HVAC Impact Evaluation FINAL Report

WO32 HVAC – Volume 1: Report

California Public Utilities Commission, Energy Division

Prepared by DNV GL
January 28, 2014

Table of Contents

Table of Contents

Table of Contents

List of Exhibits

Table of Contents

Table of Contents

Executive Summary

The 2010-12 investor-owned utility (IOU) energy efficiency program portfolios included a statewide program and various third-party programs targeting commercial and residential unitary HVAC systems. The impact measures for HVAC systems are categorized into three measure groups: Quality Maintenance (QM), Quality Installation (QI), and Upstream Equipment Incentives. Combined, the ex-ante saving estimates for these programs comprise approximately 177.2 gigawatt-hours (GWh) across investor-owned utility portfolios adopted for the 2010-12 cycle; and more importantly, they comprise 60.2 megawatts (MW) of electric demand savings claims.

Commercial Quality Maintenance

The evaluation team evaluated the installation rate and gross savings associated with the IOU Commercial Quality Maintenance (CQM) programs.[1] There were two program types for commercial quality maintenance (CQM): the statewide program, which promoted "standards-based maintenance", based on the ASHRAE/ANSI/ACCA[2] 180 standard and addressed a package of measures, and the third-party/local programs, which continued to implement individual maintenance measures similar to the 2006-08 programs. Within the statewide programs there were also claims in 2010 that were based on the previous program models with individual measures. While it was anticipated that a significant amount of savings would come from the statewide CQM packaged measures, they ultimately only accounted for 5% of total CQM ex ante energy savings and 3% of total CQM ex ante demand savings. The effort to assess installation rates for third-party/local programs was added to the scope of the evaluation after realizing the lower relative amount of savings being claimed from the statewide CQM packaged measures.

Evaluation of Statewide CQM Measures

The evaluation initially focused on the statewide programs. The statewide program workpaper assumed all units treated by the program would be brought to a performance baseline after maintenance. The conditions before maintenance vary according to the unique deficiencies present in each unit. The statewide CQM workpaper took an "expected value approach" based on assumptions about the frequency and combinations of measures that would be implemented, on average, across all projects.

[1] The evaluation team developed net to gross surveys and completed a small sample for contractors and customers, but did not calculate a net-to-gross ratio due to methodological issues.

[2] American Society of Heating Refrigeration and Air-Conditioning Engineers (ASHRAE), American National Standards Institute (ANSI), Air Conditioning Contractors of America (ACCA)

Evaluators coordinated with the statewide programs to sample and monitor units pre- and post-maintenance. Evaluation pre-maintenance site visits focused on inspecting the units and installing data loggers to collect pre-maintenance performance data. The evaluation requested schedules and updates to allow for direct observations of program technicians addressing the measures claimed in the workpaper. Sites were then monitored in the post-maintenance period. Some units may have multiple post periods due to program technicians returning as needed for multiple visits. The evaluation analyzed logger and measurement data to estimate the initial and final performance after all reported service visits.

The evaluation completed a much smaller and more clustered sample of units than originally planned. The program design and evaluation timing led to issues with gaining access for pre-measurement. The evaluation timing required pre-measurement with expedited measure installation and program inspections to confirm measures claimed by the program. Better coordination and cooperation between implementers and evaluators would have mitigated many issues associated with site access and maintenance inspection, but absent this coordination we have reported the savings based on the sample available sample. Data collection included multiple service visits by the program, but the actual tasks accomplished by the implementer at each site were not always reported or ultimately claimed in final tracking data. In the course of conducting the pre/post monitoring activities, observations of technicians performing maintenance indicated issues with the maintenance procedures and protocols, prompting an enhanced focus on observation activities and post-service inspections. Additional direct observations and inspections of units where savings claims were conducted at sample of sites that included contractors responsible for the majority of the Statewide CQM program savings claims. The results of these inspections are documented in Appendix B. Since these sites lacked pre-measurements, they were not used to estimate savings.

The evaluation team found that across all sampled units with pre/post monitoring, the gross savings per ton were negative for the statewide package of measures (Table 1). The results are from a limited sample and are considered indicative, but not definitive based on the calculated uncertainty in the measured savings. The variability in the per-unit savings was much larger than assumed by workpapers or in evaluation planning. The average savings was positive for SCE (but error bounds were greater than the estimate) and negative for SDG&E with significant uncertainty. SDG&E did not claim savings in 2010-12[3]. The overall pooled results across all sites with pre/post monitoring indicate impacts were not statistically different than zero savings with individual results at positive and negative extremes. These results indicate there is not a reliable average savings that can be deemed for the 2010-12 statewide package of measures.

[3] Savings claims were withdrawn from the final tracking data and in some cases units were further serviced and claimed in 2013-14 programs, but some units were dropped from the program and never claimed.

The evaluation team performed a detailed engineering assessment of all of the pre-post performance data[4]. Units with insufficient pre-maintenance data were eliminated from the analysis. The assessment found specific reasons leading to the outcome for each unit. For example, the assessment confirmed that maintenance has the possibility of increasing unit energy consumption. Program workpapers assume that bringing units up to a performance baseline can increase the efficiency and capacity, but do not account for the increase in energy consumption of the treated unit measured in this evaluation.[5] This evaluation also confirmed new findings from laboratory testing on significantly higher than expected measurement uncertainty for rooftop package unit performance.

Table 1: Final Estimated Gross Savings for Statewide Commercial Quality Maintenance (2010-12)

IOU Statewide Program	Ex Ante Average Savings Claimed	Ex-Post Pre-Post Average Savings Estimate	Sample Size	Standard Error	90% Confidence Interval Error Bound	Relative Precision	Lower Bound Savings	Upper Bound Savings
	Energy Savings (kWh/ton)							
SCE	307	705	17	605	994	141%	-290	1699
SDG&E	0	-1486	12	551	907	61%	-2393	-580
Total	N/A	-202	29	462	760	376%	-962	558
	Demand Savings (kW/ton)							
SCE	0.06	0.06	17	0.15	0.25	409%	-0.19	0.31
SDG&E	0	-0.43	12	0.13	0.21	50%	-0.64	-0.21
Total	N/A	-0.14	29	0.11	0.18	132%	-0.32	0.04

Note: PG&E had limited participation in the statewide program that addressed packages of measures.

The resulting energy savings estimates were driven by marginal unit efficiency changes from most maintenance activities coupled with significant increases in unit energy consumption caused by increasing the opening of the minimum outside air damper. The increased loads from opening economizer dampers diminished any efficiency improvement benefits obtained from implementation of other maintenance measures. In these cases, the economizers did not function effectively before or after maintenance, and opening of dampers resulted in additional ventilation air loads.

[4] A summary of the site by site assessments is shown in Appendix B.

[5] Increasing the capacity of the treated unit may reduce the load on an adjacent unit serving the same space. This interaction was not considered in the evaluation.

The analysis and assessments also confirmed laboratory findings on significantly higher than expected measurement uncertainty for measuring rooftop package unit performance. Changes in outside air damper position (a common maintenance treatment) changed air flow patterns in the mixing box, affecting measurement uncertainty and introducing bias in mixed air temperature and humidity measurements. While changes in measured space loads were unaffected, changes in the unit's overall performance (space load plus ventilation load) upon which unit savings are calculated were found to be unreliable. Importantly for this effort, unit power was available to assess overall program impacts, allowing a pre/post time series analysis of interval kW data at the individual unit level.

Evaluation of Third Party and Local Program CQM Measures

A verification analysis was conducted for key individual maintenance measures promoted by third party and local programs. Resulting installation rates were applied to ex ante savings estimates to provide installation rate adjusted ex post evaluation results. Economizer repair and refrigerant charge installation rates relative to program criteria across all of the IOU local programs are shown in the following tables. The installation rate for refrigerant charge measures represents the fraction of refrigerant circuits that achieved claimed benefits based on post-maintenance observations. The original planned verification analysis refrigerant charge diagnostics method which relied on refrigerant temperature and pressure measurements proved inaccurate based on field and laboratory findings, thus a refrigerant evacuation and weigh-out procedure was used to assess state of charge. The installation rate for economizer repairs represents the fraction of economizers that demonstrated basic functionality in post-maintenance observations. The evaluation developed installation rates for refrigerant charge adjustment and economizer repair measures only. Observations on other measures and additional faults were reported in Appendix B, but these observations were not used to adjust program savings claims.

Refrigerant Charge Measures

Refrigerant charge was evaluated in the field by measuring ex-post diagnostic performance on 66 units where technicians made adjustments. The results of these tests were ultimately not used to assess installation rate as new information was produced in the laboratory on the uncertainty in the accuracy of charge diagnostics. Towards the end of the evaluation the team revised the charge assessment procedure and conducted the procedure on a random sample of five single-compressor and two dual-compressor units for a total of nine refrigerant circuits. For each circuit, the EM&V master technicians recovered and weighed out charge, evacuated to 500 microns, held at or below 500 microns for 20 minutes, weighed in factory charge, and measured performance with factory charge. The average refrigerant charge adjustment reported by the local program[6] for the units evaluated was 19.4 ± 5.7%. This indicates that, on average, units had a significant amount of charge added. Recovery and weigh-out of these units indicated both over and undercharge conditions after service. Units that were still undercharged after maintenance realized some benefits, but additional potential remained. Units that had charge added and ended up overcharged had multiple outcomes including positive, zero, or slightly negative benefits.

Table 2 presents installation rate adjustments for 2010-12 refrigerant charge measure savings. A 79% install rate was developed by calculating the average achieved benefits across units based on the method used in the 2009 PG&E refrigerant charge workpaper, which was recommended by the ex ante review team as the best available data at the time. Individual units exceeded, met, or partially met the workpaper charge adjustment assumptions, with one unit showing negative benefits. Due to the small samples size, precisions are not reported by IOU. A detailed description of sample and analysis results is provided in the main report.

Table 2: Final Installation Rate and Adjusted Gross Savings for Refrigerant Charge Measures (2010-12)

IOU	Program Name	Ex Ante Energy Savings (kWh)	Installation Rate	Installation Rate Adjusted Energy Savings (kWh)
PG&E	Air Care Plus	2,339,072	81%	1,894,648
SDG&E	Non-Res HVAC Tune-up/Quality Installation	2,390,079	77%	1,840,361
TOTAL		4,729,151	79%	3,735,009

Note: Ex ante demand savings were negligible for this measure.

Recent laboratory test results from CPUC-sponsored testing and results reported by Purdue laboratories (Braun[7,8]) informed the evaluation of refrigerant charge measures. After evaluation

---

[6] Note that the amount of charge added or removed and unit factory charge was not always collected systematically in databases for the statewide program. This data was recorded systematically for the third party and local programs.

[7] Braun, Jim, David Yuill, and Howard Cheung. *A Method for Evaluating Diagnostic Protocols for Packaged Air Conditioning Equipment*. Purdue University and the California Energy Commission, 2012.

of the refrigerant charge measure claimed by the local and third-party programs the team determined that:

- Based on a limited sample, program refrigerant charge service achieved a combination of full workpaper benefits, partial benefits, and negative benefits, but the overall result was positive.
- Linking fault detection diagnostics (FDD) values to benefits is difficult. This means program FDD methods are not effective in assessing savings or bringing the unit to correct factory charge
- Manufacturer protocols perform better than generic third party protocols in determining whether a unit has achieved factory charge
- It is possible to assess savings benefits when using the weigh-out, weigh-in protocol, as long as the service adjustment is known.
- Title-24 protocol only looks at measuring subcooling for thermostatic expansion valve (TXV) units and superheat for non-TXV units and cannot diagnose faults in units with dirty coils, non-condensables, or refrigerant line restrictions. Observations of program trainings and review of training materials for 2010-12 programs indicate no specific FDD to address non-condensables and restrictions.

None of the FDD methods based on refrigerant temperatures and pressures can provide information on program benefits, since efficiency improvement calculations are not possible from the available data. However, the recovery/re-charge method does provide valuable insight on potential program benefits. This method should be considered in evaluating future iterations of the program. Once we weighed in the correct charge, the Title 24 and generic program protocols indicated 4 of 9 were incorrectly charged while manufacturer's data said 2 of 9 were incorrectly charged.

Based on the weighed charge method, 7 of 9 units were not within 5% of factory charge after program service. Therefore, the Title 24 or generic protocols were not an effective method to bring the unit to correct charge or assess program benefits. This field finding confirmed initial laboratory test results that FDD can lead to incorrect diagnosis. Additional common faults, which also affect the pressures and temperatures measured (e.g., contaminants, especially non-condensables and line restrictions), are not addressed by the Title 24 and generic methodology and lead to incorrect charge adjustment when not identified. Direct observations indicated that technicians in the field did not test for these additional faults and they were not covered in the observed 2010-12 technician trainings.

---

[8]Braun, Jim, and David Yuill. Evaluation of the Effectiveness of Currently Utilized Diagnostic Protocols. Purdue University. 2014.

This field result combined with laboratory test results supports a recommendation for the recovery/re-charge method as the only reliable method to assess program benefits. Manufacturer protocols are better than the other methods at assessing factory charge, but they are not perfect and therefore do not always result in the program-expected benefits of the refrigerant charge adjustment. Programs should use manufacturer methods for FDD and/or confirm the preferred method is at least as accurate as the manufacturer method. Recovery/re-charge should be used by program administrators for verification on a sub-sample of circuits, and EM&V should use a recovery/re-charge method going forward to assess the benefits of program charge adjustments.

Economizer Repair Measures

For economizers, the majority of units the team inspected that reported repairs were not found to be performing their primary function of changing damper position in response to temperature. The overall verification rate was estimated at 23% which assumed that due to the claimed three year EUL and timing of inspections, 20% of inspected units should have failed functional tests. In the sample, Table 3 shows adjusted gross savings for economizers resulting from the application of this verification rate. In some cases, the technicians opened the position of the dampers but the EM&V team did not observe other changes to sensors, linkages, or control settings. In one program, technicians from multiple contractors installed new controllers and sensors, but the controls were not set up properly. Additionally, although most systems with economizers required economizer repairs, none of the technicians were observed to perform them. .Non-operable economizers were the most common fault found in post-maintenance visits. In addition to the limitations of the refrigerant charge fault detection diagnostics applied in the maintenance programs, the evaluation team believes that un-diagnosed economizer faults and uncontrolled economizer leakage presents an additional challenge and future opportunity. The evaluation team made observations of economizers for the statewide programs as well, but those results were part of the CQM package and were not disaggregated. The findings of observed economizer functionality in the statewide program are included in the main report.

Table 3: Final Installation Rate and Adjusted Gross Savings for Economizer Measures (2010-12)

IOU	Program Name	Ex Ante Energy Savings (kWh)	Ex Ante Demand Savings (kW)	Sample Size	Relative Precision	Install Rate	Installed Energy Savings (kWh)	Installed Demand Savings (kW)
PG&E	Air Care Plus	6,938,405	3,488	16	20%	25%	1,734,601	872
SDG&E	Non-Res HVAC Tune-up	3,884,115	2,571	10	26%	20%	776,823	514
TOTAL		10,822,520	6,059	26	16%	23%	2,511,424	1,386

Other Program Observations

The following additional observations were not used in determining energy savings or installation rates, but these findings are presented to suggest possible causes for the high uncertainty in savings and relatively low installation rate for economizer repairs. Master technicians performed direct observations of program service. They also observed training sessions, reviewed training materials, and reviewed QC procedures. Note the review and observations were not conducted of all trainings and QC procedures within all programs.

Although both the statewide and third-party/local programs included common measures, such as refrigerant charge adjustment and economizer repair incentives, the trainings and guidance provided by programs varied. This likely led to direct observations of technicians not repairing non-functional economizers and claimed economizer repairs not working in post-maintenance inspections. For refrigerant charge the various FDD procedures were shown to not perform as well as manufacturer protocols and ultimately only direct weight out and recharge provided enough information to estimate the benefits of program charge adjustments.

Conclusions and Recommendations

The current program and commercial unitary HVAC maintenance measures in general face various inherent challenges, many of which were not addressed by 2010-12 program workpapers. For example, the programs provide incentives for adding or removing refrigerant charge without diagnosing other faults that affect refrigerant temperatures and pressures, such as restrictions or non-condensables, that can lead to incorrect refrigerant charge diagnostics.

For CQM, the evaluation team learned it is very difficult to perform refrigerant charge correction through indirect diagnostics. A comprehensive approach is needed to optimize the efficiency of an existing system and bring the unit to a performance baseline. A comprehensive approach requires that repairs be completed that may be outside of measure definitions, such as restrictions and non-condensables noted earlier. See the laboratory testing section for additional issues that may also need to be addressed. The less-than-expected savings and installation rates from both the 2006-08 evaluation and this evaluation stem from trying to address individual faults or subsets of all faults without accounting for underlying system issues. Future field and laboratory research hold the best promise to determine the impacts of addressing faults and methods to diagnose and repair them.

The recommendations provided in this report do not cover all of the changes between the 2010-12 and 2013-14 programs. These recommendations have not been tested for cost effectiveness and they are intended as areas that should be considered and, if feasible, developed further, piloted, and tested for cost effectiveness.

To address these findings, we recommend that IOU programs coordinate to clearly define measures and ensure program workpaper assumptions match the implementation process of the measures. The terminology (maintenance, repair, retro-commissioning) is less important than defining the specific actions that take place and how those actions lead to energy savings. The IOU programs should focus on activities that will lead to energy and peak demand savings, rather than trying to address all repairs and actions. A focus should also be placed on the one-time and ongoing maintenance activities that lead to more persistent savings. Other repairs and maintenance activities that improve comfort, indoor air quality, or unit reliability may increase energy consumption or peak demand, and should be accounted for if performed.

We recommend piloting approaches that have demonstrated measurable savings. For each approach with measureable savings, the program should provide a definition and examples of the effective energy efficiency measures. In 2010-12 some programs provided lists of tools that were acceptable while others did not, but these reference documents did not list the acceptable tool for each procedure. The measure definitions should include specific training requirements, approved tools based on accuracy requirements, protocols to perform and inspect the measure, and data collection requirements. We also recommend that incentives be provided for measures with the highest probable applicability, FDD repair rate, and savings. Measure recommendations include:

- **Seal Unit Leaks.** Sealing the junctions between economizers and unit cabinets may have large opportunities as they are not part of routine maintenance. Pilot efforts should measure the change in performance in the field and use reliable sealing methods such as properly rated metal taping and mastic. This recommendation requires coordination with certified test and balance (TAB) technicians to ensure minimum requirements for ventilation are met. It is not likely that economizer perimeter leakage is intended to help meet ventilation requirements. Economizer perimeter leakage will be further studied in the 2013-14 laboratory tests. The evaluation recognizes that any recommendation regarding reductions in outdoor air quantity raises liability concerns, but over-ventilation does waste energy. Some units are operated with fans that cycle on with compressors and may not meet ventilation requirements even when accounting for leakage. Outside air rates need to be evaluated and the leakage must be taken into account such that it is known and part of calculations.

- **Reconsider Generic Diagnostic-Based Refrigerant Charge Measures.** Identifying and adjusting units with incorrect charge proves to be extremely difficult on commercial units with outdoor air intakes and more difficult if multiple faults are present. A key issue becomes determining appropriate coil-entering sensor placement due to unmixed and temperature-stratified air on the inlet side of the coil, since the ideal placement changes with damper position. Evacuation and recharge is recommended as the best method to assess exact charge level. The evaluation team understands that this method may not be cost-feasible for program implementation, and thus the program recommendation is to follow unit specific manufacturer diagnostic procedures (see next bullet). We recommend that programs implementing this measure perform quality control verifications on a sample of completed jobs using recovery, weigh-out and re-charge.
- **Use Manufacturer Maintenance and Diagnostic Protocols or Compare Program Methods to them for Reference.** Manufacturer protocols provide important instructions on troubleshooting multiple faults for units including refrigerant charge, restrictions and non-condensables. These methods were found in field and laboratory testing to be more reliable than generic FDD protocols. Manufacturer protocols also may provide new methods for adjusting airflow based on static pressure measurements. Manufacturer protocols provide instructions on proper tensioning and alignment of fan belts. This recommendation does not preclude future FDD protocol development that provides better energy efficiency optimization. Any new methods should be compared to manufacturer methods to demonstrate they are better at achieving optimal unit performance. The accuracy of manufacturer and other FDD methods across ranges of airflow conditions and other faults continues to be a subject of laboratory testing.
- **Develop Criteria to Replace Rather than Repair Economizer**s. Repairing old economizers can be expensive and difficult due to limited availability of parts and the need to address old design issues. Economizer replacement offers an option that may overcome some of the fundamental economizer repair issues and should be explored. De-commissioning is not an energy efficiency measure and this recommendation calls for a replacement measure which would also entail installing Title 24 compliant controls with FDD. Unit age should also be considered in determining whether to repair or replace an economizer. Also note that integrating new economizer controls with older unit controls may present an issue. It should be considered whether this measure also

requires changes to settings or upgrades to unit controls (thermostat or energy management systems).

In the long term, HVAC manufacturers, industry associations, state licensing boards, and the U.S. Department of Energy (DOE) can help improve quality maintenance programs by supporting improved technician competency standards, FDD protocols, and service instrument standards. For utility programs in the short term, we recommend that additional field and laboratory tests and additional analyses be conducted to further research the energy saving potential of HVAC maintenance and the issues identified in this summary. The focus should be on efficiency and account for activities that may improve comfort and/or indoor air quality, but result in increased energy consumption. The program can adopt standards-based maintenance but workpapers should account for occupant comfort improvements at the expense of energy savings. Programs should focus on simple measures that can be reliably implemented and have demonstrated savings.

The evaluation also acknowledges that maintenance is a process that is unlike almost all other energy efficiency measures. This difference presents major measurement and evaluation challenges. Program workpapers assume changes lead to reduced HVAC demand and energy use. The statewide programs make those changes over a relatively long time period, while local and third party programs perform the measures as an event. The differences in program designs present different evaluation challenges. Measuring changes in unit operation over long periods associated with the state-wide program design limits the ability to provide timely feedback and introduces potential uncertainties associated with non-program issues. Estimating the impact of multiple measures installed at the same time within local programs affects the ability of evaluation to provide measure specific savings estimates. The recommendation for programs to pilot strategies also provides a framework for evaluation of specific measures or service actions in detail that are not easily accomplished in the evaluation of full scale programs.

Residential Quality Installation

The quality installation field assessments focused on residential systems because non-residential programs were not in full operation in the 2010-12 cycle. The residential quality installation workpaper assumed common practice baseline installation conditions as opposed to code minimums. The team conducted site visits at 50 program participant and 50 non-participant sites that recently installed HVAC units. Site evaluation staff followed identical data-collection protocols at participant and non-participant sites. Non-participating households were included to provide a baseline for the program evaluation. The non-participant homes spanned Pacific Gas and Electric (PG&E), Southern California Edison (SCE) and San Diego Gas and Electric (SDG&E) service areas. Because SCE was the only IOU with a Residential Quality Installation program in 2010-12, program participants could only be recruited from within the

SCE service area. Differences between participant and non-participant sites were used to establish program savings.

Participant contractors with more units installed under the program had more units in the test sample. Participant units spanned single-family homes, duplexes, and apartment and condominium complexes throughout the SCE service area. Because the non-participant sample pool was more limited (they were identified from other recent studies as being homes that had recently installed HVAC units), the team recruited non-participants first and then recruited participants from similar neighborhoods where possible. The non-participant data sources included one study that collected system age with no questions about change outs which is considered unbiased. The other source asked customers about recent change outs, which may have introduced some bias away from non-permitted installations. Potential bias was mitigated to the maximum extent possible; onsite recruiting emphasized the anonymity of test results and never mentioned compliance or permitting. This study did not investigate the issue of permitting for non-participants, but a follow-on project was conducted to determine the permit frequency and differences between permitted and non-permitted installations.[9]

Data collection for all tested units included spot measurement of unit airflow and duct leakage. Building envelope characteristics informed load-calculation models for each site for non-participants. Existing load-calculation models (a program requirement) for participants were reviewed for consistency. The evaluation team calculated energy savings for participants relative to non-participants based on system sizing, correct airflow, and reduced duct leakage. The evaluation team took the data collected for participants and non-participants and re-ran the DEER prototype DOE-2 models consistent with the method reported in the residential quality installation workpaper. The evaluation notes that the workpaper was not reviewed by the ex ante review process. If the ex ante team had reviewed the workpaper then the current DEER prototypes and calculation procedures would have been required of the workpaper.

Table 4 summarizes the field findings from the evaluation compared to the workpaper assumptions. In particular, ex ante workpapers claimed higher non-participant duct leakage than found in the baseline group. Participant duct leakage results met the QI program goal of 12% total leakage, but the baseline leakage was lower (17%) than the 24% leakage value assumed in the workpaper. This more efficient baseline resulted in lower actual savings than expected. There was not a statistically significant difference in per-ton airflow between participants and non-participants. The evaluation team modeled the mean differences to develop savings estimates as the collected data represented the best information available, since the workpaper relied on assumed values. The evaluation found that the calculated average sizing ratios to be similar between participants and non-participants. There were both under and oversized non-

[9] PG&E. *HVAC Permitting: A Study to Inform IOU HVAC Programs* . August 2014(CALMAC ID: PGE0349.01),

participant units, but there were primarily right sized and a few oversized participant units. The workpaper assumed an average amount of oversizing and used the sizing ratio to represent this in simulations.

Table 4: Comparison of Field Findings and Baseline Assumptions

Measure Parameter	Non-Participant Field Observations (mean)	Workpaper Baseline Assumptions	Participant Field Observations (mean)	QI Efficient Case Workpaper Assumptions	Non-Participant Participant Significant Difference
Total System Duct Leakage	17%	24%	12%	12%	Yes
System Airflow (CFM per ton)	300	350	338	400	No
System Oversizing	13%	20%	10%	0%	No

Table 5 provides the ex ante and ex post savings for each quality installation measure category. The evaluation only modified the sizing, duct leakage, and airflow simulation inputs between ex ante and ex post results. The overall gross realization rates were 35% for electric energy savings and 38% for demand savings. These realization rates were driven by the better-than-assumed sizing and duct leakage for non-participant units, which set the baseline for savings.

Table 5: Final Gross Savings for Residential Quality Installation Measures Installed (2010-12)

Quality Installation Measure Category	Ex Ante Energy Savings (kWh)	Ex Post Energy Savings (kWh)	Energy Real. Rate (%)	Ex Ante Demand Savings (kW)	Ex Post Demand Savings (kW)	Demand Real. Rate (%)
13 SEER Air Conditioner Quality Installation	18,417	7,124	39%	6	3	53%
14 SEER Air Conditioner Quality Installation	169,180	59,898	35%	80	37	46%
14 SEER Heat Pump Quality Installation	10,081	3,229	32%	4	1	33%
15 SEER Air Conditioner Quality Installation	273,568	94,314	34%	149	55	37%
15 SEER Heat Pump Quality Installation	28,228	9,322	33%	9	3	33%
16 SEER Air Conditioner Quality Installation	411,269	144,617	35%	227	82	36%
Total	**910,743**	**318,505**	**35%**	**475**	**181**	**38%**

A note on our realization rate calculations: The team could not separate out free-ridership from presumed non-compliance in the gross (workpaper) baseline. Since the workpaper assumed non-compliance, it essentially presumes a standard practice baseline, not a code baseline. If one assumes that participants would have met code requirements in the absence of the program, we could have estimated freer-ridership, but we had no a-priori evidence of participant code compliance nor did we think we could get an unbiased estimate of compliance by asking that question directly. The evaluation team believes that net-to-gross issues may be influencing our realization rate, but we cannot determine the extent to which this is true and cannot develop a quantitative estimate that breaks out free-ridership. Hence, there is no free-ridership estimate factored into these realization rates. The team has no way of knowing or estimating the extent to which the gross realization rates include net effects that lower the reported realization rate. Permitting of non-participants is part of a PG&E follow-up study (CALMAC ID: PGE0349.01) and will be further studied in an upcoming project.

For residential quality installation, the evaluation team learned that installations in the programs exceeded Title 24 code, on average for system design attributes covered by the program, and those installations outside the program do not perform as poorly as assumed in program workpapers. The evaluation findings showed that there was a continuous spectrum of installation efficiencies, from good to bad, outside the QI program. The team also discovered that the program workpaper does not fully capture the improved fan efficiencies present in the sampled participants, and does not model the impact on the program of applying sensible capacity standards via ACCA Manual S to system sizing.

The evaluation team has the following recommendations:

- The 2013 Title 24 code should be reviewed to determine opportunities to exceed code. The net to gross needs to be clearly defined since the gross baseline assumed is not code minimum, it is a common practice assumption below code. In the absence of the program a unit could be minimally compliant or could be at the common practice baseline. Using a common practice baseline requires specific evidence that code minimum is not appropriate and requires rethinking the use of the default net to gross value which usually applies to measures with a code baseline.
- Consider a revision to workpapers to account for the fact that baseline is comprised of a range as opposed to a point estimate. Gross savings vary by CZ and measure SEER level. In this case the net savings and baseline may also vary by code jurisdiction. Until larger studies are available the source of assumptions in workpapers should be clearly stated as well as whether baseline studies are needed to develop data to inform the inputs.
- Expand the non-participant sample to support evaluation of QI programs and 2014 "to code" pilots. Using a consistent data collection approach would allow expansion of the sample in any given climate zone, given the relatively small size of the sample in this study. The 2014 "to code" pilots will collect detailed compliance verification data. Comparing non-pilot and pilot jurisdictions may require larger samples of non-participants.
- Ultimately the programs can influence savings for actions that exceed Title 24 requirements. Participant systems did not all meet assumed performance values that were included in workpapers. Exceeding code will improve realization rates, but it is unknown if cost effective savings exist, especially relative to the new code. Common practice was worse than code in this study, but not as much as assumed in workpapers. The evaluation team recommends the IOUs explore a few areas where Title 24 can be exceeded or does not have requirements.
  - Explore more downsizing to reduce peak demand. The program sizing aligned much better with Manual J load calculations than non-participants, but did not eliminate all cases of oversizing.
  - Explore duct sealing to reach a lower target leakage, such as the 6% threshold that is set for new ducts systems.
  - Explore air handlers/furnaces, filters, and duct modifications that reduce pressure drop and improve fan system efficiency.
  - Explore whether all ANSI/ACCA 5 Quality Installation Standard elements in programs impact energy use and align with workpapers.
- Determine if there are specific locations where common practice baseline is worse than estimated in this evaluation. The non-participant sample covered large areas and many code jurisdictions. Further study could determine specific areas where common practice is worst and target program activities toward those locations. This may mean leveraging

local government partnership programs as opposed to statewide program models if it is determined that only specific areas have the greatest opportunity.
- Ensure that workpaper and evaluation modeling inputs produce HVAC energy estimates that are consistent with the Residential Appliance Saturation Study (RASS).

Commercial Upstream Program

The evaluation efforts for the upstream component of the study focused on the level of free-ridership in the Upstream HVAC program. Gross ex ante savings claims are based on workpapers that primarily use DEER estimates and were not evaluated. A participant self-report method was used for the free-ridership/net-to-gross (NTG) work, and the analysis was based on in-depth telephone interviews with 19 out of 22 participating HVAC distributors that were conducted by DNV GL in 2013. The approach took into account the program's effect on both the stocking practices and high efficiency sales of the distributors. The program offers rebates to decrease costs associated with high efficiency equipment and/or to make the price and availability equivalent to a less efficient alternative in order to steer market demand toward high efficiency equipment. One key limitation of the approach was that for distributors who did stock units, the survey did not determine if there were specific sizes or unit types that were always custom-ordered and therefore there could be little program influence on stocking.

Overall, the program is achieving a savings-weighted NTG score of 0.80, versus an unweighted NTG score of 0.58. The reason for this difference is that the distributors who account for the largest share of program savings are also the distributors who claim the highest levels of program influence. The final net to gross ratio was applied to gross energy and demand savings for 2010-12 programs. Table 6 shows the SCE program represented a majority of the statewide savings for the Upstream program.

Table 6: Upstream Net-to-Gross Ratio and Net Savings

IOU	Gross Energy Savings (kWh)	Gross Demand Savings (kW)	Net-To-Gross Ratio	Net Energy Savings (kWh)	Net Demand Savings (kW)
PG&E	17,705,130	9,100	0.80	14,164,104	7,280
SCE	64,368,795	21,952	0.80	51,495,036	17,562
SDG&E	560,854	203	0.80	448,683	162
TOTAL	82,634,780	31,255	0.80	66,107,824	25,004

Since many distributors served multiple service territories, the evaluation team developed a single net-to-gross score using the total savings from the two largest programs at SCE and

PG&E. The SDG&E program had relatively small levels of savings that were not assigned an Upstream flag in tracking data, but packaged and split system measures that mentioned Upstream in the name also received this net-to-gross adjustment.

For upstream distributor incentives, the evaluation team learned that incentives influence stocking and sales for the highest volume distributors, but have less influence on lower volume distributors. Some distributors supply units directly from manufacturers, and stocking incentives have no influence there. The team identified some opportunities beyond the efficiency incentives, given that the program does not focus on permitting and compliance, nor does it consider additional incentives for climate-appropriate systems.

The evaluation team makes the following recommendations, gleaned from the distributor interviews and analysis of program data:

- The program should consider increasing rebate levels for higher efficiency equipment to encourage more sales in the highest efficiency tiers. A 50% to 100% incentive increase could be explored for the top tier, however the evaluation team did not estimate what levels could be achieved while keeping the program cost effective. This recommendation is based on distributors indicating that the incentive is a greater percentage of the incremental cost for Tier 1 compared to the higher efficiency tiers.
- Program-provided support for how to market high efficiency equipment may be helpful to distributors, particularly those who are less successful at selling the concept of high efficiency.
- Providing a reservation system or rebate guarantee may encourage more participation, and increased high-efficiency sales, from distributors who have long sales cycles or custom build equipment. This type of rebate would only address incremental cost and would not have any influence on stocking.

Preliminary Laboratory Testing Results

The evaluation team conducted laboratory tests to evaluate economizer operation and performance and system faults for a dual compressor commercial roof top unit (RTU). Master technicians oversaw laboratory technicians who fully instrumented and tested packaged HVAC units in an Air-conditioning Heating and Refrigeration Institute (AHRI)-certified laboratory. The purpose of many tests was to evaluate systems under conditions more typical of field conditions–especially those conditions observed from pre/post evaluation sites. Findings suggest limitations with AHRI test procedures in addressing in situ efficiency; laboratory results reflect those that would be more typical of installed performance. The laboratory performed tests at a range of “outside” dry bulb and wet bulb temperature conditions to simulate temperature variations across the California coastal, inland, mountain, and desert climate zones. Additional laboratory tests continue currently and will be performed in the future on packaged units from additional manufacturers.

The laboratory tests with actual cabinet leakage and functioning economizers provided critical new findings on the diagnosis of system faults and the efficiency impacts of repairs. Based on discussions with the laboratory staff and the experience across the evaluation team, these tests appear to be the first of their kind. Critical findings from laboratory testing with an economizer include:

- The impact of economizers on system efficiency is significant and unexpected.
- The diagnosis and adjustment of charge is difficult, if not impossible, to achieve in the field since both unit total airflow and the presence of outdoor air intakes affect the reliability of fault detection diagnostics (FDD). This finding led to the field work described under Commercial Quality Maintenance that compared refrigerant charge diagnostics to a charge weigh out/weigh in procedure.
- The test results demonstrated that damper leakage in economizers likely results in over ventilation. Additionally, economizer maximum outside air rates when fully open were found to be less that is often assumed in repair measure evaluations. Leakage sources were identified and could be addressed via maintenance programs. Improving maximum outside air flow rates are not something that could be addressed in a maintenance program.

Understanding the efficiency impacts of faults presents several challenges. In reviewing the results, the team expected that the efficiency of all units would be relatively lower at higher outdoor temperature conditions and lower coil-entering wet-bulb conditions. Faults may have different impacts at different conditions, but generally should reduce efficiency relative to optimal at a given condition. Instead, we found that the effects of outdoor and indoor conditions and faults on efficiency exhibit non-linear trends and become more complex when combined and require empirical testing to derive efficiency impacts.

The following laboratory results are available based on tests of a 7.5-ton dual-compressor packaged RTU that uses a non-TXV expansion design:

- Tests with the economizer open from 10% to 30%, based on technician reported “rules of thumb,” indicated that loads and energy consumption increase 5% to 62% compared to closed dampers that deliver 15% outdoor air through damper and other sources of leakage. Assuming this leakage in unintended, then economizers impose a 5% to 62% energy penalty as a result of imperfect design.
- Outdoor airflow was 15% with closed dampers, 20% with 1-finger open, 23% with 2-fingers open, 30% with 3-fingers open, and 62% with dampers fully open. The “finger open” setting is a common installer minimum outside air setting where the width of a finger represents the width of the economizer damper opening for ventilation. Each finger width is assumed to be approximately equal to 10% of design air flow. It is

common to assume 2% outdoor airflow with closed dampers and 100% outdoor airflow with fully open dampers. The 2% closed damper setting is a required maximum damper leakage rating, while the fully open flow assumes all return air is blocked off by the return air dampers. Damper leakages (frame and actuator area leakages on the outside air damper and incomplete return duct coverage on the return air damper) make both assumptions incorrect.

1. Introduction

1.1 Program Evaluation and Overview

The 2010-12 IOU energy efficiency program portfolios include a statewide program and various third-party programs targeting commercial and residential unitary HVAC systems. The impact measures for HVAC systems are categorized into three measure groups: Upstream Equipment Incentives, Quality Maintenance, and Quality Installation. Combined, the ex-ante savings for these program claims comprise approximately 177.2 GWh across IOU portfolios adopted for the 2010-12 cycle; and more importantly, they comprise 60.2 MW of claimed electric demand savings[10].

