publications/FINAL_HVAC_Impact_Evaluation_WO32_Report_28Jan2015_Volume1_ReportES.pdf



inconsistent with the tracking claims. The economizer realization rates found in this evaluation, although low, are actually much improved from the previous (2010-12 program year) evaluation findings of 23%.

3 NTG EVALUATION

The NTGR is the estimated proportion of gross savings that can be attributed to a program. This study used a phone survey with participating contractors to estimate NTGRs for the evaluated programs based on a methodology that we described in a memorandum submitted March 15, 2016. This methodology asks concrete questions to get at quantifiable, identifiable aspects of program effect on each of the steps necessary to get energy savings from HVAC QM and tune-up programs. The NTG evaluation method was designed to be consistent with the gross methodology, where savings were broken down to a measure level, and the baseline was assumed to be no measure installed. This is also consistent with the ex ante gross savings calculations.

3.1 Contractor survey objectives

The participant survey guide is presented in APPENDIX A. The sole objective of the survey instrument was to develop reliable NTGR estimates. Framing questions were asked in support of this objective.

3.2 Contractor survey implementation

DNV GL used trained internal staff to complete the contractor surveys. As the sample design subsection explains in more detail below, DNV GL attempted to complete interviews with the full population of 93 validated participating contractors. It was able to complete interviews with 43 of these contractors.

3.2.1 Sample design

The tracking data files included energy efficiency measures completed during the 2013 and 2014 program years. For the four programs evaluated, 39% of records in the tracking data were missing the contractor name. However, implementer files allowed us to reduce this number to 7%. In the process of looking up contact information for the contractors, we discovered that a subset of the listed contractors in the remaining programs were actually dedicated program implementation contractors (e.g., Enalysys, Synergy Companies, and MATRIX). We removed records of such contractors from the sample frame as they are entirely program dependent, and therefore would be very limited in their ability to describe maintenance activities outside the program as a baseline or comparison to program maintenance activities. After removing these program implementation contractors, we arrived at a sample frame of 93 “valid” contractors, which represented about 80% of the original population.

Because we based our stratification on kWh savings per contractor, we could not expand our results back to the portion of the population where we did not have contractor information. We discuss how the difference between the sample frame and the population affects the applicable sampling error in our discussion of the results.

Table 6 presents the percentage of records matched with a contractor name by program. It shows that our overall contractor match rate was high (93%) with only the PG&E Commercial HVAC program having a match rate of less than 90% of its records. Overall, the electric energy savings represented by the matched records is 94% in PG&E, 87% in SCE and 99% in SDG&E territories.

Table 6. Tracking data records with contractor name merged

PA	Program Name	Contractor Name	Records	Percent
PG&E	AirCare Plus	Not Matched	330	2%
		Matched	15,693	98%
			16,023	100%
PG&E	Commercial HVAC	Not Matched	3,449	23%
		Matched	11,487	77%
			14,936	100%
SCE	Nonresidential HVAC	Not Matched	647	3%
		Matched	24,098	97%
			24,745	100%
SDG&E	Commercial Deemed Incentives	Not Matched	1,928	7%
		Matched	27,375	93%
			29,303	100%
Overall		Not Matched	6,354	7%
		Matched	78,653	93%
			85,007	100%

Note: Some records were originally matched, but the name listed as a contractor was actually an implementer.

The number of contractors varied substantially across the three IOUs, with fewer contractors achieving a larger share of savings for SDG&E. Table 7 presents the number of contractors and achieved savings broken out by program. Although savings are fairly equally distributed across the programs, the number of contractors achieving those savings is not. Both the Commercial HVAC and Nonresidential HVAC programs use many more contractors than the other two programs.

Table 7. Number of contractors and total savings by IOU and program

IOU	Program	Number of Contractors in program	Percent of Contractors	Ex Post kWh	Percent of Ex Post kWh
PG&E	AirCare Plus	10	11%	4,651,418	21%
PG&E	Commercial HVAC	44	48%	6,775,257	31%
SCE	Nonresidential HVAC	29	31%	5,086,757	23%
SDG&E	Commercial Deemed Incentives	10	11%	5,605,988	25%
Total		95	100%	22,119,419	100%

DNV GL attempted to complete surveys with a census of HVAC3 contractors. Weights for expansion were developed with a post-stratification of the resulting sample based on estimated *ex post* kWh savings per contractor.⁴

⁴ We limited the population to the records with contractor information as we do not know the distribution of contractor sizes for the records without a contractor matched.

The number of contractors and savings per strata are reported in Table 8.

Table 8. Contractor strata

IOU	Program	Strata Number	Number of Contractors	Percent of Contractors	Ex Post kWh	Percent of Ex Post kWh
PG&E	AirCare Plus	1	5	50%	448,315	10%
		2	5	50%	4,203,103	90%
			10	100%	4,651,418	100%
PG&E	Commercial HVAC	1	20	45%	272,435	4%
		2	6	14%	343,768	5%
		3	5	11%	444,837	7%
		4	3	7%	462,585	7%
		5	10	23%	5,251,632	78%
		44	100%	6,775,257	100%	
SCE	Nonresidential HVAC	1	19	66%	467,470	9%
		2	6	21%	716,210	14%
		3	4	14%	3,903,077	77%
		29	100%	5,086,757	100%	
SDG&E	Commercial Deemed Incentives	1	7	70%	559,629	10%
		2	3	30%	5,046,359	90%
			10	100%	5,605,988	100%

Note: Two contractors were counted with other contractors because they were determined to be the same company through the surveys.

3.2.2 Survey methods

Trained DNV GL staff conducted telephone surveys of HVAC3 contractors for two weeks from April 11 to April 27, 2016. Source data provided to DNV GL did not include contact information such as contact name, phone number, email, or address. As a result, we searched for individual firms manually. Ultimately, we completed the survey effort with a final response rate of 53%.

The full data sets originally included 98 listed firms. Of the 98, 10 firms listed were duplicates. Of the remaining 93 contractors, DNV GL completed interviews with 47 (53%), received refusals from 11 (12%), and was not able to reach 35 (39%).⁵ After several weeks of calling and a completion rate of 48% we reached out to all of the IOU program managers requesting help contacting the contractors. They all provided updated contact information for the contractors that we had been unable to reach. The updated information allowed us to obtain 5 more completes after two additional weeks of calling.

⁵ DNV GL staff attempted to contact the contractors on varying days and at varying times for up to six attempts, leaving voicemail messages to help increase survey completion rates.

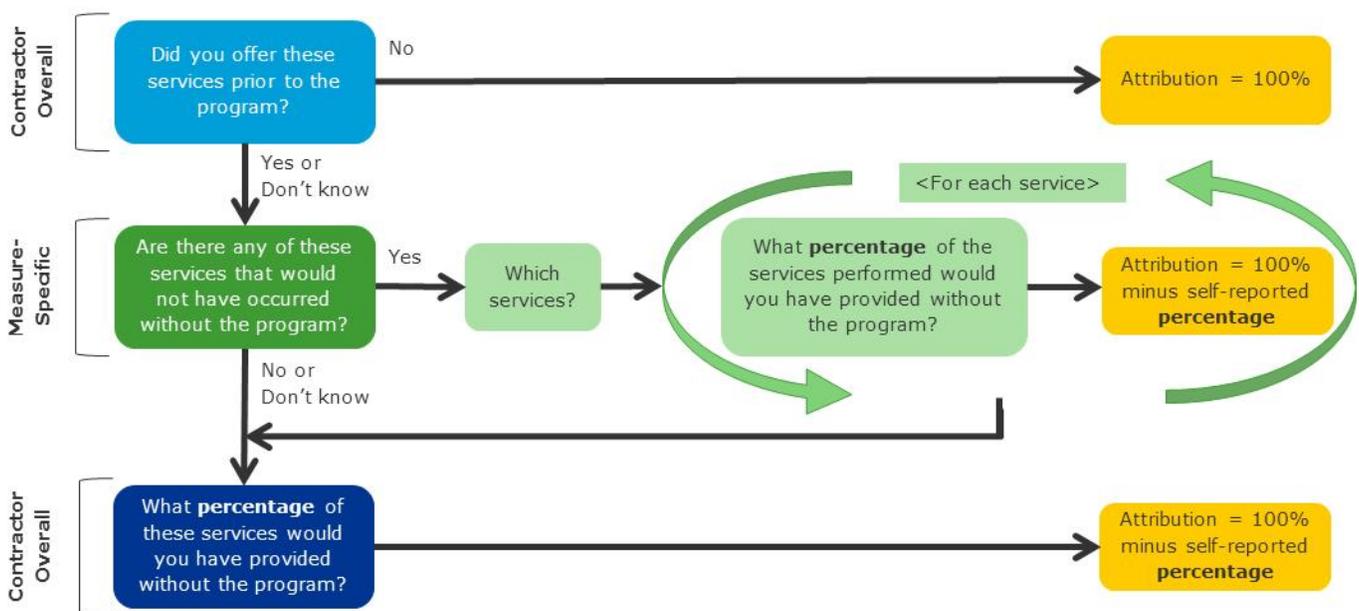
3.3 NTG analysis

The participating contractor survey instrument had a series of NTG-related questions designed to assess program influence and attribution. Our approach addressed program attribution by asking about two key program objectives:

1. Getting contractors to offer new services through program incentives
2. Increasing customer uptake of contractor offered services

This question battery consisted of three primary questions listed below in Figure , which also shows the analysis flow.

Figure 1. NTG attribution-analysis flow



DNV GL assigned a score for program attribution to each respondent that ranged from 0%-100% depending on responses to the NTG questions outlined above. The specific questions are discussed in further detail in sections 3.3.1 through 3.3.3. The first question used in attribution scoring, Frame3, addresses overall changes in the contractor's offerings. If the contractor had not offered any of the program incented services prior to the program, the program received full attribution for all of the measures installed by that contractor. If the contractor had offered at least some of the same measures prior to participating in the program, they were asked which services they had previously offered (Frame3a). If a specific measure had not been offered prior to the program, it was assigned full attribution.

For measures that had been offered prior to the program, the attribution methodology used the response to Attr3, which asks which specific measures would not have occurred without the program. If the contractor said that none of a specific measure would have occurred without the program, that measure received full attribution.

However, if the contractor said that at least some of a specific measure would have been done without the program, then the percentage that would not have been done without the program, Attr2, was used to assign attribution.

The respondent scores (for each measure) were expanded to the population using the statistical technique of ratio estimation by domains. This provided program-level NTG ratios that take into account the strata weights and savings of the individual respondents.

3.3.1 Question Frame3

One way to measure the program impacts on the maintenance practices of HVAC contractors is to ask them whether they offered any of these maintenance services before becoming involved with the programs. DNV GL designated survey questions Frame3 and Frame3a as the initial NTG questions for assigning program attribution.

- **Frame3.** Did you offer any of those services to customers before working with the <program> in <Q7 response>?
 - **Frame3a.** Which of these services did you offer before participating in the program?

We assigned an attribution rating of 100% to responses of “No” to Frame3. If respondents said “Yes” to Frame3, we asked them the follow-up question Frame3a to find out which maintenance services they had offered prior to program participation. We then assigned an attribution rating of 0% to any previously-offered services. Table 9 and Table 10 show the attribution assignments for contractor responses to questions Frame3 and Frame3a.

Table 9. NTG assignment decision based on response to Frame3 question

Response to Frame3	NTG assignment
Yes	Proceed to Frame3a
No	100% Program Attribution
Don't know	No attribution assignment
Refused	No attribution assignment

Table 10. NTG assignment decision based on response to Frame3a question

Response to Frame3a	NTG assignment
Coil cleaning	Identified as previously-offered = 0% Measure Attribution, Not identified as previously offered = 100% Measure Attribution
Economizer addition	
Economizer repair	
Thermostat adjustment	
Fan control adjustment	
Maintenance	
Don't know	No attribution assignment
Refused	No attribution assignment

3.3.2 Question Attr3

Another way to measure the program impacts on the maintenance practices of HVAC contractors is to ask whether there were any particular maintenance/tune-up measures which they thought would not have occurred without the assistance of the programs. Survey questions Attr3, Attr3a, and Attr3b explored this issue.

- **Attr3.** Are there any particular maintenance/tune-up measures that stand out to you as not likely to occur without the assistance of the program?
 - **Attr3a.** Which maintenance/tune-up measures, in particular, do you think would be **not** likely to occur without the assistance of the program?
 - **Attr3b⁶.** [IF THEY NAMED SPECIFIC MEASURES] OK, without the program training, customer leads, program endorsements, and incentives you said that you would not have performed any [list measures mentioned in Attr3a]. Of the remaining measures [list remaining services performed by this contractor], what percent would you have provided in absence of the program?
[Probe for each measure]

Table 11 shows that if the contractors responded 'No' to question Attr3, meaning that no maintenance/tune-up measures came to mind which would not have occurred without the program, DNV GL scored the attribution as 0%.