These programs were developed for residential and small commercial facilities that use packaged and split-system air conditioning systems or heat pumps. The programs include a non-residential Upstream incentive program to HVAC distributors to encourage stocking high efficiency units, a continuation from 2004-05 and 2006-08 cycles. Quality Maintenance (QM) and Quality Installation programs are also included. Both of these promote packages of measures designed to improve the efficiency of HVAC systems by correcting faults resulting from poor installation, wear and tear, and malfunction or damage, and/or by improving system control. The commercial and residential refrigerant charge and airflow (RCA) diagnostics and residential duct sealing (DTS) measures in the QM and QI packages were included in 2006-08 programs. 2010-12 programs cover a broad range of new measures, including evaporator and condenser coil cleaning, off-hour controls, and economizer repair on commercial units. The residential QM programs claimed insignificant savings and were not evaluated. Commercial QI packages are still being developed by the IOUs and no claims were made in 2010-12 programs.

Table 7 summarizes IOU programs and ex-ante energy and demand savings claims by IOU. The Upstream programs compromise the greatest savings of the three primary program types. The Upstream program claimed the largest savings among the SCE programs, but the third party and local CQM programs claimed the greatest savings for PG&E and SDG&E programs.

[10] Based on 2010-12 final tracking data and measures evaluated

Table 7: Ex Ante Savings Claims by IOU HVAC Program

IOU	Program ID	Program Name	2010-12 Savings Claims	
			Energy Savings (kWh)	Demand Savings (kW)
PG&E	**PGE21061**	SW Upstream HVAC Equipment Incentive	17,705,130	9,100
	PGE21065	SW Commercial Quality Maintenance Development	5,504,808	5,589
	PGE2181	Air Care Plus (Commercial Maintenance)	60,556,253	11,031
	All Evaluated PG&E HVAC Programs		**83,766,192**	**25,719**
SCE	**SCE-SW-007A**	SW Upstream HVAC Equipment Incentive	64,368,795	21,952
	SCE-SW-007D	ENERGY STAR Residential Quality Installation Program	910,743	475
	SCE-SW-007E	SW Commercial Quality Maintenance Development	10,269,044	6,354
	All Evaluated SCE HVAC Programs		**75,548,583**	**28,781**
SDG&E	**SDGE3161***	SW Commercial Upstream Equipment	560,854	203
	SDGE3148	SW Quality Maintenance Program (Commercial Statewide)	0	0
	SDGE3161	Non-Res HVAC Tune-up/Quality Installation (Local)	17,947,787	5,579
	All Evaluated SDG&E HVAC Programs		**18,508,641**	**5,781**
	All Evaluated HVAC Programs		***177,823,416***	***60,282***

*Measures labeled upstream were assigned to the same program ID shown for the Non-Res HVAC Tune-up/Quality Installation Local Program

1.2 Organization of the Report

This report describes the impact evaluation results for Upstream Equipment Incentives, Quality Maintenance, and Quality Installation measures. It is organized as follows:

- **Section 2: Commercial Quality Maintenance.** Pre-post monitoring and site observations provided the basis for the impact analysis and installation rates.
- **Section 3: Residential Quality Installation.** Measurements of participant and non-participant air conditioner installations established performance relative to workpaper assumptions to inform the impact analysis.
- **Section 4: Upstream HVAC Program.** Interviews with equipment distributors established net to gross ratio.
- **Section 5: Package Unit Laboratory Testing.** Laboratory efficiency tests of dual-compressor unit with and without outside air intake across a wide variety of single and combined fault conditions.
- **Appendix A: Additional Measurement Details for Commercial Quality Maintenance**
- **Appendix B: Unit Level Observations for Commercial Quality Maintenance**
- **Appendix C: Residential Quality Installation Test Results by Site**
- **Appendix D: Laboratory Instrumentation Testing**
- **Appendix E: Upstream Survey Instrument**
- **Appendix F: Upstream Individual Responses**
- **Appendix G: Upstream Incentives Levels**
- **Appendix H:Residential Quality Installation Site Instrument**
- **Appendix I: Commercial Quality Maintenance Site Instrument**

In addition to the information provided in the Appendices, all evaluation data will be compiled as part of a CPUC data warehouse effort. The Appendices contain most, but not all data collected.

2. Commercial Quality Maintenance

The evaluation team evaluated the installation rate and gross savings associated with the IOU statewide and third-party/local Commercial Quality Maintenance (CQM) program. To accomplish this, we conducted site inspections and measurements to determine evidence of measure installation for both statewide CQM and third-party/local CQM programs. Pre- and post-metering was also conducted for the statewide programs, which were the focus of the evaluation research plan.[11] The statewide CQM program workpaper took an "expected value approach," based on an assumed frequency and combinations of measures that would be implemented, on average, across all projects. The evaluation team analyzed the final measure frequencies for all CQM programs based on implementer records separately from the site inspections to clearly differentiate between the differences in the assumed and installed measure frequency and whether claimed measures were installed and operating. Site visits focused only on inspecting the measures that were performed.

The evaluation effort presented several logistical and technical challenges. Evaluators coordinated with the statewide programs to sample and monitor units pre- and post-maintenance. We monitored units before program service technicians assessed the measures needed for each unit. The team randomly sampled units at sites enrolled in the program at the time, but the outcomes included some units receiving little or no maintenance. Sites could also be enrolled and inventoried and then later dropped from programs, which was a situation that was not originally anticipated. Three units with monitoring equipment installed were replaced with completely with new air conditioners. Program implementers and contractors participating in the monitoring provided notice of service dates for direct observations in some, but not in all cases. Since the monitoring sample presented multiple logistical challenges, the evaluation team conducted additional site observations of the contractors and technicians with the greatest volume of program work in terms of units, unit tonnage, and measures implemented. The non-monitored sites could not be used to estimate energy savings but allowed the evaluators to get a representative sample of the installation rates across all CQM programs. A majority of our non-monitored sites originated with third-party /local programs, as these were the programs with the largest volume of installed units and measures.

2.1 Background

The 2010-12 QM and QI program designs and measures were developed in response to HVAC research, ASHRAE/ANSI/ACCA standards development, Strategic Plan goals, the poor net

[11] As scoped, this evaluation was not intended and did not produce ex post savings results for third-party / local programs, although these produce the majority of HVAC maintenance savings claims.

savings realization rate[12] reported in the 2006-08 evaluation of refrigerant charge and airflow (RCA) and duct test and seal (DTS) measures; and other technical issues summarized by the 2009 HVAC Maintenance Study.[13] Previously, the HVAC program was designed around a single measure approach (either RCA or DTS) that addressed each fault independently. Each measure was implemented separately, and therefore was also evaluated independently. However, HVAC faults include other issues beyond refrigerant charge or duct leakage, are not independent of one another, and cannot be properly addressed in this manner, either by program design or evaluation. Recent lab test activities[14] indicated interactions between multiple HVAC system faults and the ability of one fault to mask the existence of another fault. The 2010-12 program designs favored a more holistic and comprehensive approach as opposed to addressing only one or two potential issues affecting HVAC operating efficiency. Due to these changes and program projections, the 2006-08 evaluation results and underlying data were not applicable to the 2010-12 statewide measures and programs. The research planning for 2010-12 focused on evaluating the new program designs based on the program implementation plans.

The 2006-08 evaluation conducted by the CPUC and the 2009 HVAC Maintenance Study conducted by SCE addressed underlying technical issues in energy efficiency programs built around HVAC maintenance and provided several recommendations and areas for further study. The IOUs used these recommendations to develop the 2010-12 HVAC QM and QI programs, and to suggest changes to evaluation approaches. Based on these recommendations, the IOUs included an element in the HVAC Technology Systems and Diagnostics Advocacy (HTSDA) program, a statewide non-resource program, to address further research needs.

The 2006-08 evaluation and HVAC Maintenance Study postulated (and additional laboratory studies confirmed) technical and training issues that may mislead technicians from the proper diagnosis and measurement of faults and measurement of performance in the field. These issues can arise from inadequate or out-of-calibration sensors, poor sensor installation, diagnostic procedures that do not provide effective multi-fault repair recommendations, and lack of technician training. The IOUs subsequently designed the HTSDA and the HVAC Workforce, Education, and Training (WET) program (another non-resource component of the statewide HVAC portfolio) to help address these issues. Finally, The Energy Division review[15] of

---

[12] The savings projections for HVAC measures in 2006-08 were similar in scale to 2010-12: 2006-08 = 187 GWh vs. 2010-12 = 181 GWh, but the evaluated net realized savings were less than one third, or 53 GWh.

[13] Hunt, Marshal, Kristin Heinemeier, Marc Hoeschele, and Elizabeth Weitzel. *HVAC Energy Efficiency Maintenance Study*. Davis Energy Group and Southern California Edison, 2010. http://www.calmac.org/publications/HVAC_EE_Maintenance_Final.pdf

[14] Purdue and Chapter 5 of this report. 3

[15] CPUC Decision D.10-04-029, Ordering Paragraph 4, third bullet point, states, "Review of completed IOU workpapers regarding ex ante savings estimates are subject to Energy Division review and approval,

QM program workpapers directed the IOUs to develop justifications for ex-ante savings estimates proposed in the workpapers by June 2012. The IOUs did not provide justification, but rather relied on this study (WO32) to provide the additional data. The evaluation team provided preliminary results in a memo to the IOUs in August 2013 and addressed comments from the IOUs and Western HVAC Performance Alliance (WHPA) in September 2013.

A full comparison to previous CPUC or IOU HVAC evaluations or evaluations from other regions was not part of this study. The IOU program planners should take into account findings from all of these studies when planning for future HVAC programs.

2.1.1 Single and Packaged Measures

There were two program types for commercial quality maintenance (CQM): the statewide (SW) program that was based on the ASHRAE/ANSI/ACCA 180 standard (Standard 180), with an assumed package of measures, and the third-party/local programs that continued to implement individual measures including refrigerant charge similar to 2006-08 programs, as well as coil cleaning, economizer repair, and thermostat adjustments. Within the statewide programs there were also claims in 2010 based on the previous program models with individual measures.

The Standard 180-based SW program required a three year service agreement between the contractor and customer. Standard 180 provides a list of maintenance items to check and the frequency of those checks. Some of the maintenance activities provide energy savings while others do not. The service agreement established specific plans for units at each site. Early on in the 2010-12 programs these agreements were difficult to obtain as program processes were still being defined for detailed unit-level data collection. The impact evaluation was designed to measure energy savings resulting from maintenance activities and was not designed to study the standard, program process, or longer term market transformation or market effects. These elements could be addressed in a future study.

The three-year delivery approach presented an issue for program evaluation. The evaluation goals were stated to IOU program staff and implementers. Even if the program allows multiple years for service actions, the key actions must be taken as soon as possible to be considered first year savings. The process outlined to the evaluation team by IOU program staff and implementers was that most issues would be identified and repaired early in the process. If it was known that most actions would not be identified and addressed until the end of the three-year agreement, the evaluation team would have recommended that savings be claimed in later years and would have focused pre-post metering on later activities.

as set forth in an Administrative Law Judge's Ruling of November 18, 2009 in Application 08-07-021, et al."

The evaluation also encountered the challenge that there is no single set of fault detection diagnostic (FDD) standards. This report references studies by others on these issues. The evaluation team agrees that Standard 180 does not provide specific guidance on evaluating maintenance issues, including refrigerant charge. Because different FDD methods produce different outcomes, the charge adjustments may produce positive savings, negligible savings, or negative savings. Some tests are outlined in Title-24 for new installation verification, but do not cover all faults. Manufacturers' recommendations in service manuals represent an independent benchmark to evaluate various FDD protocols. Fortunately for this effort, the laboratory testing included in this effort allowed an evaluation of both approaches noted above along with program-specific guidelines.

The pre- and post-monitoring approach focused on the two IOUs offering statewide CQM programs with participation in 2010-12. PG&E had limited participation in 2010-12 and was not included in the statewide program sample. In the final tracking data only SCE claimed ex ante savings for CQM measures. The evaluation team developed the research plan under guidance from Commission staff, advisory consultants, and the IOUs, to focus on the statewide programs, assuming those programs would eventually be the primary offering for CQM.

The final tracking claims established that individual measures installed via third-party / local programs dominated the total savings claims from 2010 through 2012. The team then refocused efforts on evaluating the installation rate of individual maintenance measures. The team also modified its focus in light of the new 2013-14 program workpapers for individual measures. Table 8 shows that the individual measures comprise the vast majority of 2010-12 program claims. Note individual measures shown in statewide programs are holdover claims from the 2009 bridge program year and do not reference the statewide workpaper. The ex post evaluation results from pre-post monitoring summarized in this section, however, apply only to the IOU's statewide CQM program, as directed by the CPUC in the scope of this evaluation. Installation rates were developed for individual measures, but apply only to the third-party/local program savings claims. Earlier in the 2010-12 cycle the IOU's planned to transition local programs into the statewide model, but going into 2013-14 the direction is moving back to individual measures. These continuing changes in CQM program structure and the potential of examining both state-wide and local program structures adds a significant level of difficulty to program evaluation.

Table 8: IOU Ex-ante Savings Claims for Commercial Quality Maintenance

			CQM Package		Individual Measures	
Program Type	Program ID	Program Name	Energy Savings (kWh)	Demand Savings (kW)	Energy Savings (kWh)	Demand Savings (kW)
Statewide	**PGE21065**	Commercial Quality Maintenance Development	0	0	5,504,808	5,589
	SCE-SW-007E	Commercial Quality Maintenance Development	3,582,081	676	6,686,964	5,678
	SDGE3148	Quality Maintenance Program	0	0	N/A	N/A
Local / Third Party	**PGE2181**	Air Care Plus	N/A	N/A	60,556,253	11,031
	SDGE3161	Non-Res HVAC Tune-up	0	0	17,947,787	5,579
TOTAL			**3,582,081**	**676**	**90,695,812**	**27,876**

2.2 Methodology Overview

The evaluation team installed data loggers on randomly selected units at statewide participant sites to measure time-series electricity use; weather data; refrigerant temperature and pressure[16]; and outdoor air, return air, mixed air, and supply air temperatures. The team sampled refrigerant diagnostic readings from the systems and measured supply airflow and outdoor damper airflow on each unit before and after maintenance services were performed. The team performed acid tests for non-condensables on most refrigerant circuits along with spot measurements and observations. The team collected make, model, serial numbers, digital photographs, and measurements for each unit. The team also observed educational training

[16] The analysis focused on airside measurements. Time series measurements for refrigerant temperature and pressure were originally included, but due to issue with the reliability of the pressure measurement connections these were not installed later in the study and these sensors were removed on all measurement suites.

classes, and observed technician experience, certification, tools, equipment, procedures, and diagnostic protocols.

The team used on-site field observations and interviews to evaluate the behavior of technicians participating in the program. Due to difficulties encountered conducting the pre/post data logging, and in response to issues noted during field observations, the evaluation team stepped up observation and inspection activities. In total, the number of on-site field observations and inspections completed was greater than the number of sites with data loggers. The planned and final sample design is described in Section 2.2.2.The team used observation data to identify equipment faults and assess the efficacy of the technicians in identifying and repairing faults. The evaluation team conducted the following three types of field observations:

1) Pre-maintenance observations conducted on commercial packaged air conditioning systems scheduled for service by participating technicians
2) Ride-along observations and interviews conducted while maintenance services were being performed by participating technicians
3) Post-maintenance inspections and interviews conducted after program services were performed

2.2.1 Observation Methods

The evaluation consisted of field observations, surveys of technicians performing service, and monitoring of HVAC system energy use before and after the service. The evaluation team installed data loggers on 44 units in two service areas to monitor pre- and post-load impacts. The team used data loggers and master technicians conducted conduct pre- and post-observations of program participant technicians performing maintenance on units. Considering monitored and non-monitored units, the team conducted field observations on 99 participant air conditioning circuits.

The master HVAC technicians identified issues and faults during field observations. An “issue” was defined as a problem or difficulty and a “fault” was defined as a defect. Issues are problems with tools or procedures that can impact FDD and the reliability and appropriateness of subsequent repairs. Faults impact cooling system performance and can cause failures of components such as compressors, motors, economizers, sensors or controllers. Faults refer to aspects of the HVAC system that are outside acceptable manufacturer specifications and tolerances such as superheat, suction temperature, evaporator saturation temperature, condenser saturation temperature, airflow, temperature difference across the evaporator or heat exchanger, motor/compressor amps, watts, voltage, cooling/heating capacity, outdoor air damper position, cabinet/duct leakage, coil fouling, economizer change over temperature, sensors, actuators, etc. Issues cause faults and faults impact cooling system performance.

The evaluation team established metrics and criteria for FDD based on published manufacturer installation and maintenance specifications for the specific unit being serviced. Manufacturers provide instructions, specifications, and protocols for coil cleaning, airflow adjustment (belt tension/alignment, fan-belt-drive pulley turns, CFM, static pressure), refrigerant charge (protocols, tools, specifications, targets), recovery and evacuation methods, liquid line driers, and economizer setup and operation (filters, wiring, sensors, controls, dampers, operation). The same criteria were used across all observations. These criteria were used to assess overall maintenance practices, but not installation rates for measures not included in the installation rate estimates. Many points on the data collection forms for master technicians mirrored program data collection. Detailed testing methods and criteria included:

- Industry and manufacturer requirements for installing new, properly sized liquid line driers when repairing leaks, reversing valves, thermostatic expansion valves[17] (TXV), or compressors to prevent refrigerant restrictions. If water vapor is left in the system, it can combine with oil and refrigerant to form corrosive acid and sludge and produce refrigerant restrictions in the form of ice at the expansion device or filter drier (if present).
- Master technicians evaluated the refrigerant system to determine the presence of restrictions or non-condensables. This included an acid test as an indicator of non-condensables including moisture in the system. Moisture in the system may produce a partial orifice freeze-up or improper TXV tracking. Restrictions that may occur during improper installation or as a result of non-condensables in the system may include a plugged inlet screen, foreign material in orifice, filter drier restrictions, kinked or restricted liquid or suction lines, oil logged refrigerant flooding the compressor, or wax buildup in the expansion valve from the wrong oil in system. If the restriction is at the metering device, then frost or ice will develop at this location. If the restriction is at the liquid line or filter drier, then the liquid line temperature will be colder than ambient at this location. All restrictions lead to a reduction in cooling efficiency and may reduce equipment life. Depending on severity, these issues make it difficult or impossible to assess the amount of charge while using diagnostic measurements.
- A check of the air distribution system total static pressure and supply air fan-belt-drive pulley revolutions per minute (RPM) and turns and belt tension/alignment. For fan belts, the team evaluated the measure according to manufacturers' recommendations, which includes checking alignment with a straight edge and checking tension with a tension gage. Power draw is an important measurement and the inspections included this aspect.

[17] A thermostatic expansion valve is a metering device for refrigerant flow into the evaporator of an air conditioner or heat pump. The valve is placed upstream from the evaporator inlet and is connected to a temperature sensing bulb and pressure tap that are located at the evaporator outlet. As the gaseous refrigerant leaves the evaporator the TXV senses its temperature and pressure (superheat) and adjusts the flow rate to maintain the super heat at a constant value.

- Master technicians performed a cold spray test and tested the signal from the controller back to the economizer sensor. Manufacturer economizer installation instructions recommend cold spray to test sensor functionality. It has been noted this test is not recommended for repeated use on modern economizer sensors and controllers which may have other functional testing options as specified by the manufacturer. Many of the economizers in this study were older and this test was the best available method. While minimum outside air damper position was noted as part of the evaluation, minimum outside air adjustments are considered retro-commissioning[18], not maintenance measures, and are separate from economizer functionality.
- The master technician team was not aware of an industry/manufacturer procedure to evaluate the severity of dirty coils or the need for coil cleaning. Manufacturer procedures indicate cleaning inside and outside of coils by removing panels, condenser fans, or economizers as necessary, but there is no specific method to evaluate the state of coils. Manufacturer procedures vary in terms of cleaning with a solution and rinsing with water or only using water. The evaluation process included noting technician coil cleaning processes and providing a general evaluation of coil cleanliness in post-only evaluations. While not in the scope of this study, the team recommends future work to quantify the energy savings, persistence, and cost associated with following the manufacturer-recommended procedure and any alternatives that take less time or claim to provide better savings.
- Master technicians observed technician use of refrigerant system diagnostic equipment. This included noting the use of EPA 608 requirements for low-loss fittings on refrigerant hoses and de minimis purging hoses of air and water vapor prior to attaching to Schrader valves. Incorrect connection of refrigerant diagnostic gauges can create a performance issue by introducing non-condensables[19] into the system. Additionally, master technicians evaluated the placement location and attachment processes for refrigerant system and coil entering air temperature readings.

---

[18] Retro-commissioning is the application of the commissioning process to existing buildings. Retro-commissioning is a process that seeks to improve how building equipment and systems function together. Depending on the age of the building, retro-commissioning can often resolve problems that occurred during design or construction, or address problems that have developed throughout the building's life. In all, retro-commissioning improves a building's operations and maintenance (O&M) procedures to enhance overall building performance. http://cx.lbl.gov/definition.html

[19] Refrigerants are required to condense at the temperatures and pressures in the unit's condenser coil. Non-condensables found in refrigerant systems include nitrogen, oxygen, and carbon-dioxide which are common in the air surrounding unit service ports and can be inside refrigerant hoses if they are not purged before connection.

2.2.2 In Situ Performance Estimation Methods (Statewide Programs Only)

In addition to conducting ride-along observations and ex-post site inspections (sampled to represent the majority of participating contractors and measures), evaluation staff monitored a sample of sites before and after service in both the SDG&E and SCE QM programs. This monitoring component of the study aimed to characterize the performance changes caused by QM program activities on sampled units. The planned sample was relative to program claims, but the achieved sample relied on pre-measurement coordination and was more variable than expected. Table 9 shows the planned sampling and achieved sample. The PG&E program did not claim savings for the package of measures, thus pre-post monitoring was not conducted on PG&E sites. Coordination challenges prevented achievement of the full sample size for SCE. The desired sample was achieved for SDG&E, but the units sampled did not have savings claimed in the 2010-12 program cycle[20]. Precision levels did not meet expectations due to the smaller sample and also the high variability across the sample. The planned coefficient of variation (CV) for energy savings per ton was 0.6 but the results suggest a CV of 1.2 to 3.5 which is extremely high. Based on the coefficients of variation observed in the pre/post metering sample, sample sizes ranging from several hundred to over 1,000 units may be required to achieve 10% relative precision.

Table 9: Planned and Achieved Sample for Pre-Post Monitoring

Program Type	IOU	Sample Size Planned	Planned Relative Precision 90% Confidence Interval (±)	CQM Statewide Measure Savings (kWh)	Sample Size Fielded / Achieved	Expected Relative Precision 90% Confidence Interval (±)	Achieved Relative Precision 90% Confidence Interval (±)
Statewide	**PG&E**	**40**	16%	0	0	NA	NA
	SCE	**40**	16%	3,582,081	23/17	20%	141%
	SDG&E	**20**	22%	0	21/12	21%	61%

Prior to implementer technician service activities, evaluators connected metering equipment to sampled units, airflow tests were performed, and master technicians evaluated the initial state of the unit. The installed metering suites were then left in place for three to six weeks before program technicians began the maintenance process. Following service activities, airflow tests were performed again and metering suites were re-launched. Since QM program activities were

[20] Some units were eventually corrected after M&V and were claimed in the 2013-14 program cycle while others were never claimed.

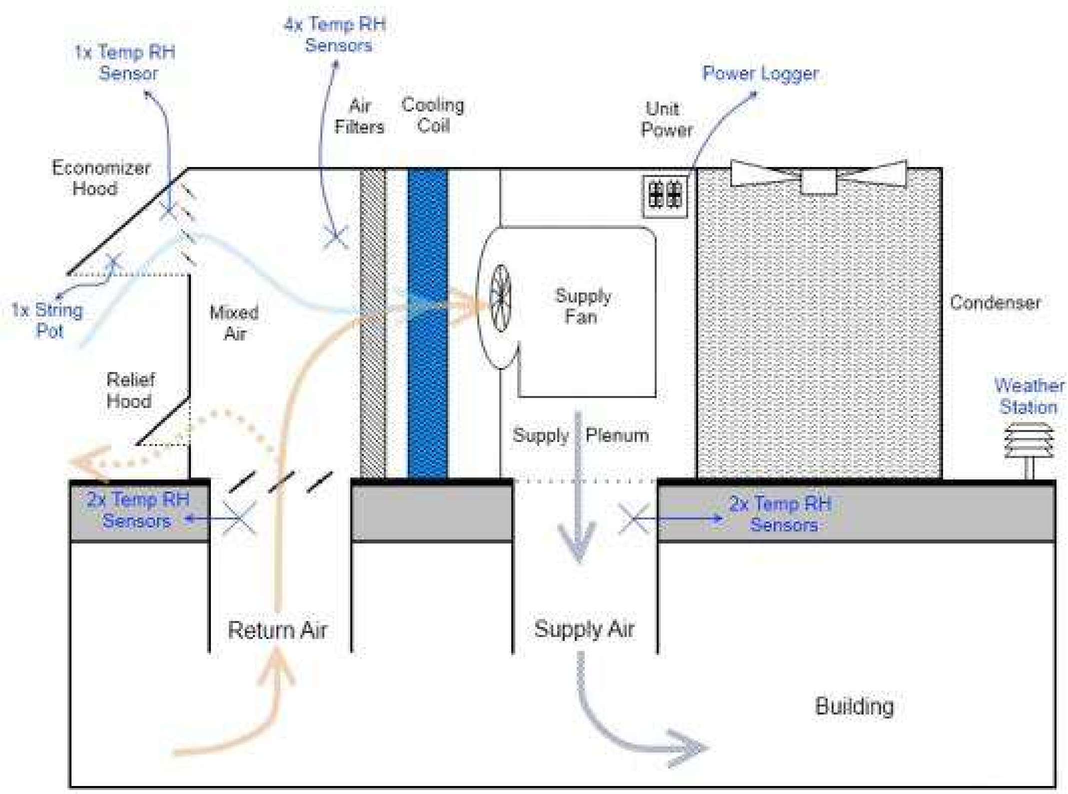
typically conducted on multiple days, often separated by months for many units, this testing and data collection process was repeated on multiple occasions for some units.

2.2.2.1 Equipment Metering

The metering suites included an array of sensors. The evaluation team installed these sensors throughout units in order to characterize airside equipment performance. Meters recorded all parameters using one (1) minute logging intervals to maximize the granularity of the recorded data.

Figure 1 provides a schematic representation of the air-side equipment measurement locations in a packaged unit. On some units we also installed refrigerant-side pressure and temperature sensors similar to the measurements taken by service technicians. These measurements were not included on all units because they used up data logger power and memory at a rapid rate, and they introduced the potential for failure resulting in refrigerant leaks, which were a safety and environmental hazard. This figure is for illustrative purposes only and does not depict an actual unit configuration. Units in the field had both economizers and fixed ventilation air intakes.

Figure 1: Packaged Unit Airside Metering Schematic



The metering suite included nine (9) temperature/relative humidity (temp/RH) sensors. In a typical installation, the evaluation team installed two (2) sensors in the return plenum just upstream of the unit's curb mount; one (1) underneath the economizer hood in an area shielded from solar radiation; four (4) in equal (as possible) quadrants inside the mixed air chamber at the face of the air filters; and two (2) in the supply air plenum just downstream of the unit curb. The team installed two sensors in the return and supply for the purposes of redundancy and four sensors in the mixed air plenum because of the stratification effects inherent to package units. Packaged units typically have little to no mixing space between the economizer hood, return air intake, and the coil face. As such, for a unit in down flow configuration (by far the most commonly observed configuration), the lower portion of the coil typically entrains return air while the upper portion entrains outdoor air. By using multiple sensors, the evaluation team aimed to capture these effects as well as possible given the limits of field instrumentation. If the multiple sensors were within the stated equipment accuracy, the values were averaged. If one sensor was consistently off, the data were reviewed to determine the correct sensor to utilize. If both sensors were close to the same but appeared erroneous, the unit was dropped from the analysis.

The metering suite also included sensors and transducers to measure power and economizer damper position. To capture unit energy use and demand, the team connected a true power pulse-count logger to the input power of each unit. The team monitored economizer damper actuation using a string-potentiometer sensor, which outputs a signal in proportion to the linear displacement of the string. The team affixed potentiometers to a static location in the economizer housing and to an economizer blade. Linear displacement registered by the potentiometer therefore translated into motion of the economizer dampers. The measurement qualitatively identified damper movement, but not actual damper position.

In addition to the monitoring equipment installed on each unit, one radiation-shielded weather station housing a temp/RH sensor was installed per site at a location removed from the immediate vicinity of the units. This logger was installed to capture ambient weather conditions to be used later in equipment performance and site load modeling.

The pre-service, post- service and sometimes "interim" service data sets were later utilized to develop DOE-2 performance maps and design conditions for each service period. The evaluation team determined that these maps were not sufficiently reliable to establish changes in capacity and efficiency. The team performed a unit-level regression analysis using sub-metered power consumption data (Section 2.2.2.2). The team also used data collected during the monitoring process to generate simplified curves that relate power draw per ton of cooling to coil and space load estimates. The team conducted an engineering assessment of the site data collected to support the planned performance map approach (Section 2.2.2.3). The engineering analysis also developed time series load profiles based on schedules and set points for the metered units and used these profiles to estimate unit-specific annualized program energy impacts. The assessments found specific reasons leading to the outcome for each unit. In some instances, the assessments determined there were insufficient pre-maintenance data to characterize unit performance. These units were eliminated from the final analysis. .

2.2.2.2 Unit Level Regression Modeling

The evaluation conducted a regression analysis on pre-post power and schedule data to estimate savings. The team originally developed performance maps and zonal DOE-2 simulation models, which proved unreliable given the measurement uncertainties. The revised primary analysis utilized the regression model published in the Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures[21].

The analysis is based on the following equation to determine unit level hourly load estimates for the pre-maintenance and each post-maintenance period:

$$L_{dh} = \alpha + \beta_{Ch} THI_{dh} + \beta_{w(d)} w(d) + \beta_{g(h)} g(h) + \beta_{2h} H_{2d} + \beta_{3h} H_{3d} + \varepsilon_{dh} \quad \text{[Equation 1]}$$

where for a particular HVAC unit;

L_{dh} = Load in kW on day *d* hour *h*, day= 1 to 365, hour = 1 to 8760

THI_{dh} = Temperature-humidity index on day *d* hour *h* in °F. For the purposes of this analysis, THI was defined as:

$$, THI = 0.5 \times OSA_{db} + 0.3 \times DPT + 15,$$

Where THI is the temperature-humidity index in °F,

OSA_{db} is the outside dry bulb temperature in °F, and

DPT is the outside air dew point temperature in °F.

w(d) = 0/1 dummy indicating day type of day *d* , Monday = 1, Sunday =7, Holiday = 8

g(h) = 0/1 dummy indicating hour group for hour *h*, hour group = 1 to 24

H_{2d} = 0/1 dummy indicating that hours in day *d* are the second hot day in a row

H_{3d} = 0/1 dummy indicating that hours in day *d* are the third or more hot day in a row

α, β_{Ch}, $\beta_{w(d)}$, $\beta_{g(h)}$ = coefficients determined by the regression

β_{2h}, β_{3h} = hot day adjustments, a matrix of coefficients assigned to binary variables (0/1) for hours defined for 2nd and 3rd consecutive hot days, the matrix variables are unique to each hour in each hot day

ε_{dh} = residual error

[21] National Renewables Energy Laboratory. *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*. Chapter 4. 2013. http://energy.gov/sites/prod/files/2013/07/f2/53827_complete.pdf

The data were filtered for the hours when the unit was on, defined by kW > "fan only" and other parameters (i.e. – specific hours per operating schedule) as determined necessary by the engineer. The inputs were run in a Statistical Analysis Software (SAS) multi-variable regression for kW, THI day-type, hour-type, and consecutive hot day adjustments. The regression used onsite weather stations, or if those data appeared to have issues local weather data available closest to the unit's location, with the intent that modeling the unit this way will yield a much more accurate model that is highly sensitive to changes in temperature and humidity inputs to THI. The estimated values from the regression represent the α, β_{Ch}, $\beta_{w(d)}$, and $\beta_{g(h)}$ coefficient values in the original modeling equation as shown in Equation 1. These values were used to model the energy consumption of the unit based on THI, day type, and hour type. For the 2nd and 3rd hot days, flags were placed in the weather file to denote whether this is the 2nd or 3rd consecutive hot day. A hot day was defined as a day when the average dry bulb temperature for all 24 hours is greater than 80 °F. An adjustment is made for both 2nd and 3rd hot days in the following manner: these hot day variables represent the β_{2h} and β_{3h} terms in the original equation. Note that these terms depend on the day of the week and hour of the day for the hour in question. For each unit-level model there are 340 regression coefficients, one for each of the variables: THI, day-type, and hour-type and each consecutive hot day adjustment results in 168 coefficients for the 0/1 dummy variables.

Finally, the model values are compared to the original metered data and a statistical analysis is done to determine the mean bias error (MBE). The MBE results for the metered time frame were used to prioritize review of the unit models by senior engineers and site staff to supplement the QC. Ultimately an engineering decision was made for each unit where the model did not match the metered data to determine if poorly modeled data was a result of variability in the actual unit usage not captured by the model, or if there was a measurement or other issue requiring exclusion of the unit from the aggregate load shape analysis. Measurement error identification was assisted by examining the residual error term in the regression. The graphing procedure for quality control also automatically calculated the mean bias error for the entire profile as well as for selected time periods. Investigating the error over the various peak hours was easier in the graphical interface. The comparative statistics were calculated for coincident data values, i.e. data values for time periods which contain data values for both the measured (base) and modeled (comparison) profiles. All time periods where data were missing for either the base or comparison profile were ignored. The MBE is the mean of the error or difference between the base and comparison profiles for all pairs of coincident data points over the mean of the coincident data points of the base profile, as shown in the equation below:

$$\text{MBE} = \frac{\frac{\sum_{i=1}^{n}(b_i - c_i)}{n}}{\frac{\sum_{i=1}^{n} b_i}{n}} \quad \text{[Equation 2]}$$

where:

b_i = data value of the base profile at time period i

c_i = data value of the comparison profile at time period i

n = number of coincident data points

The unit level model was applied to climate zone weather files (CTZ 2010)[22] to provide a weather-normalized 8760 load shape profile for the unit. This approach may introduce some spatial bias in the weather normalization, but given the relatively small sample and sites locate relatively close to the climate zone reference stations, this bias is minimal. The load predicted by the model was set to zero if the THI was less than 50°F to avoid modeling conditions where units across all samples had no compressor cooling in the meter data. This decision was made based on review of the engineering data. This restriction had no effect on summer peaks, only on off-peak and annual usage. Any information collected about when the units are activated or shut down for winter was applied when extrapolating the results. If a unit was designated as being turned on in March and off in December then no modeled usage was calculated for January and February. At this stage there is a unit level 8760 weather-normalized profile for each sample point pre and post.

2.2.2.3 Engineering Assessment

Engineers used airflow and metered unit data collected in the field to determine equipment performance before and after service activities. The team synchronized all metered data collected during a specific service interval (e.g., before service, after first service, and after second service) for a given unit, and combined this into a single data set in SAS. The team ran psychrometric calculations on parameters in the data set to determine critical engineering values for each time step. The engineers then used these values in subsequent efficiency and capacity calculations. Table 10 summarizes these critical values and notes how each was calculated.

[22] Updated Climate Zone weather available from the DEER website www.deeresources.com

Table 10: Methods of Determining Key Parameters

Calculated Parameter	Method of determination
Supply air – Average Temperature [F]	Averaged the values from the two supply air temperature sensors.
Supply air – Average Humidity Ratio [lb-water/lb-dry air]	Calculated humidity ratio for both supply air temp/RH sensors using psychrometric equations. Averaged the two values.
Supply air – Average Specific Volume [CF/lb-dry air]	Used the average supply air temperature [a], the average supply air humidity ratio [b], and psychrometric equations to determine average specific volume.
Supply air – Total Enthalpy [Btu/lb-dry air]	Used the average supply air temperature [a], the average supply air humidity ratio [b], and psychrometric equations to determine total enthalpy.
Supply air – Sensible Enthalpy [Btu/lb-dry air]	Used the average supply air temperature [a], a dry air assumption, and psychrometric equations to determine sensible enthalpy.
Mixed air – Average Temperature [F]	Averaged the values from the four mixed air temperature sensors.
Mixed air – Average Humidity Ratio [lb-water/lb-dry air]	Calculated humidity ratio for the four supply air temp/RH sensors using psychrometric equations. Averaged the four values.
Mixed air – Total Enthalpy [Btu/lb-dry air]	Used the average mixed air temperature [f], the average mixed air humidity ratio [g], and psychrometric equations to determine total enthalpy.
Mixed air – Sensible Enthalpy [Btu/lb-dry air]	Used the average mixed air temperature [f], a dry air assumption, and psychrometric equations to determine sensible enthalpy.
Return air – Average Temperature [F]	Averaged the values for the two return air temperature sensors.
Outside air mass fraction	Used the lever rule, the average supply air enthalpy [a], the average return air enthalpy [j], and the average mixed air enthalpy [f] to estimate outside air fraction as: $f_{OA} = \frac{MAh-RAh}{OAh-RAh}$
Air Mass Flow Rate – [lb-dry air/minute]	Used the supply air specific volume [c] and the spot measured volumetric flow rate [CFM] to estimate the dry air mass flow rate. Airflow measured pre and post and pre and post fan-power used to check values.

Next, engineers utilized calculated parameters to characterize two HVAC parameters: unit performance and the load imposed on the system. Unit performance refers to the mechanical cooling output and efficiency of the HVAC package, i.e., the amount of cooling provided by the cooling coil and the efficiency calculated therefrom. System load is defined by the amount of cooling provided to the space, including the building load and duct losses. To look at it another way, ***unit performance*** is calculated based on a control volume *placed around the unit* starting at the face of the coil and ending at the supply air outlet; ***system performance*** is calculated based on a control volume *placed around the building and HVAC distribution system* starting at the unit's supply air outlet and ending at the unit's return air inlet. Outdoor air load was modeled separately to account for differences in pre/post economizer operation.

Engineering assessments used a combination of SAS for raw data processing, spreadsheets to check and post-process SAS output, and the Universal Translator 3 (UT3) software to plot and analyze data.

1. The original raw data are one-minute interval data.
2. The AC unit operations are divided into five states: OFF means the whole unit is shut down; Fan On means the unit is in ventilation mode and the compressor and condenser fan are off; AC On means the unit is in initial cooling mode (unit has been running for no more than 5 minutes) and the compressor, condenser fan, and supply fan are all on; AC STABLE means the unit is in stable running conditions (unit has been running for more than 5 min but does not stop in the next minute); and Heat ON means the unit is in heating mode and fan is on (the compressor would be on for heat pump units and the compressor would be off for gas heating units).
3. To exclude high frequency cycling data, the unit should stay in one state for more than 60 consecutive seconds for the UT3 to recognize the unit as being in that specified state.

2.3 Performance Changes and Savings Estimates from Pre- and Post-Measurement

The evaluation team estimated unit savings using the pre- and post-service load and scheduling profiles developed from the metered data sets. The performance curves, load profile data, and scheduling information utilized CTZ2010 weather data to estimate unit savings.

2.3.1 Unit Changes in Energy and Demand

The evaluation team performed the analysis described in Section 2.2.2.2 on all 44 units with a wide variety of results. In some cases, there were negative savings due to increased loads from opening economizer dampers, which diminished any potential efficiency improvement benefits from other maintenance actions. When reviewing the performance changes, the team noted in data inspections that performance pre- and post-maintenance improved and diminished for different loading conditions for the same unit. This means the curves of efficiency versus load sometimes had similar shapes pre- and post-maintenance, but more often the curves had different shapes that overlapped. The data provides no measure level analysis given the numerous combinations of measures implemented and incomplete documentation by the various QM programs of when those measures were implemented across the various units.

Table 11 shows results across 29 units with reliable pre- and post-data. The units that the evaluation team could not analyze included three units that were replaced before final service, seven units on sites that were initially in the program but were later dropped and never had maintenance attempted, and five units with insufficient pre-maintenance cooling usage. Clearly the challenge of measuring field performance was a known issue prior to fielding the sample, but a roughly equal issue was the uncertainty that a pre-monitored unit would complete service in the program cycle. Despite planning for many more sites, the coordinated approach produced very few units with acceptable data. Future study options include forcing M&V units to remain in the program or obtaining more cooperation from program staff, contractors and technicians to get enough study sites to achieve an oversample.

The average savings was positive for SCE (but not statistically significant from zero) and negative for SDG&E with significant uncertainty. SDG&E did not claim savings in 2010-12[23]. The estimated savings for the statewide program with a pooled sample across both SCE and SDG&E sites was slightly negative, but not statistically different from zero.[24] The evaluation team felt the pooled sample was the most reliable due to the higher sample size (29 vs. 17) and lower standard error (462 kWh/ton vs. 605 kWh/ton). The implementation process was quite similar for both IOUs and the key difference was that SDG&E decided not to claim savings until after all maintenance and program inspection was complete. Based on the direct field observations and post-only visits for SCE and SDG&E, we determined the issues leading to negative savings for SDG&E were present within all 2010-12 statewide programs, and the pooled sample best represents the performance of the statewide program. The engineering assessments described in the next section showed there were several cases where efficiency and capacity changed minimally, but loads increased after program service due to outside air and controls changes.

Table 11: Energy and Demand Savings Estimates for Sample Commercial Quality Maintenance Statewide Program Using Regression Analysis

IOU Statewide Program	Pre-Post Average Savings	Sample Size	Standard Error	90% Confidence Interval Error Bound	Relative Precision	Lower Bound Savings	Upper Bound Savings
	Energy Savings (kWh/ton)						
SCE	705	17	605	994	141%	-290	1699

[23] Savings claims were withdrawn from the final tracking data after interim results were provided to the IOUs.

[24] As noted previously, PG&E was excluded from the statewide program evaluation due to low participation.

SDG&E	-1486	12	551	907	61%	-2393	-580
Total	-202	29	462	760	376%	-962	558
	Demand Savings (kW/ton)						
SCE	0.06	17	0.15	0.25	409%	-0.19	0.31
SDG&E	-0.43	12	0.13	0.21	50%	-0.64	-0.21
Total	-0.14	29	0.11	0.18	132%	-0.32	0.04

2.3.2 Causes for Changes Based on Engineering Assessments

In a secondary analysis, the evaluation team investigated the energy savings using an engineering assessment that utilized all available performance data. The team undertook the analysis to corroborate the regression analysis findings, but detailed analyses were limited given the sensitivity of the engineering analysis to measurement errors of temperatures, humidity, and airflow. The data was qualitatively informative in explaining findings from the regression effort. The regression analysis relied on the power consumption data, outdoor temperature conditions and schedule variables which were not as sensitive to measurement error[25].

The engineering assessment confirmed that maintenance does have the possibility of increasing unit energy consumption as illustrated in the following section. Bringing units up to baseline can increase the capacity and energy consumption of the treated unit, adding previously unmet load through increased comfort or increased outside air. The net result is a grid impact that should be taken into account. Future evaluations should consider these issues. These assessments also confirmed new findings from laboratory testing that show significantly higher than expected measurement uncertainty for rooftop package unit performance. The primary issues noted by the pre/post monitoring and site observations were:

- The pre-post change in performance did not lead to significant energy savings. Multiple compound faults made it such that the units did not perform similar to units under assumed workpaper conditions due to circuit faults not considered in workpapers and excess outside air increased after maintenance.
- When using engineering review of reliable efficiency measurement (excluding negative impacts from increased outside air), the changes in performance are small and the estimates are not statistically significant. However, the increased energy consumption from program service appears to be statistically significant as confirmed by the time series kW analysis. This means that the combination of increased loads and marginal changes in unit performance result in negative savings in some units. The average was negative for the SDG&E units and the combined sample.

[25] Note that measurement error in this case is that associated with issues such as incomplete mixing of air streams and/or spatial variation of important data, not sensor error.

- All of the economizers with time series metering data were not responding to conditions measured by temperature or CO2 sensors, meaning the dampers were essentially fixed in one position prior to maintenance. The fixed position was often increased after maintenance but the economizer still did not respond to temperature signals that would indicate proper operation, and were deemed inoperable. The findings are supported by direct observations that economizers were often not tested or repaired, but damper positions were adjusted by program maintenance technicians. According to the workpaper assumptions all non-working economizers should be repaired and minimum outside air would not change through maintenance.

2.3.3 Examples of Program Changes in Efficiency, Usage, and Outside Air

The evaluation team selected example engineering assessments of units with high, low, and negative savings to illustrate the reasons for the negative savings result. Detailed reports for each unit assessment are contained in Appendix B. More detailed reports with data plots will be made available through the CPUC data warehouse[26]. The following series focuses on three units, one of which has negative savings.

2.3.3.1 Example 1: Negative Saver (Site CC Unit 1)

Measurements indicated the OA damper position was increased by about 15% after the service. The assessment was determined by measuring return minus mixed air enthalpy compared to return minus outside air enthalpy (RAh-MAh vs. RAh-OAh) when in AC ON or AC STABLE operation. Economizer was observed to be inoperable before and after service. This and the change in unit operation schedule dominated the change in energy consumption. The unit was cycling on and off continuously during the baseline period. There were runs of 15 to 30 minutes with no continuous ventilation between cycles. During the post period, the unit was scheduled off from 10:00 pm to 4:00 am during weekdays and all-day weekends and the unit fan ran continuously between cycles. This meant that additional load was being brought in through the added ventilation. The base system efficiency (kw/ton) based on space cooling load was very close to the post kw/ton. The kw/ton values seem to be in the reasonable range of 1 to 2 kW/ton. This indicates the unit performance did not change significantly after the service. Master technician observations confirmed the most significant changes to the unit were opening of dampers with no repair to economizer functionality and changing the schedule.

2.3.3.2 Example 2: Near Zero Saver

The unit operation schedule and outside air fraction did not change significantly after program service. The damper was almost 60% open during the base and post periods with no response to temperature indicating the economizer did not function before or after service. The unit was

[26] File size limitations prevented including all site reports in this document.

cycling on and off during the baseline and post periods except for the period from 2:00 am to 6:00 am. There was no obvious operating schedule change. The base unit kW/ton is slightly better (more efficient) than the post kw/ton at the same cooling coil load. However, there was also a slight increase in space load in the post-period, but it is unclear what the source was of this slight increase. The small improvement in efficiency led to near zero positive savings for the unit. Master technician observations confirmed the most significant change to the unit was a refrigerant charge adjustment.

2.3.3.3 Example 3: Positive Saver (Site BP Unit 13)

Measurements indicated the OA damper position was decreased by about 60% after the service. The assessment was determined by measuring return minus mixed air enthalpy compared to return minus outside air enthalpy (RAh-MAh vs. RAh-OAh) when in AC ON or AC STABLE operation. Economizer was observed to be inoperable before and after service. The post-performance is close to the base-performance at the same coil load and the unit was off from 12:00 am to 6:00 am during the baseline and post periods. The weather-normalized post power was lower than the baseline power due to changes in ventilation air quantities noted. Master technician observations confirmed the most significant changes to the unit were closing dampers with no repair to economizer functionality.

2.3.4 Energy and Demand Savings Estimates

Across all sampled units for both IOUs, the average energy savings per-ton associated with running the individual pre- and post-regression models was negative (Table 11). The analysis focused on 29 units with the most accurate data available. Units were not weighted by program savings, but rather the specific results were turned into an average savings per ton and then extrapolated to the program tonnage claimed. The measured ex post energy savings estimates from the statewide pooled sample was applied to the savings claims for the SCE CQM program as shown in

Table 12. This table was developed as a result of establishing a unit energy savings (slightly negative UES) value for the pooled sample with upper and lower confidence intervals as shown in Table 11. The UES was applied to all program tonnage to get the final estimate for the upper and lower bounds. The results show a range using the measured uncertainty to illustrate the variability in savings was much larger than assumed in evaluation planning. Given the limited sample and the large degree of variability, the point UES estimate was not statistically different from zero, thus the realization rate for the SCE program was zero. SDG&E removed their claims from the final tracking data, thus the realization rate for the SDG&E program is not defined. Given the wide range of uncertainty and limited sample, the results are considered indicative of program performance, but not definitive based on the calculated uncertainty in the measured savings.

Table 12: Energy Savings Estimates for Commercial Quality Maintenance Statewide Program—Measured Ex Post

Program Type	IOU	Ex Ante Energy Savings (kWh)	Ex Post Energy Savings Range (kWh)	Energy Realization Rate (%)	Ex Ante Demand Savings (kW)	Ex Post Demand Savings (kW)	Demand Realization Rate (%)
Statewide	**SCE**	3,582,081	-11,197,680 to 6,495,120	0	676	-3,725 to 466	0
	SDG&E	0	0	N/A	0	0	N/A
	Total	**3,582,081**	-11,197,680 to 6,495,120	0	**676**	-3,725 to 466	0

2.3.5 Additional Observations for Statewide Programs

Additional contractor observations and ex-post inspections were made for the Statewide programs, but these results were not applied to energy savings claims. The goal of presenting this information in Table 13 is to illustrate specific measures that did or did not pass ex-post inspections. Some of the results presented include early 2013-14 program sites and include both data logger and non-data logger sites for SCE and SDG&E. The findings from post-only inspections seem relevant given that program workpapers assume an average frequency of pre-maintenance deficiencies and a post-maintenance condition of no deficiency for each measure. The pass and fail criteria were based on unit-specific manufacturer protocols. Some additional criteria were required for specific measures.

- Coil cleaning was assumed to pass unless the cleaning method was directly observed since use of water versus manufacturer specified methods could not be determined in post-only observations.
- The refrigerant charge pass/fail has issues as this pass rate does not account for the partial pass outcomes. This issue is discussed in greater detail relative to developing installation rates in Section 2.3.6.
- Airflow adjustment was not a program workpaper measure, but documentation existed that issues were identified by service technicians and the uncorrected deficiencies do affect assessment of charge and unit performance.

Table 13: Observation and Inspection Results for Measures Included in 2010-12 Statewide Programs

	SCE Statewide (Data Logger and Additional Observations)		SDG&E Statewide (Data Logger and Additional Observations)		PG&E Statewide (2013 Observations)	
Fault / Measure	Sample	Diagnostic Pass Rate	Sample	Diagnostic Pass Rate	Sample	Diagnostic Pass Rate
Coil Cleaning	94	96%	38	97%	41	12%
Adjust Airflow	6	17%	0	N/A	14	0%
Refrigerant System Service	17	24%	21	19%	22	23%
Economizer Functional Test	71	0%	38	13%	40	10%
Integrate Economizer Wiring	0	N/A	0	N/A	14	0%
Replace Damper Motor	0	N/A	0	N/A	18	0%
Replace Controller/Sensor	0	N/A	0	N/A	29	0%
Renovate Linkage & other components	0	N/A	0	N/A	23	0%
Replace T-stat	0	N/A	0	N/A	18	39%
Notched V-Belt Upgrade	3	100%	2	0%	0	N/A

2.4 Installation Rate Based on Ex-Post Observations

The team directly observed technicians performing maintenance for units with data loggers. The program design allowed measures to be installed over the course of the three-year service agreement, so it was not always possible to observe all maintenance as planned. The statewide programs did not record the amount of refrigerant charge added or removed or specific adjustments made to the system such as economizer minimum position, so the observations were quite valuable.

In most cases the team was not notified when additional maintenance occurred outside of the site observations; therefore direct observation of all measures was not possible. The 2010-12 SCE program team approached the M&V sampling as voluntary and recruited contractors and technicians with a relatively low volume of program jobs for pre-post monitoring. SDG&E and PG&E each had two primary participating contractors. The SDG&E program team obtained full participation in the sample but ultimately claimed no savings for CQM, while both PG&E contractors dropped from the program prior to sampling. Given these situations, the data logger sample did not include some of the top SCE contractors in the statewide programs and did not include the PG&E and SDG&E local programs.

The EM&V team concentrated new efforts in 2013 to observe and inspect work performed by the contractors accounting for the most claimed savings for the 2010-12 program cycle. The team coordinated with program implementers to perform additional contractor observations and post maintenance inspections for the statewide programs and for the local programs to supplement the initial sampling. Due to the size of the program savings claims, most of these additional observations and inspections occurred within the third-party / local programs. Since the local programs used the individual measures design, the team only performed post maintenance inspections within these programs. The refrigerant charge and economizer measures were a primary focus given the difficulty in assessing the coil cleaning measures, since coils may have gotten dirtier or have been cleaned since initial service and prior to evaluation inspections.

Details for each unit assessed are provided in Appendix B. These details are extremely important since the measures varied across the programs. Manufacturer's recommended maintenance procedures, including manufacturer's refrigerant charge diagnostic protocols were used as the basis of comparison across all sites sampled. The EM&V team determined late in the process, after the last round of laboratory testing, that the most reliable method for assessing refrigerant charge was complete removal and weigh-in to factory charge. This process was completed on a select number of units since the process is more expensive than indirect diagnostics and impact evaluation did not have budget and time to re-assess a larger sample. The master technicians described the details of the results compared to manufacturer and Title-24 diagnostics in the Appendix B. Lab testing indicated that manufacturer refrigerant charge protocols provided the best set of diagnostics for all possible faults, but still provided false diagnostics in some cases.

2.4.1 Fault Detection Diagnostics Role in Installation Rate

Correct FDD and repair are important for HVAC maintenance. If technicians cannot correctly perform FDD and repairs, then HVAC maintenance programs will not realize ex ante savings. In addition to the laboratory findings presented later in this report, another set of recent reports indicate problems correctly identifying faults using various RCA protocols. Yuill and Braun indicated that the Title 24 RCA protocol identifies faults in 46% of cases without faults, misdiagnoses in 25% of cases with faults, and does not detect faults in 32-55% of cases with faults. [27,28]

---

[27] Braun, Jim, David Yuill, and Howard Cheung. *A Method for Evaluating Diagnostic Protocols for Packaged Air Conditioning Equipment*. Purdue University and the California Energy Commission, 2012.

[28]Braun, Jim, and David Yuill. Evaluation of the Effectiveness of Currently Utilized Diagnostic Protocols. Purdue University. 2014.

Are technicians correctly following a poor protocol or are errors compounded due to incorrectly implementing a poor protocol? The evaluation found problems with both protocols and implementation of those protocols. Additionally, the team found a lack of awareness and training regarding manufacturer protocols that are provided in installation manuals.[29] Manufacturers provide unit specific service procedures in the installation manuals with trouble-shooting information on many but not all faults. The current Title 24 RCA protocol is based on superheat measurement, subcooling measurement, and airflow diagnostic protocols from at least three different manufacturers that have been available for decades.[30] These protocols have been applied in previous program cycles and were still used in some 2010-12 cycle programs. The Title 24 protocols help technicians evaluate proper RCA when no other faults are present. The Title 24 and at least some manufacturer protocols are not designed to account for or diagnose other faults and were developed under conditions with no combinations of faults.

In addition, many protocols depend on accurate measurement of the temperature and humidity of air entering the evaporator coil. Unlike residential systems, common rooftop packaged units have outside air intakes that bring in ambient air that does not fully mix with the return air. The coil entering temperature is only equal to the return air temperature with the outdoor air intakes are completely sealed. Service technicians measuring only the return air or a single point measurement near the coil produce inaccurate coil entering temperature measurements leading to misdiagnoses of system faults. Protocols that do not require direct measurement of coil-entering temperature and humidity are also affected by outside air because these conditions change the unit loading that changes the target values for the diagnostics.

The instrumentation used for FDD can introduce significant measurement uncertainty, thus confounding the ability of the FDD protocol to accurately diagnose faults. The measurements include air temperatures and the pressure and pipe surface temperatures of the refrigerant circuit. The 2006-08 HVAC Evaluation and previous studies show that the target tolerances cannot be reliably achieved without an accurate and relatively expensive combination of instrumentation.

Program technicians do not evaluate FDD of outdoor airflow through economizer or make-up air dampers. Additionally, outdoor airflow is not included in efficiency tests used to develop diagnostics for packaged units. The database used by the Yuill and Braun study did not include

---

[29] Carrier 2005. 48HJD/HJE008-014, 48HJF008-012 Single-Package Rooftop Gas Heating/Electric Cooling Units. Installation, Start-up, and Service Instructions. Form 48HJ-32SI. Fig. 57 – Cooling Charging Charts. http://www.docs.hvacpartners.com/idc/groups/public/documents/techlit/48hj-32si.pdf

[30] Carrier 1986. R22 Superheat, Subcooling, and Airflow Calculator. GT24-01 020-434

Carrier 1998. R410A Superheat, Subcooling, and Airflow Calculator. GT58-01A 020-517

Trane 1996. Air Conditioning Charging Calculator. Pub. No. 22-8065-07

York 1991. Superheat, Subcooling, and Airflow Calculator. Form 501.00-PM5Y (5/91).

units with fixed outdoor air intakes or economizers. This critical gap was filled by laboratory testing conducted by the evaluation team, where units were tested with economizers, as described in Section 5 of this report.

2.4.2 Results for Recovery and Re-Charge

The team evaluated refrigerant charge in the field by measuring ex-post performance on units where technicians made adjustments. The evaluation selected a random sample of packaged rooftop air conditioners towards the end of the evaluation as new information from laboratory testing was released on the uncertainty in the accuracy of charge diagnostics. The detailed charge assessment sample included five single-compressor and two dual-compressor units for a total of nine refrigerant circuits. For each circuit, the EM&V master technicians recovered and weighed out charge, evacuated to 500 microns, held at or below 500 microns for 20 minutes, weighed in factory charge, and measured performance with factory charge. The average refrigerant charge adjustment reported by the local program for the units evaluated was 19.4 ± 5.7%. This indicates that, on average, units had a significant amount of charge added. Recovery and weigh-out of these units indicated cases of both over and undercharge after service. Units that were still undercharged after the program-added charge realized benefits, but additional potential remained. Units that had charge added and ended up overcharged had multiple potential outcomes including positive, zero, or slightly negative benefits.

The team looked at the refrigerant charge benefits using the efficiency estimate at various charge conditions in the 2009 PG&E workpaper, which the ex ante team recommended based on the available data. The workpaper provided estimated EER at various charge conditions for TXV and non-TXV units. The team calculated program assumed EER's based on claimed initial charge, actual initial charge and final charge state. From this the team calculated claimed, actual and correct savings percentages based on average EER's. Installation rates were then calculated as the actual benefits divided by the claimed benefits since these rates were developed and applied to the third-party and local programs. Results are summarized in Table 14. Note that the change in EER values in this table show both TXV and non-TXV, but for the installation rate calculation, program level total tonnage of TXV and non-TXV units were used given the frequency of metering devices in the sample was not likely to be representative. Total tonnage associated with TXV units was 16% for the AirCare Plus program and 46% for the SDG&E local program.

Table 14: Installation Rate Based on Recovery and Re-charge Method

Unit #	Circuit #	Program Charge Adjustment	Unit Weighed Charge Compared to Factory Charge after Program Service	Disposition - Did Program Adjustment Have a Benefit based on direction and amount	Claimed Change in EER (TXV/Non-TXV)	Actual Change in EER (TXV/Non-TXV)

Unit 1	Circuit 1	Added 7.1%	Charge within -2.6%	Partial benefit, high discharge pressure due to failed condenser fan reduced efficiency by 27%. Undercharge had 5% impact.	0.06	0.33	0.08	0.42
Unit 1	Circuit 2	Added 6.5%	Charge within -1.1%	Partial benefit, also see above	0.05	0.29	0.06	0.33
Unit 2	Circuit 3	Added 14%	Overcharged 6.5%	Partial benefit (negative for TXV), Overshot	0.18	0.87	-0.09	0.17
Unit 2	Circuit 4	Added 13%	Overcharged 18.3%	Negative benefit	0.16	0.78	-0.32	-0.38
Unit 3	Circuit 5	Removed 17%	Undercharged 31.4%	Negative benefit	0.41	0.50	-0.62	-2.30
Unit 4	Circuit 6	Added 33%	Undercharged 9.4%	Increased benefit exceeds range of DEER charge adjustment assumptions, representing additional opportunity	0.89	3.49	1.20	4.39
Unit 5	Circuit 7	Added 19.1%	Undercharged 17%	Increased benefit exceeds range of DEER charge adjustment assumptions, additional opportunity if condenser kept clean.	0.32	1.42	0.80	2.90
Unit 6	Circuit 8	Added 16.8%	Undercharged 15.8%	Benefit consistent with DEER, additional opportunity	0.26	1.16	0.65	2.37
Unit 7	Circuit 9	Added 48%	Overcharged 6.1%	Increased benefit exceeds DEER, over by 6.1%	1.29	4.87	1.14	4.69
Unit 8	Circuit 10	Added 13%	Overcharged 13%	Negative benefit	0.16	0.78	-0.32	-0.38
Installation Rate - Combined based on program savings								**79%**
Installation Rate – PG&E AirCarePlus (14% TXV)								**81%**
Installation Rate – SDG&E Local (46% TXV)								**77%**

If only diagnostics are used after program service then most units would be indicated as not passing, and the resulting installation rate would be very low. However, as shown above the program adjustments did have benefits. The outcome would be much different if the installation rate was based on pass/fail diagnostics as indicated in Table 15.

Table 15: Comparison of FDD Methods for Recovery and Recharge Sample – As Found Conditions

		Diagnostic outcome after program service (as found condition)		
Unit #	Compressor #	2008 Title-24 / 06-08 Program (SH or SC)	Generic Program FDD (SH, SC, COA, EST)	Manufacturer FDD (DP, SP, AT, LT, ST, SH. SC)
Unit 1	Circuit 1	Fault	Fault	Fault
Unit 1	Circuit 2	Fault	Fault	Fault
Unit 2	Circuit 3	Fault	Fault	Fault
Unit 2	Circuit 4	Fault	Fault	Fault
Unit 3	Circuit 5	False Pos.	Fault	Fault
Unit 4	Circuit 6	Fault	Fault	Fault
Unit 5	Circuit 7	Fault	Fault	Fault
Unit 6	Circuit 8	False Pos.	Fault	Fault
Unit 7	Circuit 9	Fault	False Pos.	Fault
Unit 8	Circuit 10	False Pos.	Fault	Fault
		Fault	**Fault**	**Fault**
		7 of 10	**8 of 10**	**10 of 10**

Finally, the team assessed diagnostics after charge additions brought units back to factory conditions. Results are shown in Table 16. This analysis illustrates that diagnostics can cause technicians to add or remove charge for units already properly charged. It also indicates that technicians will have difficulty knowing when to stop adjusting charge for units with deficiencies. The false alarm rate was lowest for manufacture protocols and the manufacturer and generic program protocols provide valuable information on other faults besides charge amount.

Table 16: Comparison of FDD Methods for Recovery and Recharge Sample – Factory Charge Conditions

		Diagnostic indicates False Alarm (FA) at Factory Charge (Diagnostic Failed) - Partial means diagnostic indicates other issues		
Unit #	Compressor #	2008 Title-24 / 06-08 Program (SH or SC)	Generic Program FDD (SH, SC, COA, EST)	Manufacturer FDD (DP, SP, AT, LT, ST, SH. SC)
Unit 1	Circuit 1	False Alarm	Partial Pass	Partial Pass
Unit 1	Circuit 2	False Alarm	Partial Pass	Partial Pass
Unit 2	Circuit 3	False Alarm	False Alarm	Partial Pass
Unit 2	Circuit 4	False Alarm	False Alarm	Partial Pass
Unit 3	Circuit 5	False Alarm	False Alarm	False Alarm
Unit 4	Circuit 6	Pass	Pass	Pass
Unit 5	Circuit 7	Pass	Pass	Pass
Unit 6	Circuit 8	NA	NA	NA
Unit 7	Circuit 9	False Alarm	Partial Pass	Pass
Unit 8	Circuit 10	Pass	Pass	Pass
		False Alarm %	**False Alarm %**	**False Alarm %**
		67%	**44%**	**11%**
		Pass %	**Pass %**	**Pass %**
		33%	**22%**	**44%**
		Partial Pass %	**Partial Pass %**	**Partial Pass %**
		0%	**33%**	**44%**

The evaluation team found problems both with the recommended refrigerant charge (RCA) fault detection procedures and with incorrect implementation of the procedure, which led to a large number of undiagnosed faults. Refrigerant charge FDD methods seek to assess the amount of refrigerant in the system indirectly through measurements of operating pressures and temperatures. The Title-24 approach could not be guaranteed to determine correct refrigerant charge faults under ideal conditions (Braun[31,32]) and were more problematic when applied to commercial field conditions. In the sample examined as part of this effort, the Title-24 approach correctly assessed charge in only 1/3 of the cases. Additional common faults, which also affect

[31] Braun, Jim, David Yuill, and Howard Cheung. *A Method for Evaluating Diagnostic Protocols for Packaged Air Conditioning Equipment*. Purdue University and the California Energy Commission, 2012.

[32]Braun, Jim, and David Yuill. Evaluation of the Effectiveness of Currently Utilized Diagnostic Protocols. Purdue University. 2014.

the pressures and temperatures measured (e.g., contaminants, especially non-condensables and line restrictions), are not addressed by the Title-24 methodology and lead to incorrect charge adjustment when not identified. This section focuses on the third-party and local programs. Similar issues existed in the statewide programs, but separate installation rates were not developed.

After third-party and local program service and evaluation, the team determined that:

- Program refrigerant charge service achieved full benefits, partial benefits, and negative benefits based on the weighed-out charge and reported adjustments.
- Linking FDD values to benefits is difficult. This means program and typical verification FDD methods frequently fail in assessing potential savings or achievement of factory charge. FDD protocols are inherently pass/fail, making it difficult to quantify the benefits of incremental but suboptimal improvements.
- Manufacturer protocols perform better than the Title 24 or generic program FDD protocols in determining whether a unit has achieved factory charge
- Linking results of weigh-out/weigh in charge testing to benefits is plausible if correct charge adjustment is known. However, correctly adjusting charge based on common protocols remains challenging.
- The Title-24 protocol only looks at superheat for non-TXV units and subcooling for TXV units and cannot diagnose faults with high/low heat transfer, pressure, restrictions, or non-condensables

2.4.3 Installation Rates for Third-Party and Local Programs

The CPUC defines installation rate as measures being installed and capable of saving energy as intended. Initially, the study estimated installation rates by sampling units completed by the contractors in each IOU third-party and local program that comprised 70% or more of combined savings. The samples were relatively small compared to the size of the programs since they were not the initial focus of the evaluation. The evaluation team did not use the diagnostic assessment to evaluate the installation rate of refrigerant charge measures based on the discussion in Section 2.4.2. The installation rate for economizers is based on the diagnostic tests presented in this section. Other measures are included in the report tables below but no installation rate adjustments were made for non-economizer measures.

The evaluation team found that a majority of economizers were not functional after servicing, and that a number of other issues were not addressed by the maintenance effort. The evaluation team also considered that some economizer repair measures had an effective useful life (EUL) claimed of three years and inspections for some sites occurred close to that timeframe. The expectation would still be that half of the economizers would still be working based on the EUL definition. Table 17 and Table 18 show the results for all observed measures for PG&E AirCare

Plus and SDG&E local programs. Overall, only 5% of economizer measures in the PG&E AirCare Plus sample and none in the SDG&E Local Program sample passed the post diagnostic assessment. Considering the EUL of the measures and timing of inspections the final economizer installation rate assumes that an average of 20% of the measures would have failed at the time of the inspection. Therefore the installation rate was estimated to be 25% for AirCarePlus and 20% for SDG&E Local program. The pass and fail criteria were based on unit-specific manufacturer protocols. Some additional criteria were required for specific measures.

- Coil cleaning was assumed to pass unless the cleaning method was directly observed since use of water versus manufacturer specified methods could not be determined in post-only inspections.
- The refrigerant charge pass/fail has issues as this pass rate does not account for the partial pass outcomes. This issue is discussed in greater detail relative to developing installation rates in Section 2.3.6.

The majority of economizers with reported repairs were not found to be performing their primary function of changing damper position in response to temperature. In some cases, the program technicians opened the position of the dampers, but did not observe other changes to sensors, linkages, or control settings. In one program, new sensors and new controllers were installed by the program technicians but were not set up such that the controller or sensor actually changed the damper position. Additionally, although most systems required economizer repairs, none of the technicians performed them in direct observations, and they were the most common fault found in post-maintenance visits. In addition to the limitations of the fault detection diagnostics (FDD) applied in the CQM programs, the evaluation team believes that un-diagnosed economizer faults and inherent leakage of economizers present an additional challenge and future opportunity.

Table 17: PG&E AirCare Plus Measure Observations

Measure	Sample	Diagnostic Pass Rate
Basic Package	33	100%
Refrigerant Charge and Airflow Service	27	52%
Functional Economizer Test and Repair	33	6%
Economizer Adjustment	33	0%
Economizer Control Package	19	11%
Programmable T-stat Replacement	11	100%
Programmable T-stat Adjustment	25	44%

Table 18: SDG&E Local Program Measure Observations

Measure	Sample	Diagnostic Pass Rate
Condenser Coil Clean	39	100%
Refrigerant Test	39	100%
Refrigerant Service	37	35%
Economizer Repair	8	0%
Replace Economizer Controller/Sensor	2	0%
Evaporator Coil Clean	23	100%
Comb Condenser Coil Fins	2	100%

Further study is ongoing on whether the program criteria and other methods can reliably be used to diagnosis and adjust refrigerant charge as described in the preceding section. See Section Appendix B for detailed summaries of all units observed.

The evaluation team applied the installation rates for refrigerant charge and economizer measures to the ex ante savings reported by the local programs. Similar data were collected for statewide programs but an additional adjustment was not made to those estimates which were based on pre/post measurements across a package of measures. Table 19 shows the adjusted gross savings for refrigerant charge after applying the installation rates developed in Table 14. It should be emphasized that the savings result occurred somewhat randomly given the poor performance of the diagnostics in assessing factory charge. Determining the installation rate for charge required program tracking of charge adjustment amounts and the recovery and re-charge method of EM&V. Both programs used proprietary refrigerant system analyzers. Individual units were found to be above and below factory charge after service in the weighed charge sample. The installation rate may change dramatically in future evaluations given the wide variation in estimated actual benefits from charge adjustment. The evaluation team recommends using a more conservative installation rate for future programs if the recovery and recharge cannot be performed due to budget constraints. Additional EM&V using this method is strongly recommended for 2013-14 evaluation given the small sample available to date.

Table 19: Final Installation Rate and Adjusted Gross Savings for Refrigerant Charge Measures (2010-12)

IOU	Program Name	Ex Ante Energy Savings (kWh)	Ex Ante Demand Savings (kW)	Install Rate	Installed Energy Savings (kWh)	Installed Demand Savings (kW)
PG&E	Air Care Plus	2,339,072	0	81%	1,894,648	0
SDG&E	Non-Res HVAC Tune-up/Quality Installation	2,390,079	1	77%	1,840,361	1
TOTAL		4,729,151	1	79%	3,735,009	1

Table 20 shows the installation rate applied to economizer repair measures. The criterion for the installation rates was basic functionality of the economizer and the short EUL for some economizer repairs was considered.

Table 20: Final Installation Rate and Adjusted Gross Savings for Economizer Measures (2010-12)

IOU	Program Name	Ex Ante Energy Savings (kWh)	Ex Ante Demand Savings (kW)	Install Rate	Installed Energy Savings (kWh)	Installed Demand Savings (kW)
PG&E	Air Care Plus	6,938,405	3,488	25%	1,734,601	872
SDG&E	Non-Res HVAC Tune-up	3,884,115	2,571	20%	776,823	514
TOTAL		10,822,520	6,059	23%	2,511,424	1,386

2.5 Additional Observations

The following additional observations were not used in determining energy savings or installation rates, but these findings are presented to suggest possible causes for the high uncertainty in savings and relatively low installation rate for economizer repairs. Observations of training sessions, review of training materials, and review of QC procedures was performed by master technicians who also performed direct observations of program service as well as post-maintenance inspections. Note that review and observations were not conducted of all trainings within all programs.

2.5.1 Technical Training Issues

In the post measure assessments, the evaluation team found a number of issues with systems related to program measures, especially systems where the economizer minimum outdoor air position was systematically increased, as well as lack of economizer functionality. The program workpapers intended to repair non-functioning economizers, but in the direct observations process, service technicians also opened economizers beyond their previous minimum positions. An SCE training video obtained through a data request directed contractors to increase minimum outdoor air position; this behavior was confirmed by field observations and in the pre-post data logger analysis.

Master technicians attended additional program training and reviewed training material and observed that the training for programs in 2010-12 did not include functional testing of economizers or repair processes. This is a likely root cause of the observed low installation rate for economizer repairs in the third-party and local programs. These findings on economizer repair issues are similar to other evaluation studies.[33]

Observations of training classes indicate that the 2010-12 statewide and local programs consistently provided training on how to enter data into the program database, but there was variation in procedures and information for technicians to diagnose and repair faults that lead to improved energy efficiency. Given that the program training observed varied, it is not surprising that the observed procedures used varied across technicians. In late 2013 and early

[33] Wang W, S Katipamula, Y Huang, and MR Brambley. 2013. "Energy savings and economics of advanced control strategies for packaged air conditioners with gas heat." Energy and Buildings 65:497-507. doi:10.1016/j.enbuild.2013.06.033

Katipamula S, and MR Brambley. 2005. "Methods for Fault Detection, Diagnostics and Prognostics for Building Systems - A Review Part I." HVAC & R Research 11(1):3-25.

California Energy Commission (CEC). 2003. Small HVAC Problems and Potential Savings Reports. 500-03-082-A-25. http://www.energy.ca.gov/2003publications/CEC-500-2003-082

2014, program contractors were contending that repairs of faults are not covered under their maintenance contracts, and need to be bid and implemented separately as "retro-commissioning." This interpretation is not consistent with workpaper assumptions for the statewide programs.

2.5.2 Technical Tool Issues

Approximately 92% of program technicians had issues with tools or procedures. None had proper tools to evaluate economizers or outdoor air damper position. Approximately 50% did not have EPA low-loss fittings on their refrigerant hoses. Those that did have low-loss fittings often did not purge hoses of air and water vapor prior to attaching to the system. Lack of low loss fittings and failure to purge hoses causes non-condensables or contaminants to enter the system when adding refrigerant or attaching hoses. The weight of non-condensables introduced through the use of non-EPA fittings was not measured in the field. Field tests indicated the presence of non-condensables, but it was not possible to quantify the amount. Current lab tests are attempting to quantify the impact of non-condensables introduced by attaching typical service equipment, but quantification may be difficult, even under laboratory conditions.

In almost all direct observations, program technicians did not use proper tools or procedures to measure relative humidity, economizer operation, or damper position. Manufacturer recommendations for coil cleaning were only followed in one direct observation. Most technicians did not install fan belts with proper tension or alignment, which led to reduced airflow, efficiency, and premature failure.