Table 11. NTG assignment decision based on response to Attr3 question

Response to Attr3	NTG assignment
Yes	Proceed to Attr3a
No	0% attribution
Don't Know	No attribution assignment
Refused	No attribution assignment

If the response to Attr3 was 'Yes', meaning there were some measures that might not occur without the program, DNV GL then asked respondents question Attr3a which probed for which measures they believed would not occur without the program. If a measure was specifically mentioned, then DNV would ask them question Attr3b, which asked the respondent to estimate what proportion of those maintenance/tune-up measures they still would have offered in the absence of the program. Table 12 and Table 13 show how we calculated program attribution depending on the responses to questions Attr3 and Attr3b.

⁶ Original wording of Attr3b proved too confusing for respondents and was reworked by the survey team. Question wording originally read: You just said that if these program trainings, customer leads, program endorsements and program incentives had not been available, you *still* would have provided approximately [X%] of maintenance/tune-up services that you provided in [YEAR]. What % of [MEASURE X] would you have provided in the absence of the program?

Table 12. NTG assignment decision based on response to Attr3a question

Response to Attr3a	NTG assignment
Coil cleaning	If measure mentioned, proceed to Attr3b. If not, Attr2 score.
Economizer addition	
Economizer repair	
Thermostat adjustment	
Fan control adjustment	
Maintenance	
Don't know	No attribution assignment
Refused	No attribution assignment

Table 13. NTG assignment decision based on response to Attr3b question

Response to Attr3b	NTG assignment
Coil cleaning	Measure Attribution = 100% - percentage
Economizer addition	
Economizer repair	
Thermostat adjustment	
Fan control adjustment	
Maintenance	
Don't know	No attribution assignment
Refused	No attribution assignment

3.3.3 Question Attr2

In developing our NTG methodology, we assumed that many respondents might not be able to provide measure-specific estimates of program attribution. Therefore we included survey question Attr2 which asked respondents to estimate the percentage of their program-eligible maintenance/tune-up services which they would have still offered absent the program.

- **Attr2.** In [YEAR] you implemented [#] maintenance/tune-up measures through [PROGRAM NAMES] including [LIST OF QUANTITIES OF TOP 5 MEASURES]. If these program trainings, customer leads, program endorsements, program market materials and program incentives had not been available, approximately what % of maintenance/tune-up services would you still have provided in [YEAR]?

Table 14 shows how DNV GL calculated program attribution for responses to question Attr2. We applied the percentage estimate uniformly to all maintenance/tune-up services they offered through the program.

Table 14. NTG assignment decision based on response to Attr2 question

Response to Attr2	NTG assignment
Percentage	Attribution = 100% - percentage
Don't Know	No attribution assignment
Refused	No attribution assignment



DNV GL was able to apply the NTG method described in this section to all survey responses.

4 FINDINGS

This section presents our findings for *ex post* gross and net savings for both electric and gas.

4.1 Net savings results

DNV GL estimated NTG savings ratios using the contractor survey responses to the evaluation’s NTG questions that we described in the previous section. We then applied the *ex post* NTG ratio to the *ex post* gross energy and demand savings to arrive at *ex post* net program energy and demand savings. All values can be found below, in section 4.1.1.

For the NTG surveys, DNV GL attempted a census of contractors from populations of finite size. With this sampling approach a finite population correction (FPC) is used to define the standard error of the estimate. The FPC becomes important to apply when the sample size (*n*) is more than 5% of the population (*M*). FPC reduces the error more when the sample size is a large proportion of the population.

While FPC is typically used when reporting sampling errors for program evaluations, it is not appropriate for all situations. The FPC error is only applicable when the population studied is the same as the population to which the estimate will be applied. The NTG study sampled from a population of contractors who implemented measures in the 2013-15 program years. The results of this study may be applied to the 2015 program year alone and/or prospectively for future program years. When applying the study results to program timeframes other than the 2013-15 population we sampled from, we recommend that the non-FPC errors reported be considered as the most appropriate estimates of error to consider when applying these results.

All error and precision estimates reported in the body of this report are non-FPC errors (FPC off). APPENDIX B reports the FPC (FPC on) standard errors to aid comparison with the errors reported in most other studies.

4.1.1 NTGRs

The NTGRs and gross realization rates are applied to the *ex ante* savings in Table 15. The NTGRs for all programs combined are used because the individual program NTGRs have high uncertainty as illustrated in the later part of this section. Contractor responses indicated that 64% of verified program kWh savings would have occurred without the programs.

Table 15. Program kWh savings with gross realization rate and NTGR applied, PY2013-14

Savings	PG&E Commercial QM	PG&E AirCare Plus	SCE QM	SDG&E Deemed Incentives
Ex Ante kWh	11,018,555	6,347,479	4,439,795	11,794,978
GRR	0.86	0.75	1.32	0.51
NTGR	0.36	0.36	0.36	0.36
NTGR RP at 90% confidence	±0.17	±0.17	±0.17	±0.17
Ex Post kWh	3,415,847	1,712,365	2,115,775	2,161,592

Generally, we found that the NTGRs were low across most programs and measure groups. The exceptions were the refrigerant charge adjustment (RCA) measure group that scored a high NTGR across all programs,

and the QM measure group, only performed in the SCE QM program. The QM measure group is the only measure in the SCE QM program (it is comprehensive and includes all of the other measures), resulting in a high NTGR for that program. The fan control and thermostat measures had the lowest NTGRs among the measure groups.

Table 16 also shows that three of the four programs had overall NTGRs in a similar range. The one exception was PG&E's Commercial QM program which had a much lower overall NTGR. This result surprised us somewhat because this PG&E program had a similar design to Southern California Edison's (SCE) QM program, which had a much higher NTGR. A closer examination of these results revealed that the differences in NTGRs between the PG&E and SCE programs were mostly driven by different assessments of program attribution by a small number of contractors who accounted for a very large volume of program savings in each of the programs. With these contractors removed the program results were very similar.

Table 16. Measure-level kWh NTGRs across programs, PY 2013-14

Measure Group	PG&E Commercial QM	PG&E AirCare Plus	SCE QM	SDG&E Deemed Incentives
Coil Cleaning		0.00 ±0.00		0.48 ±0.63
Economizer	0.21 ±0.30	0.55 ±1.19		
Fan Control	0.05 ±0.08	0.00 ±0.00		
Maintenance			0.49 ±0.28	
RCA	1.00 ±0.00	0.37 ±1.06		0.84 ±0.37
Thermostat	0.21 ±0.19	0.31 ±1.00		
Overall	0.16 ±0.16	0.29 ±0.67	0.49 ±0.28	0.54 ±0.49

Another contributing factor for the very different NTGRs between the PG&E and SCE QM programs may have been methodological. Because the SCE program did not track individual measures separately, we could not evaluate the effectiveness of the program in promoting different measures.

Tables 17 through 20 present the NTG ratios and the FPC off confidence intervals for the PG&E Commercial QM and PG&E AirCare Plus; SCE QM; and SDG&E Deemed Incentives programs. The attribution scoring results by program with FPC on are presented in APPENDIX C.

PG&E QM NTGR

For the PG&E QM program overall, respondents accounting for only 7% of the weighted ex post savings indicated that they had not installed program measures prior to the program. The only measure where contractors attributed more than 25% of savings to the program was RCA. However, because RCAs only made up a small portion of program savings, the overall program had a low NTG ratio.

Nearly all contractors indicated that they had offered at least some of the services included in the program prior to participating in the program the program (Frame3), though none of the contractors indicated that they had offered RCAs prior to the program (Frame3a). Most (55% of weighted *ex post* savings) of the overall NTGR results are based on a response pattern of "yes" the contractor offered these services prior to participating in the program and "no" there are not any services that would not have occurred without the program leading to respondents indicating the percentage of maintenance/tune-up services they would have

installed overall in the absence of the program. Nine contractors indicated that they would have completed the same projects without the program, six indicated that they would have completed over half of the projects without the program, and two indicated that they would have completed a quarter or less of the projects without the program.

Overall ratios for each of the savings types (kWh, kW, therms) vary due to the different distribution of savings of each among the measures. For instance, RCA is the only measure with claimed demand savings causing an extremely high electric demand NTGR while thermostat and fan control are the only measures with claimed gas savings. The overall kWh NTGR is 0.16, while the kW NTGR is 1.00 and therm NTGR is 0.21.

Table 17. PG&E Commercial QM NTGRs

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (therms)	
		NTGR	90% CI	NTGR	90% CI	NTGR	90% CI
Coil Cleaning*	0						
Economizer	14	0.21	±0.30	N/A	N/A	N/A	N/A
Fan Control	21	0.05	±0.08	N/A	N/A	0.07	±0.09
RCA	16	1.00	±0.00	1.00	±0.00	N/A	N/A
Thermostat	18	0.21	±0.19	N/A	N/A	0.21	±0.19
Overall	22	0.16	±0.16	1.00	±0.00	0.21	±0.18

*None of the contractors interviewed for PG&E's Commercial QM program were matched with coil cleaning measures in the tracking data. Coil cleaning was performed through the program, but not claimed or reported in the tracking data so it was impossible to link the activity to contractors.

After asking contractors what percent of the projects completed through the program would have been completed without the program (Attr2), the interviewers asked them to explain their response (Attr2a). Some of these responses included:

- *The number of projects completed did not change:*
 - One contractor said that the “trainings helped them understand the QM program better but not how to do the work.”
 - A second contractor responded that “It’s what the business does and is part of the services they offer.”
 - A third contractor said the “offerings didn’t really affect their business.”
 - A fourth contractor responded that “they have contracts they have to follow, which is a visit every 3 months.”
- *The trainings helped them to identify more things to do:* One contractor said “Due to the trainings, they were able to spot more things. Trainings from the liaison focused on certain avenues of the business world which helped them get more work.”
- *Incentives encouraged more customers to complete more economizer repairs and thermostats:*
 - One contractor said they “might not have done as much economizer or thermostat work if rebates weren’t offered.”
 - Another contractor indicated that “economizer customers don’t want to pay for the service by itself and the program helps with that [through] the rebates.”

PG&E AirCare Plus NTGR

DNV GL completed interviews with three of ten contractors participating in the PG&E AirCare Plus program. The NTGRs are presented in Table 18. None of the measure specific or overall NTGR are statistically different from zero. However, due to the small sample sizes, the confidence intervals are large.

Table 18. PG&E AirCare Plus NTGRs

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		NTGR	90% CI	NTGR	90% CI	NTGR	90% CI
Coil Cleaning	3	0.00	±0.00	0.00	±0.00	N/A	N/A
Economizer	3	0.55	±1.19	N/A	N/A	N/A	N/A
Fan Control	2	0.00	±0.00	N/A	N/A	0.00	±0.00
RCA	3	0.37	±1.06	0.43	±1.12	N/A	N/A
Thermostat	3	0.31	±1.00	N/A	N/A	0.28	±0.95
Overall	3	0.29	±0.67	0.18	±0.44	0.28	±0.94

SCE QM NTGR

The data for the SCE QM program did not allow for results at the measure level. The program as a whole scored 49%. The NTGR attribution pathway for the large majority of responses (82% of weighted ex post electric savings) was a “yes” response to the question about whether they offered these services prior to program participation and a “no” response to the question about whether there were any services that would not have occurred without the program. (see APPENDIX C). All contractors indicated that they had offered a similar service provided through the program prior to working with the program (Frame3). However, when asked which measures they thought would not have occurred without the program, 11 out of 16 indicated general maintenance (Attr3a). This is not surprising since it is common for HVAC systems to only be serviced when they break.⁷

- A quarter of respondents indicated that 75% or more of the projects they completed through the program would have been done without the program (Attr2).
- Five respondents indicated that between 50% and 70% of the projects would have been completed without the program, and
- Seven respondents indicated that less than half of the projects would have been completed (Attr2).

Table 19. SCE QM NTGRs

Measure Group	N	Electric (kWh)	
		NTGR	90% CI FPC Off
Maintenance	16	0.49	±0.28
Overall	16	0.49	±0.28

After asking contractors what percent of the projects completed through the program would have been completed without the program (Attr2), respondents were asked to explain their response (Attr2a). Some of these responses included:

- *They increased the frequency of maintenance done for existing customers:*
 - One contractor indicated that they “would have done all the tasks (measures) but the frequency would have dropped on some tasks, maybe to once a year rather than twice.”
 - A second contractor said that “it made some customers do maintenance quarterly instead of less frequently.”

⁷California HVAC Behavior Study states that very few contractors focus their business on maintenance services.

- They expanded their customer base through the program incentives:
 - One contractor said “Incentives. [The] marketing material [is] good too. The program has been very helpful in acquiring new business.”
 - Another contractor indicated that the “incentive gets them to table - then they have the ability to sell the program and benefits.”
 - A third contractor said “some of these things are a hard sell to customers without incentives. Incentives make them move in the direction of wanting to do it.”
 - A fourth contractor said “Incentives - some customers were in it for the money.”
 - A fifth contractor said “Incentives - people love free money.”