2.6 Conclusions and Recommendations

2.6.1 Conclusions

Ultimately, we determined that the overall results were zero but highly uncertain realization rates and total program savings for 2010-12 statewide CQM program. This result is applicable to SCE claims only, since SDG&E withdrew claims from the final tracking data and PG&E sites were not included in the study sample.[34] The evaluation team reviewed all of the detailed data and observations and concluded that the underlying reasons for the result were that the current program has various inherent challenges, many of which were likely unknown until the field and laboratory research under this study became available. The reasons center on the program

[34] PG&E claims under the statewide program were measure-based holdover claims from a previous cycle.

implementers changing economizer minimum outside air positions without fixing their functionality. There were negligible changes in unit and system efficiency from RCA measures, but significant improvement in unit operation was noted for thermostat and fan control measures. The lack of efficiency improvement likely stems from the program providing incentives for adding or removing refrigerant charge without diagnosing other faults such as low airflow and refrigerant restrictions or non-condensables.

Multiple HVAC faults can present the same diagnostic readings depending on the combination of faults and operating conditions. This reality means that it can be unknown which of the faults or which combination is present for a given diagnostic result, leading to mis-diagnosis and incorrect remediation when a service technician focuses on a single fault. The programs also assumed economizers would be repaired, but these repairs were often considered non-maintenance repairs or retro-commissioning measures by the program contractors and, as such, outside the scope of program services. Statewide program training for SCE in 2012 guided technicians to increase minimum damper positions rather than making non-functional economizers functional. The increase of minimum outside air may be the largest driver of increased energy consumption for the monitored units. For all programs, outdoor air intake at the unit makes diagnosing other faults such as airflow and refrigerant charge difficult, if not impossible through FDD.

The team also found very few economizer repairs working which led to a conservative estimate of 23% mean installation rates. For refrigerant charge, estimating installation was much more complicated than originally expected and only a small sample could be conducted which showed issues with FDD and a wide range of positive and negative benefits. The installation rates present the same underlying issues as discussed above. Laboratory tests results, which are presented in Section 5, indicate challenges associated with relying on refrigerant charge adjustments to improve efficiency. For economizers, the programs focus on installation of sensors or changing outside air minimum position instead of testing and repairing economizer functionality. Logically, reducing minimum damper position to code levels instead of increasing minimum position based on rules-of-thumb identified in program training might be a reliable measure to reduce energy consumption (i.e., easier to diagnose and repair).

These issues demonstrate some of the reasons why commercial HVAC maintenance programs have failed to achieve potential savings as documented in this and previous evaluations.[35]

---

[35] KEMA. 2010. Evaluation Measurement and Verification of the California Public Utilities Commission HVAC High Impact Measures and Specialized Commercial Contract Group Programs, 2006-2008 Program Year, Volume 1 and Volume 2. California Public Utilities Commission (CPUC). http://www.calmac.org/publications/Vol_1_HVAC_Spec_Comm_Report_02-10-10.pdf

Footnote 13 - Hunt et al. 2010

2.6.2 Economizer Findings Relative to CQM Programs

In 2010-12, the programs did not provide consistent training or incentives regarding economizer minimum outdoor air damper position to improve energy efficiency. Maintenance programs require retro-commissioning or test and balance (TAB) service prior to other repairs to address minimum outside air regardless of the whether new sensors or controllers are added. The primary functional components such as linkages and dampers need to be working as a prerequisite to adding or improving sensors and controllers. While programs provide incentives for repairing non-functional economizers, none of the technicians performed economizer repairs while being observed. If contractors view repairs as "retro-commissioning" not maintenance, then these measures could be replaced by others to reduce unnecessary outside air, such as sealing the junctions of economizers to units, or expanding the program to include both retro-commissioning and maintenance items.

Participating technicians in one program received incentives for 1,136 economizer tests in 2010-12. The IOUs paid a total of 438 repair incentives for wiring, damper motor, sensor/controller, and linkage repairs. Although it is not indicated in the tracking data, an individual unit likely required multiple repairs. Assuming all repaired units received new sensors, we estimate repairs were made on 148 units or 13% of those tested. In another program, 11 out of 16 economizers at one site were reported as repaired, but only 5 were still working one year later while 11 were not. These finding indicate that there are significant economizer savings available.

Replacement parts for old economizers with analog sensors and controllers are no longer manufactured or available from local distributors.[36] It can be more expensive to repair a broken economizer than to install a new one, and many repairs don't work. Field observations of one program found more than 70% of economizer repairs were not working one year later. These findings are similar to other evaluation studies.[37] New economizers have improved digital sensors and controllers with on-board diagnostics and are relatively easy to install.[38] However, a past study of newly constructed buildings in California with new economizers reported 64% failures after 2 to 3 years of operation.[39]

---

[36] The programs provide incentives of $1,125 to $1,485 to repair broken economizers.

[37] Pacific Northwest National Laboratory study found 62.5% of economizers not working at a university laboratory building and 100% not working at a hotel in San Francisco (EDR 2011). An evaluation in New England found 44% of economizers failed within 2 years (HEC 1993).

[38] See http://www.micrometl.com/economizers.aspx and http://beyondinnovation.honeywell.com/products/jade.

[39] California Energy Commission (CEC). 2003. Small HVAC Problems and Potential Savings Reports. 500-03-082-A-25. http://www.energy.ca.gov/2003publications/CEC-500-2003-082

The primary issue occurs when new controllers are installed on older economizers that have mechanical issues. The new sensors and control algorithms send signals to components that cannot function and respond. The 2010-12 programs offered incentives to decommission non-functional economizers. If economizer dampers are closed when units are decommissioned, laboratory tests summarized in chapter 5 indicate cooling efficiency can be improved by 5% to 140%, depending on the pre-existing damper position, while maintaining Title 24-required outdoor ventilation air quantities for many occupancies.

2.6.3 Recommendations

The recommendations provided in this report do not cover all of the changes between the 2010-12 and 2013-14 programs. These recommendations have not been tested for cost effectiveness and they are intended as areas that should be considered and if feasible, developed further, piloted, and tested for cost effectiveness.

To address these findings, we recommend that IOU programs coordinate to clearly define measures and ensure program workpaper assumptions match the implementation process of the measures. The terminology (maintenance, repair, retro-commissioning) is less important than defining the specific actions that take place and how those actions lead to energy savings. The IOU programs should focus on activities that will lead to energy and peak demand savings, rather than trying to address all repairs and actions recommended in Standard 180. A focus should also be placed on the one-time and ongoing maintenance activities that lead to more persistent savings. Other repairs and maintenance activities that improve comfort, indoor air quality, or unit reliability may increase energy consumption or peak demand, and should be accounted for if performed.

We recommend piloting approaches that have demonstrated measurable savings. For each approach with measureable savings, the program should provide a definition and examples of the effective energy efficiency measures. In 2010-12 some programs provided lists of tools that were acceptable while others did not, but these reference documents did not list the acceptable tool for each procedure. The measure definitions should include specific training requirements, approved tools based on accuracy requirements, protocols to perform and inspect the measure, and data collection requirements. We also recommend that incentives be provided for measures with the highest probable applicability, FDD repair rate, and savings. Measure recommendations include:

- **Seal Unit Leaks.** Sealing the junctions between economizers and unit cabinets may have large opportunities as they are not part of routine maintenance. Pilot efforts should measure the change in performance in the field and use reliable sealing methods such as properly rated metal taping and mastic. This recommendation requires coordination with certified test and balance (TAB) technicians to ensure minimum requirements for ventilation are met. It is not likely that economizer perimeter leakage is intended to help meet ventilation requirements. Economizer perimeter leakage will be further studied in the 2013-14 laboratory tests. The evaluation recognizes that any recommendation regarding reductions in outdoor air quantity raises liability concerns, but over-ventilation does waste energy. Some units are operated with fans that cycle on with compressors and may not meet ventilation requirements even when accounting for leakage. Outside air rates need to be evaluated and the leakage must be taken into account such that it is known and part of calculations.

- **Reconsider Generic Diagnostic-Based Refrigerant Charge Measures.** Identifying and adjusting units with incorrect charge proves to be extremely difficult on commercial units with outdoor air intakes and more difficult if multiple faults are present. A key issue becomes determining appropriate coil-entering sensor placement due to unmixed and temperature-stratified air on the inlet side of the coil, since the ideal placement changes with damper position. Evacuation and recharge is recommended as the best method to assess exact charge level. The evaluation team understands that this method may not be cost-feasible for program implementation, and thus the program recommendation is to follow unit specific manufacturer diagnostic procedures (see next bullet). We recommend that programs implementing this measure perform quality control verifications on a sample of completed jobs using recovery, weigh-out and re-charge.
- **Use Manufacturer Maintenance and Diagnostic Protocols or Compare Program Methods to them for Reference.** Manufacturer protocols provide important instructions on troubleshooting multiple faults for units including refrigerant charge, restrictions and non-condensables. These methods were found in field and laboratory testing to be more reliable than generic FDD protocols. Manufacturer protocols also may provide new methods for adjusting airflow based on static pressure measurements. Manufacturer protocols provide instructions on proper tensioning and alignment of fan belts. This recommendation does not preclude future FDD protocol development that provides better energy efficiency optimization. Any new methods should be compared to manufacturer methods to demonstrate they are better at achieving optimal unit performance. The accuracy of manufacturer and other FDD methods across ranges of airflow conditions and other faults continues to be a subject of laboratory testing.
- **Develop Criteria to Replace Rather than Repair Economizer**s. Repairing old economizers can be expensive and difficult due to limited availability of parts and the need to address old design issues. Economizer replacement offers an option that may overcome some of the fundamental economizer repair issues and should be explored. De-commissioning is not an energy efficiency measure and this recommendation calls for a replacement measure which would also entail installing Title 24 compliant controls with FDD. Unit age should also be considered in determining whether to repair or replace an economizer. Also note that integrating new economizer controls with older unit controls and motors presents an issue. It should be considered whether this

measure also requires changes to settings, upgrades to unit controls (thermostat or energy management systems), and/or upgrades to motors.

In the long term, HVAC manufacturers, industry associations, state licensing boards, and the U.S. Department of Energy (DOE) can help improve quality maintenance programs by supporting improved technician competency standards, FDD protocols, and service instrument standards. For utility programs in the short term, we recommend that additional field and laboratory tests and additional analyses be conducted to further research the energy saving potential of HVAC maintenance and the issues identified in this summary. The focus should be on efficiency and account for activities that may improve comfort and/or indoor air quality, but result in increased energy consumption. The program can adopt standards-based maintenance but workpapers should account for occupant comfort improvements at the expense of energy savings. Programs should focus on simple measures that can be reliably implemented and have demonstrated savings.

The evaluation also acknowledges that maintenance is a process that is unlike almost all other energy efficiency measures. This difference presents major measurement and evaluation challenges. Program workpapers assume changes lead to reduced HVAC demand and energy use. The statewide programs make those changes over a relatively long time period, while local and third part programs perform the measures as an event. The differences in program designs present different evaluation challenges. Measuring changes in unit operation over long periods associated with the state-wide program design limits the ability to provide timely feedback and introduces potential uncertainties associated with non-program issues. Estimating the impact of multiple measures installed at the same time within local programs affects the ability of evaluation to provide measure specific savings estimates. The recommendation for programs to pilot strategies also provides a framework for evaluation of specific measures or service actions in detail that are not easily accomplished in the evaluation of full scale programs.

The IOUs have made changes for 2013-14 programs and continue to receive input from the WHPA to further improve the programs. Feedback on drafts of this report has included stakeholder recommendations for the future programs and evaluations. Additional program strategies and EM&V Recommendations may include, but are not limited to the following:

- Continue ACCA 180 efforts as non-resource market transformation effort. Focus on measures and providing savings in resource programs to maximize savings. Measures that are part of routine maintenance such as changing air filters and cleaning coils with water could be part of non-resource training efforts to ensure they are done.
- Do more process evaluation. The project was designed as an impact evaluation and did not include a process evaluation. Some process work was done for the SCE program based on interim findings of WO32. This recommendation will be raised within the HVAC EM&V Project Coordination Group.

- Study new economizers and controllers that meet current standards to assess whether technicians are able to better understand the operation and setup. This effort would help to assess the previous recommendation of replacing entire economizer assemblies given efforts made to add automated fault detection in new economizer controllers and new written and hands on certification processes for technicians.
- Track and support industry-wide efforts. The Consortium for Energy Efficiency is currently supporting a combined system efficiency laboratory testing approach for these packaged units and economizers as an improved energy rating. AHRI has formed a committee to work on a combined test standard. This would provide a test method and metrics to fully evaluate the RTU with Economizer.
- There are plans for both field and laboratory efforts in the 2013-14 evaluation cycle. Given the uncertainties in field measurements, laboratory measurements are critical to providing empirical evidence of the potential benefits. Lab tests should be coordinated between the CPUC and IOUs and coordinated with field tests to establish realistic test conditions. Field measurements are also made to establish in-situ efficiency changes within the limitations of field measurement accuracy. Field limitations may limit some measurements to being acceptable for qualitative assessment, but not reliable enough for quantitative analysis.

3. Residential Quality Installation

This section presents the methodology and results of the residential quality installation program evaluation. The evaluation approach relied on field measurements of key parameters for a sample of participants and non-participants, focusing on HVAC system oversizing, airflow, and duct leakage. Field results were run through a DOE-2 prototype simulation model to develop annual estimates of energy use, and participant use was compared against non-participant use to determine savings.

The overall gross realization rates were 35% for electric energy savings and 38% for demand savings. These realization rates were driven by the finding that non-participant baselines for system sizing and duct leakage were better than those assumed in ex ante program calculations.

3.1 IOU Workpaper Assumptions for Quality Installation

Energy savings in the program workpaper are usage reductions from quality installation (QI) practices that exceed Title 24 requirements or common installation practices for attributes not addressed by code. The savings from the system replacement with a higher efficiency unit are relatively simple to determine and are documented in the DEER database. However, the savings from quality installation practices present unique challenges. Interactive effects between QI practices and installation of high-efficiency equipment further complicate the issue.

The IOU workpaper identified the basis for quality installation savings as:

- **Reduction in HVAC system oversizing**. The workpaper reflects the understanding that HVAC units are typically oversized by 20%. This is excessive and can lead to lowered efficiency and shortened useful life. Participants are assumed to be sized to the load while installations outside the program are assumed to be oversized 20%. The workpaper did not specify if the target size is determined via Manual J load calculations or Manual S system sizing calculations.
- **Airflow Correction**. The workpaper reflects the belief that fan settings are typically incorrect, resulting in airflow of only 350 CFM/ton assumed for units outside the program compared with the workpaper-recommended 400 CFM/ton assumed to be achieved by participants.
- **Duct leakage sealing**. Ducts are often not properly sealed (DEER 2005 Measure D03-458 was used as the basis for savings). Participants are assumed to have 12% leakage; with an assumed baseline leakage of 24% for installations outside the program.

3.2 Methodology Overview

The quality installation field assessments focused on residential systems since non-residential QI programs were not in operation in 2010-12. Onsite visits were conducted at non-participating homes identified using known installations from other general population studies such as the CLASS (California Lighting and Appliance Saturation Study) conducted under WO21. These non-participants were recruited from PG&E, SCE, and SDG&E service areas to provide a common-practice baseline. Because SCE was the only IOU with a residential quality installation (RQI) program, program participants could only be recruited from within the SCE service area. The team conducted site visits at 50 program participant and 50 non-participant sites, all of which were located in climate zones (CZs) 8 through 16. The significance of climate zones is explained in Section 3.3.1. Identical data-collection protocols were followed at participant and non-participant sites.

The program workpaper established aspects of energy code that, on average, are presumed to not be fully met outside the program. The aspects affect installed performance and included duct leakage, airflow, and sizing. The baseline for unit efficiency was a new minimum efficiency unit, not the existing system since the program assumed a replacement upon burnout of the existing system. The pre and post energy consumption is not equivalent to the savings from quality installation since the savings claims were not early replacements.

3.2.1 Participant Selection

The evaluation team sampled QI participant sites for extrapolation of the results to the program population. The evaluation team recruited participants proportional to contractor participation in the RQI program. The evaluation sampled the participant population at the site level; if more than one new unit existed at a given site; one new unit was randomly selected within the site for metering.

This program is subject to self-selection effects at two different levels, and at each level it is more likely that the self-selection effects will lead to increasing bias in the savings results. The two levels of potential self-selection effects leading to greater bias are: (1) participants seeking out contractors with good installation practices; and (2) contractors who already have good installation practices tending to self-select into the program. It is not certain that self-selection exists or that it operates in this particular manner, but the nature of the program is such that it has a degree of potential for self-selection, and that there is an unusually strong potential for self-selection to lead to more bias in the results given the particular quasi-experimental design used here. The results should be interpreted as most likely being an upper bound for net savings rather than a point estimate.

This study acknowledges that it may be difficult to compare participants with general non-participants since the customers and contractors differ. We think this is bias within the program, not within the evaluation. The question of what the customer would have done without the program is mitigated by selecting non-participants from similar geographic and socio-economic areas. This approach is good for program evaluation and not so good for estimating statewide compliance rates. The non-participant customers represent baseline conditions for 2010-12 participants, but likely do not represent statewide average conditions.

3.2.2 Non-Participant Selection

Non-participant installations were identified through WO54 and CLASS and characterized by climate zone, with 18 out of 50 installations coming from SCE service territory CZs 8-16. In WO54, a market assessment study, telephone surveys asked about recent HVAC system change-outs. The study approach was vulnerable to self-selection bias. The evaluation team could not quantify the extent of this bias and there were no other means to identify additional recent non-participant installations within the evaluation budget. CLASS, a random site survey that did not ask about HVAC change-outs during recruiting, had no similar self-selection bias. All available contacts were exhausted so there was no subsampling. Recruiters offered a $200 incentive and made up to seven attempts until acceptance or refusal. Non-participants were recruited to allow DNV GL field staff to perform onsite diagnostic measurements and observations to establish a common-practice baseline performance. The field team used the same tests, test protocols, and observations as site visits for QI participants. The aggregate non-participant performance was used as the baseline for each participant unit level analysis.

3.2.3 Onsite Data Collection

Residential QI programs seek to provide savings through right-sizing systems, testing and sealing ducts, adjusting fan airflow, and optimizing fan wattage draw. The evaluation team established data collection protocols to determine the effects of these measures. Data collection protocols were identical for program participants and non-participants. Data collection forms can be found in Appendix G and site level data can be found in Appendix C.

3.2.3.1 Demographic and Usage Data

Upon arrival at the site, the DNV GL staff interviewed the resident to gather information about the number of permanent residents, thermostat setpoints, hours of use (if not using a programmable thermostat), and the age of the residence. If a programmable thermostat was used, setpoints and hours were taken directly from the thermostat.

3.2.3.2 Temperature and Relative Humidity Measurements

Rather than waiting for the unit to reach steady state to take spot measurements, the strategy was to collect time series data that could be cleaned and averaged off-site with data loggers. Hobo Micro Stations were installed onsite to measure outdoor air, supply air and return air temperature and relative humidity. The outdoor measurements were made near the air conditioner condensing unit. The loggers were set up to record measurements at one-minute intervals throughout the test period. They were removed before the team left the site.

3.2.3.3 Airflow Testing

A unit's airflow was measured in cooling mode using The Energy Conservatory (TEC) TrueFlow meter kits. Whenever possible the TrueFlow plate was placed at the air handler to eliminate the effect of return duct leakage. There is a standard adjustment for return duct leakage of 4% when remote return grills are the only access point. When time permitted, after measuring the air flow in cooling mode, the team switched the unit to either fan only or heating mode depending on the effect on residents' comfort, allowed it to stabilize, and measured airflow again. The evaluation made additional measurements in each mode to assess whether cooling airflow was set differently in other modes, meaning there would be potential to easily increase flow through changes in fan settings. Static pressure was also measured.

3.2.3.4 Refrigerant Charge

The refrigerant liquid line and suction line temperatures and pressures were measured with the unit in cooling mode[40]. Superheat and subcooling were estimated, but this element was not included in the final analysis.

3.2.3.5 Power Measurements

For split systems, the cooling mode power and power factor for the entire condensing unit and the entire air-handler unit were measured with Fluke 345 clamp-on true RMS power meters. For package units, the cooling mode power and power factor for the entire unit were measured, and where possible, the supply fan alone.

3.2.3.6 Infiltration

TEC Minneapolis Blower Door kits were used to depressurize the houses. Where possible, the house was depressurized to 25 Pa and 50 Pa below ambient pressure. Some houses could not be depressurized to 50 Pa below ambient; in these cases the team used lower pairs of values such as 30 Pa and 15 Pa, or 24 Pa and 12 Pa, and used flow correction factors to obtain the infiltration in cubic feet per minute that would have resulted from the CFM25 and CFM50 test. CFM testing at two pressures allowed calculation of a proper flow exponent to better model the impact of infiltration on building loads.

---

[40] Similar high grade instrumentation as was used in 2006-08 evaluation, including industrial grade Crystal pressure gauges and platinum thermistor temperature sensors. CALMAC ID: CPU0028.01;

3.2.3.7 Duct Leakage

TEC Duct Blasters were used to measure duct leakage to outside (LTO), which is leakage outside the conditioned envelope, and total duct leakage. First, all supply air registers and all but one of the return air registers were taped off and sealed. Next, the Duct Blaster fan was then attached to the untaped return register and used to pressurize the conditioned space to 25 Pa above ambient pressure. Then the Duct Blaster fan was used to pressurize the duct system to 0 Pa relative to the conditioned space. With the conditioned space and duct system at equal pressure, there should be zero net flow between them; that is, there should be no leakage between ducts and conditioned space, thus, any air moving through the Duct Blaster was leaking outside the conditioned space. After noting the LTO at 25 Pa, the team repeated the test at 50Pa. As with infiltration testing, there were cases where the team used 30 Pa and 15 Pa or 24 Pa and 12 Pa because they were unable to pressurize the space to 50 Pa. For total leakage, the blower door was removed and the duct system was tested at 25 Pa and 50 Pa with respect to the conditioned space.

3.2.3.8 Building Shell Characteristics

Site and building information sufficient to allow calculation of heating and cooling loads was gathered. This included creating floor plan sketches; determining wall, floor, and ceiling construction types and dimensions; recording attic and crawlspace insulation types and thicknesses; observing window types and sizes; and building orientation.

3.2.3.9 Notes on Data Collection

Other than the data the team collected onsite, there was no consistent information about what scope of installation work was done or how equipment was selected for non-participant sites. It was unknown if the installations included other changes to duct systems or if furnaces were included in air conditioning replacements.

3.3 Results

3.3.1 Title 24 Testing Requirements

HVAC replacements within California must comply with relevant sections of Title 24, Part 6, which define the baseline regulatory codes for the program.[41] However, due to low code

[41] California Energy Commission, *2008 Building Energy Efficiency Standards: Residential Compliance Manual,* August 2009. Section 8.4

compliance rates observed in other studies[42], the IOU workpaper established a "common practice" baseline for this study. In addition, baseline assumptions used in the QI program workpapers submitted by IOUs were compared with field results of non-participants from this study. Participants and non-participants alike are required to meet Title 24 requirements for replacements, which are described in the following section.

3.3.1.1 Requirements

Units located in climate zones in hotter areas of the state (CZ 2 and 8 through 16) need more cooling than units located elsewhere. Since HVAC systems in these zones operate more of the time and use more electricity to provide more cooling, it is important that these units operate as efficiently as possible. As a result, Title 24 imposes more stringent requirements on systems operating within these zones. The individual requirements impose specific climate zone restrictions. These include duct testing and sealing, duct insulation, and refrigerant charge verification. However, for new systems serving newly constructed homes or additions to existing homes, 2008 Title 24 duct sealing requirements apply to homes in all zones. Per *§151(f)10* such new construction must be shown by a HERS rater to leak less than 6% of system fan airflow.[43]

Per *§152(b)1E*, in zones 2 and 9 through 16, if a major component of the HVAC system (air handler, outdoor condensing unit, cooling or heating coil, or furnace heat exchanger) is replaced or installed, the duct system must meet 2008 insulation and air distribution requirements.[44] Per *§152(b)1D,*[45] if this replacement or installation requires the replacement or addition of more than 40 feet of ductwork in unconditioned space, the ducts must also be tested and sealed so that leakage is no more than 15% of nominal system airflow.[46] Because this is an evaluation of QI measures, every site included installation or replacement of at least one major component and is subject to this requirement. However, if the entire duct system is replaced, it must meet the same 6% leakage requirement as new construction.

In zones 2 and 8-16, refrigerant charge testing is required for new systems and when any major component of the HVAC system is replaced. The QI program workpaper further describe aspects

---

[42] G. William Pennington, *"Underground Economy: Contractors Failure to Pull Permits for Residential HVAC Replacements"* Testimony to the Little Hoover Commission, 27 March 2014 pg.1-17

[43] Op. cit.

[44] Op. cit.

[45] Op. cit.

[46] If the rater or installing contractor tries but is unable to reduce leakage to 15%, then any of these options can be used: (1) contractor or HERS rater must verify that leakage to the outside is less than 10% of system fan airflow; (2) contractor or HERS rater must verify that leakage is reduced by more than 60% compared to before the alteration and a smoke test shows that all accessible leaks have been sealed; or (3) HERS rater must verify that all accessible leaks have been sealed.

where participants exceed Title 24 minimums and assumptions are made about non-participant systems. These assumptions are described in Section 3.1.

Load calculations are required for all replacements, but evidence of the calculations was not available for non-participants.

3.3.2 Field Findings on System Sizing

Data collected onsite informed the development of an ACCA Manual J-based system-sizing model for all participants and non-participants.[47] The primary analysis compared the calculated size to the installed tonnage to determine the amount of over- or under-sizing. Residential-sized units are typically available in half-ton increments, so the analysis looked at distributions of units sized within half ton increments higher or lower than calculated. Units sized at 4.5 tons are rare, but the actual sizing ratio was used in the analysis and the bins were created for display purposes. The QI programs require the use of both Manual J and Manual S[48] for equipment sizing; therefore participants' project folders contained a calculation in accordance with Manual J and Manual S. Participant sampled sites averaged 1,880 square feet of the floor area served by the tested unit. Non-participants averaged 1,770 square feet of the floor area served by the tested unit.

The data in Table 21 show oversized and undersized units in both the participant and non-participant samples. Both groups tended to have oversized units with a small difference in mean sizing ratio, but non-participants had a wide distribution with more cases of significant oversizing. The evaluation used program approved Manual J software in the analysis.[49]

47 ACCA Manual J is a standard for producing air conditioning and heating load calculations for single-family homes, small multi-unit residential structures, condominiums, town houses, and manufactured homes. More information is available from the Air Conditioning Contractors Association (www.acca.org)

48 ACCA Manual S provides sizing requirements for cooling and heating equipment, allowing the selection of equipment based on sensible and latent loads and ensuring the selected equipment will be properly matched to the local climate.

49 The team used Rhavc from Elite Software for the Manual J calculations.

Table 21: Preliminary System Sizing based on Manual J Calculations

Bin	Participants	Non-Participants
Oversized > 1 ton	2	15
Oversized 0.5-0.9 ton	13	12
Sized within 0.5 ton	33	15
Undersized 0.5-0.9 ton	2	3
Undersized > 1 ton	0	5
Average Manual J tons	3.31	3.09
Average Installed tons	3.64	3.54
Average Manual J / Installed	0.91	0.87
Percent of sample sized within 0.5 ton	66%	30%

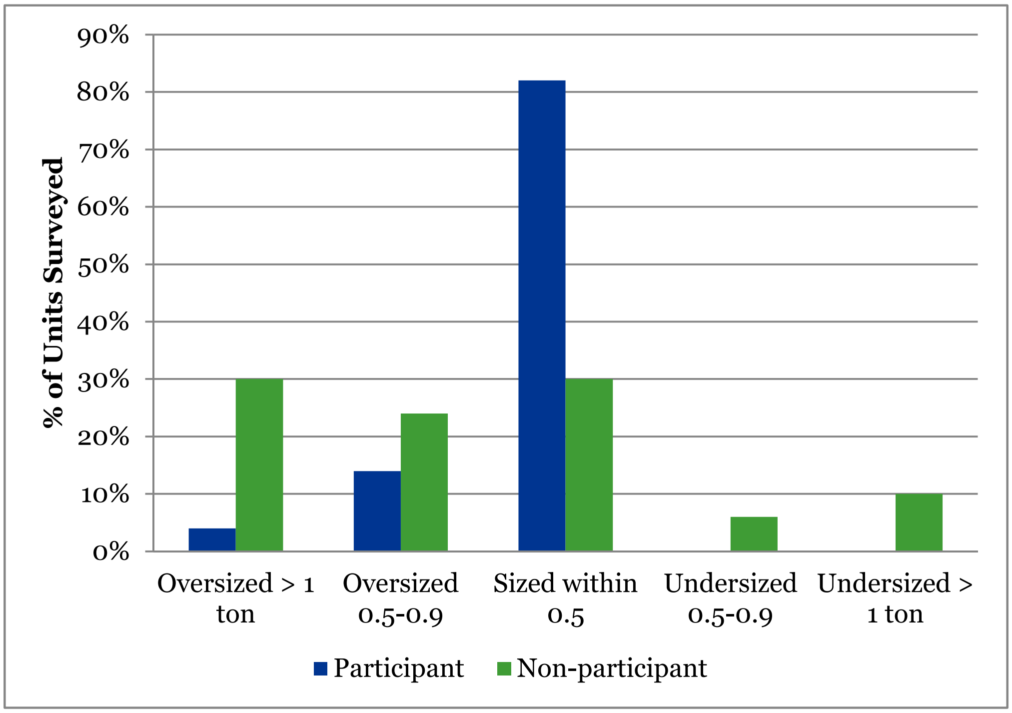
The evaluation team applied Manual S site-specific calculations provided by the participants' contractors. The distribution of participant sizing ratios varied, but participant and non-participant sizing ratios were fairly similar. Consequently, the non-participant sizing ratio remained at 0.87 while participants, taking advantage of Manual S sizing, had ratio of 0.90 (Table 22). The average target capacity using Manual S changed slightly from 3.31 tons for participant Manual J to 3.28 tons for participant Manual S. The difference between participants and non-participants is not statistically significant using Manual J or Manual S as the program sizing target.

Table 22: System Sizing based on Manual S Total Capacity Calculations

Bin	Participants	Non-Participants
Oversized > 1 ton	2	15
Oversized 0.5-0.9 ton	7	12
Sized within 0.5 ton	41	15
Undersized 0.5-0.9 ton	0	3
Undersized > 1 ton	0	5
Average Target Size -Manual S or Manual J, tons	3.28	3.09
Average Installed, tons	3.64	3.54
Average Manual S/Installed	0.90	0.87
Percent of sample sized within 0.5 ton	82%	30%

Non-participant results indicate that the oversizing ratio observed in the sample was 13% compared with the 20% workpaper assumption. It should also be noted that about 30% of non-participant systems were oversized by one ton or more and 10% were undersized by one ton or more. The QI program appears to eliminate extreme cases of improper sizing because achieving increased savings requires participant oversizing to be eliminated and undersizing to be allowed where possible. Figure 2 provides the sizing distribution of units surveyed.

Figure 2: System Sizing based on Manual S Total Capacity Calculations



3.3.3 Field Findings on Duct Leakage

All participants were required to undergo duct testing and sealing, and most of the non-participant systems were located in climate zones where Title 24 requires duct testing and sealing for system replacements. Five non-participants were in climate zones where duct testing and sealing were not required; these sites were excluded from the duct leakage findings. Both participant and non-participant duct leakage could be compared with Title 24 requirements and the workpaper assumptions on duct leakage. Data for the excluded sites is included in Appendix C.

If the workpaper assumptions were correct, non-participants would have an average of 24% total leakage relative to nominal airflow and participants would have 12%. Most of the systems sampled did not significantly extend or replace ducts, so the 15% leakage requirement from Title 24 generally applied; however, some systems had new or significantly renovated ducts and were required to meet the 6% duct leakage requirement. Table 23 shows that the mean duct leakage of participant units was 11.5% and non-participant units were 16.6%. The difference in total duct leakage between the two groups was statistically significant at the 90% confidence interval.[50]

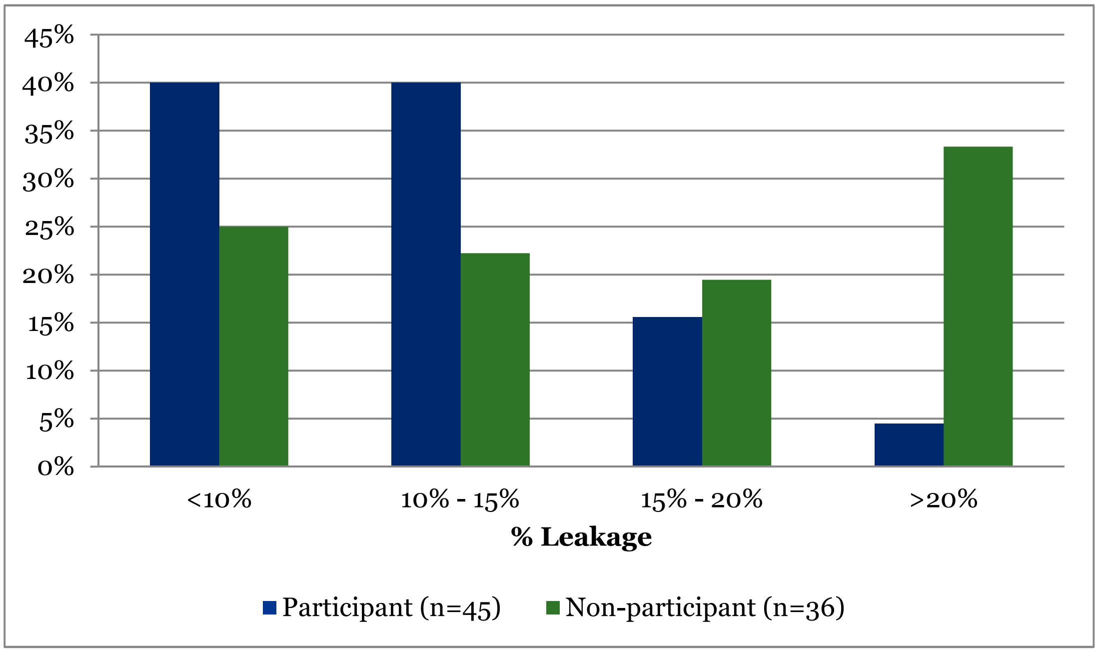
Table 23: Total Duct Leakage for Recent Residential Installations

Group	Sample Size	Mean Duct Leakage	Standard Deviation	90% CI Error Bound	Relative Precision
Participants	45	11.5%	0.045	± 0.011	± 9%
Non-Participants	36	16.6%	0.085	± 0.023	± 14%

Figure 3 illustrates the percentage of total duct leakage for recent participant and non-participant installations. The distribution shows that there were participant systems that had leakage greater than 15% which exceeds program requirements. While the average shown in Table 23 meets the workpaper assumption, eliminating cases of leakage greater than program requirements will increase savings. The workpaper assumed non-participant ducts would be twice as leaky on average at 24%, but the average was 16.6%. More importantly almost half of the tested systems had leakage meeting program requirements of 15% or less. Also note that the non-participant test results were limited to locations where Title 24 required duct leakage less than 15%.

[50] Given the research design and the use of standard practice (rather than Title 24 energy code) as baseline in the workpaper, the evaluation team was not able to separate gross and net savings; that is, the team was unable to separate free-ridership from presumed non-compliance in the gross workpaper baseline. If we assumed that participants would meet code requirements in the absence of the program we could estimate free-ridership, but we have no evidence that was the case nor could we get an unbiased estimate by asking that question directly because of participant self-selection bias.

Figure 3: Participant and Non-participant Total Duct Leakage



The total leakage above fed all additional calculations. The evaluation also measured the leakage outside the conditioned space (LTO-leakage to outside) relative to nominal unit airflow. Total duct leakage is the sum of leakage into conditioned spaces and leakage to outside of conditioned spaces. For example leaky ducts in an attic or crawlspace will experience leakage outside the conditioned space, while leaky ducts in a conditioned mechanical closet will not. Leakage to outside of conditioned space is wasted cooling. These LTO measurements reflect the actual delivered cooling and amount of leakage, which has a direct impact of energy lost. The LTO results provide additional information since the correlation between total leakage and leakage to outside is a fixed assumption in the prototype simulations.

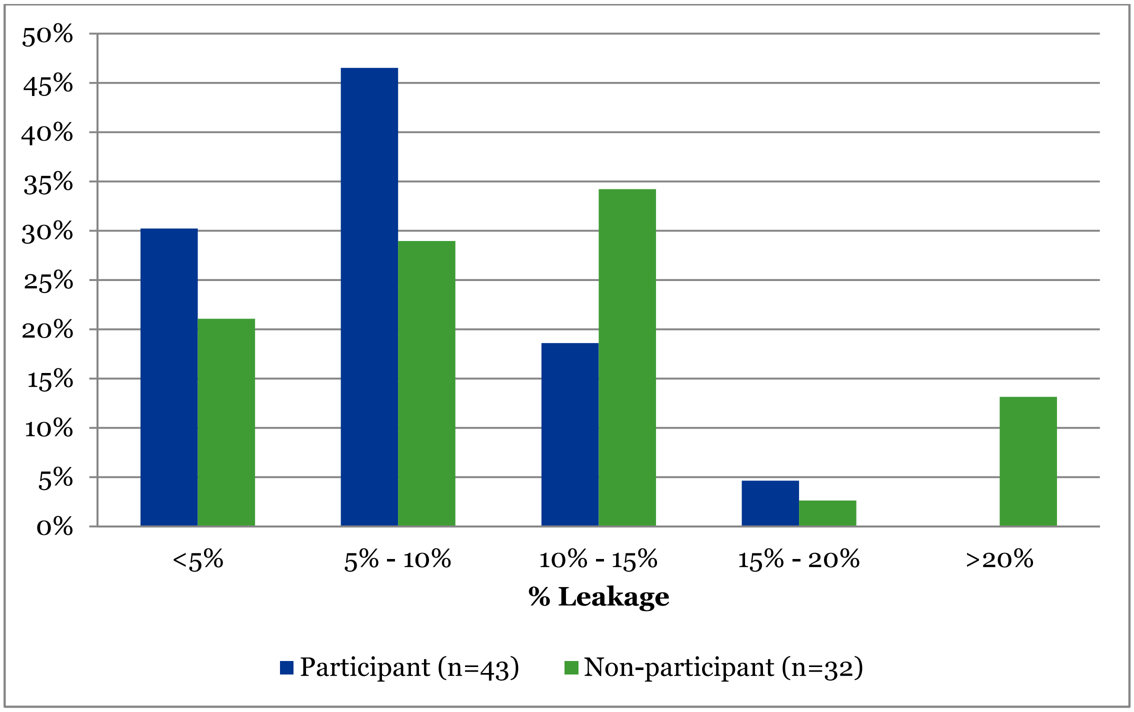
Table 24 shows program participants averaged 7.4% LTO, while non-participants averaged 10.7% LTO. Qualitatively, some non-participant sites had extremely leaky ducts while the spread in the amount of duct leakage for participants was much tighter, indicating the influence of the program on reducing duct leakage. The evaluation team found that the difference in LTO between the two groups is statistically significant at the 90% confidence interval.

Table 24: Duct Leakage to Outside for Recent Residential Installations

Group	Sample Size	Mean LTO	Standard Deviation	90% CI Error Bound	Relative Precision
Participants	43	7.42%	0.037	± 0.009	± 13%
Non-Participants	32	10.73%	0.067	± 0.019	± 18%

Figure 4 shows the distribution of LTO for participants and non-participants. LTO of 10% is a Title 24 threshold that can substitute for total leakage of 15%. The distribution of participant and non-participant units above and below 10% LTO were similar to the total leakage results.

Figure 4: Duct Leakage to Outside for Recent Residential Installations



3.3.4 Field Findings on Fan Airflow

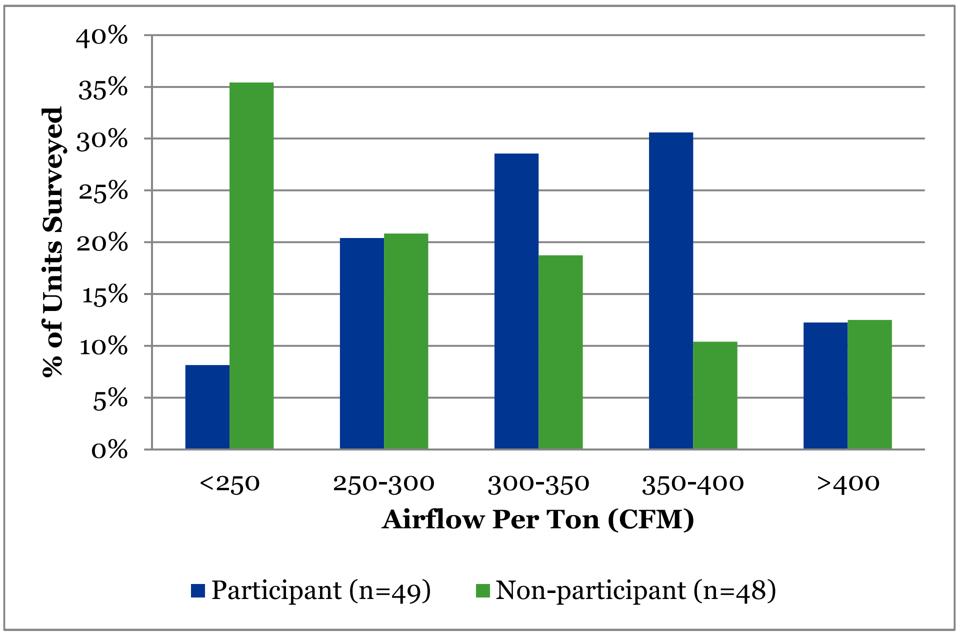
The QI workpapers assumed that non-participant units provide airflow of 350 CFM per ton and participant units provide 400 CFM per ton, which resulted in the claimed benefits. The evaluation used nominal cooling tons established by AHRI ratings for each unit. The collected data showed that the averages were closer to 300 CFM per ton for non-participants and 338 CFM per ton for participants. These values are within the 300–350 CFM/ton range for Title 24 compliance. The 10% difference between participant and non-participant airflow was similar to the workpaper assumptions. The participant units mean CFM/ton was 337.5 and the nonparticipant units' was 299.7 (Table 25). This difference is not statistically significant at the 90% confidence interval.

Table 25: Airflow per Installed Ton

Group	Sample Size	Mean CFM per Ton	Standard Deviation	90% CI Error Bound	Relative Precision
Participants	49	337.5	68.37	± 16.07	± 5%
Non-Participants	48	299.7	85.74	± 20.36	± 7%

Figure 5 shows the distribution of participant and nonparticipant unit airflows. The most important finding was that just over a third of non-participant systems had very low airflow under 250 CFM/ton. There were participant systems with low airflow as well and improving the airflow of those systems would increase program savings.

Figure 5: Airflow per Installed Ton



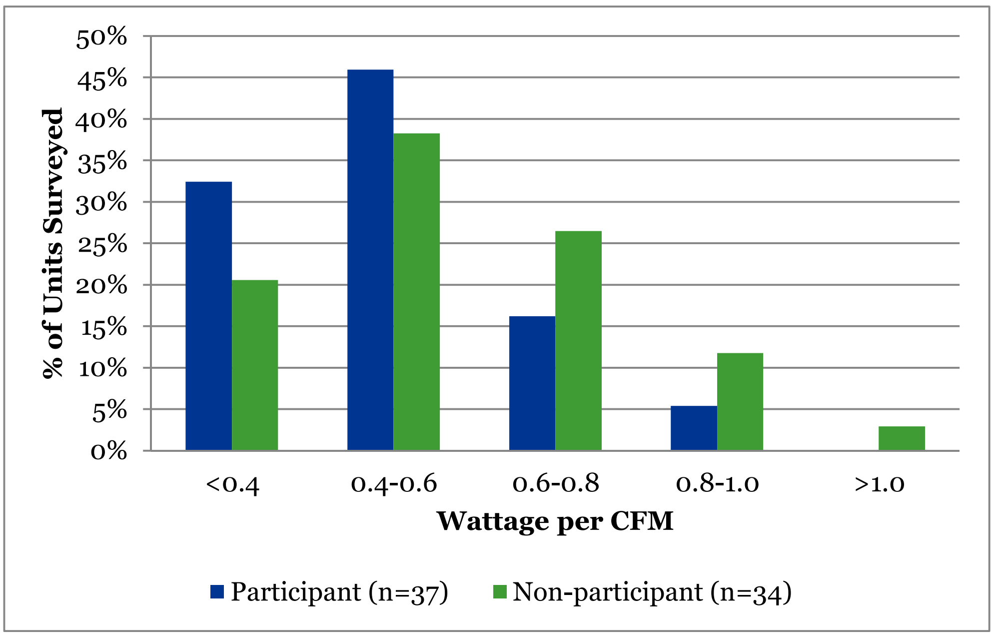
Perhaps even more important than airflow is the difference in the total fan power between participants and non-participants. If the high airflow associated with participant sites requires additional fan power then energy savings will be reduced or eliminated. Wherever possible, the field teams measured fan power in cooling and either heating or fan-only modes. Table 26 shows the mean power consumption per ton was lower for participants (0.486 W/CFM) than for non-participants 0.569 W/CFM); the difference is not statistically significant at the 90% confidence interval. Note that this difference may be partially due to the fact that QI participants also installed high efficiency units with more efficient fans, but this aspect was not studied as the focus was on the QI aspects not the unit efficiency and fan motor efficiency. Additional information on static pressure, fan settings, and design airflow were not part of the analysis, but these data were collected may be useful for additional future analyses (Appendix C). Table 26 lists the average watts per CFM for participants and non-participants.

Table 26: Fan Watt Draw Relative to Measured Airflow

Group	Sample Size	Mean Watts per CFM	Standard Deviation	90% CI Error Bound	Relative Precision
Participants	37	0.486	0.171	± 0.046	± 10%
Non-Participants	34	0.569	0.211	± 0.060	± 10%

The fan watt draw relative to airflow for both groups is shown in Figure 6. The distribution of fan power relative to airflow was influenced by the lower measured airflows for non-participants.

Figure 6: Fan Watt Draw Relative to Measured Airflow



3.4 Measure Level Analysis (UES)

The gross unit energy savings estimates for RQI measures were developed through simulations to extrapolate and normalize observed savings to first year annual savings. The evaluation focused on data collection using participant and non-participant studies for RQI that focused on the efficacy of the program measures. Onsite data collection activities were combined with calibrated simulation modeling to estimate impacts. This approach supports the evaluation and measurement of residential HVAC systems in the current cycle; it also provides field data to improve ex-ante estimates used in future program cycles.

3.4.1 Simulation

As described in detail below in Section 3.5, the analysis team used the test results and building characteristics gathered from each non-participant site to generate building heating and cooling load calculations at the level of the space served by that system. These test results included spot measurements of HVAC unit input power; supply, return, and outdoor temperatures; airflow rates; and duct leakage rates. The results were then input into the same single-family DEER eQuest prototypes used in the program workpaper, and results were developed for all measure categories, climate zones, and vintages and then re-applied to tracking data that listed the square footage served by each unit in the program population[51]. We note the DEER simulation models used for both the 2010-12 workpaper and this evaluation were outdated[52], but properly re-running the analyses with an updated prototype was not within the scope of the evaluation.

3.5 Rhvac Sizing Calculations for Non-Participants

3.5.1 Software

For non-participant sites, design cooling loads were calculated with Rhvac Version 9.01.129.[53] For participant sites, the QI program required contractors to submit models in their choice of Rhvac or Right-Suite.[54] The analysis team reviewed each of the participant models and confirmed that the input values were consistent with the data the evaluation collected at that site and the non-participant modeling process.

3.5.2 General Project Data Section

The General Project section of Rhvac stores information on the ambient design conditions as well as certain ductwork characteristics. Ambient conditions were modeled by selecting a reference city based on geographic proximity to the site or, in the absence of a nearby reference city, on the most climatically similar reference city[55] available. Within the General Project Data Section, modelers selected duct roughness factors based on the observed predominant duct material (e.g., aluminum, fiberglass, steel), as well as minimum and maximum air velocity based

---

[51] Documentation of DEER models and assumptions as well as non-DEER workpapers are available at www.deersources.com

[52] Simulated residential energy consumption exceeded current RASS estimates.

[53] Rhvac is a Manual J residential heating/cooling load calculation application from Elite Software Development (www.elitesoft.com).

[54] Right-Suite is a Manual J residential heating/cooling load calculation application from WrightSoft (www.wrightsoft.com).

[55] The reference cities are taken from Table 1A, ACCA Manual J, 8th edition, version 2 and aligned with Title 24 reference cities in Table J2. A comparison of altitude deviations was not conducted

on the observed predominant duct type (rigid or flexible). . Specifying the duct material was done for completeness and does not affect the final results.

3.5.3 System Data Section

The System Data section of the Rhvac program contains information on the HVAC system as well as information on system design conditions. Data include thermostat setpoints for heating and cooling, supply air temperature in heating and cooling modes, and the building shell infiltration rate.

The winter (heating) and summer (cooling) thermostat set-points were entered from the observed heating and cooling schedule.[56] Where schedule information was not available, default-modeling setpoints were utilized: 68°F for winter and 80°F for summer.

The difference between the air entering the rooms and the inside design temperature of the system is the winter or summer leaving coil-room delta T value (in degrees). The inputs for these modeled values were taken from the difference in supply and return temperature logger readings during the cooling and heating mode operations. In two cases the supply and return temperature information were unavailable for the heating or cooling modes, the default values of 70°F for winter and 20°F for summer were used.

The building shell infiltration data were entered into the Rhvac single-point blower door method section. For this method, modelers input the wind shielding class[57], the number of stories in the building being modeled as well as a calculated flow exponent "n," and recorded flow rate for the blower door test at 50 Pa. The flow exponent was calculated using the results of the two blower door tests and the following equation:

$$n = \frac{\ln\left(\frac{CFM_{50}}{CFM_{25}}\right)}{\ln\left(\frac{Pa_{50}}{Pa_{25}}\right)} \quad \text{[Equation 3]}$$

In cases where insufficient data were collected to calculate the value of n, a default value of 0.65 was used. Using the input flow exponent, flow rate, and pressure differential for the blower door test at 50 Pa, Rhvac calculated an equivalent leakage area in square inches. For the four sites where blower door testing was not possible, the Manual J Air Changes per Hour (ACH) method

[56] The field teams often observed that the thermostat was being used as off/on switches ("I turn it on when I get hot and I turn it off when I cool down.") For those sites we used 68°F/80°F after confirming with the residents that these were reasonable values. For sites using thermostats as intended, there was little variation; these sites used 78°F -82°F setpoints for cooling.

[57] The default selection was "4 – substantial obstruction," unless there was a compelling case to use "3- more obstructions," or "5 – surrounded by structures."

was used as an alternative to model infiltration. For this method, the ACH rate was determined by selecting a wind shielding class range, conditioned floor area range, qualitative construction tightness, and quantity of fireplaces.

Input data for Rhvac's HVAC System Design section include airflow rates and detailed ductwork characteristics. When TrueFlow data on system air flow in cooling or heating modes were available, the modeler changed the default system air type from "auto" to "fixed" and entered the average recorded airflow rate for the cooling or heating mode. When TrueFlow data were not available, the default 'auto' system air type was used, and Rhvac calculated the required airflow rates. For ductwork characteristics, modelers selected the inclusion of sensible and latent cooling duct gains.

For the calculation of duct load factors, modelers input information on duct location, insulation R-values, duct surface area, and leakage rates. For both supply and return ducting, modelers entered the observed duct location, such as attic, crawlspace, garage, or conditioned space. The observed duct insulation R-values were entered into the duct load factor calculation window. When duct insulation was observed but the R-value was unknown, a default R-value of 4.2 was used. Duct surface area was modeled using the Rhvac duct surface area calculator, which employs the ASHRAE Standard 152 duct surface calculation method. This method calculated supply duct area as the conditioned floor area (in ft^2) multiplied by a factor of 0.27, and return duct area as the conditioned floor area (in ft^2) multiplied by the number of returns and factor of 0.05[58].

The duct leakage rate was calculated as the quotient of the flow rate (in CFM) result from the DuctBlaster duct leakage to the outside at 25 Pa and the Rhvac calculated total duct surface area to produce an overall ductwork leakage rate in CFM/ft^2. For one case when the calculation produced overall leakage rates that exceeded 0.35 CFM/ft^2, the highest Rhvac default leakage rates of 0.35 CFM/ft^2 supply and 0.7 CFM ft^2 return were input in place of the calculated leakage rate.

3.5.4 Room Data Section

The Room Data section of Rhvac covers information on the building dimensions, materials, insulation levels, fenestration, and internal loads. When possible, sites were modeled as one large rectangular room with representative length, width and height dimensions that create a

[58] If there were no remote returns (instrument showed return duct is 0% of total ducting): return count was 0 but modelers input a "Return Duct Surface Area" 20 ft2.

If airflow was tested at the returns, modelers input the number of test locations from the instrument.

If airflow was tested at the unit, modelers assumed 2 returns for a two-story home or a unit with capacity >= 4 tons; modelers assumed 1 return for a single-story building or a unit with capacity < 4 tons.

simplified building shell. Heat transfer surfaces such as floors, walls, and roofs/ceilings were modeled by selecting the appropriate types, materials, and insulation levels.

When modeling slab flooring, modelers selected the passive heat, heavy dry or light moist soil type.[59] If modeling a floor above a garage, modelers selected the appropriate material type and assumed R-13 insulation level.

For modeling of roofs and ceilings, modelers selected a vented attic (unless otherwise noted in the site observations) with the appropriate roofing material type and color. Modelers selected attic insulation R-values corresponding to the type and amount of insulation observed onsite. If the attic R-value was unknown, the modelers selected insulation values based on the vintage of the home and the historical Title 24 attic insulation R-value code requirements[60], which are presented in Table 27 below:

Table 27: Title 24 Attic Insulation R Values

Period	R-value
<1978	11
1978-1992	30
1992-Present	38

When modeling 2x4 framed walls, an R-13 insulation value was entered, while an R-19 insulation value was used for 2x6 framed walls. For shaded heat transfer surfaces such as partitions or garage walls, modelers entered -1 for the summer temperature difference, which calculates the appropriate temperature difference between a shaded exterior wall exposed to ambient air and the indoor conditioned space.

For modeling of fenestration, modelers selected the appropriate number of glass panes, frame material, and glazing treatment (low-E/ no low-E[61]). The following window characteristics were used for modeling all window fenestration except skylights: internal shading blinds 45 degree light color and 50% drawn, insect screens outside with 50% coverage, ground coefficient = 0.23 (green grass), and no overhangs or projection included in models.

For the modeling of doors, the appropriate door material (wood or metal) and type (solid or hallow core, polystyrene, etc.) were used. For doors recorded as metal insulated, modelers assumed polystyrene insulation.

---

[59] This option was the middle-of-the-road choice and refers to soil under the slab. No thickness info was available for concrete slabs at non-participant sites and floor-covering data were not captured.

[60] The team sourced historical attic insulation R-value code requirements from the California Energy Commission website: http://www.energy.ca.gov/title24/standards_archive/

[61] The presence of low-e treatment was determined using ETEKT + (AE1600) low-e detectors.

Internal equipment and occupant loads were also input into the room data section of Rhvac. For all modeled sites, internal equipment loads of 2,400 Btuh sensible heat and 350 Btuh latent heat[62] were used. Internal occupant loads were accounted for by entering the reported year-round occupancy into Rhvac, which then calculated occupant loads based on 230 Btuh of sensible heat and 200 Btuh of latent heat per occupant. The evaluation reviewed program assumptions in models which did include site variation as well as variation in detail based on which of the two approved software packages were used. The site collected data was used to develop assumptions for non-participants that were consistent with the level of detail and assumptions in program models.

3.6 Savings Based on Participant and Non-Participant Comparison

The savings per square foot from the program workpaper were applied to the square footage of sites in the program population to get ex ante savings in the tracking data. Table 28 shows the distribution of savings and units by measure category which indicates the efficiency level and equipment type. The methodology used to develop the inputs and savings results were described in previous sections.

Table 28: Ex Ante Savings for Quality Installation Measures Installed 2010-12

Quality Installation Measure Category	Ex Ante Energy Savings (kWh)	Ex Ante Demand Savings (kW)	Number of Units
13 SEER Air Conditioner Quality Installation DX Equipment	18,417	6	137
14 SEER Air Conditioner Quality Installation DX Equipment	169,180	80	1,045
14 SEER Heat Pump Quality Installation DX Equipment	10,081	4	39
15 SEER Air Conditioner Quality Installation DX Equipment	273,568	149	1,401
15 SEER Heat Pump Quality Installation DX Equipment	28,228	9	110
16 SEER Air Conditioner Quality Installation DX Equipment	411,269	227	2,152
Grand Total	**910,743**	**475**	**4,884**

[62] Lighting loads are entered separately from equipment loads in Rhvac. 2,400 Btuh was the default recommendation for appliance loads in the Manual J "speed sheet". Rhvac accounts for latent heat from human occupants; we included 350 Btuh to account for houseplants that were typically present.

The evaluation team took the HVAC test results collected for participants and non-participants and re-ran the DEER single-family residential prototypes consistent with the method used in the program workpaper.[63] The evaluation noted that the workpaper was not reviewed by the ex ante review process. During the simulation process it was clear that within workpaper calculations the total cooling capacity was adjusted in the DEER prototypes based on the 20% oversizing assumption.

Across measure categories the gross realization rates were reviewed at the zone level per square foot, consistent with the workpaper. The duct leakage, airflow, and sizing inputs modified in the workpaper were the only variables modified in the ex post analysis. The evaluation team reapplied the results by measure and climate zone. The table below describes the realization rates of each climate zone and is followed by ex post savings by measure category, which varied based on the climate zone mix of each measure. The precision of the input values are reported previously in this section while the energy savings after extrapolation through the prototype simulations would be difficult to assign a precision estimate. The results in Table 29 present only the comparison between ex ante and ex post means. The climate zone results follow the workpaper assumed mix of building vintages, which varies by climate zone. Another set of drivers of variation by climate zone are the effects of the sizing, airflow, and duct leakage inputs on the simulation models.

Table 29: Gross Realization Rates by Climate Zone

Climate Zone	Energy Savings Realization Rate (kWh)	Demand Savings Realization Rate (kW)
6	23%	61%
8	33%	23%
9	41%	77%
10	32%	27%
13	30%	58%
14	11%	24%
15	43%	54%
16	9%	35%

Table 30 shows the energy and demand savings for each measure description from the program tracking data. The workpaper used the same measure descriptions and the results reported are

---

[63] Energy savings due to system "right sizing" were calculated by modifying "right sized" measure runs of the single family DEER prototypes. The total cooling capacities (COOLING-CAPACITY) were multiplied by the ratio of the observed average installed system size to the average calculated Manual J system size for participants and non-participants (see Table 22). The non-participant ratio was applied to the "baseline" run while the participant ratio was applied to the "measure" run.

based on the volume of measures claimed and the realization rates were applied by measure and climate zone combination. The results show that many installations in the program were of high SEER units.

Table 30: Ex Post Electric Savings for Quality Installation Measures Installed 2010-12

Quality Installation Measure Category	Ex Ante Energy Savings (kWh)	Ex Post Energy Savings (kWh)	Ex Ante Demand Savings (kW)	Ex Post Demand Savings (kW)
13 SEER Air Conditioner Quality Installation DX Equipment	18,417	7,124	6	3.2
14 SEER Air Conditioner Quality Installation DX Equipment	169,180	59,898	80	37.0
14 SEER Heat Pump Quality Installation DX Equipment	10,081	3,229	4	1.3
15 SEER Air Conditioner Quality Installation DX Equipment	273,568	94,314	149	55.0
15 SEER Heat Pump Quality Installation DX Equipment	28,228	9,322	9	3.0
16 SEER Air Conditioner Quality Installation DX Equipment	411,269	144,617	227	81.9
Grand Total	**910,743**	**318,505**	**475**	**181**

The overall gross realization rates were 35% for electric energy savings and 38% for demand savings. The final realization rates are based on the final savings and similar to a weighted average of the realization rates by climate zone. These realization rates were driven by the better than assumed sizing and duct leakage for non-participant units.

The evaluation determined that the ex-ante estimates need to better define how the gross savings and the net to gross ratio work together. The evaluation focused on measuring the primary gross savings assumptions in the workpaper for participants and non-participants. The final ex-post savings result is likely a combination of a realization rate and a net-to-gross factor.[64] The evaluation did not estimate a net to gross ratio, but the workpaper value of 0.7 should be reviewed. It is important to note that it was not possible to identify free-ridership from presumed non-compliance in the gross workpaper baseline. Since the workpaper assumed non-compliance, it essentially presumed a standard practice baseline, not a code baseline. If it

[64] 2010-12 net implies net of freeriders – spillover is important for 2013-14 and the spillover effect of utility and CEC trainings on code compliance is an interesting topic to research in the future.

was assumed that participants would have met code requirements in the absence of the program it would have been possible to estimate free-ridership. The evaluation team believes that net-to-gross issues may be influencing the realization rate though it is not possible to determine the extent to which this is true.

3.7 Conclusions and Recommendations

The QI program did result in energy savings, but the realization rates were generally low. Table 31 summarizes the field findings compared to the workpaper assumptions.

Table 31: Comparison of Field Findings and Baseline Assumptions

Workpaper Assumption	Non-Participant Field Observations (mean)	Baseline Assumptions	Participant Field Observations (mean)	QI Efficient Case Assumptions	Non-Participant Participant Significant Difference
Total System Duct Leakage	17%	24%	12%	12%	Yes
System Airflow (CFM per ton)	300	350	338	400	No
System Oversizing	13%	20%	10%	0%	No

In particular, ex ante workpapers claimed higher duct leakage than found in the baseline group. Duct leakage—a major source of HVAC energy loss—was lower in participating households than in non-participating households by a statistically significant amount, with a mean total leakage of 11.5% of nominal airflow for participants versus 16.6% for non-participating homes.[65] While this meets the QI program goal of 12% total leakage, the average baseline leakage was lower than the 24% leakage value assumed in the workpaper, although some non-participants had very high duct leakage. This resulted in lower actual savings than expected.

There was not a statistically significant difference in per-ton airflow between participants and non-participants, nor was there a significant difference in per-CFM power draw between the two groups. The evaluation team modeled the mean differences to develop savings estimates as the collected data represented the best information available, since the workpaper relied on assumed values. The evaluation team measured greater airflow for participants, but the non-participant results had high variation contributing to the lack of statistical significance of the difference. Some non-participants had very low airflow. Power draw (in watts per CFM of

[65] This includes sampled non-participants only in those CZs where duct sealing is required (2 and 9-16).

measured airflow) was below the recommended maximum of 0.58W/CFM for participants and non-participants alike.

The evaluation team found that the QI program resulted in the majority of systems (41 out of 49) sized within one-half ton of the size recommended by Manual S calculations. This compares to only 15 out of 50 non-participant systems being "right-sized." However, the average sizing ratios were calculated to be similar as there were both under and oversized non-participant units, but there were primarily right sized and a few oversized participant units. The workpaper assumed an average amount of oversizing and used the sizing ratio to represent this effect in simulations. The evaluation team reports the number of over and undersized cases, but did not separate the effects of under and oversizing in the ex post calculations. Future studies could be conducted to separately assess the effects and impacts of under and oversizing.

The evaluation team has the following recommendations:

- The 2013 Title 24 code should be reviewed to determine opportunities to exceed code. The net to gross needs to be clearly defined since the gross baseline assumed is not code minimum, it is a common practice assumption below code. In the absence of the program a unit could be minimally compliant or could be at the common practice baseline. Using a common practice baseline requires specific evidence that code minimum is not appropriate and requires rethinking the use of the default net to gross value which usually applies to measures with code baseline.
- Consider revisions to workpapers to account for the fact that baseline are comprised of a range as opposed to a point estimate. Gross savings vary by climate zone and measure SEER level. In this case the net savings and baseline may also vary by code jurisdiction. This includes under and oversizing and a range of duct leakage and airflow after installation. The findings suggest that about half of the installations outside the program perform similar to the program participants and the other half perform poorly. Until larger studies are available the source of assumptions in workpapers should be clearly stated as well as whether baseline studies are needed to develop data to inform the inputs.
- Baseline system oversizing was closer to 13% than the assumed 20%. Manual S allows up to 140% sizing relative to Manual J which leads to participant systems that don't meet the workpaper assumption that all participant systems are sized to the load. Increasing savings due to sizing requires that participant oversizing be eliminated and undersizing be allowed where possible to create a significant difference with common practice and recommended code.
- Revise workpaper assumptions to reflect the findings with respect to baseline values in addition to reconsidering the meaning of the assumed net-to-gross ratio.
    - Baseline total duct leakage was around 16.6% rather than the assumed 24%. Participant total duct leakage was near the claimed 12%.

  - We found that, directionally, the non-participant airflow per ton of cooling was about 300 CFM, lower than the baseline value of 350 CFM assumed in the workpaper. The difference between participants and non-participants was not statistically significant, but since both groups had lower airflow than assumed there may be additional opportunity to capture airflow savings.
- Investigate program overlaps between Quality Install and the Energy Upgrade California Home Upgrade Program. Highest available efficiency installations made up a large portion of QI program participants and some of these installations may be appropriate to include in Energy Upgrade California Home Upgrade Advanced path jobs. The strategy could capture pre-retrofit deficiency savings and load reduction leading to smaller installed QI units.
- Expand non-participant sample to support evaluation of QI programs and "to code" pilots. Using a consistent data collection approach would allow expansion of the sample in any given climate zone, given the relatively small size of the sample in this study. The "to code" pilots will collect detailed compliance verification data. Comparing non-pilot and pilot jurisdictions may require larger samples of non-participants.
- Ultimately the programs can influence savings for actions that exceed Title 24 requirements if they exist. Participant systems did not all meet assumed values which could be improved by program inspections. Exceeding code will improve realization rates, but it is unknown if cost effective saving remain, especially relative to the new code. Common practice was worse than code in this study, but not as much as assumed in workpapers. The evaluation team recommends the IOUs explore a few areas where Title 24 can be exceeded or does not have requirements:
  - Explore downsizing to reduce peak demand. The program sizing aligned much better with Manual J load calculations than non-participants, but did not eliminate all cases of oversizing. Non-participants did not systematically oversize, but rather there was a wide range of sizing relative to load calculations. A focus on reducing the installed size compared to the unit that is being replaced would have peak demand benefits if the reduction in size can be documented.
  - Explore duct sealing to reach a lower target leakage, such as the 6% threshold that is set for new ducts systems.
  - Explore air handlers/furnaces, filters, and duct modifications that reduce pressure drop and improve fan system efficiency.
  - Explore whether all ANSI/ACCA 5 Quality Installation Standard elements in programs impact energy use and align with workpapers. The standard does not call out items specific to energy efficiency so the program workpapers should reference which items in the standard are tied to energy savings and assess whether some elements may have non-energy benefits and determine if they are energy neutral or add additional load.

- Determine if there are specific locations where common practice baseline is worse than estimated in this evaluation, potentially using AMI or other data[66]. The non-participant sample covered large areas and many code jurisdictions. Further study could determine specific areas where common practice is worst and target program activities toward those locations. This may mean leveraging local government partnership programs as opposed to statewide program models if it is determined that only specific areas have the greatest opportunity.
- Ensure that workpaper and evaluation modeling inputs produce HVAC energy estimates that are consistent with Residential Appliance Saturation Study (RASS)

[66] Comparative Use programs, such as Opower, should be able to define regions to focus on without violating confidentiality – maybe targeting specific counties, cities, or even ZIP codes

4. Upstream HVAC Program

4.1 Summary

The evaluation efforts for the upstream component of the study focused on the level of free-ridership in the Upstream HVAC program. Gross ex ante savings claims are based on DEER estimates and were not evaluated. A participant self-report method was used for the free-ridership/net-to-gross (NTG) work, and the analysis was based on in-depth telephone interviews with 19 out of 22 participating HVAC distributors that were conducted by DNV GL in 2013. The approach took into account the program's effect on both the stocking practices and sales practices of the distributors.

Overall, the program is achieving a savings-weighted NTG score of 0.80, versus an unweighted NTG score of 0.58. The reason for this difference is that the distributors who account for the largest share of program savings are also the distributors who claim the highest levels of program influence.

The evaluation team makes the following recommendations[67], gleaned from the distributor interviews and analysis of program data:

- The program should consider increasing rebate levels for higher efficiency equipment to encourage more sales in the highest efficiency tiers.
- Program-provided support for how to market high efficiency equipment support may be helpful to distributors, particularly those who are less successful at selling the concept of high efficiency.
- Providing a reservation system or rebate guarantee would encourage more participation, and increased high-efficiency sales, from distributors who have long sales cycles or custom build equipment.
- Future evaluation should consider data from the Market Effects and Commercial Market Share Tracking studies in addition to self-reported distributor information. The parallel 2010-12 studies did not include such an overarching analysis, but this is recommended for 2013-14.

In the remainder of this section, we describe the upstream program, report on the distributor surveys, present the NTG analysis and results, and provide conclusions and recommendations for the upstream program.

[67] The recommendations were not tested for cost effectiveness

4.2 The Upstream Program

From program years (PY) 2010 through 2012, the California IOUs funded an Upstream HVAC non-residential rebate program. The program implementation contractor is Energy Solutions and the California utility program partners include PG&E, SCE, SDG&E, and SMUD. The Upstream program has been in operation since 1998. More information can be found on Energy Solution's website http://energy-solution.com/index.php/case-studies/upstream-hvac.

The program has three primary goals:

- Encourage participants (distributors) to increase their stock of high-efficiency non-residential equipment in order to be readily available to customers (contractors).
- Encourage participants to up-sell equipment to customers (e.g., explaining to customers the technical benefits of the efficient option and calculating the payback or net present value when possible).
- Encourage the purchase and installation of the most efficient equipment available.

The Upstream HVAC program enlists HVAC distributors who are willing to participate under the program's terms and conditions to sell high efficient heating and cooling equipment for non-residential (commercial) use. Rebate amounts paid to distributors are based on performance tiers.[68] Tiers vary based on equipment type, capacity, and efficiency (SEER, EER, or IEER ratings). Distributors' customers are typically licensed HVAC contractors (C-20) or mechanical design engineers.

Equipment eligible for rebates must be installed in non-residential buildings within the sponsoring IOU's service territory and must meet program-specified efficiency requirements; distributors must provide information on the location of the installation. The most common rebated equipment includes:

- Three-phase package and split equipment (air-cooled and water-sourced heat pumps, water and evaporative-cooled AC)
- Single-phase equipment (air cooled)
- Single-phase ductless equipment (mini and multi-split equipment)

Distributors are also entitled to receive rebates for these less common equipment types:

- Three-phase air-cooled chiller equipment
- Three-phase water-cooled chiller equipment
- Three-phase variable-refrigerant-flow equipment (VRF heat pumps and heat recovery)

[68] See Appendix E for 2010-2012 SCE Qualifying Minimum Equipment Efficiencies & Incentive Levels for Commercial Air Conditioners

4.2.1 Utility Level Claimed Savings for 2010-12

PY 2010-12 energy savings claims for each IOU are detailed in

Table 32. The SDG&E savings were not classified in the original program ID associated with upstream incentives, but some measures in their Non-Res HVAC program were clearly labeled as a part of the Upstream HVAC program.

Table 32: Ex Ante Savings for Upstream Programs

Program ID	Program Name	2010-12 Savings Claims	
		Energy Savings (kWh)	Demand Savings (kW)
PGE21061	Upstream HVAC Equipment Incentive	17,705,130	9,100
SCE-SW-007A	Upstream HVAC Equipment Incentive	64,368,795	21,952
SDGE3147	Commercial Upstream Equipment	0	0
SDGE3161	Non-Res HVAC	560,854	203
TOTAL		82,634,780	31,255

4.3 Survey with Participating Distributors

During the two-year program cycle, 22 HVAC distributors received rebates for the sale of high-efficiency equipment, and DNV GL was able to interview 19 of these participating distributors. The distributor interview length was, on average, slightly longer than one hour.

The distributor survey was designed to address these objectives:

- **Characterize participating distributors** including distributor geographic territories, exclusive rights to serve in a region, number of manufacturers represented per distributor, quantity of equipment maintained in stock, equipment stock turnover rates (i.e., period of time before needing to replenish stock) and distributor market shares.
- **Estimate program free-ridership** for high-efficiency equipment sold and/or stocked by distributors.
- **Identify sales volume** by type, amount, and efficiency levels of equipment currently being sold.
- **Identify availability of a tracking database** for company sales and willingness to share data on non-program-qualified HVAC units.

The last two survey objectives (sales volume and tracking database) were part of a joint effort with the CPUC Market Effects Report (Work Order 54).

4.3.1 Characteristics of Participating Upstream Distributors

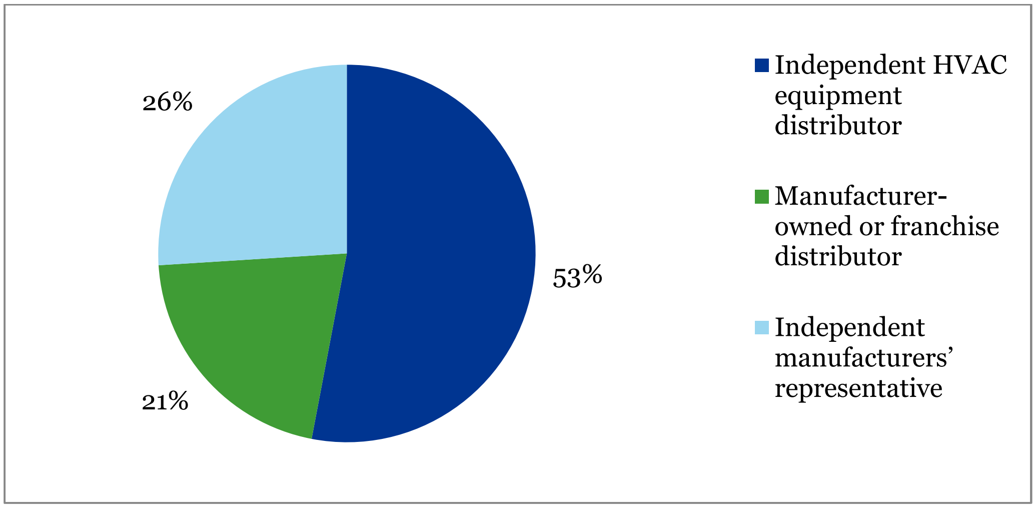
Evaluators asked distributors a battery of survey questions to better understand their business practices, the products and services they offer, and their use of rebates. The evaluators presented distributors with three distinct distribution business models and asked them to identify which model most closely relates to their practice. The models presented are as follows:

- **Independent HVAC equipment distributors** are distributors who and buy and sell equipment from various HVAC manufacturers and are not linked to a single manufacturer.
- **Manufacturer-owned or franchise distributors** are distributors that are-manufacturer-owned or franchised by a manufacturer. These distributors sell equipment directly from the manufacturer; they generally do not represent more than one manufacturer unless the manufacturer has multiple brands.
- **Independent manufacturer's representatives** are distributors who are independent sales agents that have exclusive rights to sell equipment for a specific manufacturer. Manufacturer representatives generally represent more than one manufacturer. Often, the products offered are complementary and do not compete with each other.