SDG&E Deemed Incentive NTGR

DNV GL staff completed interviews with seven of ten contractors participating in the SDG&E Deemed Incentive program. Overall, respondents accounting for 43% of the *ex post* electric savings (2 out of 7 respondents) indicated that they had not installed the offered measure types prior to the program (Frame3; see APPENDIX C). Four indicated that they had not offered RCAs prior to the program and two indicated that they had not offered coil cleaning prior to the program (Frame3a). When asked which measures would not likely occur without the program, two contractors indicated coil cleaning and one indicated RCAs (Attr3a). When asked what percent of projects would have been completed without the program, five out of seven responses were less than 50% (Attr2).

Table 20. SDG&E Deemed Incentive NTGRs

Measure Group	n	Electric (kWh)		Electric Demand (kW)	
		NTGR	90% CI FPC Off	NTGR	90% CI FPC Off
Coil Cleaning	7	0.48	±0.63	0.50	±0.62
Economizer*	0				
RCA	6	0.84	±0.37	0.79	±0.47
Overall	7	0.54	±0.49	0.55	±0.49

*None of the contractors interviewed for SDG&E's Deemed Incentives program were matched with economizer measures in the tracking data.

After asking contractors what percent of the projects completed through the program would have been completed without the program (Attr2), respondents were asked to explain their response (Attr2a). Some of these responses included:

- They expanded their customer base through the program incentives:
 - One contractor said that they “would have done all the work for their customers they already had, but the new customers they wouldn't have had without the program.”
 - A second contractor responded that “Only some of their customers that went through the program were existing customers, and they would not have gotten the other customers' business without the program.”
 - A third contractor indicated that they were able to expand the commercial side of their business through the program.
- They locked in service contracts for a longer period of time: One contractor indicated that with the incentive, they were able to lock in customers for three years and increase customer awareness of how maintenance relates to energy.

4.2 Spillover and Other Market Effects

This section describes both the methodologies we used to estimate spillover savings and possible market effects and the findings we derived from the survey responses.

4.2.1 Methodology

We estimated potential spillover for the commercial QM programs. While these programs will receive the global 5% spillover adder, the CPUC was still interested in an estimate of program-specific spillover, in part to help assess the reasonableness of the global adder.

The rationale for estimating spillover for participating contractors is that QM and tune-up programs provide contractors with tools and abilities that can be applied to jobs outside of the program as well as inside the program. Contractors may choose not to submit claims for rebates for a number of reasons including: 1) seeking to avoid paperwork burden and other “hassle” costs, 2) being aware that a given project did not completely follow program protocols; or 3) their customers chose not to pursue the rebates, for whatever reason.

We calculated spillover based on the following two survey questions:

- **Spill1.** Do you also use QM or tune-up program training, checklists, tools, and/or protocols for HVAC maintenance jobs that do not receive a QM or tune-up program incentive, but are eligible for the QM or tune-up program?
- **Spill2.** <if Spill1=yes> What percent of HVAC maintenance jobs that you perform with these methods and/or tools receive incentives through the QM or tune-up programs?

Table 21 and Table 22 show how we use the responses to questions Spill1 and Spill2 to calculate spillover.

Table 21. Spillover assignment decision based on response to Spill1 question

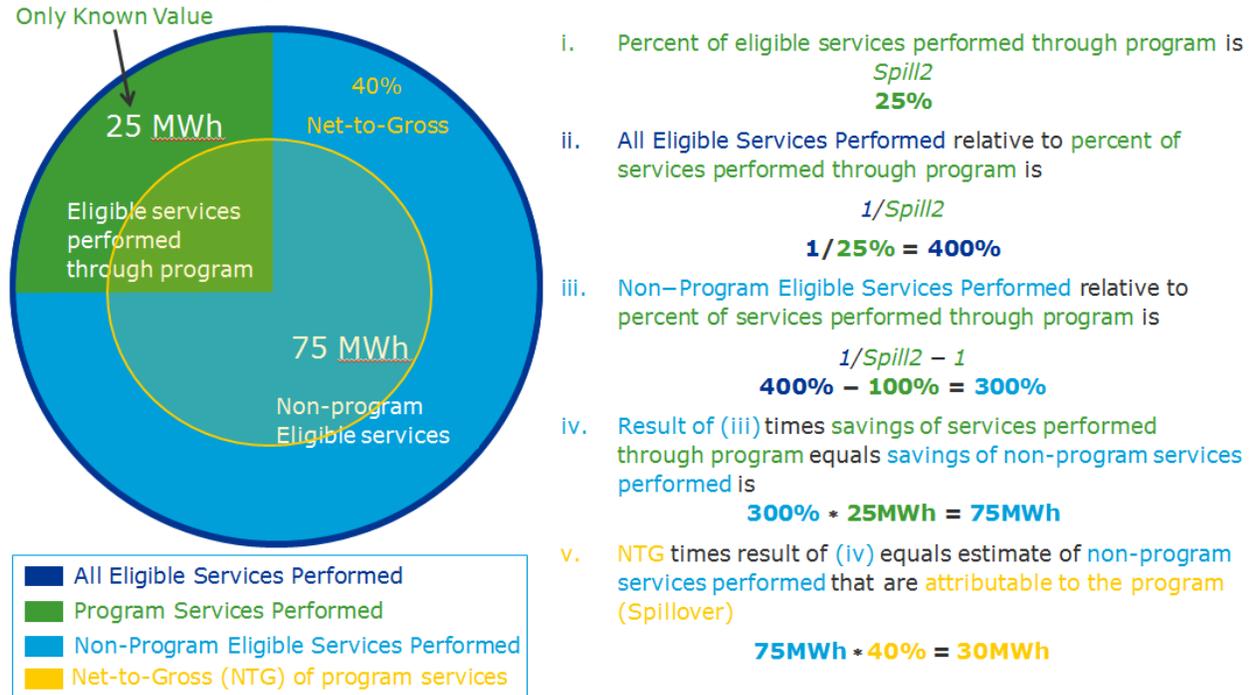
Response to Spill1	NTG assignment
Yes	Proceed to Spill2
No	0% Spillover
Don't know	0% Spillover
Refused	0% Spillover

Table 22. Spillover assignment decision based on response to Spill2 question

Response to Spill2	NTG assignment
% in program	Spillover = $[(1/ \% \text{ in program}) - 100\%] * \text{NTG}$
Don't know	0% Spillover
Refused	0% Spillover

The NTG assignment in Table 22 first estimates the total amount of services that the contractor performed (within and without the program) relative to the percent performed within the program. Subtracting 100% from the total amount provides an estimate of non-program activities using program methods relative to the amount performed as part of the program. Multiplying by the NTG provides an estimate of non-program activities attributable to the program.

Figure 2. Estimating spillover for each contractor based on response to question Spill2



This method likely over-represents spillover. The formula in Table 22 applies the same level of program attribution (the NTG ratio) to their non-program HVAC maintenance jobs as it does to the maintenance jobs which were claimed through the program. This is a simplifying assumption that overstates the program attribution. This overstatement is due to the fact that the NTGR for the program activities reflects the impact of both the program financial incentives and the non-financial program activities (e.g., training, checklists, tools, etc.) whereas the program’s influence over non-program activities would be limited to non-financial program activities only.

In addition to these more direct quantifications of spillover, other questions in the survey instrument also explore the issue as to whether participation in the program changed the standard practices of the HVAC contractors. Such changes in standard practices could potentially generate additional spillover savings in the future, even with the cessation of the QM programs, if there was evidence that they were 1) program-induced; and 2) sustainable in the absence of the programs.

These survey questions exploring potential changes in standard practices included:

- **Frame3.** Did you offer any of those services to customers before working with the <program>?
- **Frame3a.** Which of these services did you offer before participating in the program?
- **Frame6.** Has the <program> caused you to change your HVAC maintenance or tune-up service offerings?
- **Frame 6b.** How has the program caused you to change your HVAC maintenance and tune-up service offerings?

We did not use the responses to these questions to generate any additional estimates of spillover energy savings. However, we did use the responses to these questions to gain a better understanding of the possible market transformative impacts of these programs both quantitatively (e.g., What percentage of the

HVAC contractors changed their maintenance/tune-up practices as a result of participating in the programs?) and qualitatively (e.g., Which of their maintenance/tune-up practices changed as a result of program participation?)

4.2.2 Findings

Spillover was found to be negligible in all four programs. Table 23 shows the spillover estimates for both the QM and Tune-up program groups. For both program types the spillover is very small, and indistinguishable from zero given the error bound.

Table 23. Combined program spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		Potential Spillover	90% CI FPC Off	Potential Spillover	90% CI FPC Off	Potential Spillover	90% CI FPC Off
All Programs	48	0.01	±0.01	0.01	±0.01	0.02	±0.03

Table 24 through Table 27 present spillover estimates and the confidence intervals for each of the programs individually. As noted in section 4.2.1, because the CPUC has a separate team of consultants estimating a more generic cross-program spillover “adder,” we estimated potential spillover for the commercial QM programs only for informational purposes. It will not be included in the NTG score to avoid double counting with the global spillover adder. Also, as noted in that subsection, our spillover method likely overstates the spillover savings somewhat because it applies the same assumption of program attribution (the NTGR) to the non-program HVAC maintenance activities as it does to the program HVAC maintenance activities, even though the non-program activities did not receive the benefits of the program incentives.

Table 24. PG&E Commercial QM spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		Potential Spillover	90% CI FPC Off	Potential Spillover	90% CI FPC Off	Potential Spillover	90% CI FPC Off
Coil Cleaning*	0						
Economizer	14	0.00	±0.01	N/A	N/A	N/A	N/A
Fan Control	21	0.00	±0.00	N/A	N/A	0.00	±0.00
RCA	16	0.35	±0.43	0.29	±0.35	N/A	N/A
Thermostat	18	0.05	±0.08	N/A	N/A	0.04	±0.08
Overall	22	0.03	±0.04	0.29	±0.36	0.04	±0.07

*None of the contractors interviewed for PG&E’s Commercial QM program were matched with coil cleaning measures in the tracking data.

Table 25. PG&E AirCare Plus spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		Potential Spillover	90% CI FPC Off	Potential Spillover	90% CI FPC Off	Potential Spillover	90% CI FPC Off
Coil Cleaning	3	0.00	±0.00	0.00	±0.00	N/A	N/A
Economizer	3	0.00	±0.00	N/A	N/A	N/A	N/A
Fan Control	2	0.00	±0.00	N/A	N/A	0.00	±0.00
RCA	3	0.00	±0.00	0.00	±0.00	N/A	N/A
Thermostat	3	0.00	±0.00	N/A	N/A	0.00	±0.00
Overall	3	0.00	±0.00	0.00	±0.00	0.00	±0.00

Table 26. SCE QM spillover estimates

Measure Group	n	Electric (kWh)	
		Potential Spillover	90% CI FPC Off
Maintenance	16	0.01	±0.02
Overall	16	0.01	±0.02

Table 27. SDG&E Deemed Incentives spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)	
		Potential Spillover	90% CI FPC Off	Potential Spillover	90% CI FPC Off
Coil Cleaning	7	0.00	±0.01	0.00	±0.01
Economizer*	0				
RCA	6	0.00	±0.01	0.00	±0.01
Overall	7	0.00	±0.01	0.00	±0.01

*None of the contractors interviewed for SDG&E's Deemed program were matched with economizer measures in the tracking data.

As noted in subsection 4.2.1, in addition to the spillover questions (Spill1 and Spill2) we also asked other survey questions (Frame3, Frame3a, Frame6, and Frame6b) which explored the possible market transformative impacts of these programs both quantitatively and qualitatively. These questions explored the issue as to whether participation in the program changed the standard practices of the HVAC contractors and how these practices might have changed. A fuller analysis of these survey responses appears in APPENDIX D.

In summary, these survey responses indicated that the vast majority of the respondents had offered coil cleaning of the condensers, RCA services, economizer repair, or fan control adjustment before working with the program. The two maintenance services which the contractors were less likely to say they had offered before working with the programs were the coil cleaning of the evaporator and the thermostat measure.

When we asked the HVAC contractors whether the program caused them to change their HVAC maintenance or tune-up service offering, less than a third (31%) said that they had. Those that reported changes in their offerings mostly cited either: 1) the program incentives encouraging them to promote maintenance services they had not previously promoted; or 2) the program trainings enhancing their existing maintenance services.

The first category of changes in offerings is clearly program-dependent since it relies on the availability of program incentives. The second category of changes holds out some hope of more sustainable changes in practices. For example, some contractors mentioned using more diagnostic equipment in their maintenance practices or taking a more disciplined and consistent approach to their maintenance. However, the survey responses also raised questions as to whether the higher quality of these maintenance practices would continue without the program incentives. For example, one of the contractors said that they stopped participating in the program because they realized how much time the extra quality control was taking and when the incentives were reduced, the costs ended up outweighing the benefits. In addition, some of the contractors explicitly said that they started expanding services like economizer repairs and thermostat adjustments because the incentives made it more appealing for both them and their customers.



One of the contractors reported that because of the program trainings, they were better able to explain the benefits of maintenance services to their customers. Such improvements in salesmanship could potentially increase customer uptake of maintenance services absent program support. However, only one contractor mentioned this. In addition, one wonders how effective these sales pitches would be without the “sweetener” of the program rebate. Finally, the absence of the rebates could also discourage HVAC contractors from making such a sales pitch in the first place, due to the diminished likelihood of success.

In conclusion, it is unlikely that these programs are producing any long-term market effects among the population of participating HVAC contractors. The vast majority of the contractors are reporting that they offered these maintenance services before joining the program. So, with a few exceptions, the program is not teaching them to offer new services which they did not offer before. There was some limited evidence that the program had encouraged a few of these contractors to improve the quality of their maintenance services. However, there was also evidence that the higher quality of these maintenance services were dependent on the program incentives. Finally, the generally low spillover savings casts further doubt on the prospect for market effects. If the contractors viewed offering these maintenance services as good value proposition even without the program incentives, one would expect the program spillover results to be a lot higher.

5 CONCLUSIONS AND RECOMMENDATIONS

The error bounds for the program-level NTGRs were high due to large variation in contractor responses and a low number of responses in some programs. Due to this, we decided to produce NTG results shown in Table 1 where the program-level NTGRs were combined for all programs.

Even the error bounds for the blended NTGRs, are relatively high, so we are cautious about drawing conclusions or making recommendations for program redesign. However, the fact that the kWh NTGRs are less than 50% for most of the measures, three of the four programs, and for the blended program designs indicates that these HVAC programs are having only modest impacts on the maintenance practices of these HVAC contractors. Other evidence from the contractor surveys – such as the small percentage of HVAC contractors reporting changes in their maintenance service offerings due to program participation – support the conclusion that these effects are modest. A non-participant baseline study of HVAC maintenance standard practice would be extremely valuable, and we would recommend doing such a study before making any drastic program changes. It could be that the QM programs are attracting above average contractors who were already performing the program measures, but whether these activities are being done in the marketplace outside of the program contractors is still an open question.

The HVAC contractor surveys also revealed little evidence for spillover savings or any long-term market effects. The electric spillover savings ranged from zero to two percent. The vast majority of the contractors are reporting that they offered these maintenance services before joining the program. So, with a few exceptions, the program is not teaching them to offer new services which they did not offer before. There was some limited evidence that the program had encouraged a few of these contractors to improve the quality of their maintenance services. However, there was also evidence that the higher quality of these maintenance services were dependent on the program incentives. Finally, the generally low spillover savings casts further doubt on the prospect for market effects. If the contractors viewed offering these maintenance services as a good value proposition even without the program incentives, one would expect the program spillover results to be a lot higher.

We hypothesized that there may have been response bias where the more progressive program contractors may have been more likely to respond to our survey. Though we did not have any way to gauge the progressiveness of a contractor, we were able to correlate contractor responses to whether they continued to participate in the program in 2015, thinking that less progressive contractors may drop out of the program. All of SCE's program contractors continued to participate in 2015, giving no information about a potential response bias, but does say something about program contractor retention. In PG&E and SDG&E programs the percent of complete surveys with contractors active through 2015 was about 10% higher for our survey respondents than for the program overall. This represents a slight bias but not an overwhelming effect.

There were a few encouraging results. As noted, the NTGRs for the RCA measure were much higher than those for other maintenance measures. These NTGRs were also higher across three different programs. These results suggest that the programs are influencing the adoption of this measure group. However, the



gross savings report also found low gross realization rates for this measure. Therefore this measure is unlikely to contribute significant savings despite the higher NTGRs.⁸

As described in the methodology, the NTG evaluation method was designed to be consistent with the gross methodology, where savings were broken down to a measure level, and the baseline was assumed to be the pre-maintenance state. If future evaluations were to consider the issue of service maintenance quality improved through the QM program then it would be important for both the gross savings and the NTG assessments to consider the issue. The impact evaluation would need to assess quality by comparing ex ante and ex post participant system operation to that of systems that did not participate in the program. We would need to characterize both groups, and if the contractors in the program have significant firmographic differences from all other contractors it would be more difficult to attribute program effects. Regardless, achieving a representative sample of maintenance outside the program would be subject to potential contractor self-selection bias and would need to address the issue of contractors who participated in past programs, but not in the current program. The scope and scale of the effort and methodological issues to overcome appear greater than what is justified by current levels of program activity. Therefore the evaluation team is not recommending this approach. If this is a future priority, then additional funding will need to be allocated to the evaluation effort for the additional surveys and recruitment to complete monitoring.

Recommendations include:

- Recommend improving the quality of the tracking data. The lack of well-populated customer and contractor contact information in the tracking database impacted this study, and also the gross impact study. It has not only cost the evaluation more time (and budget) doing additional data requests and looking up missing information, but also impacted the sample size (and result precision) since some contractors could not be reached at all.
- Recommend a non-participant baseline study of HVAC maintenance standard practice. It would be important to coordinate with the gross impact study, and for both studies to incorporate standard practice baseline studies during the same evaluation cycle for consistency. As mentioned, care will need to be taken to create a non-participant contractor sample with similar characteristics to the program participant contractors.

⁸ In addition, experts from the CPUC HVAC Laboratory Project Coordination Group now suggest that refrigerant charge should only be adjusted when the system is significantly over or under-charged, HVAC Lab Testing Meeting. PG&E San Ramon Advanced Technologies Center, Thursday, November 12, 2015.

APPENDIX A. Installation Contractor Survey

**CPUC
HVAC 3
Contractor NTG survey**

Note: This document is currently designed for trained DNV GL staff to execute surveys.

Introduction	
	Name
Intro4	Position
Intro5	Years in position
Framing	
	What are <program> basics
Frame1	When did you start with the <program>?
Services	
Frame2	Offered through program
Frame3	Offered before program
Frame4,5	Offered outside program
Frame6	Changes to...
Frame7	If program ended, would you change...
Sales Practices	
Frame8	Offered through program
Frame9	Offered before program
Frame10	Offered outside program
Frame11	Changes to...
Frame12	If program ended, would you change...
Impression of Customers	
Frame13	Customers informed about QM/TU
Frame14	Importance of QM/YU
Satisfaction with	
Frame15	Training
Frame16	Leads
Frame17	Incentives
Frame18	Marketing
Attribution	
Frame3	QM/TU offered before the program
Frame3a	If so, were services different and how
Attr1	Any participating customers received services prior to program
Attr1a	If so, what %
Attr2	[ASSESSMENT ATTRIBUTION QUESTION]
Spillover	
Spill1	Program used outside of incentive jobs
Spill2	What %

Introduction

Hello, my name is ___ and I'm calling from DNV GL on behalf of the <program> that is sponsored by <IOU>. We are calling for research for the California Public Utilities Commission about the <program> that your company participates in, not for any sales purposes. Is <contact> available?

Intro:

According to records from <IOU>, your company participated in the <program>. <IOU> wants to ask your company a few questions that will allow them to improve their program.

[IF NECESSARY, ADD: "We're not selling anything, this is purely for research purposes to help <IOU> improve this <program>"]

[IF NECESSARY, ADD: "All your responses will be kept confidential."]

Intro1. Are you familiar with the <program> program?

1	Yes	Goto Intro4.
2	No	Intro1a
98	Don't know	
99	Refused	

a) Is there somebody else with your company who might be familiar with this program?

1	Yes	Goto Intro2
2	No	Thank and terminate
98	Don't Know	Thank and terminate
99	Refused	Thank and terminate

Intro2. What is the name and contact information of the person you suggest?

- a) Name
- b) Title
- c) Phone
- d) Call back date
- e) Call back time

Intro3. And can you give me your name, so I can mention it when I call?

- a) Name
- b) Title

Thank and terminate.

[Keep following up until the right person is on the phone, then start at beginning]

Intro4. Could you please tell me what your position is at <company>?

Intro5. And how long have you been working in this position at <company>?

Framing

Services

Frame1. Approximately how many years ago did your company first begin participating in the <program>?

1	Record # Years	Frame2
98	Don't know	
99	Refused	

Frame2. What services do you offer to customers through the <program>?

[If necessary: we are asking about any and all potential services that you offer to customers through <program>.]

(Select *all* that are mentioned. If they say that they offer a package of maintenance services, record and ask them to specify which services this package includes.)

1	A package of maintenance services	Frame3
2	Coil cleaning – condenser	
3	Coil cleaning – evaporator	
4	Refrigerant charge adjustment RCA	
5	Economizer repair	
6	Thermostat adjustment	

7	Fan Control adjustment	Intro1a
77	Other: (Record)	
98	Don't Know	
99	Refused	

Frame3. Did you offer any of those services to customers before working with the <program> in <Q7 response>?

1	Yes	Frame3a
2	No	Frame4
98	Don't Know	
99	Refused	

- a) Which of these services did you offer before participating in the program?
(Select *all* that are mentioned. If they say that they offered a package of maintenance services before program involvement, ask them to specify which services this package included.)

1	Coil cleaning – condenser	Frame4
2	Coil cleaning – evaporator	
3	RCA	
4	Economizer	
5	Thermostat	
6	Fan Control	
77	Other: (Record)	
98	Don't Know	
99	Refused	

Frame4. Do you currently offer any HVAC maintenance or tune-up services to customers that are not participating in the <program>?

1	Yes	Frame5
2	No	Frame6
98	Don't Know	
99	Refused	

Frame5. What types of HVAC maintenance or tune-up services do you offer these non-program customers?
(Select all that are mentioned. If they say that they offer a package of maintenance services, ask them to specify which services this package includes.)

1	Coil cleaning – condenser	Frame6
2	Coil cleaning – evaporator	
3	RCA	
4	Economizer	
5	Thermostat	
6	Fan Control	
77	Other: (Record)	
98	Don't Know	
99	Refused	

Frame6. Has the <program> caused you to change your HVAC maintenance or tune-up service offerings?

1	Yes	Goto Frame6b
2	No	If Frame3≠1, Goto Frame6a, else goto Frame7
98	Don't Know	

99	Refused	
----	---------	--

- a) Can you please clarify? You previously said that prior to participating in the program, you did not offer <maintenance/tune-up services mentioned Frame4>.

(Change responses to appropriate questions, OR Record explanation)

(This is a consistency check – be sure to be polite and to not “badger the witness”)

1	Record explanation:	Frame7
98	Don't Know	
99	Refused	

- b) How has the program caused you to change your HVAC maintenance and tune-up service offerings?

1	Record:	Frame7
98	Don't Know	
99	Refused	

Frame7. If the program were to end tomorrow, would you continue to offer all of the same HVAC maintenance and tune-up services that you are currently offering?

1	Yes	Frame8
2	No	Frame7a
98	Don't Know	Frame8
99	Refused	

- a) Which HVAC maintenance or tune-up services would you stop offering if the program went away? (Select *all* that are mentioned)

1	Coil cleaning – condenser	Frame7b
2	Coil cleaning – evaporator	
3	RCA	
4	Economizer	
5	Thermostat	
6	Fan Control	
77	Other: (Record)	Frame8
98	Don't Know	
99	Refused	

- b) Why would you stop offering these maintenance or tune-up services if the program went away?

1	Lack of customer interest	Frame8
2	Unavailability of rebates	
77	Other: Record	
98	Don't Know	
99	Refused	

Sales Practices

Frame8. I would like to ask about your sales practices. How do you sell <program> services to customers?

Probe for:

- How new customers are identified
- Sales pitch
- Feature energy efficiency
- Feature QM or TU branding?
- Mention programs specifically by name?

Frame9. [Ask if Frame4 =1, else go to Frame11]

You indicated earlier that you also offer HVAC maintenance or tune-up services to customers who are not in this program. How do you sell non-program maintenance or tune-up services to customers? [If it is a different sales method than the program sales method (see response to Frame8), probe for reasons for these differences]

1	Record:	Frame10
98	Don't Know	
99	Refused	

Frame10. [Ask if Frame3 =1, else go to Frame11]

You indicated earlier that you offered HVAC maintenance or tune-up services to customers *before* you joined the program? What were your sales practices for these HVAC maintenance or tune-up services before participating in the program?

[If it is a different sales method than the program sales method (see response to Frame8), probe for reasons for these differences]

1	Record:	Frame11
98	Don't Know	
99	Refused	

Frame11. Has the program caused you to change your sales practices for HVAC maintenance or tune-up services?

1	Yes	Goto Frame11b
2	No	If sales practices in response to Frame8 (with program) are different than those for Frame 10 (before program), then Goto Frame11a, else goto Frame12
98	Don't Know	Frame 12
99	Refused	

- a) Could you please clarify? You previously said that prior to participating in the program, you did not offer <maintenance/tune-up services mentioned Frame4>, but your sales practices have not changed. How is it that you are offering more services, but your sales practices are unchanged?

(Change responses to appropriate questions, OR Record explanation)

(This is a consistency check – be sure to be polite and to not “badger the witness”)

1	Record explanation:	Frame12
98	Don't Know	
99	Refused	

- b) How has the program caused you to change your sales practices for HVAC maintenance or tune-up services?