The evaluation team found it was fairly common for distributors to have more than one business model. For example, a distributor may be a manufacturer's representative for certain product lines and an independent HVAC equipment distributor for other product lines; or the distributor may be manufacturer-owned but sell other equipment as an independent distributor for product lines the manufacturer does not produce. We found such hybrid business models for about 25% of the distributors interviewed.

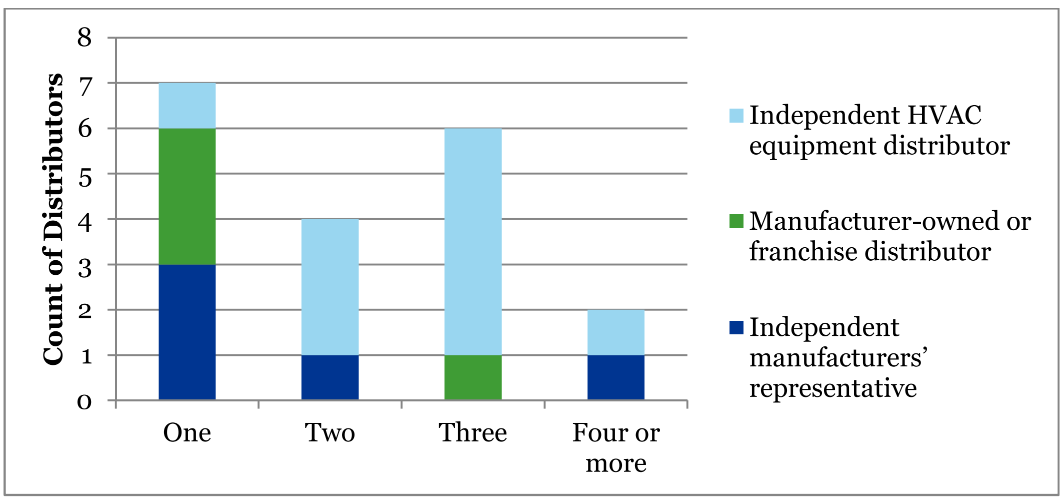
Distributors were asked which model best represented their practice. The majority (53%, as illustrated in Figure 7) stated they were independent HVAC equipment distributors, 26% were independent HVAC equipment distributors, and the remaining 21% were independent manufacturers' representatives.

Figure 7: Business Models of Upstream Distributors (n=19)



Distributors were also asked how many different packaged and split system manufacturers their company carries. As illustrated in Figure 8, independent distributors carry units from more manufacturers (most offer two or three) while manufacturer-owned/franchise distributors and independent manufacturers' representatives typically offer units from a single manufacturer.

Figure 8: Number of Unique Manufacturers Distributors Represent (n=19)



Another characteristic captured in the survey was the concentration of distributors in the state's major metropolitan areas. As shown in Table 33, the San Francisco bay area and the Central Valley had the highest concentration of distributors covering a geographic area, while the lowest concentration was in the southern part of the state: Los Angeles (LA) Inland Empire, LA Coastal, and San Diego. The "other areas" were most often cited as coastal cities throughout the state. Approximately three quarters of distributors also serve markets in states other than California. None were aware of nationwide Upstream programs and few were aware of mid- or downstream programs at the national level.

Table 33: Distribution Areas Served by Participant Distributors*

Geographic Areas	n=19	%
San Francisco Bay Area	15	88%
Central Valley (Modesto, Fresno, Bakersfield)	13	76%
Northern Valley (Redding	12	71%
Sacramento Area	12	71%
Desert (Palm Springs, Imperial Valley)	7	41%
LA-Coastal (LA, Orange County)	6	35%
LA-Inland Empire	6	35%
San Diego	5	29%
Other Areas	8	47%

*(multiple-response)

Distributors were also asked if they were knowledgeable about market and sale trends outside of California, and eleven of the 19 respondents said they were knowledgeable or somewhat knowledgeable. They were then asked how demand for high efficiency equipment in California compared to other states. Most respondents said demand in other states was lower due to electricity costs being cheaper (with Hawaii noted as an expectation[69]); lower due to California having Title 24, and some also mentioned the IOUs incentive programs.[70] [71] One distributor cited CA as having lower demand for high efficiency cooling than other states and extreme efficiency demands occur in: "Minnesota and Wisconsin, and less so in Oregon and Washington and to a degree Iowa."

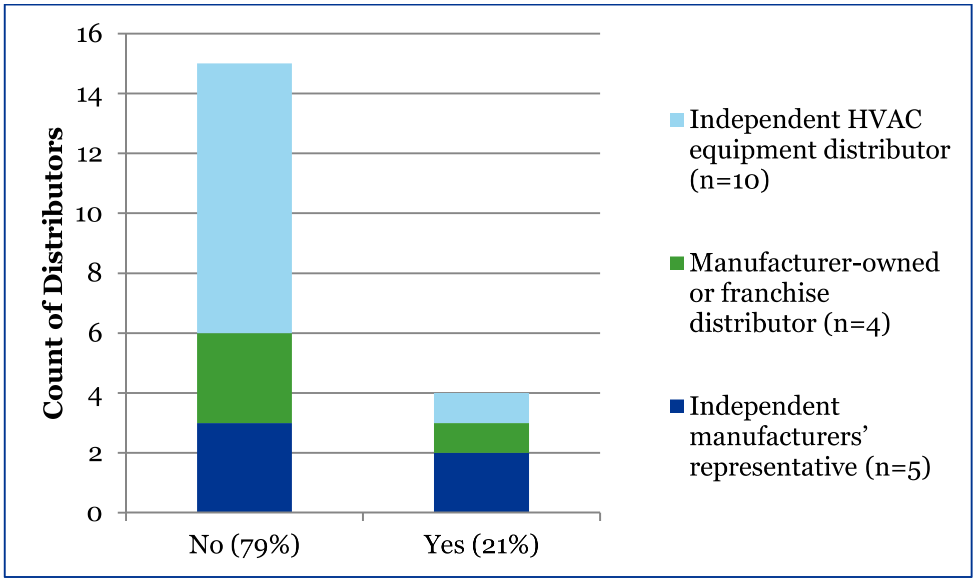
The target participants of the Upstream program are HVAC equipment distributors that sell commercial HVAC equipment, not HVAC companies that mainly install HVAC equipment and only do some distribution. To identify whether the program was engaging the right mix of distributors, we asked respondents if their company installed equipment. For this question, we anticipated few companies offering installation services. We found this to be the case. As illustrated in Figure 9, the majority of distributors (79%) do not offer installation services.

69 "High efficiency cooling equipment is proportional to the electric rates of in those areas. So Hawaii for example has some of the highest rates in the country and so there is a large demand there".

70 "Higher for CA because we have more emphasis on energy-efficiency. I wish the other states were as high as CA we specifically target the western states but nothing compares with your program".

71 "Demand is lower (in other states), my sense is based on Title-24, energy costs in California, and we sell more high efficiency equipment in the state of CA. The incentives are also factors you have to include that are driven through these utility programs (SCE, SMUD & PG&E) and incentives. There are localized programs, Oncor has a program in Dallas, and it's regional and local. I don't know that there are other states that have as good of a quality of as CA when it comes to rebates."

Figure 9: Does Distributor Install Equipment? (n=19)

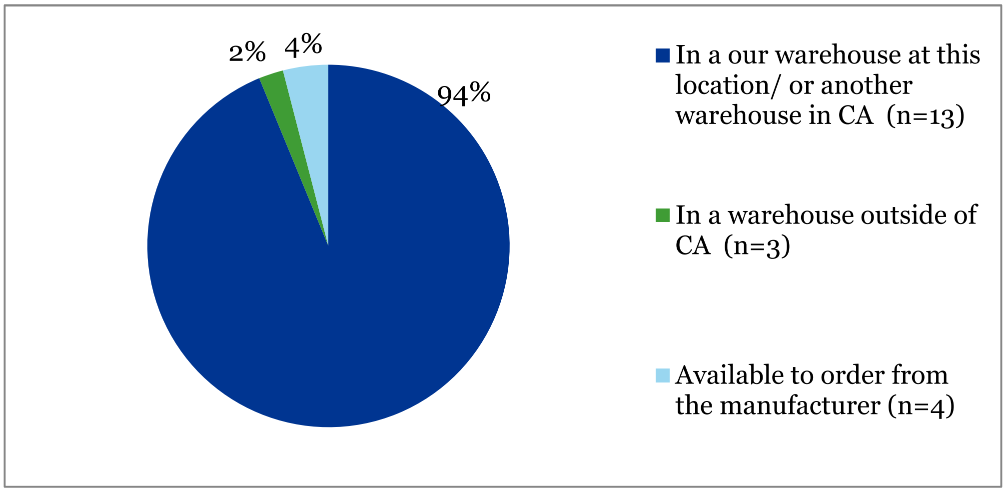


Also as shown in Figure 9, the majority (90%) of independent distributors do not install equipment; installations are more evenly divided among the other two types of distributors with slightly more independent manufacturers installing equipment. DNV GL found that manufacturer representatives often engage in sales of custom equipment (such as ultra-large-tonnage units), which appear to require installation by distributors.

If the respondent stated they distribute and install HVAC equipment, we then asked if they were more of a distributor, more of an installer, or if their activities were evenly divided. Among the 21% of respondents stating their business included both distribution and installation, only one distributor responded their activities were "evenly divided" while the remaining indicated they mainly engaged in distribution.

Distributors were asked where they source their shipments from. Evaluators applied distributor's share of program savings to their responses of where shipments come from. As illustrated in Figure 10, 94% of program saving shipments come from a warehouse in California, 2% come from warehouses outside of CA, and 4% are available to order from the manufacturer.

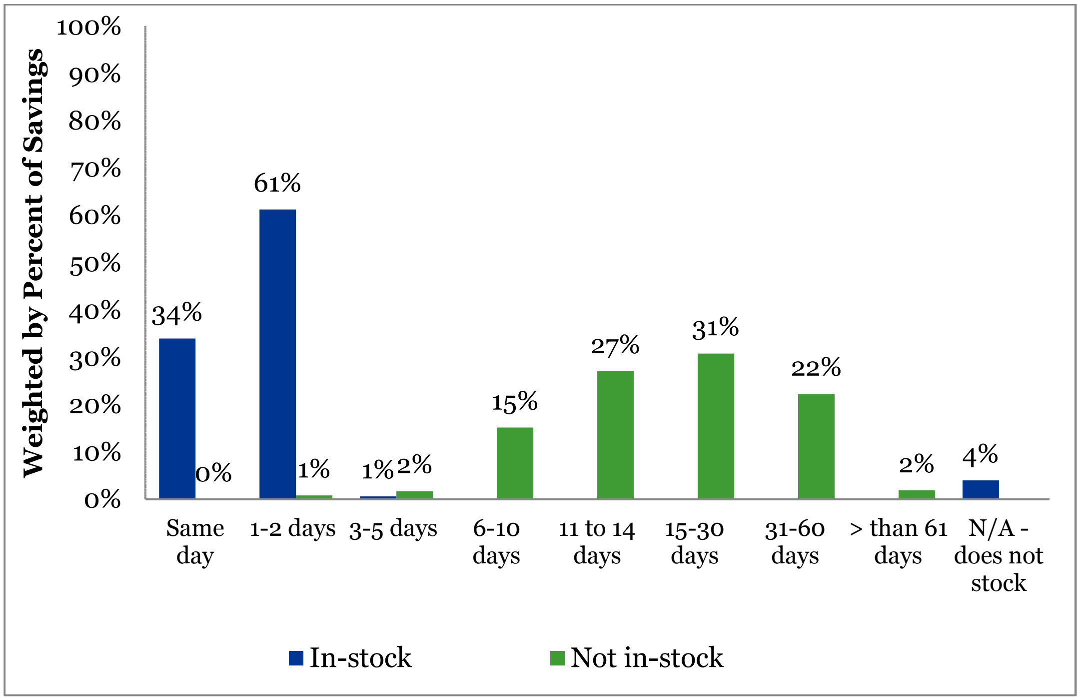
Figure 10: Shipment Sources for HVAC Equipment by Program Savings (n=19)



One of the program's theories is if high efficiency equipment is not readily available, a purchaser will select a lower efficiency unit, especially for units serving the emergency replacement market. In a few cases we found distributors citing this factor. In their cases, they made the decision to stock higher inventories of high efficiency equipment to differentiate themselves from competitors. Their theory is, by retaining a larger stock of high efficiency units, they can upsell to the next best model, but if the high efficiency models are out of stock then they don't have anything to sell.

Figure 11 compares the average length of time to receive a shipment for in-stock and out-of-stock equipment. The results were then weighted by distributors' shares of program savings. As illustrated, 95% of the shares are readily available and can be sold/received within 2 days or less. And 95% of out-of-stock equipment can be received within one week to two months. Without program intervention driving high efficiency, and following the theory that most distributors would have limited inventories of high efficiency, one would expect high efficiency delivery timeline would follow the out-of-stock timeline. According to the distributor's, the market is highly competitive. As such, the distributors experience a constant sense of urgency to deliver products quickly, if not immediately. Having what consumers want, when they want it, is relatively important as compared to other factors such as price and service.

Figure 11: Average Length of Time to Receive Equipment (n=19)



Balancing equipment supply and demand is a challenge for distributors; unless there is consistent demand, a distributor may only carry a very limited supply of a particular efficiency level or may only offer the equipment as a special order.

Distributors describe this assessment of supply and demand as their "turns rate," or how many times does a piece of equipment sell (turn over) in a given time period (e.g., one year, one month). The turnover rate is constantly re-evaluated and is of significant importance to distributors.

As one distributor indicated:

> *"The number one factor is the turns, (the number of times you sell that product per year). If you only sell it once a year there is no reason to keep it in stock; but if you can sell what you have in stock multiple times per year then it's a product you want to have in stock. We look at what people are buying and we also look at what rebates are out there that help us pay for the additional stock."*

The pressure to stock equipment that has a high turnover rate is a major barrier for the program, since higher efficiency units still have less demand than standard units.

4.3.2 Upstream Program Participation

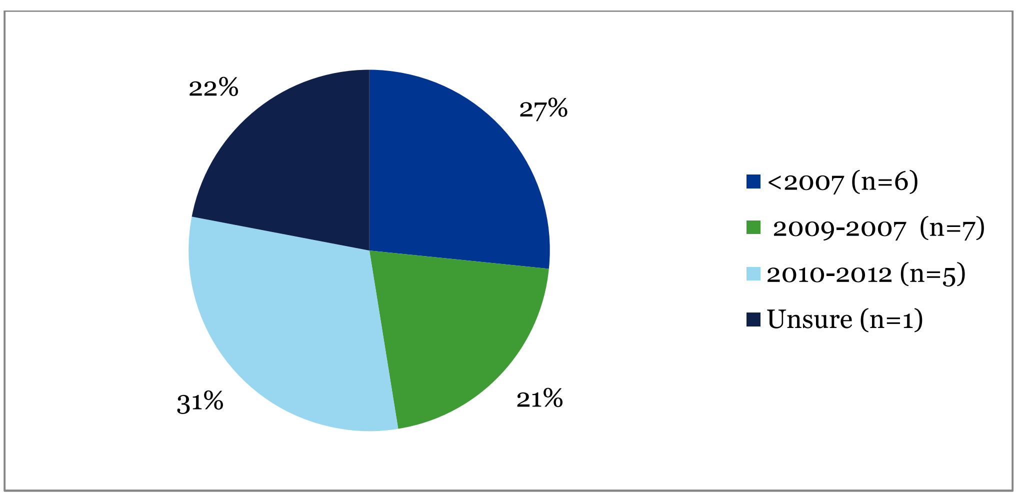
We solicited survey responses from high-level company decision-makers, most of whom were company presidents, vice presidents, or commercial sales managers. The role respondents had at their company enhanced the reliability of the survey results as they were likely to be more knowledgeable of the direct effects of the program and could provide the company's historical perspective on their participation. As illustrated in Table 34, the majority of respondents interviewed were personally involved in the decision to participate in the program. One respondent stated he/she was not personally involved in the decision to participate but oversaw the program implementation.

Table 34: Was Respondent Personally Involved in the Decision to Participate in the Upstream Program? (n=19)

Response	n=19	%
Yes	14	74%
No	5	26%

Figure 12 illustrates distributors' share of program savings, for the 2010-12 period, by the year they first enrolled in the program. Participation has grown steadily, with new enrollments peaking in the year 2010. During that year, six distributors joined the program, the most significant year of growth in the program's history. On average, the program has grown by a measure of two new distributors each year since 2006.

Figure 12: Program Share of Savings by Year Distributors Engaged in Upstream Program (n=19)



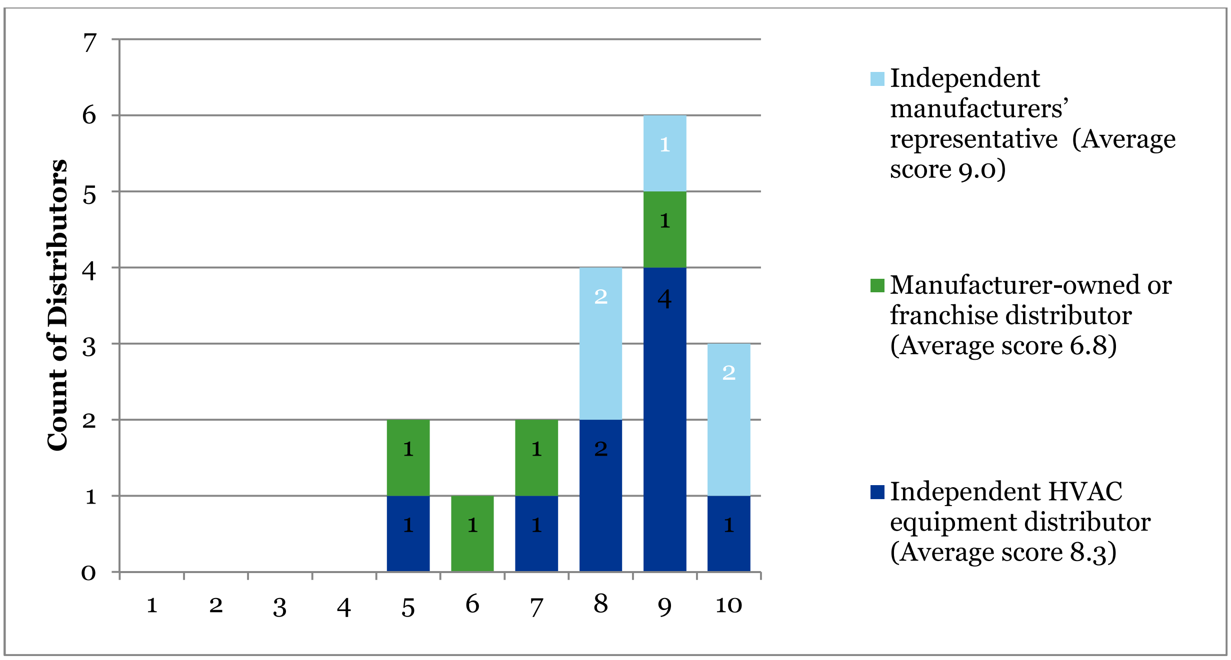
Unfortunately, one distributor could not recall when the year the company enrolled. Its share of program savings accounts for slightly more than 1/5 of the total share. The lack of this information does impact our peak savings by year of enrollment. This information would have likely informed us on the program cycle where savings reached its peak. As illustrated, the savings peak at 31%, (absent that distributor) occurred in the same program cycle as the enrollment-peak occurred.

4.3.3 Importance of Energy Efficiency

To identify whether distributors prioritize energy-efficient products in their marketing efforts, the evaluation team asked if energy-efficient HVAC is a particular focus of their marketing efforts or product offerings. The overwhelming response was "yes" with 18 of the 19 respondents agreeing. Additionally, 18 of the 19 respondents indicated their company tries to upsell higher efficiency product lines.

Interviewers asked distributors who rated energy efficiency as important just how important efficiency is in their marketing efforts and product offerings. Eighteen distributors rated importance on a Likert scale of 0 to 10 where 0 is not at all important and 10 is the most important. Figure 13 shows, independent manufacturer representatives rated energy efficiency the highest, with an average score of 9.0. Independent equipment distributors rated importance of energy efficiency slightly lower, with an average score of 8.3, and manufacturer- owned distributors rated it the least important with an average 6.8 score.

Figure 13: Importance of Energy Efficiency Rated 0-10 (n=18)



Independent manufacturer representatives rated efficiency the highest with an average score of 9. To the follow-up question of why energy efficiency is important, these distributors, with specialty products, reported that they market their brand as very efficient and often work directly with design engineers to customize their equipment. Below is a sample of comments to the follow-up question, “In what ways does your company focus on the sale of energy-efficient HVAC equipment?”

> *“Well by the products that we represent, we are an exclusive [manufacturer’s] representative, we tend to be leaders in the industry as far as efficiency is concerned.”*

> *“We generally get mechanical engineers on projects they are designing that an owner and architect have brought to them. They call us requesting a certain type of equipment we always push them into the highest efficiency and not just efficiency but also pushing the operation, the sequences, the controls to make those things happen, a lot of our work, 70% of our orders, are from engineers.”*

> *“By the design of the project and working with the engineering firm, laying out the job, evaluating our equipment making sure it’s meeting the mark.”*

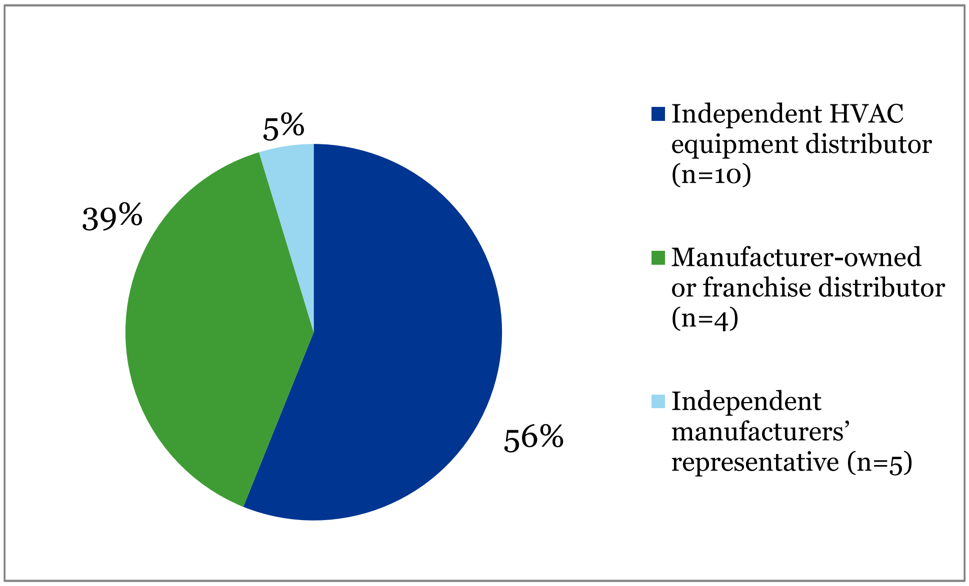
Manufacturer Owned/Franchise distributors rated efficiency the lowest, just slightly lower than independent distributors. These distributors promoted energy efficiency by giving consumer and dealer incentives, by working with design consultants to promote lifecycle costs vs. first cost, and through product design and development.

Independent Equipment Distributors: These distributors promoted energy efficiency by: upselling when the order is placed, marketing and advertising, training, guaranteeing or maintaining inventory for high efficiency stock, promoting return on investment or lifecycle, knowledge sharing via annual meetings with dealers, instant rebates (Upstream dollars), and selling the high efficiency unit for the low efficiency price.

4.3.4 Claimed Savings by Distributor Model Type

As illustrated in Figure 14, during the 2010-12 cycle 10 independent HVAC equipment distributors accounted for 56% of the program savings. Four manufacturer-owned or franchise distributors accounted for 39% of the program savings, and independent manufacturers’ representatives accounted for 5%. Evaluators were unable to interview three distributors. These non-respondents claimed less than 1% of the total claimed rebates during the 2010-12 cycle, a minimal share of program savings.

Figure 14: Distributors Percent of Claimed Savings



4.4 Measuring Program Influence on Upstream Distributors to Develop Net to Gross

4.4.1 Program Theory and Distributor Practice

Historically, distributors believed that to compete in the marketplace they had to stock the least efficient equipment in order to offer the lowest-priced products to the market, while higher-efficiency equipment theoretically sold less often due to higher cost (for both the distributor and customer) and therefore was not stocked and required a special order. Upstream rebates were designed to influence sales and stocking practices towards higher efficiency. So one might ask, how can the program influence stocking such that high efficiency equipment is readily available? The method that appears to have the greatest impact is for the program to provide funding to allow distributors to discount units or to offer a higher-efficiency unit at the same price as a lower-efficiency unit. As one distributor articulately stated: *"When you eliminate the price differential between a high efficiency and a standard efficiency unit you eliminate the only real objection to buy a high efficiency unit."* Some distributors also attempt to influence the market through knowledge sharing: *"We have nine people that specialize in that line alone and their sole job is to support contractor sales and engineering sales department in that line of product. There is a cost associated with training contractors and engineers on that type of product."*

Influencing stocking is a complex issue because stocking inventory is ultimately a demand-based decision or as distributors call it their "turns-rate," the number of times a product sells in a given time period. The turns-rate is of great importance, as reported by distributors: *"How* *often you can turn over the inventory, the "turns-rate", looks like cash so you need to turn it over a lot, you need to move it;"* and "*If we sit on inventory for more than one year there is a penalty; you never want to be penalized by headquarters; if you sit on a unit for more than a year, it's a write off."* If a product has few turns (little demand) then there is no reason to keep an item in stock, and as the above comments illustrate, having the right mix is critical to business operations. Stocking decisions are tactical and evidence based; products must have a sufficient demand to warrant storage at a local warehouse.

To fully embrace the program offerings, distributors must be willing to take risks by stocking equipment that doesn't have significant demand. *"People thought it might be risky; there was a little push back... Everyone is a bit change-averse... We found very quickly that any resistance or risk was made up for; that we were moving high efficiency faster than we could keep them in the factory."*

However some distributors may not be willing to take the risk and some may find they are ineffective at creating demand even after they took the risk. *"There will be no changes now or in the future because we don't plan on doing the promotions again since they were not successful;" and "The fact that we were flat using our skills or lack of, however you want to look at it, we were not able to change the marketplace."*

Ultimately, to be successful the program distributors must find a way to create demand: *"We supply and try and create demand;"* and *"The market has to be developed, in other words there is no market demand for high efficiency equipment; we have to bring that message to our customers."*

Therefore, it appears participant distributors must be willing to take risks and stock more high efficiency units than the market is currently requesting with the expectation they can direct demand toward high efficiency units. To change the market, the unit cost and availability need adjusting. The expectation is rebates will create market demand by increasing availability and making higher-efficiency units cost-competitive alternatives.

This outcome may be facilitated by defraying distributors' cost of carrying higher-efficiency units. This in turn will allow for distributors to reduce their equipment and/or storage cost for high efficiency systems, and any remaining rebates can be used to drive market demand towards the high efficiency equipment by reducing the purchase price.

According to some distributors, rebates are also used to fund knowledge sharing or for compensating commission-based sales staff when they effectively sell the benefits of high-efficiency equipment. Under this application it is difficult to be sure whether the rebates are actually improving a service offering or if the rebates are simply increasing a distributor's profit on the sale of a high-efficiency product that would otherwise sell without a rebate. If there are no service-enhancement offerings (such as spending more time with customers, increased staffing, and utilizing tools such as energy-savings models) or offerings are actually reduced, the rebate could simply be padding the pockets of distributors or paying for employees' wages. According to an article published by HARDI that referenced this particular program, distributors will "come up with creative ways to motivate sales staff rather than just lowering the price of the equipment." [72]

The evaluation found that rebate handling by individual distributors was discretionary and managed in a myriad of ways, such as: compensating commission-based sales staff for selling higher-efficiency equipment, covering equipment carrying costs, or passing on a portion of the rebate to the customer to drive sales. While the program model may appeal to many distributors because they are seeing the benefits through increased staffing levels and selling high efficiency equipment, the lack of transparency emphasizes the potential for misuse.

The industry-accepted definition of a free rider is a program participant who would have implemented the program measure or practice in the absence of the program[73]. If we attempt to frame the Upstream Program into this standard free-ridership paradigm, one could point to examples of free ridership. A distributor could be considered a free rider when they have made no changes to their practice and naturally-occurring market demand is the only driver to account for high-efficiency equipment sales.

Some quotes from surveyed distributors highlight cases where the program has been both ineffective and effective.

Skepticism of the Upstream Model:

- Distributor B: *"I think because the simple fact is the rebate is not finding its way to the person who is spending the money on the product."*

Examples of handling of the rebate in line with program theory:

- Distributor K: *"The incentive program does drive our decision, to a certain extent. It allows us to have carrying cost and sale price cost with a reasonable delta vs. standard efficient equipment. It flattens the playing field and drives the market in that direction."*
- Distributor M: *"When we don't have the inventory costs we 'pass it through' to the buyer."*
- Distributor O: *"As I said the rebate is what helps us with manpower to promote this product and without that it makes it very difficult."*

---

[72] Angela D. Harris, "Distributors Play Key Role in Efficiency Programs" HARDI Net, 15 March 2012, http://www.hardinet.org/distribution-center,

[73] The TecMarket Works Team, *California Energy Efficiency Evaluation Protocols,* April 2006, pg. 226

- Distributor S: *"The element of the rebate pass through, in order to remain competitive the Upstream programs have become apparent and public knowledge."*

4.4.2 NTG Score Algorithm

As previously mentioned the NTG algorithm relied on a self-reported approach. The evaluation team developed a self-reported scoring methodology that quantifies program influence to sales and stocking, given these are two factors that the program aims to effect. In a two-part[74] question, distributors quantified program influence. The two-part questions were repeated for sales and stocking. The evaluation team assigned influence-weights for the program influence on stocking and influence-weights for the program influence on sales to reflect the direct overall influence of the program. In order for the program to increase the efficiency of systems being installed, there needed to be a strong influence on the sale of high-efficiency systems. The team determined that sales were of primary importance; therefore sales were given a higher influence-weight of 70%, while stocking was assigned a lower influence-weight of 30%. Each distributor received an individual score ranging from 0-100% based on the scoring algorithm.

To obtain full program credit in the scoring, the distributor must state that the rebate was influential to both sales and stocking and they must state the rebate was highly influential relative to other factors affecting high efficiency equipment sales. All influential factors where rated on a Likert 0-10 scale, where is zero is not influential and ten is very influential; the distributor's assigned 0-10 ratings determined the quantified program influence.

To avoid leading respondents during the interview process, the evaluation interviewers did not directly ask the question "did the rebate influence your equipment sales and stocking practice" but rather asked "what are the primary factors that influence sales and stocking for high-efficiency equipment?" Asking respondents directly if the rebates were an influence could introduce bias and respondents may want to give the "right" answer even if no influence exists; for this reason the survey asked for the top influential factors. By asking for the primary factors of influence, the respondents must state on their own that the program was influential, and furthermore they must rate the importance of the rebate in addition to other factors that may have also influenced their sales and stocking practices.

Figure 15 and Figure 16 present the NTG questioning sequence and scoring approach that address both stocking and sales influences of the program, and subsequent text further describes the process.

---

[74] NTG Stocking question: "What specific factors influence your company's decision to stock a certain efficiency level for packaged and split system units? Please cite as many factors as you think are relevant. And for each of the factors you cited [in Q1] could you rate on a 0 to 10 scale how influential each individual factor is in your stocking decision? Where 0 means not at all influential and 10 means extremely influential.

Figure 15: Flow Diagram for Upstream NTG Scoring on Stocking

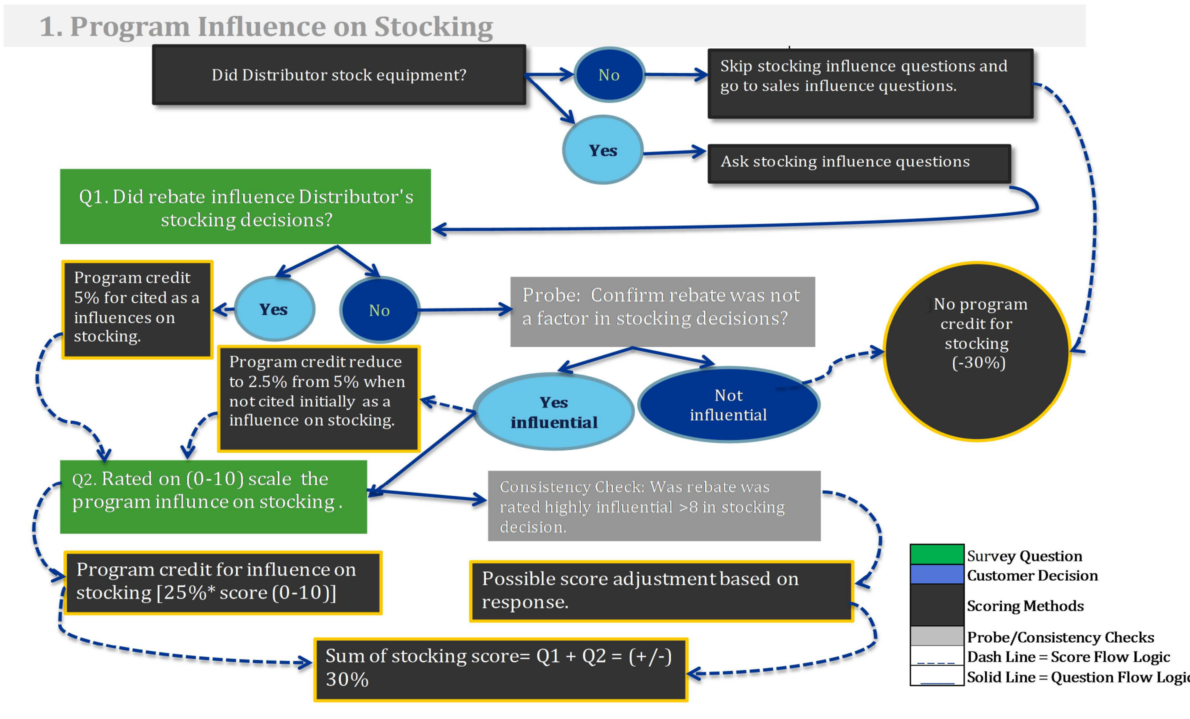
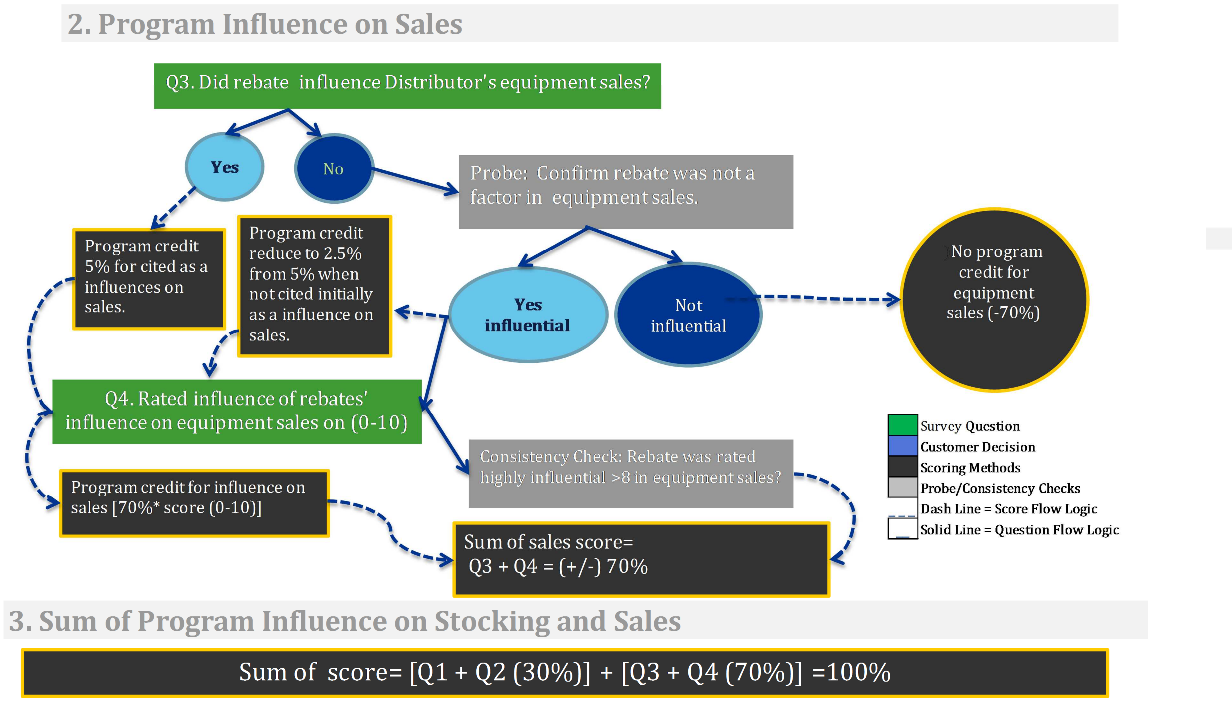


Figure 16: Flow Diagram for Upstream NTG Scoring on Sales



The program received full credit (net-to-gross ratio of 1.0) when all of the following conditions are met:

- **Stocking—**when the distributor claimed that the rebate influenced their stocking practices, the program received 5% credit in the scoring algorithm. The evaluation team thought it was appropriate to give nominal credit when respondents cited on their own, unprompted, that the program was one of the factors that influenced stocking. Had the questions been designed to lead or directly ask if the rebate was a factor then the degree of importance would have been the only consideration for weighting program influence. In this survey, the team combined the citation (or acknowledgement that the program was an influence) and degree of importance to generate the stocking score. The respondent was then asked to rate the degree of importance *for all factors* on a Likert 0-10 scale. If the rebate was cited as a factor, the score assigned to the rebate was then multiplied by 25%. The sums of these scores resulted in the following equation for program influence on stocking (under maximum credit):

  [5% + (0-10*0.25)% =30%].