1	Record:	Frame12
98	Don't Know	
99	Refused	

Frame12. [Ask if IF Frame7=1 (THEY WILL CONTINUE TO OFFER HVAC MAINTENANCE/TUNE-UP SERVICES IF THE PROGRAM WENT AWAY) else skip to Frame13]

If the program were to end tomorrow, would you continue to use the same sales practices for your maintenance and tune-up services that you are currently using?

1	Yes	Frame12a
2	No	Frame13
98	Don't Know	
99	Refused	

a) If yes, how?

1	Record:	Frame13
98	Don't Know	
99	Refused	

Impressions of Customers

I would now like to your thoughts about how your customers are reacting to this program.

Frame13. In your opinion, how informed are your customers about the HVAC maintenance/tune-up services which the program supports? On a scale of 1 to 10, where 10 is 'completely informed' and 1 is 'not informed at all,' how informed are your customers about these services?

1	Record Score:	If <8, Frame13a Else Frame14
98	Don't Know	
99	Refused	

a) If <8: What do you think could be done to increase customer awareness of program services?

1	Record:	Frame 14
98	Don't Know	
99	Refused	

Frame14. In your opinion, how much do you think your customers value these HVAC maintenance or tune-up services which the program supports?? Please use a scale of 1 to 10 where 10 is "extremely valuable" and 1 is "not valuable at all".

1	Record Score:	If <8, Frame14a Else Frame15
98	Don't Know	
99	Refused	

a) If <8: What do you think could be done to make customers value these services more?

1	Record:	Frame 15
98	Don't Know	
99	Refused	

Satisfaction & Program Impressions

The program offers various services designed to help contractors sell HVAC maintenance or tune-up services. I'm going to mention four of these. For each one I mention please let me know: a) if you have received that program service 2) And, if yes, how satisfied you have been with it. For your satisfaction we're going to use a 10-point scale where a 10 is "very satisfied" and 1 is "very dissatisfied"

a.

		a. Have you received this program service?	b. How satisfied were you with it?(Use 10 point satisfaction scale)	c. Why do you say that? (If satisfaction <8)
Frame15.	Maintenance Training (provide brief description of service)	If Yes, go to 15b. All other responses go to Frame16		
Frame16.	Sales Leads provide brief description of service)	If Yes, go to 16b. All other responses go to Frame17		
Frame17.	Financial Incentives (Rebates) provide brief description of service)	If Yes, go to 17b. All other responses go to Frame18		
Frame18.	Program Marketing Materials and Advertisement provide brief description of service)	If Yes, go to 18b. All other responses go to Attr1		

Attribution

Attr1. [Ask IF Frame3=1, else go to Attr2]

You indicated earlier that you had offered HVAC maintenance or tune-up services before joining this program. Of the customers you have offered program services to, did any of them receive your maintenance/tune-up services prior to program participation?

1	Yes	Attr1a
2	No	
98	Don't Know	
99	Refused	

a) Approximately what % of your customers that have participated in the program also received your maintenance/tune-up services prior to program participation?

1	Record %	Attr2
98	Don't Know	
99	Refused	

Attr2. In [YEAR] you implemented [#] maintenance/tune-up measures through [PROGRAM NAMES] including [LIST OF QUANTITIES OF TOP 5 MEASURES]. If these program trainings, customer leads,

program endorsements, program market materials and program incentives had not been available, approximately what % of maintenance/tune-up services would you still have provided in [YEAR]?

1	Record %	Attr2a
98	Don't Know	Attr3
99	Refused	

a) Why do you say that?

1	Record:	Attr3
98	Don't Know	
99	Refused	

Attr3. Are there any particular maintenance/tune-up measures that stand out to you as not likely to occur without the assistance of the program?

1	Yes	Attr3a
2	No	Spill1
98	Don't Know	
99	Refused	

a) Which maintenance/tune-up measures, in particular, do you think would **not** likely to occur without the assistance of the program? [Record all identified]

1	Coil cleaning – condenser	Attr3b
2	Coil cleaning – evaporator	
3	RCA	
4	Economizer	
5	Thermostat	
6	Fan Control	Spill1
98	Don't Know	
99	Refused	

b) [IF THEY NAMED SPECIFIC MEASURES] You just said that if these program trainings, customer leads, program endorsements and program incentives had not been available, you **still** would have provided approximately [X%] of maintenance/tune-up services that you provided in [YEAR]? What % of [MEASURE X] would you have provided in the absence of the program?

[Note to reviewers: Data collection instrument will include note to ensure interviewers understand to look for contradictions between responses to A and B Also, collection instrument will automatically fill responses from previous questions in here.]

1	Coil cleaning – condenser	Spill1
2	Coil cleaning – evaporator	
3	RCA	
4	Economizer	
5	Thermostat	
6	Fan Control	
77	Other: (Record)	
98	Don't Know	
99	Refused	

[NOTE TO INTERVIEWERS: IF THE RESPONDENT UNDERSTOOD THE QUESTION, THE MEASURE-SPECIFIC % (b) SHOULD BE LOWER THAN THE PROGRAM-SPECIFIC QUESTION. IF THIS IS NOT THE CASE, PLEASE CLARIFY THE QUESTION AND RESPONSE WITH THE RESPONDENT

IF THE RESPONDENT MENTIONS MULTIPLE MEASURES WITH DIFFERENTIAL EFFECTS, LIMIT THE MEASURE-SPECIFIC QUESTIONS ONLY TO THREE MEASURES [TO BE DETERMINED BY EVALUATION TEAM BASED ON GROSS SAVING IMPACTS]

Spillover

Spill1. Do you also use QM or tune-up program training, checklists, tools, and/or protocols for HVAC maintenance jobs that do not receive a QM or tune-up program incentive, but are eligible for the QM or tune-up program?

1	Yes	Spill2
2	No	F1
98	Don't Know	
99	Refused	

Spill2. What percent of HVAC maintenance jobs that you perform with these receive methods and/or tools receive incentives through the QM or tune-up programs?

1	Record %	F1
98	Don't Know	
99	Refused	

Firmographics

F1. Does your company have more than one location?

1	Yes	F2
2	No	F3
98	Don't Know	
99	Refused	

F2. Do you work out of the main office or is this a satellite or local branch?

1	Main office	F3
2	Satellite office/ Local branch	
3	Other: Specify	
98	Don't Know	
99	Refused	

F3. About how many full time employees work at this location?

[IF THEIR COMPANY HAS MORE THAN ONE LOCATION, ADDITIONALLY ASK ABOUT HOW MANY EMPLOYEES AT ALL LOCATIONS]

1	Record # at location	F4
2	Record at all locations (if applicable)	
98	Don't Know	
99	Refused	

F4. Approximately what percentage of your HVAC and water heating equipment sales occurred in the residential versus commercial markets? [%s SHOULD ADD UP TO 100%]

1	Residential %	F5
2	Commercial %	
3	Other	
98	Don't Know	
99	Refused	

F5. I'm going to read a short list of services. Please tell me if your company performs the service for residential customers.

- a. Repairs
- b. Replacements
- c. New Installations

F6. Do

yo	1	Yes	End
u	2	No	Goto F7
off	98	Don't Know	End
er	99	Refused	End

these same services to your commercial customers?



F7. How are your commercial offerings different?

1	Record # at location	End
98	Don't Know	
99	Refused	

End: Thank and terminate

APPENDIX B. Retrospective NTGR and Potential Spillover Confidence Intervals

Table 29 through Table 32 present the NTGRs and confidence intervals that should be used if the results are applied to the program years studied.

Table 28. Program Type retrospective NTGRs

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		NTGR	90% CI FPC On	NTGR	90% CI FPC On	NTGR	90% CI FPC On
QM	38	0.31	±0.11	1.00	±0.00	0.21	±0.11
Tune-up	10	0.43	±0.17	0.52	±0.05	0.28	±0.79

Table 29. PG&E Commercial QM retrospective NTGRs

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		NTGR	90% CI FPC On	NTGR	90% CI FPC On	NTGR	90% CI FPC On
Coil Cleaning*	0						
Economizer	14	0.21	±0.19	N/A	N/A	N/A	N/A
Fan Control	21	0.05	±0.04	N/A	N/A	0.07	±0.04
RCA	16	1.00	±0.00	1.00	±0.00	N/A	N/A
Thermostat	18	0.21	±0.12	N/A	N/A	0.21	±0.11
Overall	22	0.16	±0.10	1.00	±0.00	0.21	±0.11

*None of the contractors interviewed for PG&E's Commercial QM program were matched with coil cleaning measures in the tracking data. Coil cleaning was performed through the program, but not claimed or reported in the tracking data so it was impossible to link the activity to contractors.

Table 30. PG&E AirCare Plus retrospective NTGRs

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		NTGR	90% CI FPC On	NTGR	90% CI FPC On	NTGR	90% CI FPC On
Coil Cleaning	3	0.00	±0.00	0.00	±0.00	N/A	N/A
Economizer	3	0.55	±0.94	N/A	N/A	N/A	N/A
Fan Control	2	0.00	±0.00	N/A	N/A	0.00	±0.00
RCA	3	0.37	±0.88	0.43	±0.94	N/A	N/A
Thermostat	3	0.31	±0.79	N/A	N/A	0.28	±0.75
Overall	3	0.29	±0.56	0.18	±0.37	0.28	±0.79

Table 31. SCE QM retrospective NTGRs

Measure Group	n	Electric (kWh)	
		NTGR	90% CI FPC On
Maintenance	16	0.49	±0.14
Overall	16	0.49	±0.14

Table 32. SDG&E Deemed Incentive retrospective NTGRs

Measure Group	n	Electric (kWh)		Electric Demand (kW)	
		NTGR	90% CI FPC On	NTGR	90% CI FPC On
Coil Cleaning	7	0.48	±0.05	0.50	±0.05
Economizer*	0				
RCA	6	0.84	±0.04	0.79	±0.03
Overall	7	0.54	±0.05	0.55	±0.05

*None of the contractors interviewed for SDG&E's Deemed Incentives program were matched with economizer measures in the tracking data.

Table 33through

Measure Group	n	Electric (kWh)	
		Potential Spillover	90% CI FPC On
Maintenance	16	0.01	±0.01
Overall	16	0.01	±0.01

Table 37 present the potential spillover estimates and confidence intervals that should be used if the results are applied to future program years.

Table 33. Program type retrospective spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		Potential Spillover	90% CI FPC On	Potential Spillover	90% CI FPC On	Potential Spillover	90% CI FPC On
QM	38	0.02	±0.02	0.29	±0.26	0.04	±0.05
Tune-up	10	0.00	±0.00	0.00	±0.00	0.00	±0.00

Table 34. PG&E Commercial QM retrospective spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		Potential Spillover	90% CI FPC On	Potential Spillover	90% CI FPC On	Potential Spillover	90% CI FPC On
Coil Cleaning*	0						
Economizer	14	0.00	±0.01	N/A	N/A	N/A	N/A
Fan Control	21	0.00	±0.00	N/A	N/A	0.00	±0.00
RCA	16	0.35	±0.27	0.29	±0.23	N/A	N/A
Thermostat	18	0.05	±0.05	N/A	N/A	0.04	±0.04
Overall	22	0.03	±0.03	0.29	±0.26	0.04	±0.05

*None of the contractors interviewed for PG&E's Commercial QM program were matched with coil cleaning measures in the tracking data.

Table 35. PG&E AirCare Plus retrospective spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)		Gas (Therms)	
		Potential Spillover	90% CI FPC On	Potential Spillover	90% CI FPC On	Potential Spillover	90% CI FPC On
Coil Cleaning	3	0.00	±0.00	0.00	±0.00	N/A	N/A
Economizer	3	0.00	±0.00	N/A	N/A	N/A	N/A
Fan Control	2	0.00	±0.00	N/A	N/A	0.00	±0.00
RCA	3	0.00	±0.00	0.00	±0.00	N/A	N/A
Thermostat	3	0.00	±0.00	N/A	N/A	0.00	±0.00
Overall	3	0.00	±0.00	0.00	±0.00	0.00	±0.00

Table 36. SCE QM retrospective spillover estimates

Measure Group	n	Electric (kWh)	
		Potential Spillover	90% CI FPC On
Maintenance	16	0.01	±0.01
Overall	16	0.01	±0.01

Table 37. SDG&E Deemed Incentives retrospective spillover estimates

Measure Group	n	Electric (kWh)		Electric Demand (kW)	
		Potential Spillover	90% CI FPC On	Potential Spillover	90% CI FPC On
Coil Cleaning	7	0.00	±0.00	0.00	±0.00
Economizer*	0				
RCA	6	0.00	±0.01	0.00	±0.00
Overall	7	0.00	±0.00	0.00	±0.00

*None of the contractors interviewed for SDG&E's Deemed program were matched with economizer measures in the tracking data.

APPENDIX C. Attribution Scoring

Table 38 through Table 40 present the attribution results by question and program. A “yes” indicates that the respondent attributed some or all of the decision to install the measure to the program. We did not include the PG&E AirCare Plus attribution results to preserve respondent confidentiality. As contractors often installed more than one measure and may have provided different responses across measures, the number of contractors should not be summed across rows.

Table 38. PG&E Commercial QM attribution

Attribution			# of Measures	# of Contractors	% of Weighted Savings (kWh)
Frame3	Attr3	Attr2			
Yes	Yes	Yes	0	0	0%
Yes	Yes	No	1	1	<1%
Yes	No	Yes	11	8	4%
Yes	No	No	11	8	3%
No	Yes	Yes	8	4	14%
No	Yes	No	1	1	<1%
No	No	Yes	13	7	55%
No	No	No	24	11	22%

Table 39. SCE QM attribution

Attribution			# of Measures	# of Contractors	% of Weighted Savings (kWh)
Frame3	Attr3	Attr2			
Yes	Yes	Yes	1	1	2%
Yes	Yes	No	0	0	0%
Yes	No	Yes	0	0	0%
Yes	No	No	0	0	0%
No	Yes	Yes	7	7	82%
No	Yes	No	0	0	0%
No	No	Yes	8	8	16%
No	No	No	0	0	0%

Table 40. SDG&E Deemed Incentives attribution

Attribution			# of Measures	# of Contractors	% of Weighted Savings (kWh)
Frame3	Attr3	Attr2			
Yes	Yes	Yes	0	0	0%
Yes	Yes	No	0	0	0%
Yes	No	Yes	6	4	43%
Yes	No	No	0	0	0%
No	Yes	Yes	1	1	15%
No	Yes	No	0	0	0%
No	No	Yes	5	4	43%
No	No	No	1	1	<1%

APPENDIX D. Possible Market Effects

In addition to these more direct quantifications of spillover, other questions in the survey instrument also explored the issue as to whether participation in the program changed the standard practices of the HVAC contractors. Such changes in standard practices could potentially generate additional spillover savings in the future, even with the cessation of the QM programs, if there was evidence that they were 1) program-induced; and 2) sustainable in the absence of the programs.

These survey questions exploring potential changes in standard practices included:

- **Frame3.** Did you offer any of those services to customers before working with the <program>?
- **Frame3a.** Which of these services did you offer before participating in the program?
- **Frame6.** Has the <program> caused you to change your HVAC maintenance or tune-up service offerings ?
- **Frame 6b.** How has the program caused you to change your HVAC maintenance and tune-up service offerings?

The interviewers asked the contractors:

- **Frame3.** Did you offer any of those services to customers before working with the <program>? ⁹

Only three of the 49 responding contractors said they had *not* offered *any* of the program services before working with the programs (one respondent did not know). This was not too surprising since some forms of equipment maintenance have long been standard service offerings in the HVAC industry and the word “any” encompasses a broad range of possible services. In addition, as discussed below, this does not account for the fact that the program may have improved the quality of existing maintenance services.

The interviewers then asked the 45 contractors who said that they had offered program-qualifying maintenance services before working with the program the following question:

- **Frame3a)** Which of these services did you offer before participating in the program?

The vast majority of the respondents said that they had offered coil cleaning of the condensers, RCA services, economizer repair, or fan control adjustment before working with the program. The two maintenance services which the contractors were less likely to say they had offered before working with the programs were the coil cleaning of the evaporator and the thermostat

The interviewers also explored the influence of the programs on standard HVAC maintenance practices through another survey question:

- **Frame6.** Has the <program> caused you to change your HVAC maintenance or tune-up service offerings?

Fifteen of the 49 responding contractors (31%) said that the program had caused them to change their HVAC maintenance or tune-up services offerings. The interviewers asked them the follow-up question: “Frame 6b) How has the program caused you to change your HVAC maintenance and tune-up service offerings?.” The contractors gave a wide variety of responses to these questions, but the most common

⁹ “Those services” in this question refers back to the previous question: Frame2. What services do you offer to customers through the <program>?



themes were that the program either encouraged them to expand their maintenance services or encouraged them to improve their salesmanship and delivery of existing maintenance services.

Some of the categories of program influence included:

- *The program incentives encouraged them to promote maintenance services they had not previously promoted:*
 - For example, one contractor said that prior to program participation, when they would sell maintenance services to a customer and noticed that the economizer or thermostat were not working optimally, they normally would not offer to repair or adjust these because they assumed that the customers would not want to pay for these services. However, when the QM programs provided incentives for economizers and thermostats, it made it easier for the contractor to offer these services and for the customers to accept them.
 - A second contractor also said that his company focuses more on economizers than they did before joining the program.
 - A third contractor said that before the program his company would just see whether the thermostat works and now they are adjusting and tuning the thermostats.
 - A fourth contractor said that prior to program participation, their maintenance services were limited to changing filters and coils on an annual basis.
 - A fifth contractor said that the program had “improved the range of their offerings.”
 - A sixth contractor said that they had not offered any maintenance services before becoming involved with the program.
 - A seventh contractor said that the initially expanded their maintenance offerings due to the program but stopped these enhanced offerings due to the paperwork burden and reductions in program rebate levels.
- *The program trainings enhanced their existing maintenance services:*
 - One contractor said that because of the PG&E trainings, his company was able to do more diagnostics than they could before and that after program participation “they do better quality control of the unit and especially the economizer end of it.”
 - A second contractor said the programs did not change what services they offered but “now they have higher standards, they are more disciplined, more critical in how they perform their maintenance services.”
 - A third contractor said that while they are offering the same maintenance services that they offered before the program, they are now “more sophisticated in what they can offer or how they present it to customer.”
 - A fourth contractor said that because of the program trainings, they were better able to explain the benefits of maintenance services to their customers.

- 
- *The program raised customer expectations about what constituted quality maintenance:* One contractor said that the QM programs helped increase “the importance of maintenance” among their customers and made “a higher level of maintenance” more visible within their customer base.
 - *Shifting their customers focus:* One contractor said that the program had encouraged them to focus more on commercial customers than they had before.
 - *Mentioning the rebates in their sales pitch:* One contractor noted that they now mention the program rebates in their sales pitches to customers.

APPENDIX E. Public Comments & Responses

Numerous comments were received from members of the public. The DNV GL NTG team responses are provided in Table 41.

Table 41. Draft report comments from members of the public and DNV GL team responses

No.	Commenter	Subject	Sect./Pg.	Comment	Response
1	Abram Conant	Uncertainty	Overall	Combining highly uncertain results, ranging from 0 to 1, for many dissimilar measures, into a single NTG value does not eliminate the uncertainty. Instead of producing an extremely uncertain value that is likely to be misinterpreted by readers of the study as being far more meaningful than it actually is, we suggest concluding that the methods applied were not successful in identifying NTG with acceptable confidence. Recommendations to improve future evaluations would be more useful than forcing a numerical result with little confidence that it's actually representative of the NTG for these programs.	The report does recommend a cautious interpretation of the NTG results: e.g., page 2, ""Because the error bounds, even for the blended NTGRs in Table 1, are relatively high, we are cautious about drawing conclusions or making recommendations for program redesign." We agree that it would be useful to find ways to improve these evaluations going forward. We did suggest a few of these in the report. For example, we did mention that the IOUs did not initially provide us with contact information for the participating HVAC contractors and this negatively impacted our interview completion rate which, in turn, negatively impacted our error bounds due to smaller samples sizes. We have added emphasis on this in the report as a lesson learned for future evaluations.

No.	Commenter	Subject	Sect./Pg.	Comment	Response
2	Robert Mowris	Summary	pg. 1	<p>The HVAC3 NTG study provides flawed research methodologies, biased survey questions/responses, and analysis representing yet another “false alarm” evaluation of HVAC maintenance programs. The study failed to differentiate between “generic” and “quality” HVAC maintenance measures, and findings are not supported by previously published research reports and currently available evidence. The study should have included unbiased survey questions to evaluate the difference between generic services and quality maintenance services offered by the program and asked these questions of more participants using better analysis and reporting with greater transparency, oversight and accountability in order to properly evaluate NTG ratios. The study should have included a collaborative process and pretest to solicit comments and suggestions from colleagues, program implementers, and industry experts to examine each question and the entire questionnaire to produce reliable and credible results per AEA guidelines. Based on the lack of adherence to the CPUC California Evaluation Framework and American Evaluation Association (AEA) guidelines and biased survey questions, the HVAC3 NTG study should be rejected and the NTG ratios for all CQM measures should be established at 1.0.</p>	<p>Our process for conducting this evaluation was very transparent and received approval from a wide range of stakeholders. The methodology used for this report, as well as the contractor interview guides, were posted on BaseCamp and reviewed by the IOU's and available to their implementers. The methodology documents were reviewed and approved by the CPUC and the CPUC consultants. Direct responses and adjustments were made at each stage of the process. Like all deliverables the final report also includes public stakeholder review which we are responding to now. The feedback provided is used not only in the current report review, but is also considered by the CPUC and IOUs to inform long term research needs and future studies. These consultants and stakeholders included individuals with decades of HVAC and net savings assessment experience. In addition to this extensive review process, before beginning the contractor interviews, our evaluation partners at the Western Cooling Efficiency Center at UC Davis also conducted a focus group-type meeting with some of the participating contractors to make sure we were covering the relevant issues.</p> <p>The commenter makes a number of assertions about “biased” interviews questions but doesn’t explain why he thinks the questions were biased beyond claiming that “they failed to differentiate between ‘generic’ and ‘quality’ HVAC maintenance measures.” Furthermore, the interview guide had over a half dozen questions that allowed HVAC contractors to distinguish the maintenance services and related sales practices that they routinely offer from those they offered through the HVAC3 programs. Therefore, if the participating HVAC contractor thought that the program had changed these maintenance services or sales practices, they had ample opportunity to state this. The relevant interview guide questions included:</p> <ul style="list-style-type: none"> • Frame6. Has the <program> caused you to change your HVAC maintenance or tune-up service offerings? • Frame6. b) How has the program caused you to change your HVAC maintenance and tune-up service offerings? • Frame7. If the program were to end tomorrow, would you continue to offer all of the same HVAC maintenance and tune-up services that you are currently offering? • Frame 9. You indicated earlier that you also offer HVAC maintenance or tune-up services to customers who are not in this program. How do you sell non-program maintenance or tune-up services to customers? [If it is a different sales method than the program sales method (see response to