  If the respondent stated that they did not stock equipment, the team skipped the questions on stocking and reduced the score by 30%. The team based the decision to automatically reduce a respondent's score based on the program theory component that in order to increase sales the equipment must be readily available or in-stock.

- **Sales—**the evaluation team repeated the same set of questions to address program influence on sales. The methodology was the same; the rebate was given nominal credit (5%) if cited since the respondent was not prompted to address the rebate. The team then asked the respondent to rate the degree of importance for all factors mentioned. If the rebate was one of the influential factors cited, the degree of importance (rated 0-10) assigned to the rebate was multiplied by 65%. The sum of those two scores resulted in the following equation for program influence on sales (under maximum credit):

  [5% + (0-10*0.65)% =70%].

- **The full equation** was, then:

  stocking [5% + (0-10*0.25)%] + sales [5% + (0-10*0.65)%] =100%

Partial Credit— To make sure the evaluation team explicitly accounted for how the rebate may or may not have influenced their decisions and practices and to clear up any misinterpretation the respondent may have had about the questions, the team also prompted the respondent with an aided question about the rebate, to confirm whether it was a factor or not. This prompt occurred after importance of all other factors was rated and *only* if the respondent did not previously cite the program as an influential factor.

If the respondent did not cite the rebate as a factor that influenced stocking and/or sales, the team asked them to confirm that the rebate was not an influence. If the evaluation team needed to prompt the respondent to confirm or deny the rebate influence, the team reduced their score slightly. The team asked the question: "In the previous questions you did not cite the rebate program as a factor that influenced in your stocking decisions; I'd like to confirm, did the rebate have an influence you're stocking decisions?" If yes, then, "Could you rate on a 0 to 10 scale how influential the rebate was?" If the evaluation team aided the respondent on either stocking or sales, the team reduced the score overall by just 2.5%. If the evaluation team aided the respondent on both stocking and sales the team reduced the overall score by 5%.

No Credit—the program received no credit if the respondent did not cite the rebate as a factor that influenced stocking and/or sales.

Detecting Inconsistencies— The evaluation team designed questions to address inconsistencies in responses to stocking and sales influences. For example, if a respondent did not initially state the rebate was influential but later provided a high score (>8) after being prompted, the team treated this as an inconsistency to be explored. If the respondent provided a high score after being prompted about the rebates, then the team asked the following consistency question: "You did not previously mention the rebate as a factor that influenced stocking. By providing a score of [>8] I would interpret this to mean the program was very influential. Could you please explain why you gave this rating? And why you did not mention the rebate initially?" The evaluation team either increased or decreased scores depending upon the responses. Out of the 19 respondents, the team asked the consistency question to five distributors. Responses to this consistency question were some of the more interesting ones captured in the survey. Examples of responses that exemplify how powerful the rebates were include:

> "*The rebate is always the underlying factor. I do not mention it but it is always the common denominator, it's there. If there were no tiers, 2, 3, or 4 you would definitely [have] sales occur in tier 1 and below. To promote a low-efficiency piece of equipment takes no time, no effort, and our stock of that equipment can be a lot less expensive than our high efficiency equipment. We can live in a commodity world of low-efficiency equipment; we can do that, and we have done that in the past but if we want to elevate ourselves to a high efficiency and stock more expensive high efficiency equipment the only way we can do that is to have the rebate offset and subsidize that.*"

> *"It [rebate] does, I feel like it does in terms of allows us to drive some market demand and differing our carrying costs. The program definitely has influence in those areas. Because the rebate program being in effect has allowed us to defray carrying costs such that there is little to no difference in carrying costs between a high efficiency and a standard performance unit. It's influential in allowing us to make the switch to 60% HE. It also allows us the opportunity to put high efficiency in front of the customer at an attractive enough price point. That drives market demand. If there is little to no difference in the initial costs then high efficiency becomes a determinant; its rare people will go with the less efficient unit."*

> *"Yes in the way I described it is a primary driver in the demand. There are only a few things that a distributor does to sell equipment. You have customer service and sales, price, delivery. And that's what customers are going to use on their decision on what to buy. The biggest factor in the price of the equipment is the rebate. So in the incremental price to go to high efficiency, if the majority of it is covered by the rebate and the payback on energy is 2 years they'll almost always buy it. So the number one factor in selling high efficiency is the rebate. My sales guy can be the greatest but if it's a 10 year payback no one is going to buy it."*

Adjusted Scores -Distributor's scores were adjusted by the evaluation team to apply credit accordingly. The result of the scoring algorithm underwent any necessary adjustments accounting for the verbatim responses that revealed inconsistencies in the scoring. Adjustments were made in both increasing and decreasing directions based on the consistency of scored question responses coupled with the open ended responses. The adjustments made by evaluators decreased the entire program score by 3%.

4.4.3 Applying Weights to NTG Score Algorithm

Below in the results subsection, we present both unweighted and weighted NTG scores. These are defined as follows:

Unweighted Scores - This score was generated by taking the simple average score across all 19 participant distributors.

Weighted Scores - To determine the weighted NTG score, the evaluation team took the influence score and then multiplied it by distributor's share of program savings, then summing. The share of program savings took into account distributors' 3-year average savings claimed from 2010 through 2012. Large participants dominated the total program savings. For example, the lowest share of program savings was less than 1% and the highest share was almost 25%. The majority of participant distributors (84%) each accounted for less than 5% of program savings during this 3-year cycle. In other words the program is made up of a few distributors (13%) who claim a significant share of savings. Distributor's total energy (kWh) and demand (kW) savings as reported by the IOUs (PG&E and SCE) formed the basis of the net to gross weights. Since many (63%) distributors served multiple service territories a single net to gross score was developed using the total savings from SCE and PG&E.

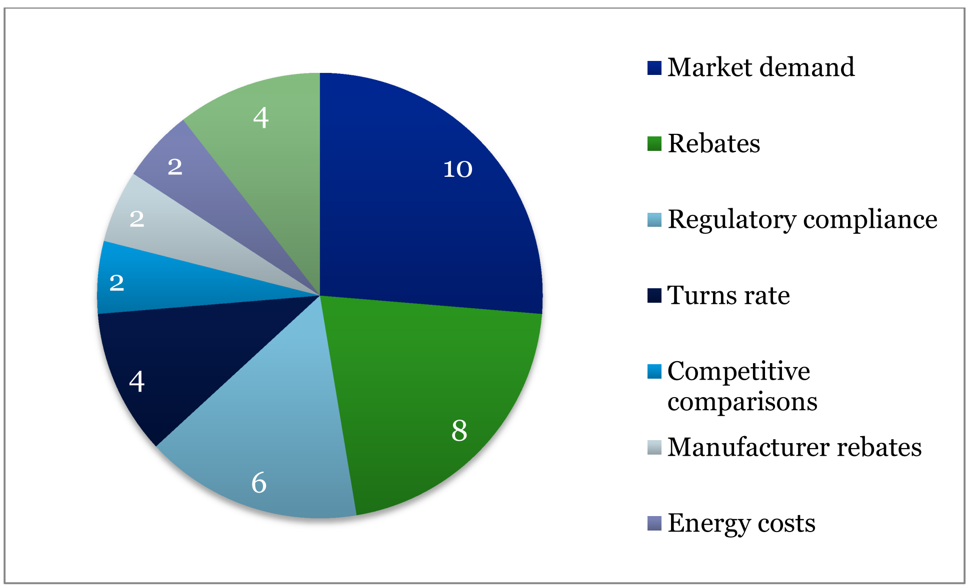
4.4.4 Program Influence on Stocking

A primary program goal was to influence distributors to stock high efficiency non-residential HVAC equipment. To determine whether the program was successful at meeting this goal, evaluators asked a series of questions about distributor's stocking practices. Given the level of importance that the IOUs have attached to program influences on stocking practices, it was interesting that several (21%) of the distributors do not stock units themselves but rather order equipment directly from manufacturers. From the onset of the survey design the evaluation team anticipated the program would only accept distributors who stocked equipment; this is why the scoring algorithm took into account both stocking and sales.

In most cases, the non-stocking distributors were selling custom equipment that was sometimes very large and therefore could not be kept in-stock. Nearly all of the non-stocking distributors were independent manufacturer's representatives. As previously illustrated these distributors rated energy efficiency important (9 on a 10 point scale). Efficiency is important because these distributors rely on sales staff to sell their products and components. It appears non-stocking distributors can influence their market through sales and design support that incorporates higher efficiency equipment components and/or by directly interacting with design engineers to assist them in their selection process.

Among stocking distributors, the evaluation team asked what factors influenced their stocking decisions. Distributors typically cited 3 or more factors; the number of respondents citing a given factor is illustrated in Figure 17. As shown, distributors most often stated (10 of 15) that market demand was the primary factor that influenced their stocking of high-efficiency equipment, followed by the rebate (8 of 15) and regulatory compliance requirements (Title 24). The "other" reasons as cited in Figure 17 included: new products, physical limitations of space, cost to carry, and tax rebates. One key limitation of the approach was that for distributors who did stock units, the survey did not determine if there were specific sizes or unit types that were always custom-ordered and therefore there could be little program influence on stocking. One unique spillover effect the program had for a distributor was improving the physical placement of their inventory: *"The accessibility of sales data that includes installation addresses has provided us with the ability to target our stocking on a warehouse by warehouse basis in a way that we didn't have before. Based on this data we have more efficient inventory and more efficient placement of that inventory as the different locations have different size requirements."*

Figure 17: Factors Often Cited as Influential to Stocking among Upstream Distributors (n=15)*



*multiple response

4.5 Upstream NTG Results

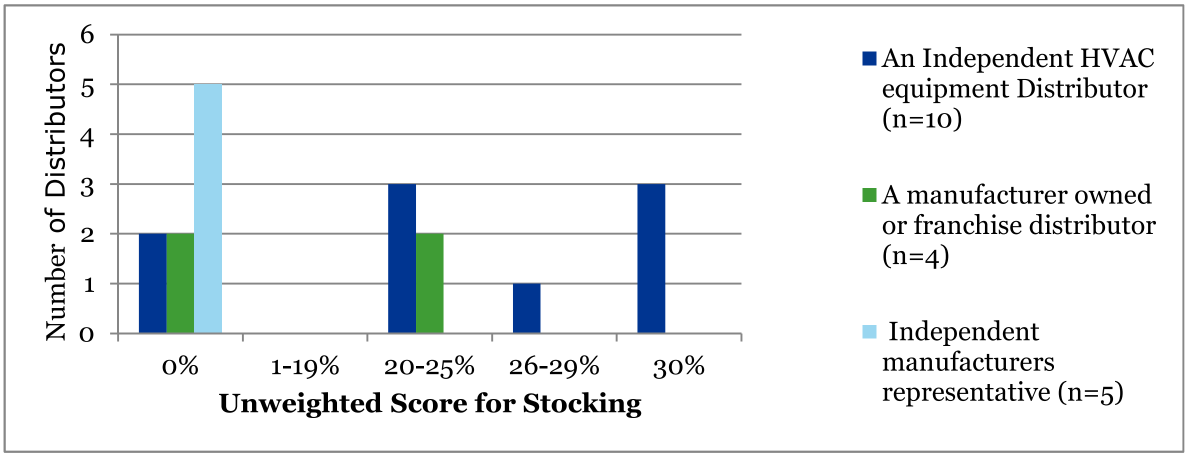
This subsection presents the upstream distributors NTG results for both un-weighted and weighted values.

4.5.1 Un-weighted NTG Results

The evaluation team first reviewed the un-weighted results stocking and sales, respectively, prior to combining the stocking and sales influence with the savings claimed by distributors.

Figure 18 displays the un-weighted results of program influence on stocking by type of distributor. The NTG scores are grouped in bins: 0%, 1-19%, 20-25%, 26-29%, and 30%. (As previously described the evaluation team assigned program attribution for stocking a maximum score of 30% and assigned sales attribution a maximum score of 70%for a total of 100%) To attain the maximum 30% for stocking, respondents must have cited the program as a factor that influenced their stocking practices and given it an influence score of 10.

Figure 18: Un-weighted NTG Results on Stocking by Type of Distributor (n=19)



The average un-weighted score for all distributors, including those that did not stock, was 14% out of a maximum score of 30%[75]. By removing those distributors that do not stock, the average score increased to 18%. As previously described, independent manufacturer's representatives did not stock. One-third of stocking distributors claimed the program had no influence on stocking, and most cited low or no program influence on sales.

As for why the program did not change distributors stocking and sales, two distributors explained they had ineffective marketing campaigns to sell program qualified high efficiency equipment.

One distributor found the promotions they used simply did not generate sufficient demand to change their stocking practices (<5% of their customers took advantage of their promotions) and have since given up on promotions and do not anticipate any changes to stocking practices for high efficiency units.

- Distributor B – "Prior to us taking the money off at the front end we relied on selling jobs that were already specified as high efficiency. Our marketing approach to offer the rebate

[75] Score is determined as 30% assigned to stocking and 70 assigned to sales for a total of 100%. The full equation is as follows: stocking [5% + (0-10*0.25)%] + sales [5% + (0-10*0.65)%] =100%.

once we earn our money from the utility never really took off. If the customer asked for it we would sell it but we were not upselling any equipment then."

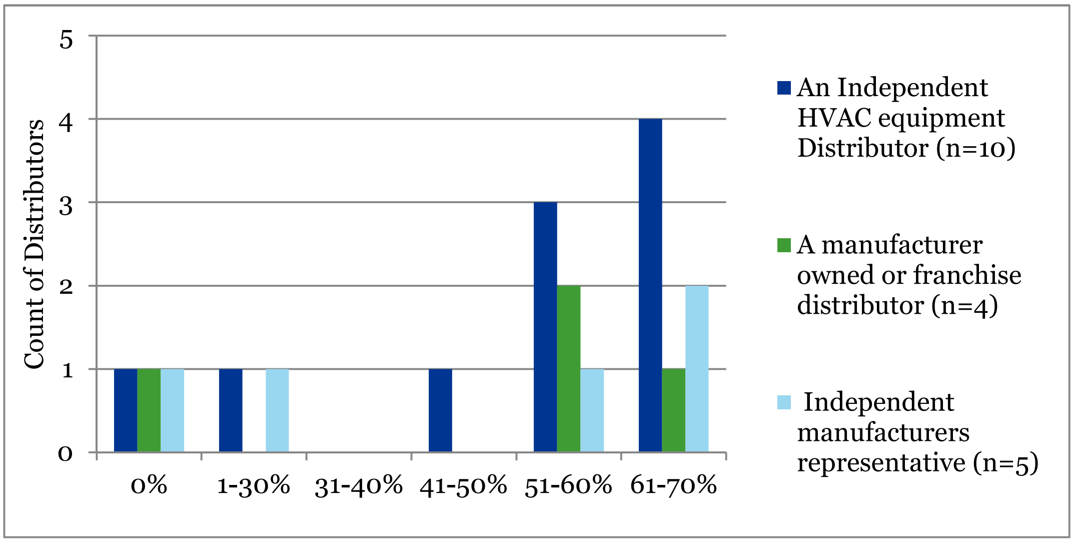
- Distributor D – *"There was a lack of focus; we are at the mercy of the customer and what the customer wants. If a customer says they want a 13 SEER that's what I'm going to give them and that's what the majority of them want."* DNV GL: *"You didn't have a strategy to upsell correct?"* Respondent: *"Right. There were a couple of promotions where we offered a discount if the customer bought a 14-16 SEER unit but in my opinion it was not successful; of the entire customer base maybe 5% took advantage of that promotion."*

A different distributor explained that after changing the way the rebates were provided to their customers they significantly increased their high efficiency stocking volume by a measure of 7,500 high efficiency units, which over time influenced the volume of high efficiency units they sold. This change came only after they tried other approaches that did not work. (It should be noted that while a change occurred from a per-ton rebate to matching the cost of a lower-efficiency unit, no program credit was cited for this evaluation cycle because the change occurred in 2013. The findings suggest low-rebate-claiming distributors could benefit from program sponsored coaching or customized methods to market direct sales towards rebate qualified products.)

The team repeated a similar set of questions to parse out factors that influence the sale of high efficiency program-qualified equipment. Figure 19 displays un-weighted results of program influence on sales by type of distributor. The scores were grouped into bins of 0%, 1-30%, 31-40%, etc. The average un-weighted score among all respondents was 47% out of a total of 70%[76].

[76] Recall again, the score is determined as 30% assigned to stocking and 70% assigned to sales for a total of 100%. The full equation is as follows: stocking [5% + (0-10*0.25)%] + sales [5% + (0-10*0.65)%] =100%.

Figure 19: Un-weighted NTG Score on Sales by Type of Distributor (n=19)



At one or more points in the survey, distributors cited how the rebates were used. Low-scoring distributors were often non-stocking distributors with a long sales cycle. They cited that they paid incentives to commission-based sales staff, "tracked rebates" (which we would interpret as relying on naturally occurring high efficiency sales), choosing not to lower high efficiency equipment purchase price and/or ineffective promotions.

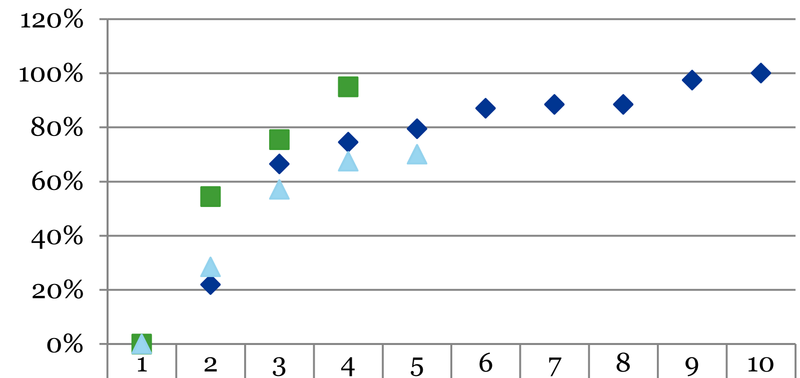
Distributors with long sales cycles (primarily due to custom-built equipment) may benefit in a different rebate framework or distribution model such as a reservation system. Due to a sales cycle that could span months or years, some distributors cited the uncertainty of claiming a rebate as reason for not incorporating the rebate into their purchase price. The uncertainty of the rebate availability and uncertainty of obtaining installation address information (a requirement) were both challenges for distributors.

- Distributor P – *"[Utility] has been good about saying there is limited funds; we never count on [the rebate] when we're quoting a project because we don't know when it will close/ship. Our sales cycle is lengthy; it could be several years long before we ship any equipment. We don't know whether the rebates will be there or have money."*
- Distributor B – *"We can't guarantee the rebates will be available when we bid the job nor can we guarantee we will get the necessary information to submit the rebate. Occasionally a rebate could be used on an immediate replacement job but we really don't use them on most jobs."*

Distributors with high scores generally used rebates to generate demand; the high-efficiency market had to be developed by incorporating rebates into price offerings, or by offering a high-efficiency unit at the same price as a lower-efficiency unit.

Prior to weighting and final adjustment, the simple average overall program score was 61%. The individual overall scores by distributor are illustrated in Figure 20.

Figure 20: Un-weighted Individual NTG Scores (n=19)



	1	2	3	4	5	6	7	8	9	10
◆ 1-An Independent HVAC equipment Dist	0%	22%	67%	75%	80%	87%	89%	89%	98%	100%
■ 2-A manufacturer-owned or franchise distributor	0%	55%	76%	95%						
▲ 3- Independent manufacturers' representative	0%	29%	57%	68%	70%					

4.5.2 Weighted and Adjusted NTG Results

The evaluation team determined the final net-to-gross results using the methodology described above in Section 4.4. The team tabulated individual scores and weighted the results of individual respondents using the proportion of program savings each respondent accounted for. (Energy and demand savings reported by the IOUs formed the basis of the net to gross weights.) The team made a final adjustment based on the open-ended responses to the survey (adjustments are described in detail in Appendix D). The goal of the final adjustment was to catch any inconsistencies in the responses and adjust the score up or down if a distributor's algorithm-based score did not line up with their responses to the open-ended questions.

Table 35 contains the program's un-weighted and savings-weighted (kWh and kW) net-to-gross scores. The un-weighted NTG score was 61%, and the energy-savings and demand-savings weighted scores were both80 %. Weights were determined by distributor's share of total kWh and kW savings.

The table shows, by distributor, the results of the scoring NTG algorithm and adjustments made by the evaluation team based on verbatim responses (supportive evidence) as described in Appendix D. The team made adjustments in both increasing and decreasing directions based on the consistency of scored question responses with the open ended responses. All adjustments were independent of the savings weights. The overall adjustments decreased the programs un-weighted scores by 3%: from 61% to 58%.

The final adjustments included:

- **Four distributors** whose scores were reduced (none of which received full program credit before the adjustments were applied); reductions applied to distributors of all business models.
- **Five distributors** whose scores were increased (all of which whose scores were already high (75% or higher) before the adjustments). Only independent HVAC distributors' scores increased.
- **Ten distributors** whose scores were unchanged.

Savings values were based on the 2010-12 tracking data for PG&E and SCE. Since numerous distributors (63%) serve both service territories, the team developed a single net-to-gross score using the total savings from the two programs. The SDG&E program had relatively small levels of savings that were not assigned an Upstream flag in tracking data, but measures that indicated packed or split systems and mentioned Upstream in the name also receive this net-to-gross adjustment, but these measures were not considered in the weights as they could not be directly assigned to the interviewed distributors.

The evaluation team estimated the final net-to-gross as 0.80 for energy savings and 0.80 for demand reduction. The distributors with large shares of program savings reported the most program influence on high efficiency equipment sales. These are slightly less than the ex-ante net-to-gross of 0.85. As a comparison, the results are also less than the scores reported in the 2006-08 impact evaluation, which had values greater than 0.9.

Table 35: Net-to-gross Results by Distributor

Distributor	Un-weighted Original NTG Score	Un-weighted Adjusted NTG Score	Energy Savings Weight (kWh)	Demand Savings Weight (kW)	Energy Savings-Weighted NTG Score	Demand Savings-Weighted NTG Score
A	0%	0%	1.8%	1.7%	0.0%	0%
B	0%	0%	0.1%	0.1%	0.0%	0%
C	0%	0%	0.2%	0.3%	0.0%	0%
D	22%	0%	0.2%	0.2%	0.0%	0%
E	29%	0%	2.8%	2.7%	0.0%	0%
F	55%	55%	22.0%	19.7%	12%	11%
G	57%	57%	0.5%	0.6%	0.3%	0.3%
H	67%	74%	0.2%	0.3%	0.1%	0.2%
I	68%	0%	0.7%	0.9%	0.0%	0.0%
J	70%	70%	0.5%	0.4%	0.0%	0.3%
K	75%	100%	18.4%	16.5%	18%	17%
L	76%	76%	15.2%	14.9%	12%	11%
M	80%	80%	0.2%	0.2%	0.1%	0%
N	87%	100%	0.7%	1.0%	1%	1%
O	89%	100%	3.2%	4.3%	3%	4%
P	89%	100%	5.5%	7.4%	6%	7%
Q	95%	85%	2.0%	2.7%	2%	2%
R	98%	98%	19.2%	17.6%	19%	17%
S	100%	100%	6.8%	8.6%	7%	9%
Total	**61%**	**58%**	**100.0%**	**100.0%**	**80%**	**80%**

The results by type of distributor provide some addition insights. As illustrated in Table 36, the program is getting very little savings or program attribution from independent manufacturers' representatives. The manufacturer-owned and franchise distributors reported less influence on stocking driving the stocking score and overall score.

Table 36: Adjusted Weighted Scores for Stocking and Sales

Type of Distributor	Stocking Score	Sales Score	Overall Score	Proportion of Program Savings
Independent HVAC equipment distributor	16.2%	37.5%	53.7%	56.1%
Manufacturer-owned or franchise distributor	4.1%	21.3%	25.4%	39.2%
Independent manufacturers' representative	0.0%	0.6%	0.6%	4.7%
Total (Maximum Potential Total)	20.3% (30%)	59.4% (70%)	79.7% (100%)	100%

The final net to gross ration was applied to gross energy and demand savings for 2010-12 programs. Table 37 shows the SCE program represented a majority of the statewide savings for the Upstream program.

Table 37: Upstream Net-to-Gross Ratio and Net Savings

IOU	Gross Energy Savings (kWh)	Gross Demand Savings (kW)	Net-To-Gross Ratio	Net Energy Savings (kWh)	Net Demand Savings (kW)
PG&E	17,705,130	9,100	0.80	14,164,104	7,280
SCE	64,368,795	21,952	0.80	51,495,036	17,562
SDG&E	560,854	203	0.80	448,683	162
TOTAL	82,634,780	31,255	0.80	66,107,824	25,004

4.6 Conclusions and Recommendations

The following conclusions and recommendations are based upon evaluators' observations and participating distributors input.

4.6.1 Conclusions

- The largest distributors in the state of California are active participants in the Upstream program. (This information was confirmed through responses from fellow distributors when asked who the largest distributors were.)
- The program attracted a variety of distributors, many of whom have non-competing product offerings.

- Approximately ¾ of distributors serve markets in states other than California[77]. None were aware of nationwide Upstream programs and few were aware of mid or downstream programs at the national level.
- How distributors utilize Upstream rebates is discretionary. The program in its current design allowed distributors (if they so choose) to steer market demand towards high efficiency equipment by lowering equipment costs, by covering equipment carrying costs, or by increasing stocking and sales attention given to high efficiency units.
- Non-stocking distributors had a very small share of the claimed rebates and had less opportunity to use them due to uncertainty of rebate availability because of long sale cycles.
- One resourceful distributor, with multiple distribution centers in California, found by analyzing the program's required data they were able to more efficiently manage their stock placement. Knowledge sharing of using program-required data such as this could be marketed to current and prospective distributors as a way to more efficiently manage their supply.
- There was some evidence of distributors enhancing their service offerings (they were able to hire more staff, take more time to communicate the lifecycle benefits to customers, etc.). (See example citation in Appendix D under "Distributor P")
- The program obtained a relatively high weighted NTG score (0.80) because a few distributors with significant market shares of high efficiency equipment and program rebates rated the program as influential, while in contrast the adjusted unweighted score for all distributors was just 58%.
- There were no program process-related questions included in the survey, but we did find participants complimenting the program's simplicity. Although for some distributors, the required unit-level installation data posed a challenge, as their customers (contractors) generally do not provide distributors this information. (This issue was not the case for distributors who required the information for warranty purposes.)
- Based on responses to the questions on sales volume by tier levels (examined more closely in Market Effects Work Order 54), we found that most sales occur in Tier 0

[77] The program tracks and provides addresses where units are installed and there is a program inspection process. Only units installed in California are claimed and verification is required for rebates.

(standard efficiency) and Tier 1 categories.[78] The lack of sales above Tier 1 could be due rebate levels and/or due to manufacturer's limited product offerings. [79].

- As identified in the sales volume questions (examined more closely in Market Effects Work Order 54), for a limited number of distributors, standard efficiency (Tier 0) still makes up the vast majority of their sales[80].
- We found to reap the program benefits, distributors must be willing to take risks by stocking more high efficiency (HE) units then the demand is currently requesting with the expectation they can direct demand toward high efficiency units.

4.6.2 Recommendations

The evaluators determined that the program succeeded, based on the savings weighted net to gross values. The first recommendation may apply to all participants and would increase gross savings. The other two recommendations target the lower volume distributors that had lower individual net to gross scores, but these enhancements may have marginal impact on program results. Note that the evaluators did not test these recommendations for cost effectiveness.

The program may want to re-evaluate incentive caps and rebate levels for higher efficiency equipment above Tier 1. The program may want to expand to ultra-large tonnage[81] (in excess of 60 tons) and offer larger rebates for the higher tier levels[82] [83] [84]. The larger units may not benefit from any influence on stocking, but rebates could influence sales of high efficiency large units. Based on the data collected from the sales volume questions, distributors' sales primarily occur in Tier 0 (non-program qualifying) and tier one, the first level

78 Baseline Characterization Market Effects Study of Investor-Owned Utility Residential and Small Commercial HVAC Quality Installation and Quality Maintenance Programs in California ,February 2015. Section 2.4, pg. 19 11"The number of tiers and tier standards, defined based on minimum unit SEER, EER, or IEER ratings, vary by HVAC unit type and capacity. Almost two-thirds of single-phase air-cooled HVAC units (60%) sold in 2011 and 2012 fall into Tier 0 and do not meet minimum program standards for this type of unit. Most of the rest of this unit type (40%) are rated at Tier 1, the lowest qualifying efficiency level for single-phase air-cooled equipment. At the other end of the spectrum, most ductless multi-split units sold are higher-efficiency units with over three-quarters (78%) of units sold rated at Tier 2 and most of the remaining units rated at Tier 1. Air-cooled three-phase and-water cooled units were sold over a wide range of efficiency levels from Tier 0 to Tier 4."

79 "Distributor Q: "Tier one, if you get into the higher tiers you get into a very select limited product offering for any manufacturer. The tier one is what sell the most because the price jump from tier zero to a tier one is an increase but not an overwhelming increase so it makes sense for the customers to do purchase tier one vs. two, three, and four."

80 Distributor A: "Our greatest increase is in the standard baseline equipment."

81 Distributor J – "One of the things I would love to see done is to create additional tiers for the larger tonnage equipment that would incentivize the higher efficiency (larger than 60 tons). We have designed our product lines to beat these efficiencies already."

82 Distributor O – "They have come out with more tiers of higher efficiency between 2 and 3 and 3 and 4 they have really racked up the efficiency but the dollars that they offer for that efficiency increase is not bridge[ing] the gap to get to that efficiency for those top tier levels."

83Distributor N – "The rebate has caused us to have more high efficiency then less. The highest tiers are the most expensive and slowest selling pieces and we had a whole bunch because we were buying blindly, but we have been doing a much better inventory analysis and metrics and we may have 2 years vs. 60 days so we would cut back on those very high efficiency unit."

84 Distributor I: "I don't know that the programs are high enough to offset participation to greater efficiency levels. We will stock whatever is selling. It's how we run the business. We tend to stock the same things. We probably stock a little higher with high efficiency because with standard efficiency we can always give them the next best. If you are out of stock on the best you don't have anything else to give them. So our tendency is less turns of inventory on the high efficiency materials. Standard efficiency accounts for approximately 60% of the inventory and 70% of the volume."

of participation. Evidence of this was identified through the sales-volume questions as reported in the Market Effective WO54.

Individual marketing support may help distributors, particularly those who are less successful at claiming rebates. It appeared some distributors don't know how to "go to market" with the rebate. Some of the challenges included obtaining unit installation information from contractors and how to market a rebate-discounted unit. Distributors may benefit from a marketing or inventory analysis and sales plan. The program should not assume distributors necessarily know how to use the rebates to their advantage, particularly given some may be risk averse. Marketing to increase demand for the highest tier of efficient equipment could create more demand for these units.

Distributors who have long sale cycles or custom built (non-stocked) equipment would likely benefit in a reservation system or some guarantee, such as bridge funding. The absence of a guarantee[85] [86] for a limited number of distributors was cited as a challenge. Note that this assumes different program logic and there is no stocking barrier, only a cost barrier to install high efficiency systems that are custom built.

85 Distributor B – "We can't guarantee the rebates will be available when we… bid the job nor can we guarantee we will get the necessary information to submit the rebate. Occasionally a rebate could be used on an immediate replacement job but we really don't use them on most jobs."

86 Distributor J – "Certainly we don't advertise the rebate to our customers because you know the program incentivizes us to promote those products and we do in many cases pass on those savings to make jobs more viable but we don't want that to be particularly visible by the customer because it's not a 100% guarantee that we are going to get the incentive so we'll sometimes hedge our bets with it but we don't advertise that to our customers."

5. Packaged Rooftop Unit Laboratory Testing Preliminary Results

The evaluation team conducted laboratory tests to evaluate field measurement instrument accuracy, economizers, and a dual-compressor commercial roof top unit (RTU). Tests were performed at Intertek Testing Services, Inc., Plano, Texas. The laboratory is AHRI-certified. Additional laboratory tests continue currently and will be performed in the future on packaged units from other manufacturers. The team initially conducted tests in a vertical configuration to simulate field conditions. Subsequent tests and AHRI tests were performed horizontally for safety and time constraints. R-22 systems were tested as part of this effort because they currently dominate systems in the CQM program. Additionally, new R-22 systems were sought out for testing while a few were still available. Tests on R410a systems are planned.

The evaluation team tested a 7.5-ton two-compressor commercial packaged unit with an economizer at varying damper positions, refrigerant charge, and airflow.[87] The team chose the 7.5-ton unit for testing because its manufacturer family of units[88] has 15%program market share,[89] higher than any other manufacturer family of units. The 7.5-ton unit has six fixed-orifice metering devices, one on the header of each circuit. The unit was fully instrumented and tested in an AHRI-certified laboratory per ANSI/AHRI Standard 210/240-2008 (cycling tests) and ANSI/AHRI 340/360-2007 (efficiency tests). The test setup conformed to ASHRAE Standard 37. Most tests were performed at a range of outside dry-bulb and wet-bulb temperatures found in AHRI standards and current and past CPUC field studies (i.e., 55/51 °F, 60/54 °F, 65/57 °F, 70/60 °F, 82/68 °F, 95/75 °F, and 115/80 °F) to simulate various diurnal temperature swings across the California coastal, inland, mountain, and desert climate zones.

Based on discussions with the laboratory staff and the experience of the EM&V team, the laboratory tests with actual cabinet leakage and functioning economizers are the first of their kind. These tests provided critical new findings on the diagnosis of system faults and the efficiency impacts of repairs. In future program cycles, the evaluation team will perform additional laboratory tests on packaged units from other manufacturers.

---

[87] These tests are in progress. A ton is a unit of cooling capacity equivalent to 12,000 British thermal units (Btu) per hour. The Btu is the energy required to raise one pound of water one degree Fahrenheit.

[88] A family of units includes units of multiple sizes with the same initial characters in its model number and all of the units are covered in the same installation and service manual.

[89] Program market share is based on 2006-08 and 2010-12 tracking data. The general population includes more small units.

Critical findings from laboratory testing with an economizer include:

- The impact of economizers on system efficiency is significant and unexpected. All tests show that even optimally adjusted systems with an economizer perform significantly lower than their expected system efficiency.

- The diagnosis and adjustment of charge is difficult to achieve in the field since both unit total airflow and the presence of outdoor air intakes affect the reliability of fault detection diagnostics (FDD). This finding led to the field work described under Commercial Quality Maintenance of refrigerant charge diagnostics compared to a charge weigh out/weigh in procedure.

- The test results demonstrate that eliminating typical unit air leakage can be a more reliable measure to improve cooling efficiency than what is typically obtained from refrigerant-side repairs. See the commercial quality maintenance program recommendations presented in chapter two on additional field requirements and considerations for this strategy.

5.1 Role of Lab Testing in Impact Evaluation

Understanding the efficiency impacts of faults can be challenging. In reviewing the results, the expectation should be that the efficiency of all units will be relatively lower at higher outdoor and indoor temperatures. Faults may have different impacts at different conditions, but generally should reduce efficiency relative to optimal at a given condition. The effects of outdoor and indoor conditions and faults on efficiency are non-linear and become more complex when combined, requiring empirical testing to derive efficiency impacts. Once known, the efficiency impacts can be used in engineering analyses to estimate energy and demand impacts of faults.

The field measurements needed for program impact evaluations have physical constraints and other inherent limitations, while laboratory testing relieves the limitations to the greatest extent possible. The testing laboratory controls all of the temperature conditions and faults and uses sensitive instrumentation to measure unit performance. The influence of system issues on the unit can also be eliminated. In the field however, the weather and occupants control the conditions, faults are diagnosed but are not known directly, and instrumentation must sacrifice sensitivity for reliability and cost per sample point. The evaluation found that detailed laboratory testing became essential to understanding the findings from fieldwork.

5.2 Out-of-Box Factory Charge, Optimal Charge, and Diagnostic Tests

Out-of-box tests were conducted to evaluate rated and field efficiency based on factory charge, optimal charge, and diagnostic conditions. Test results are for a unit placed on a platform to simulate rooftop unit conditions. These tests were all performed under standard AHRI rating conditions. Economizers were not installed. Lab results and field findings from this study indicated that adjustments for air flow and supply fan power are relevant for DEER assumptions and that using values obtained in the AHRI rating process are misleading. These tests were all performed under standard AHRI rating conditions without economizers. Measured efficiency increased by 7% at the "A" rating point when manufacturer's charging diagnostics were used to establish "correct" charge levels in lieu of nameplate charge values.