No.	Commenter	Subject	Sect./Pg.	Comment	Response
					<p>Frame8), probe for reasons for these differences]</p> <ul style="list-style-type: none"> • <u>Frame 10</u>. You indicated earlier that you offered HVAC maintenance or tune-up services to customers before you joined the program? What were your sales practices for these HVAC maintenance or tune-up services before participating in the program? [If it is a different sales method than the program sales method (see response to Frame8), probe for reasons for these differences] • <u>Frame 11</u>. Has the program caused you to change your sales practices for HVAC maintenance or tune-up services? • <u>Frame11b</u>. b) How has the program caused you to change your sales practices for HVAC maintenance or tune-up services? • <u>Frame 12</u>. If the program were to end tomorrow, would you continue to use the same sales practices for your maintenance and tune-up services that you are currently using? <p>In addition to these more direct questions, there were also other questions such as the open-ended follow-up questions (e.g. "Why do you say that?") to the program attribution questions which also allowed respondents to comment on how the program might have changed what they typically did for maintenance. It is also important to point out that the report did discover that the HVAC3 programs did encourage a few of the contractors to improve the quality of their maintenance practices. However, these program effects were very limited.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
3	Robert Mowris	Methodological Issues	pg. 2	<p>There are three significant problems with the NTG survey. The first problem is only asking participating HVAC contractors the questions instead of participating and non-participating contractors, customers and technicians. The second problem is with using self-reported responses to three simple questions to evaluate the NTG ratio for each measure in the Commercial HVAC maintenance programs. The third and most important problem is using biased questions to lead contractors to believe they might have offered exactly the same services prior to the program.</p>	<p>As explained above, the chosen NTG methodology was only implemented after a lengthy and rigorous review and approval process involving a wide range of stakeholders. Our NTG methodology memo explained why we chose not to interview nonparticipating contractors or participating customers. We have added to the report those paragraphs from said memo that explained our reasons for not interviewing these two groups:</p> <ul style="list-style-type: none"> • While budgetary constraints were a consideration, the main concern with the nonparticipating contractors was the risk of self-selection effects. • The NTG methodology memo also explains why we did not interview end users: "The midstream design of the QM and tune-up programs means that participating customers do not have a complete picture of program effects. The 2015 HVAC QI/QM decision-making study found that customers had little awareness of or in interest in QM and mostly just cared whether or not the equipment worked." Also, customers wouldn't know whether the program had encouraged their local HVAC contractor to adopt higher quality maintenance service than they would have otherwise. Nor would they be able to tell that the maintenance service they had received through the program was of higher quality than the maintenance service that they would have received if the program did not exist. <p>The previous response to comment #2 explains that the questions were not biased because the HVAC contractors were given ample opportunity to distinguish their normal maintenance services from those required by the program.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
4	Robert Mowris	Biased / Leading Questions	pg. 2	<p>Asking HVAC contractors simple biased and leading questions will likely produce misleading results that contractors would have provided most, if not all HVAC maintenance services in exactly the same way with exactly the same tools, training, software, and frequency without the program. Contractors businesses are based on HVAC maintenance services. Despite assurances of confidentiality, many are unlikely to concede a lack of competence or ability to provide similar services without program support, which may include fault detection diagnostic (FDD) tools, training, software, equipment, and incentives.</p>	<p>As indicated in our response to comment #2, the questions were not biased because the HVAC contractors were given ample opportunity to distinguish their normal maintenance services from those required by the program.</p> <p>The commenter implies that many HVAC contractors are saying that there is no difference between the program maintenance practices and their routine maintenance practices not because there are no differences, but because the HVAC contractors are reluctant to admit that their normal maintenance practices are of lesser quality. While this kind of bias is possible, it would be very difficult to verify (one would have to secretly observe HVAC technicians doing both program and non-program maintenance visits and compare the differences or maybe audit non-program invoices as the commenter later suggests).</p> <p>Furthermore the report reveals some significant variation in the NTG of the various maintenance measures. For example, the NTG ratios for the measure RCA are very high across most programs while the NTG ratios for other measures like Fan Controls are very low across all programs. If it was just a simple case where the HVAC contractors were unwilling to cite program influence because they "are unlikely to concede a lack of competence or ability to provide similar services without program support," as the commenter contends, then the NTG ratios for all the measures would be low, which is not the case.</p> <p>In addition, it is possible to imagine biases moving in the opposite direction. For example, it would be plausible that some contractors might exaggerate the impacts of the program on their maintenance practices so that they could insure that the program incentives would continue. This is sometimes called the "don't kill the golden goose" bias and has been recognized in other California program evaluations. The bottom line is that, while it is likely that some biases exist, it is very difficult to measure their magnitude and directionality. Absent this information, we have chosen to take the contractors at their word.</p> <p>Finally it is unclear how the NTG of this program could even be evaluated without talking to participating contractors. As discussed in response to comment #3, there are serious problems with trying to evaluate the program based on interviews with nonparticipating contractors or participating customers.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
5	Robert Mowris	Response Bias	pg. 3	<p>"The Frame 3 and Frame 3a questions are missing important information to describe "services" including "using comparable tools, training, software, frequency, and potentially requiring three year service agreements" before working in the program. Without this information the question creates bias towards a response based on prior "generic" services being "exactly the same" as program services, when in fact, this is not the case. Prior "generic" services are not similar to "quality" services provided by the program. If the services were exactly the same, then there would be no need for the extensive program quality control and Workforce Education and Training (WE&T) requirements for participation. This difference is substantial and causes significant bias that cannot be ignored. None of the generic measures described in the HVAC3 questions and responses have been previously offered at the same level of quality or frequency as required by the programs."</p>	<p>As discussed above, the HVAC contractors were given ample opportunity to distinguish their normal maintenance services from those required by the program. Questions such as those that follow are open-ended enough to allow the contractors to bring up examples of how the program maintenance services differ from their routine maintenance service:</p> <ul style="list-style-type: none"> • Frame6. Has the <program> caused you to change your HVAC maintenance or tune-up service offerings? • Frame6. How has the program caused you to change your HVAC maintenance and tune-up service offerings?

No.	Commenter	Subject	Sect./Pg.	Comment	Response
6	Robert Mowris	Historical Maintenance Schedules	pg. 3	Regular HVAC maintenance has typically only included changing air filters. More frequent air filter replacement, coil cleaning, refrigerant charge adjustment (RCA), economizer repair, smart thermostats, and fan control are required by the ACCA 180 Standard upon which the programs are based. Generic air filter replacement and coil cleaning have been historically performed less frequently based on visual inspections without quarterly, semi-annual, or annual frequency. RCA has only been historically performed when systems lost refrigerant charge due to a leak or where technicians added refrigerant only if the suction line was not "six-pack" cold. Given the high cost of refrigerant and lack of time, there would be no reason for technicians to spend valuable time checking and correcting refrigerant charge using state-of-the-art FDD tools and software without training and incentives offered by the programs. Generic economizer repair has historically not been offered by any contractors in California (or anywhere else).	Once again the commenter is insisting that the program encouraged HVAC contractors to provide a higher quality maintenance service than they typically provide. Yet the interviews with participating HVAC contractors indicated that these effects on their typical practices were very limited.
7	Robert Mowris	Contradictory findings	pg. 3 - 4	1. Internet searches of "HVAC Economizer Repair Services" indicate only one HVAC contractor is currently advertising economizer repair services in California. While this contractor is participating in the CQM program, they did not advertise or perform economizer tests or repairs in the 2012-13 program where they performed less than 0.04% of total work.	Citing the lack of economizer repair services in online advertising is problematic since we know that many of the HVAC contractors programs were offering economizer repair services through their participation in the IOU QM programs. Yet, with the singular exception of the example cited by the commenter, none of the participating contractors who offered economizer repair services chose to advertise it, based on the commenters' web search. This only suggests that the economizer services are rarely advertised rather than that they are rarely offered. It is possible that most contractors offering these services did not consider them as "selling points" or as ways to differentiate themselves from their competitors. It is also possible that some HVAC contractors are late adopters when it comes to web advertising and may be relying on more traditional ways of advertising their services such as mailers and non-web yellow page ads.

No.	Commenter	Subject	Sect./Pg.	Comment	Response
8	Robert Mowris	Contradictory findings	pg. 4	2. Furthermore, 85% of total economizer incentives in 2012-13 programs were for testing economizers indicating very few contractors were performing economizer repairs in the 2012-13 program. Clearly, the SCE QM maintenance program has played a substantial role in motivating this contractor and others to provide more comprehensive quality maintenance services including economizer repairs. In spite of this evidence the HVAC3 NTG study only gives the SCE program a NTG of 0.49 +/- 0.28 without any indication of how the NTG ratio applies to economizer repairs or other measures specifically. Research studies indicate approximately 64% of economizers do not function properly and 30% of fan schedules operate continuously during unoccupied periods.	<p>The reason for the lack of measure-level granularity for the NTG in the SCE program is mentioned in the report—the SCE QM program did not report claims on a measure level. That said, the report does provide some evidence that the program was impacting economizer and RCA services as described:</p> <ul style="list-style-type: none"> • <u>Economizer services.</u> In the report, one contractor said they “might not have done as much economizer or thermostat work if rebates weren't offered.” Another contractor indicated that “economizer customers don't want to pay for the service by itself and the program helps with that [through] the rebates.” In addition, for the PG&E AirCare Plus program, the economizer had the highest NTG of all the measures (although the HVAC contractor sample size was very small). • <u>RCA services.</u> RCA had a NTG ratio of 100% for the PG&E Commercial QM program (n=18), 84% for the SDG&E Deemed Incentive program (although HVAC contractor sample sizes were very small n=6). Four of the eight responding contractors in SDG&E program said they had not offered RCA services before joining the program. However, the NTG ratio for the RCA was much lower (0.37) for participants in the PG&E AirCare Plus program.
9	Robert Mowris	Contradictory findings	pg. 4	3. ContractingBusiness.com indicates “economizer repair opportunities are everywhere, but few if any contractors are actually providing economizer repair services”.	The quote from ContractingBusiness.com is presumably generalizing about national practices which may not be relevant to California where years of QM programs may have had some market transformation effects.
10	Robert Mowris	Contradictory findings	pg. 4	4. Interviews of participating contractors and technicians in 2010-14 CQM programs indicated that most were unwilling to attempt economizer repair services. Those that did perform economizer repair services would not have performed these services without program training and incentives.	As the response to the previous comment indicates, there was evidence in the report that some of the HVAC contractors who participated in the program only started offering economizer repairs after joining the program. It also should be noted that the gross realization rates for economizer repairs were only in the 43%-56% range as reported.

No.	Commenter	Subject	Sect./Pg.	Comment	Response
11	Robert Mowris	Methodological Issues (Sample)	pg. 4	<p>Customers and technicians are important market actors in the decision making process. Most customers are not aware of the ANSI/ASHRAE/ACCA Standard 180 or the quality and frequency of services required for optimal HVAC equipment efficiency. Participating customers are required to sign a maintenance service agreement in order for the customer and contractor to receive program incentives. Technicians with a 2-year degree from accredited HVAC trade schools are taught how to provide generic maintenance services, but expediency, lack of FDD tools and software, and pressure by typical HVAC contractors often makes it difficult, if not impossible, for technicians to provide the quality and frequency of maintenance services required for optimal HVAC equipment efficiency. Not including customers and technicians in the HVAC3 NTG study represents a fundamental problem with the HVAC3 NTG study since the programs cannot function without influencing both of these two critical program participants.</p>	<p>Once again the commenter is insisting that the program encouraged HVAC contractors to provide a higher quality maintenance service than they typically provide. Yet the interviews with participating HVAC contractors indicated that these effects on their typical practices were very limited. If it is sometimes the case that the contractors are saying that there is no difference between the program-required maintenance services and their typical maintenance services because they could not recognize a real difference, then this would point to a need for the programs to do more education of contractors so that they can better distinguish these services.</p> <p>The explanation as to why the evaluators did not interview participating customers appears above. In terms of interviewing "technicians", we did interview dozens of HVAC contractor representatives and some of these might be field technicians although we have not checked whether this was indeed the case. But regardless, it is unclear why the commenter thinks that the junior field technicians might have a different perspective than other HVAC contractor representatives we have interviewed.</p> <p>Finally one of the commenter's arguments is that the program encourages higher quality maintenance services than are typically offered. But even if this could be proven, it's important to note that in cases where the counterfactual is not "no maintenance service" but instead "a cheaper, lower quality maintenance service," in such cases the NTG should not be applied to the full gross savings (e.g., the "no maintenance service" baseline), but instead only to that incremental portion of added energy savings that might result (if it could be proven) of using a more expensive, higher-quality, program-supported maintenance service over a less expensive, presumably lesser-quality, typical maintenance services. See more discussion of this below in the discussion of the FDD tools.</p> <p>Although a pretest of the survey instrument would have helped to ensure we were interviewing the right actors, we were not able to complete a pretest due to budget constraints.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
12	Robert Mowris	Methodological Issue (Survey Design)	pg. 4	<p>The HVAC3 NTG study should have included a collaborative process to solicit comments and suggestions from program implementers. The study should also have included a pretest of the survey questions to examine each question and the entire questionnaire to produce reliable and credible results. According to Salant and Dillman (1994), pretesting needs to answer the following questions.</p> <ul style="list-style-type: none"> • Does each question measure what it is intended to measure? • Do respondents understand all the words? • Are questions interpreted similarly by all respondents. • Does each closed-ended question have an answer that applies to each respondent? • Does the questionnaire create a positive impression – one that motivates people to answer it? • Are the answers respondents can choose from correct? • Does any aspect of the questionnaire suggest bias on the part of the researcher? <p>Colleagues, utility personnel, and program implementers should have been asked to participate in the questionnaire design well before field implementation per AEA guidelines.</p>	<p>Our process for conducting this evaluation was very transparent and received approval from a wide range of stakeholders. The methodology used for this report, as well as the contractor interview guides, were posted on BaseCamp and reviewed by the IOU's and available to their implementers. The methodology documents were reviewed and approved by the CPUC and the CPUC consultants. Direct responses and adjustments were made at each stage of the process. Like all deliverables, the final report also includes public stakeholder review to which this appendix is a response. The feedback provided is used not only in the current report review, but is also considered by the CPUC and IOUs to inform long term research needs and future studies. These consultants and stakeholders included individuals with decades of HVAC and net savings assessment experience. In addition to this extensive review process, before we started our contractor interviews, our evaluation partners at the Western Cooling Efficiency Center at UC Davis also conducted a focus group-type meeting with some of the participating contractors to make sure we were covering the relevant issues.</p>
13	Robert Mowris	Recommendation (Survey Design)	pg. 4	<p>Multiple initial questions are required to determine if participants believe there is any difference between the "generic" services offered by contractors without the programs and tools, training, software and incentives provided by the programs for comprehensive services and/or 3-year maintenance agreements. The survey should include a series of yes/no, multiple choice, numerical scale, and text open end questions to understand the perspective of all participants in the programs including customers, technicians, and contractors.</p>	<p>As noted in a previous response, the interview guide had over a half dozen questions (including y/n questions and open-ended questions) that allowed HVAC contractors to distinguish the maintenance services and related sales practices that they routinely offer from those they offered through the HVAC3 programs. Therefore, if the participating HVAC contractor thought that the program had changed these maintenance services or sales practices, they had ample opportunity to state this.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
14	Robert Mowris	Influence of CQM Program	pg. 5	The statewide CQM programs are based on the ANSI/ASHRAE/ACCA Standard 180 which most contractors, technicians and customers would not be aware of without the programs. Third party programs are focused on delivering HVAC maintenance measures to improve energy efficiency, which would not be offered without the program since specific FDD tools are required to perform services in the programs and participating contractors would generally not purchase these tools without the program.	See previous responses to similar comments.
15	Robert Mowris	Recommendation (Survey Design)	pg. 5	Answers to multiple choice questions should be used to cross check responses to other questions with numerical answers. For example, contractors who indicate they offer the same "generic" services without the program, but indicated they do not use comparable FDD, tools, training, equipment, software, materials, procedures, or frequencies required by the programs, then another set of questions would need to be asked to probe deeper into the first response to clarify the response and obtain a more accurate assessment.	See previous responses to similar questions regarding the number and variety of interview questions. While recommending additional probing is reasonable, it must be remembered that with all interview guides there is a danger of respondent fatigue. Adding further probes may necessitate eliminating other interview questions to reasonably manage the duration.
16	Robert Mowris	Recommendation (Survey Design)	pg. 5	The HVAC3 NTG study should have included a collaborative process to solicit comments and suggestions from program implementers. The study should also have included a pretest of the survey questions to examine each question and the entire questionnaire to produce reliable and credible results. Colleagues, utility personnel, and program implementers should have been asked to participate in the questionnaire design per AEA guidelines. ... Findings presented in the study cannot be relied upon to evaluate NTG ratios for the programs. Based on biased questions and lack of rigor in the study, all CQM measures should be assigned a NTG ratio value of 1.0.	See previous responses to similar comments.

No.	Commenter	Subject	Sect./Pg.	Comment	Response
17	Robert Mowris	Confidence Intervals (Coil Cleaning)	pg. 5	<p>The HVAC3 NTG study provides wide variations with average values and even wider variations with confidence intervals for all NTG findings across different programs for the same measures (see table 16 on page 18 of the HVAC3 report). These findings indicate significant problems with questions asked of participating HVAC contractors, analysis of the responses, and the validity of asking questions of only participating HVAC contractors. The Coil Cleaning NTG varies from 0 to 48% with confidence intervals of 0 to +/-63% indicating a range of -15 to +108% for the NTG. By definition, without including spillover the NTG can only vary from 0 to 1 (i.e., 0 to 100%). The NTG cannot be negative and without including spillover, the NTG cannot exceed 1.0. The confidence intervals indicate problems with the sample size (i.e., large confidence intervals indicate sample size is too small).</p>	<p>We agree that difficulties completing in-depth interviews with many contractors led to reduced sample sizes and negatively impacted the confidence intervals. We attempted to determine a census of all contractors who participated in the program. As discussed above, lack of contractor contact info from the IOUs severely hampered our ability to complete interviews in the initial data collection efforts. We have added appropriate caveats to the report about the interpretation of these NTG results due to these high CIs. Additionally, while the relative precision is not great, it is relative to some pretty small point estimates. As a result, at a statistical level we can be fairly confident the true population values are not particularly high. For example, for kWh, with a point estimate of 0.36 and CI of ± 0.17, it's fairly unlikely that the true population value is higher than one standard deviation above the point estimate of 0.53. Despite the sampling uncertainty, these results provide a better estimate than was previously available.</p> <p>While the smaller sample sizes negatively impacted the precisions of our NTG estimates, it is important to note that other factors—such as inherent variability of pre-existing maintenance practices within the population of participating contractors—might also have contributed to the variability of the NTG ratios.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
18	Robert Mowris	Contradictory findings (Coil cleaning)	pg. 6	<p>The coil cleaning NTG of 0 to 48% indicates contractors believe their customers would have paid them to perform coil cleaning services at the same level of quality and frequency as required by the programs. Field inspections and interviews of participants at school sites found that school maintenance personnel generally do not perform coil cleaning and participating HVAC contractors would not be able to provide coil cleaning services to schools at the same level and quality required by the programs without incentives from the programs. Therefore, the coil cleaning NTG should be higher than what is reported in the HVAC3 study and possibly 1.0.</p>	<p>See previous responses to comments concerning:</p> <ul style="list-style-type: none"> • The fact that HVAC contractors were given ample opportunity to distinguish their normal maintenance services from those required by the program • The program impacts for economizer services <p>For the coil cleaning measure, the report indicates "the two maintenance services which the contractors were less likely to say they had offered before working with the programs were the coil cleaning of the evaporator and the thermostat measure."</p> <p>For the fan control measure, the commenter writes: "While HVAC contractors (or customers) can and should turn off fans or change thermostats to not operate during unoccupied periods, most would not perform the Fan Control measure without the program. Therefore, the Fan Control NTG should be 1.0." However, the commenter does not cite any evidence for this assertion. The report cites examples of maintenance measures—economizer repairs, RCA, coil cleaning, etc.—which contractors said they started doing only after joining the program. However, the fan control was not one of these measures.</p>
19	Robert Mowris	Confidence Intervals (Economizer)	pg. 6	<p>The Economizer measure NTG varies from 21 to 55% with confidence intervals of 30 to +/- 119% indicating a range of -64 to +174% for the NTG. The confidence intervals don't make sense. Findings from internet searches of "HVAC Economizer Repair Services", ContractingBusiness.com, and contractor/technician interviews in 2010-14 CQM programs are inconsistent with these results. Furthermore, most customers do not know what an economizer is or how economizers work so customer would not be able to seek economizer repair services when hiring an HVAC contractor to provide maintenance services. Therefore, the Economizer NTG should be 1.0.</p>	<p>See previous responses to similar comments.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
20	Robert Mowris	Fan Control NTG Findings	pg. 6	The Fan Control NTG varies from 0 to 5% with a range of -3 to 13%. The Fan Control measure involves contractors turning off fans accidentally left on and operating 24-hours per day or thermostats accidentally operating the HVAC system during unoccupied periods. While HVAC contractors (or customers) can and should turn off fans or change thermostats to not operate during unoccupied periods, most would not perform the Fan Control measure without the program. Therefore, the Fan Control NTG should be 1.0.	Please see previous responses to similar comments.
21	Robert Mowris	Maintenance NTG Findings	pg. 6	The Maintenance NTG is 48 +/- 28%. These findings indicate 48% of contractors would have performed HVAC maintenance services for free in the SDG&E QM program when, in fact, they would not have done so. They only offer Maintenance to specific customers (primarily schools) with incentives provided by programs. Field inspections and interviews of participants at schools have documented that school maintenance personnel generally do not perform maintenance services, and the participating HVAC contractors would not offer maintenance services for free to schools without incentives from the programs. Therefore, the Maintenance NTG should be 1.0.	Not sure what the commenter means by: "These findings indicate 48% of contractors would have performed HVAC maintenance services for free in the SDG&E QM program when, in fact, they would not have done so." Presumably, if the contractors were offering a similar QM service outside the program, they would not offer it for free but would need to cover their costs with an added profit.

No.	Commenter	Subject	Sect./Pg.	Comment	Response
22	Robert Mowris	RCA NTG Findings	pg. 6 - 7	<p>The RCA NTG varies from 37 to 100% NTG with a range of -69% to 143%. The lowest NTG is 37% for the PG&E AirCare Plus (ACP) program where technicians are required to use specific FDD RCA tools and software to perform services and the specific tools are only provided by the program. Without the PG&E ACP program the FDD RCA tools and software would not be available to participants without a significant monetary investment, which indicates a lack of awareness of the FDD RCA tools and software by the EM&V surveyors. If the majority of participants owned and used the FDD RCA tools before the program, then the response might be believable, but the participants do not own the tools and a follow-up question should have been asked to determine how participants would perform the same RCA services without the program. Clearly, if a follow-up question had been asked about the FDD RCA tools and software, then the impossible response of 37 +/- 106% would have been ignored. It should have been ignored anyway since a negative response is impossible.</p>	<p>Whenever a program requires participants to use a tool (e.g., FDD RCA tools) or buy a product (e.g., the CEC-spec LEDs required by the California Upstream Lighting program) that may be used much less frequently in the non-program world due to their very high cost (thus the need for program incentives), there is a temptation for program advocates to urge a high NTG for such measures or products because they are not being used much outside the program. However, it is possible that the contractors using non-FDD RCA tools, or the customers buying LEDs which are Energy Star®-rated but do not meet the CEC spec, do not view the program-prescribed tools or products as sufficiently superior to their cheaper EE tool/product alternatives. Otherwise they would be using/buying them outside the program without seeking program incentives. Even if the program advocates could make the cases that the program-incentivized tools/products were superior to those being used outside the program—e.g., that they produced a 10% savings compared to the cheaper non-program EE tools/products—it could be argued that any NTG ratio should be applied not to the whole gross savings (e.g., EE product compared to baseline) but only to that incremental portion of gross savings (e.g., the 10%) that one gained from going from the cheaper EE tool/product alternative to the more expensive, program-required EE tool/product.</p>

No.	Commenter	Subject	Sect./Pg.	Comment	Response
23	Robert Mowris	Thermostat NTG Findings	pg. 7	The Thermostat NTG varies from 21% to 31% with a range of -69 to +131%. As noted above, the confidence interval variation indicates problems with the sample size (i.e., large confidence intervals indicate sample size is too small). The HVAC maintenance programs provide incentives to install new smart thermostats with the following features: 1) 3 to 5F dead band between occupied cooling and heating set point, 2) unoccupied 60F heating and unoccupied 85F cooling, 3) unoccupied fan auto, 4) schedule start no earlier than 1 hour prior to and no later than ½ hour after occupancy schedule, 5) 2-stage cooling control to enable economizer operation, or 6) smart features for efficient mobile access. While HVAC contractors can and should install smart 2-stage thermostats, they would not install smart thermostats compatible with economizers (which they are unable to repair) without incentives from the program. Therefore, the thermostat NTG should be 1.0.	See previous responses to similar comments.
24	Robert Mowris	Recommendation (NTG values)	pg. 7	Based on historical field research studies and currently available evidence, all commercial quality maintenance program measure NTG values should be set to 1.0.	See previous responses to similar comments.
25	Robert Mowris	AEA Guidelines	pg. 7	The HVAC3 NTG study did not meet the requirements of the American Evaluation Association (AEA) guidelines regarding data-based systematic inquiry, competence, integrity, respect, and responsibility for general and public welfare.	See previous responses to similar comments.

No.	Commenter	Subject	Sect./Pg.	Comment	Response
26	Robert Mowris	Recommendation (AEA guidelines)	pg. 8	CPUC-funded NTG studies should adhere to the AEA guidelines for data-based systematic inquiry, competence, integrity, respect, and responsibility for all stakeholders. Based on the lack of adherence to the CPUC California Evaluation Framework and AEA guidelines, lack of vetting of survey questions and respect for stakeholders, use of biased questions, and not including customers and technicians in the sample, the HVAC3 NTG study should be rejected and the NTG ratios for all CQM measures should be established at 1.0.	See previous responses to similar comments.
27	Kristin Heinemeier, Ph.D., P.E., ASHRAE Fellow	Recommendation (Survey Design)	Overall	Pg 25 says "... survey responses indicated that the vast majority of the respondents had offered coil cleaning of the condensers, RCA services, economizer repair, or fan control adjustment before working with the program. The two maintenance services which the contractors were less likely to say they had offered before working with the programs were the coil cleaning of the evaporator and the thermostat measure." I believe that when asked, many contractors might respond "Yes, I provided economizer repair before the program", when in actuality they do not consistent offer to all customers, do more than "offer" services, provide QM that reduce energy use for customer, or provide service with consistent definition. I recommend, in preparation for the next cycle, conducting a survey or other study to identify the distinction between the different levels of service. We must define two "pots" of services: quality and not quality, and have some criteria for deciding, for each measure, which pot a particular maintenance practice belongs into.	The commenter's suggestion for future interview guides to do more probing to better understand why so many contractors consider the program QM services not much different than their typical maintenance is reasonable. However, to accommodate it, it may be necessary to sacrifice other interview questions to keep the interview at a reasonable length. In addition, as addressed in previous responses to similar comments, even if it can be proven that the program maintenance services are superior to the typical maintenance services that a contractor would otherwise offer, in such cases the program should only get credit for the incremental savings between the program maintenance service and typical maintenance service rather than between the program maintenance service and no maintenance, at all.

No.	Commenter	Subject	Sect./Pg.	Comment	Response
28	Joe Schmutzler	Maintenance Assumptions	Overall	<p>As the previous manager of the Premium HVAC Service program for PSE (similar to AirCare Plus), I have serious reservations about the contractor claims as presented in the report. The fact that DNV-GL relied solely on contractor self-assessments for this part of the review leads me to believe these numbers are highly suspect. We had new contractors claim they offered the services, but after going through the training none made the same claim. At Transformative Wave, we also found most sites need extra work to be done to complete QM programs (revealing the service protocols were not being done). It would be best to review actual documented service invoices to see what was billed to the customer for equipment and time to validate this data.</p>	<p>We appreciate the comment and understand that the findings are contrary to the commenter's previous experience with the Premium Cooling Program. As discussed above, we believe the methodology used in this study was sound and that the results are valid for the contractors surveyed. As previously mentioned, the contractors were given ample opportunity to discuss differences between their previous practices and practices through the program. The large error ratio indicates high uncertainty partly due to small sample, but also due to the large response variation within the sample. In other words, some contractors are changing their practices due to the program and others are not. All the contractors we interviewed had participated in the program and been through the training.</p>