Table 38 provides the rated EER versus "out-of-box" factory charge and laboratory optimal charge for 2-stage cooling (i.e., both compressors), 1-stage cooling, no economizer, and airflow of 3,000 standard cubic feet per minute (CFM).[90] The out-of-box factory charge was 7.6 pounds (lbs.) in circuit 1 (C1) and 8.1 lbs. in circuit 2 (C2). The laboratory optimal charge was 8.59 lbs. in C1 and 10.35 lbs. in C2. The optimal charge was established using manufacturer charging charts where target suction temperature (ST) is a function of suction pressure and outdoor air temperature as shown in The total static pressure, change in pulley size, and cabinet sealing represent conditions allowed in rating tests that are not feasible in field conditions. Additional testing for faults was based on more typical field conditions.

Figure 21 (Carrier 2005)[91]. The total static pressure, change in pulley size, and cabinet sealing represent conditions allowed in rating tests that are not feasible in field conditions. Additional

[90] EER is the cooling capacity in thousand Btu per hour (kBtuh) divided by total air conditioner electric power (kW) including indoor fan, outdoor condensing fan, compressor, and controls. This is the manufacturer AHRI rating.

[91] Charging chart conditions presume 400 CFM/ton at 0.25 in WC static at near sea level conditions

testing for faults was based on more typical field conditions. The measured ST must be within ±5°F of the target ST. These tests were all performed under standard AHRI rating conditions without economizers. Measured efficiency increased by 7% at the "A" rating point when manufacturer's charging diagnostics were used to establish "correct" charge levels in lieu of nameplate charge values.

Table 38: Rated EER vs. Tested EER for Out-of-Box Factory and Laboratory Optimal Charge – Vertical Flow Configuration

Test	Manufacturer Rated EER	Out of Box Factory Charge EER	Laboratory Optimal Charge EER
Verification "A" Test Sealed Cabinet 10" Pulley	11.0	9.6	
2 Stages "A" Test 95°F OD 80/67°F ID No Seal, 7" Pulley	11.0	8.4	9.0
2 Stages "B" Test 82°F OD 80/67°F ID No Seal, 7" Pulley	13.5	9.8	10.9
1st Stage "A" Test 95°F OD 80/67°F ID No Seal, 7" Pulley		6.0	6.5
1st Stage "B" Test 82°F OD 80/67°F ID No Seal, 7" Pulley		6.8	8.0

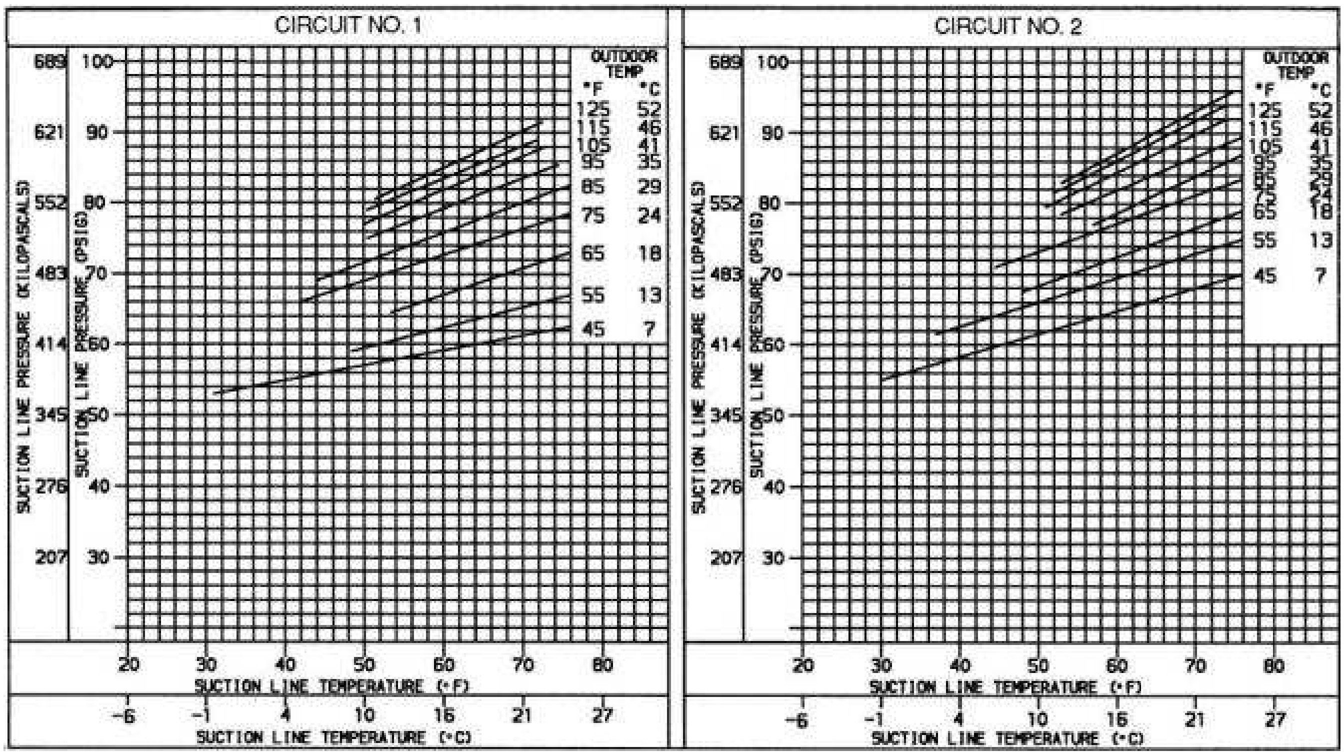
Source: Intertek, Plano, TX
We were able to replicate the manufacturer EER within the allowable tolerance when tested with a horizontal flow setup.

Indoor (ID) conditions were 80°F dry-bulb/67°F wet-bulb and outdoor (OD) conditions were 95°F dry-bulb for the "A" tests and 82°F for the "B" tests. The verification "A" test was performed with cabinet seams taped to eliminate leakage, 10-inch diameter motor-drive pulley, A55 V-belt, and 0.25 inches water column (in WC) total static pressure to reduce blower fan power and achieve maximum EER per AHRI test protocols.[92] The AHRI standards consider verification tests to meet claimed ratings if they are within 5%. The out-of-box verification "A" test with sealed cabinet and 10 inch pulley was 13% less efficient than the manufacturer-rated EER. The out-of-box factory and optimal charge tests with no seal, 7 inch pulley[93], and A49 V-belt were 18% to 27% less efficient than the manufacturer-rated EER. The total static pressure, change in pulley size, and cabinet sealing represent conditions allowed in rating tests that are not feasible in field conditions. Additional testing for faults was based on more typical field conditions.

[92] The inch of water column is defined as the pressure exerted by a column of water of 1 inch height at 39°F (4°C) at standard acceleration of gravity. 1 inWCin. WC is approximately 0.0361 pounds per square inch (psi) or 249 Pascals at 0°C.

[93] The 7 inch pulley was supplied with the unit.

Figure 21: Manufacturer Charging Charts for Dual-Compressor Unit (Carrier 2005)



Tests of refrigerant charge diagnostic conditions versus economizer outdoor air damper position for C1 and C2 are provided in Table 39. Tests were conducted at 2500 CFM airflow. The charge level corresponding to maximum efficiency was 13% and 28% greater than factory charge for C1 and C2, respectively[94]. The superheat (SH) protocol was based on the same manufacturer superheat table used in the Title 24 RCA protocol (CEC 2008).[95] Diagnostics for C1 were

[94] Charge was increased from factory charge up to 60% overcharge incrementally. The optimal charge was determined after review efficiency results for the incremental data. An optimal charge was not sought for all fault conditions by the evaluation team given the iterative process required.

[95] In the 2010-12 statewide programs the service technicians could use any method and Title-24 verification procedures were observed to be used in the absence of program protocols.

acceptable for all damper positions since ΔSH and ΔST are within ±5°F. Diagnostics for C2 indicated undercharge (i.e., "fail") for all outdoor air damper positions and both protocols (ΔSH and ΔST were greater than +5°F). C2 received more warm outdoor air and less cool return since it is above C1 and more aligned with the outdoor air dampers. This increased the cooling load on C2 relative to C1, as indicated by higher actual superheat and suction temperatures (pressures). The diagnostic numbers in italics failed, while other numbers indicate pass diagnostics.

Table 39: Refrigerant Charge Diagnostics versus Outdoor Air Damper Position

Test	OA %	Outdoor DB/WB °F	Indoor DB/WB °F	Mfr ΔST C1	T24 RCA ΔSH C1	Mfr ΔST C2	T24 RCA ΔSH C2
Laboratory Optimal Charge with No Economizer	4.5	95/75	80/67	-3	2	-1	3
Closed Economizer Dampers (Rated at 10 CFM/ft^2)[96]	15.4	95/75	75/62	1	2	*25*	*27*
Economizer Dampers Open 1 Finger "10% OA"	19.5	95/75	75/62	-1	1	*26*	*26*
Economizer Dampers Open 2 Fingers "20% OA"	23.2	95/75	75/62	-1	1	*26*	*28*
Economizer Dampers Open 3 Fingers "30% OA"	30.1	95/75	75/62	-1	1	*26*	*29*
Economizer Dampers Fully Open	62.1	95/75	75/62	-2	1	*26*	*30*

Source: Intertek, Plano, TX

Table 40 provides laboratory test results of system[97] efficiency, designated as EER-star (EER*) versus refrigerant charge and outdoor air damper position for the 7.5-ton two-compressor packaged unit. Relative EER* is the cooling capacity in thousand Btu per hour (kBtuh) supplied to the space divided by total air conditioner electric power (kW) including indoor fan, outdoor condensing fan, compressor, and controls. This is not the manufacturer EER rating which includes the same power inputs, but only includes the coil load and would not consider outside air as an inefficiency. Compared to the approximate (~) factory charge, the +20% charge (laboratory optimal) improved efficiency by 8% to 13% for closed damper, and by 9% to 39% for 1-finger open. Under charge of -20% reduced efficiency by 27% to 55% for closed dampers and 27% to 95% for 1-finger open. Tests of +60% overcharge relative to factory yield diagnostics within manufacturer ST specifications for both. However, +60% overcharge yielded negligible efficiency improvements of -8% to +2% when compared to +20% overcharge laboratory optimal.

---

96 Closed damper outdoor air fraction is approximately 15% consistent with ASHRAE 62.1.

97 The system efficiency is that seen by the space and differs from the appliance efficiency obtained in AHRI tests. The significant difference is that O/A (either through dampers or cabinet leakage) is treated as a system inefficiency rather than a space load.

Table 40: System Efficiency versus Refrigerant Charge and Outdoor Air Damper Position

Test	Refrig. Charge c1/c2 (lb)	Airflow (CFM)	EER* 82/68°F OD 75/62°F ID	EER* 95/75°F OD 75/62°F ID	EER* 115/80°F OD 75/62°F ID
No Economizer AHRI Rated Efficiency	7.6/8.1	3000	12	10	7.5
~Factory Charge Closed Damper	7.07/8.73	2500	7.9	5.3	2.9
~Factory Charge 1-Finger Open		2500	7.5	4.4	1.8
-20% Charge Closed Damper	5.55/7.11	2500	5.8	3.3	1.3
-20% Charge 1-Finger Open		2500	5.5	2.3	0.1
+20% Charge Laboratory Optimal Closed Damper	8.59/10.35	2500	8.6	6.0	3.2
+20% Charge Laboratory Optimal 1-Finger Open		2500	8.2	5.1	2.5
+40% Charge Closed Damper	10.11/11.97	2500	8.7	6.0	3.2
+40% Charge 1-Finger Open		2500	8.3	5.2	2.2
+60% Charge Closed Damper	11.63/13.59	2500	8.7	5.8	2.9
+60% Charge 1-Finger Open		2500	8.4	5.1	2.0

Source: Intertek, Plano, TX, Factory charge tests with economizer are pending.

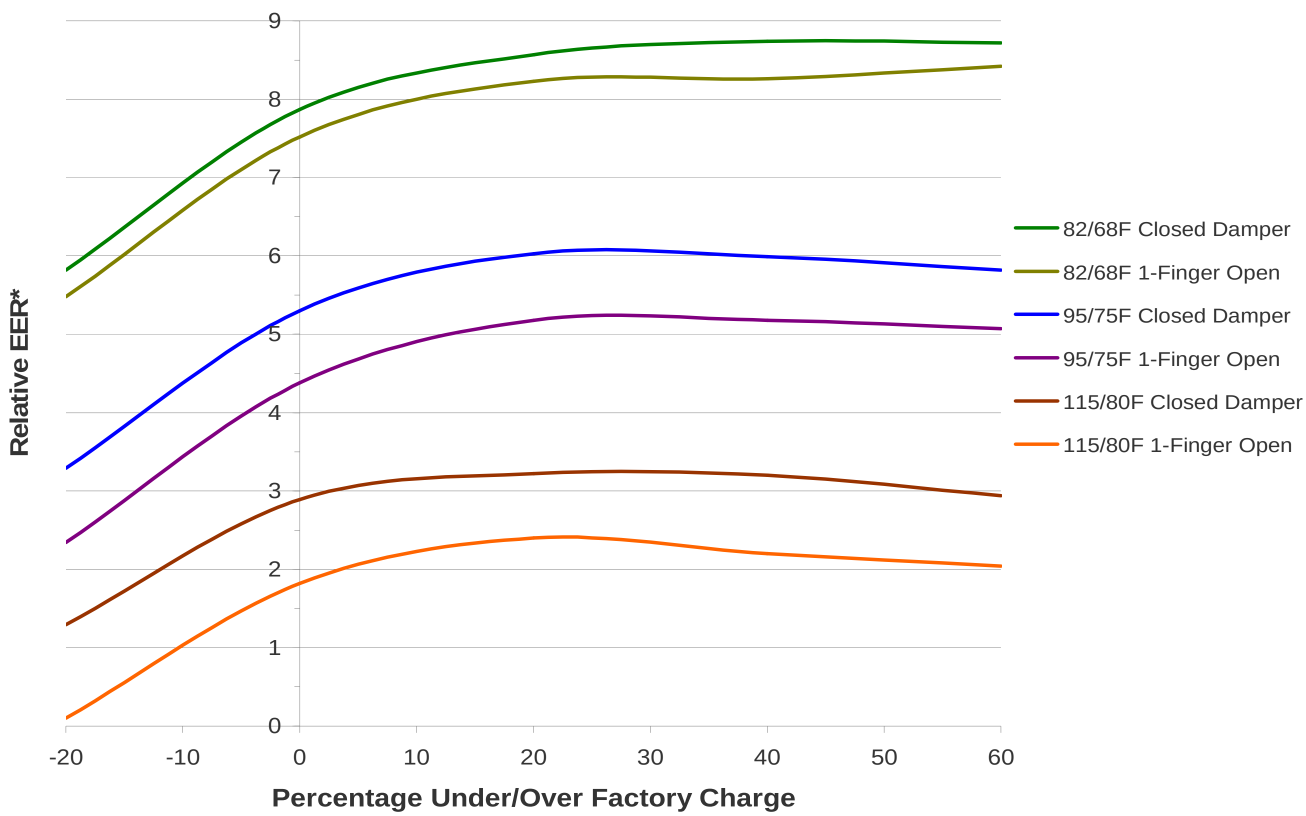
Outdoor air damper position is important for FDD, especially for multiple circuits. The efficiency-optimal charge is difficult to achieve in the field since the economizer should be replaced with a panel to eliminate outdoor air coming into the unit and allow an accurate measurement of coil entering temperature. Technicians must be careful to follow manufacturer's charging charts when adding refrigerant to a fixed-orifice system since overcharging can cause liquid refrigerant to flood the compressor at startup or when refrigerant incompletely vaporizes in the evaporator and enters compressor during operation.[98] Repeated flooding during normal off cycles or excessive flooding during steady-state operation can dilute oil in the compressor and cause inadequate bearing lubrication and premature failure.[99] Ultimately, factory charge is assumed to provide the best value, which is close to optimal efficiency and eliminates potential liquid flooding. It can reliably be achieved by evacuation and re-charge using a refrigerant scale, but the use of diagnostic readings could not establish factory charge via diagnostic readings and two typical protocols for the unit tested.

[98] Tomczyk, J. 1995. Troubleshooting and Servicing Modern Air Conditioning and Refrigeration Systems. ESCO Press. Mt. Prospect, Ill.: Educational Standards Corporation.

[99] Emerson 2010. AE-1280 Application Guidelines for Copeland® Compliant Scroll Compressors (ZR*1 Models). Emerson Climate Technologies. 1675 West Campbell Road, Sidney, OH 45365.

These tests show that even optimally adjusted systems with an economizer don't perform anywhere close to their expected system efficiency. Outside air that is brought in for ventilation requirements is an appropriate unit load, but unintended leakage acts an efficiency loss. Figure 22 shows laboratory test results of relative EER* versus percentage under/over factory charge for varying damper positions and outdoor drybulb/wetbulb temperature conditions. The -20% factory charge tests indicated a -26% efficiency impact at 82/68°F, -38 to -47% impact at 95/75°F, and -55 to -95% impact at 115/80°F. Increased unintended outdoor airflow through an economizer reduced system efficiency across all charge conditions. The impact was greatest at the hottest outdoor conditions. The +20% overcharge tests (i.e., laboratory optimal) indicated an efficiency improvement of 9% to 14% with closed dampers and 9% to 32% improvement with 10% open damper setting. The 40% to 60% overcharge tests indicated negligible or negative impacts on efficiency of +2% to -8% depending on damper position. These test results illustrate the challenge associated with trying to improve efficiency with refrigerant charge adjustments. Units serviced in a program might already be slightly overcharged, and simply adding more charge can reduce efficiency. The field evaluation found many units were in fact overcharged after service, and as stated above, overcharge can lead to compressor flooding and damage. The test results demonstrate that reducing unintended air leakage can be a more reliable measure to improve cooling efficiency. In the results presented, the underlying cause for overcharge insensitivity is the metering device used in the system. The metering device for the particular manufacturer produced results similar to a thermostatic expansion valve (although it operates differently) which also mitigates the impact of overcharge on system efficiency.

Figure 22: Relative EER* versus Percentage Under/Over Factory Charge for Varying Damper Positions and Outdoor DB/WB Temperature Conditions



5.3 Impact of Economizer Damper Position and Airflow

Outdoor airflow through economizer or make-up air dampers is not measured by program technicians and not included in efficiency tests used to rate packaged units. The following "rule-of-thumb" is assumed by most technicians: closed dampers 2% outdoor air (OA), 1-finger 10% OA, 2-fingers 20% OA, 3-fingers 30% OA, and fully open 100% OA.[100] Approximately 74% of observed units in the programs after maintenance had economizer or make-up air dampers set to one or more fingers open after maintenance was completed. Observations at a post-maintenance site with small packaged units found 75% of make-up air dampers fully open. Typical as-found conditions varied, but closed dampers were the most common for the pre-maintenance baseline.

---

[100] One-finger open is approximately 0.7 inches – typical finger width (1.8 cm), 2-finger is 1.3 inches (3.3 cm), and 3-finger is 2 inches (5.1 cm).

Minimum outdoor air damper position can have a significant impact on energy efficiency. Table 41 provides laboratory tests of the relative energy efficiency ratio (EER*) versus outdoor air damper position for a 7.5-ton two-compressor packaged unit and rated airflow of 3,000 CFM. Indoor conditions were 75°F dry-bulb/62°F wet-bulb for all tests. Outdoor conditions varied as follows: 95°F dry-bulb/75°F wet-bulb, 82°F dry-bulb/68°F wet-bulb, and 115°F dry-bulb/80°F wet-bulb. The laboratory tests were performed with no economizer installed as well as with an ASHRAE 90.1 compliant economizer rated at 10 CFM/ft^2 or 67.1 CFM outdoor air leakage with closed dampers per ANSI/AMCA Standard 500-D-12 (ANSI/AMCA 2012). The actual outdoor air leakage with closed dampers was 462 CFM or seven times greater than the 67.1 CFM rated damper leakage.

Table 41: System Efficiency (EER*) versus Outdoor Air Damper Position for 7.5-ton Unit

Test	Outdoor Air %	Outdoor Air (CFM)	EER* 82/68°F OD 75/62°F ID	EER* 95/75°F OD 75/62°F ID	EER* 115/80°F OD 75/62°F ID
No Economizer Factory Charge AHRI Rating	4.5	135	12.0	10.0	7.5
Closed Economizer Dampers (Rated at 10 CFM/ft^2)	15.4	462	8.7	6.4	3.7
Economizer Dampers Open 1 Finger "10% OA"	19.5	585	8.3	5.7	2.9
Economizer Dampers Open 2 Fingers "20% OA"	23.2	696	8.0	5.1	2.1
Economizer Dampers Open 3 Fingers "30% OA"	30.1	903	7.8	4.5	1.4
Economizer Dampers Fully Open	62.1	1577	6.6	1.9	-1.4

Source: Intertek, Plano, TX

Note the damper leakage measurement used in the lab is for the entire economizer unit and did not follow the ANSI/AMCA damper test procedure. The leakage includes any leakage around the perimeter of the economizer where it is bolted on to the packaged unit, plus any other cabinet leakage sites. The team set test conditions to replicate field conditions determined at the time lab tests were initiated. The external static pressure (actually return plenum negative pressure) is known to be a factor in economizer operation. Additional testing at varied return plenum pressure is underway and planned.

Table 41 also shows that with economizer dampers closed, the 7.5-ton unit is 38% to 104% less efficient than the same unit tested without an economizer per the ANSI/AHRI 340/360 and

ANSI/AHRI 210/240 test procedures (ANSI/AHRI 2007, ANSI/AHRI 2008, and Carrier 2001).[101]

Tests were performed in compliance with ANSI/ASHRAE 340/360 for a specific set of test conditions. The lab tests expanded the test envelope (ambient conditions) beyond that of the standard, but tested in a manner consistent with the standard. With economizer dampers open from 10% to 30% (1 to 3 fingers), efficiency is reduced by 5% to 62% compared to closed dampers. Reduced efficiency is, of course, expected at the test conditions where the outside air is warmer than the thermostat set-point temperature. Economizers save energy when the outdoor air is cooler than the indoor cooling set-point, and they replace fixed outside air dampers to provide fresh air to satisfy ventilation standards in CA Title 24 and ASHRAE 62. These standards are written in terms of CFM/person and CFM/building square feet, requiring a calculation specific for each building. HVAC technicians rarely know the required building ventilation, and often err on the side of providing too much outside air or closing off economizers. Providing excess outdoor air will reduce efficiency and increase space cooling and heating energy use.

Building occupancies require different minimum amounts of outside air. Laboratory tests indicate closed dampers provide 15.4% outdoor air, 1-finger open provides 19.5%, 2-fingers open provides 23.2%, 3-fingers open provides 30.1%, and fully open provides 62.1%. The SCE statewide program provided three training videos about economizers. The trainer recommended 3-fingers open to achieve 15% outdoor air, but this damper setting not only provides twice the stated outside air but will also reduce cooling efficiency by 11% to 62% if 15% is the appropriate target. Test and Balance (TAB) certified technicians are required to make the appropriate measurements and settings prior to ongoing program maintenance. However, a more appropriate maintenance suggestion would be to seal known cabinet and economizer assembly leakage areas.

5.3.1 Economizer Impacts with Low Airflow

Table 42 provides laboratory test results of relative EER versus airflow and outdoor air damper position for the 7.5-ton two-compressor packaged unit. Compared to optimal refrigerant charge and airflow and closed damper position, low airflow of 16% to 33% reduced efficiency by 2% to 23%. For 1-finger open damper position, low airflow reduced efficiency by 1% to 24%. Note that these findings do not necessarily indicate potential benefit associated with a change in air flow via a maintenance effort. The tested unit can increase air flow via a change in the pulley opening position. Doing so on an existing system would result in an increase in supply air fan power.

[101] Manufacturer EER ratings based on the AHRI 340/360 do not include cabinet or economizer outdoor air damper leakage.

Tests were all performed at a fixed external total pressure. Increasing air flow in the field via pulley adjustments would produce an increase in external pressure and thus fan power.

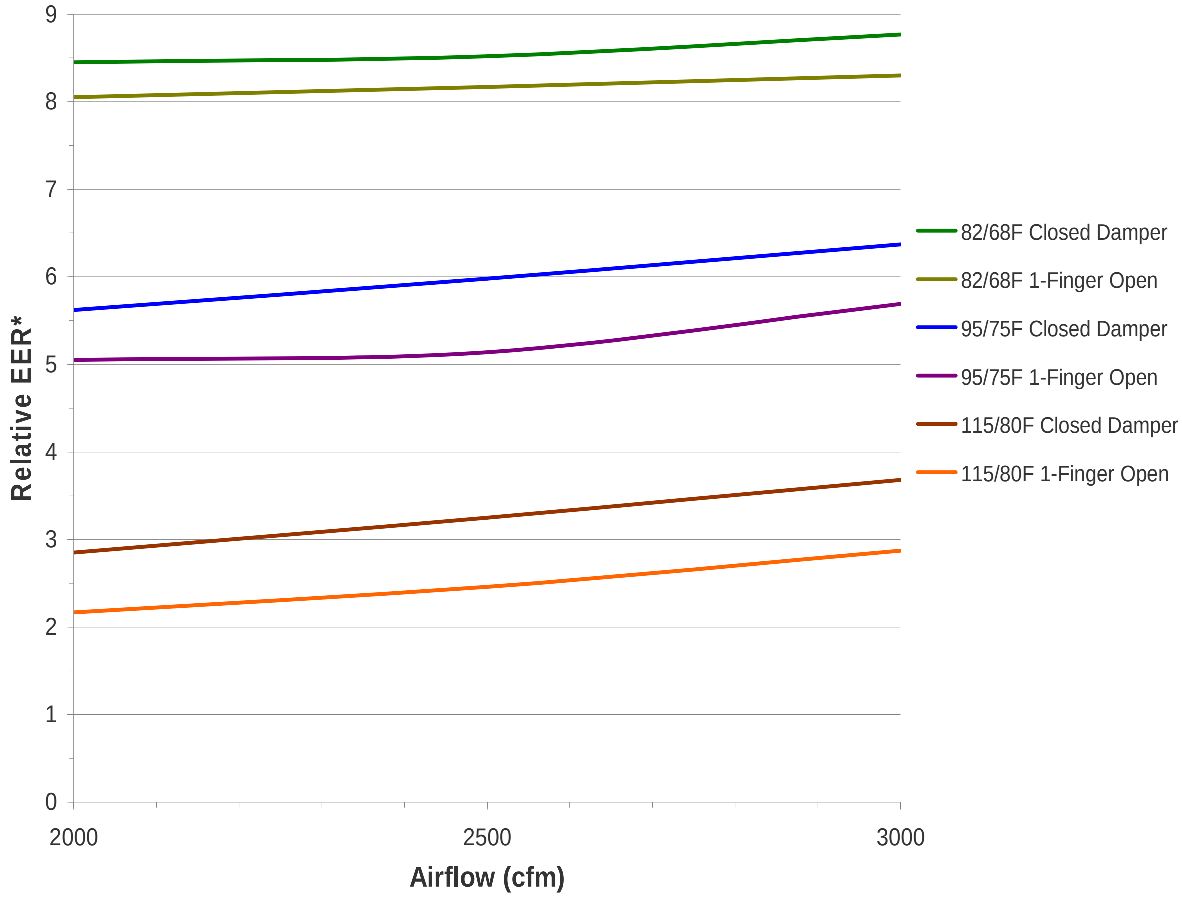
Table 42: System Efficiency versus Airflow and Outdoor Air Damper Position

Test	Refrig. Charge c1/c2 (lb)	Airflow (CFM)	EER* 82/68 OD 75/62 ID	EER* 95/75 OD 75/62 ID	EER* 115/80 OD 75/62 ID
No Economizer Factory Charge AHRI Rating	7.6/8.1	3000	12.0	10.0	7.5
Laboratory Optimal RCA Closed Damper	8.59/10.35	3000	8.7	6.4	3.7
16% Low Airflow Closed Damper	8.59/10.35	2500	8.5	6.0	3.2
33% Low Airflow Closed Damper	8.59/10.35	2000	8.4	5.6	2.9
Laboratory Optimal RCA 1-Finger Open	8.59/10.35	3000	8.3	5.7	2.9
16% Low Airflow 1-Finger Open	8.59/10.35	2500	8.2	5.1	2.5
33% Low Airflow 1-Finger Open	8.59/10.35	2000	8.0	5.0	2.2

Source: Intertek, Plano, TX
*System efficiency, Not EER at AHRI conditions of return air only, relative EER includes outside air.

Figure 23 shows relative EER* versus airflow (CFM) for varying damper positions and outdoor DB/WB temperature conditions. Low airflow has more impact on efficiency in extreme weather conditions, even when dampers are closed. One-finger open dampers have a larger impact on EER* than low airflow.

Figure 23: System Efficiency (EER*) vs. Airflow for Varying Damper Positions and Outdoor DB/WB Temperature Conditions



5.4 Preliminary Conclusions

The laboratory testing provides new insights into unit performance under conditions similar to those found in the field for units in QM programs. Additional laboratory results and conclusions will be included in a separate report as part of the 2013-14 evaluation projects. Results of the 7.5-ton dual-compressor packaged roof top unit tests provided the following preliminary conclusions:

- Unit rating conditions include horizontal flow, cabinet sealing, and no outside air intake. This differs from field conditions where systems often are installed on rooftops with sharp bends to vertical flows, unsealed cabinets, and outside air intake that mixes with return air. In a controlled laboratory experiment the impacts of these conditions on efficiency can be assessed. In the field, the maximum achievable measured efficiency will be impacted by these physical differences.
- Tests of out-of-box efficiency without an economizer indicated an 8.4 EER, which was 24% less efficient than the AHRI rating of 11 EER at outdoor conditions of 90°F dry-bulb

and indoor conditions of 80°F dry-bulb and 67°F wet-bulb. Adding charge improved efficiency but was still 22% lower than the unit's rating. It should be noted that these tests were made with the unit in a down-flow position rather that the standard horizontal-flow arrangement of AHRI tests. Additional tests are being conducted in the horizontal-flow arrangement of AHRI tests. Out of the box tests with AHRI conditions and horizontal flow replicated the manufacturer EER within the allowable tolerance.

- Tests of the economizer open from 10% to 30% indicated that efficiency is reduced by 5% to 62% compared to closed dampers (which deliver 15% outdoor air).
- Outdoor airflow was 15% with closed dampers, 20% with 1-finger open, 23% with 2-fingers open, 30% with 3-fingers open, and 62% with dampers fully open. Designers, technicians, and program implementers incorrectly assume 2% outdoor airflow with closed dampers and 100% outdoor airflow with fully open dampers.
- Minimum damper position of 3-fingers open (30% outdoor air) reduced efficiency by 10% to 62% and reduced economizer savings by approximately 50% compared to closed damper.
- Factory charge tests with and without an economizer indicated that C1 and C2 were above manufacturer charging chart specifications for suction temperature. This indicated an undercharge condition.
- Tests of +60% overcharge (relative to factory) yielded diagnostics within manufacturer ST specifications. However, +60% overcharge yielded negligible efficiency improvements of -8% to +2% compared to +20% overcharge (laboratory optimal).
- Tests of 16% to 33% low airflow indicated a 2% to 24% reduction in efficiency (depending on damper position). Low airflow had more impact on efficiency in hot weather conditions.
- Tests indicated that unintended leakage can have a larger negative impact on EER* than improper refrigerant charge or low airflow. Additional testing should provide more definitive results relative to the impacts of various faults.
- Appendix C provides the results of service instrument accuracy tests. Tests of measurement instruments indicated that it can take 5 to 10 minutes for sensors to measure refrigerant temperatures. Tests were conducted with eight sensors on liquid and suction lines. The smallest differences between reference probe and tested sensors were with specific Type-K clamp probes with accuracy ranging from 1.1 ± 0.6 °F on suction lines at 115°F. Some Type-K clamp probes had suction line accuracy ranging from 6.8 ± 1.0°F when tested at 115°F outdoor conditions. Differences in accuracy are attributable to design and manufacturing. The team found the largest differences relative to reference probe with Type-K bead probes and thermistors. Bead probes had differences of 10.7 ± 3.3 °F, cylindrical thermistors had differences of 9.7 ± 7.1°F, and clamp thermistors had differences of 5.4 ± 2.1°F. The largest differences were with the suction line measurements where tube temperatures were 25°F to 40°F less than ambient. The liquid line temperature is typically 8°F to 12°F above ambient so there were smaller variations from measured temperatures to actual tube temperatures. The recommended pipe

temperature instruments are low tolerance Type-K clamps. Pressure measurement instruments and sensors available for future evaluation measurement require further testing.

Additional testing and studies should evaluate potential savings. A thorough examination should be made that considers an efficiency estimate of "free cooling" that acknowledges the small amount of cooling provided relative to fan power consumption when outdoor conditions are close to the economizer changeover setpoint.

6. Public Comments and Responses

Note that table, section, and page numbers refer to the draft for comment and may have changed. Comments are as posted on the Public Display Area.

#	Name	Comment	Response
1	Adrienne Thomle, Honeywell ECC, Commercial Buildings	The requirement of proper on-going maintenance of the unit is clearly needed. A unit can be operating at peak performance and if not properly maintained can slowly become an "energy waster" rather than an energy saver. We spend a lot of time training technicians on the refrigerant section of a RTU we need to spend more time training the technicians on energy saving devices.	We agree on the needed focus on training for energy savings
2	Adrienne Thomle, Honeywell ECC, Commercial Buildings	There are numerous field studies that have been completed over the past 25 years that prove if economizer controls are applied as recommnded by the manufacturer, the sensors are located in the recommended locations, the dampers are equiped with side and edge seals and the system is properly set up for the RTU then the system will save energy and provide comfort for the occupants including proper ventilation. The issue we have in the field is many technicians do not understand the operation of an economizer as most learn on the job and not in their training. There are a high number of technicians who do understand the operation and provide good results. As the demand for energy savings and codes become more defined for proper application of the economizer systems, suppliers and technicians will continue to improve. Economizer housings will be assembled to low leakage and the controls are now simple for the technicians to understand the setup and operation. I would like to suggest a follow up study completed on unts that have been installed in the past 2 years by technicians that have passed a written and hands on certification.	Recommendation noted on studying new installations. This study focused on maintenance of existing commercial systems.
3	Adrienne Thomle, Honeywell ECC, Commercial	The practice of using "cold spray" to force the OA damper open is not a good practice as it tends to create a "wild" economizer affect and will damage	Most economizers in QM programs that need repair are older than 6 years old.

	Buildings	the sensor with repeated refrigerant spray on the electronic components. The practice of using cold spray was used on electro-mechanical economizers which have not been manufactured for 6 years.	We would appreciate other functional test suggestions
4	Adrienne Thomle, Honeywell ECC, Commercial Buildings	Most RTUs in the years between 2010 and 2012 had old analog economizers installed with leaking dampers. What wa the age of these units? Did they comply with Title 24 - 2013? The requirements were changed to allow for the short comings of the analog economizers with leaking dampers. The new digital economizer controls with low leakage dampers would definitely show better results. With the requirements in Ttile 24 -2013 for Demand Control Ventilation - I suggest the OA damper be set at 0% open and allow the CO2 or occupancy sensor to control increased ventilation based on the number of occupants or allow the use of using cfm input from the technician to set the minimum damper position. Every technician has their own idea on how to set the damper for ventilation and the one, two or three finger or listening for air flow is not the answer but this is what they use today. Technician training and certification is needed in the field.	Exact unit ages are not known in all cases. Existing systems in QM would not comply with Title 24-2013 We agree the rules of thumb are not ideal for setting outside air flow rates, but appreciate the confirmation that those methods are used by some.
5	Adrienne Thomle, Honeywell ECC, Commercial Buildings	There are many reasons the OA damper could be opening in the system and performance of the economizer system is based on the application and location of the sensors. Were the MAT sensors on the old analog units before or after the coil? The location of the coil can make a huge difference in the operation of the system.	The primary issue reported are economizers that do not operate after maintenance. The sensor location is one possible issue. The location varied by unit.
6	Adrienne Thomle, Honeywell ECC, Commercial Buildings	The work order stated that opening the economizer reduced the energy savings, the question is was the economizer damper opening due to a call for cooling and OA conditions were good for economizing or was the damper opening to provide more ventilation based on a call from a CO2 sensor in the sapce?	The statement can be clarified. The economizers were not responding to conditions measured by temperature or CO2 sensors, the dampers were essentially fixed in one position prior to maintenance. The fixed position % open was increased after

			maintenance but the economizer still did not respond to temperature signals that would indicate favorable operation.
7	Mark Lowry WHPA Chief of Staff	Attached is a letter submitted on behalf of six industry organizations represented on the Western HVAC Performance Alliance (WHPA) Executive Committee. In short, it requests an extension of the comment period (to September 22, 2014) to allow additional technical and methodological review of the report. http://energydataweb.com/cpucFiles/0/comments/comAttach_2504.pdf	Extension was granted
8	Richard Hatlen, Joint Journeyman Apprentice Training Center (JJATC)	I was listening in on Monday with all the data and comments and questions. I few things come to mind I know that a great deal of work has gone into putting out a Energy program that will benefit the end user and the folks involved in doing the work. There has been a great deal of data analized and research done. I dont feel there has been a fair sampling of evaluation of the processes. I also feel that and I mean no disrespect but arent we making something more complex then it needs to be? If we want an effective program with a process that works you have to make it less cumbersome to implement and you will get much better results , resultanting in True Energy savings to the customers and the utility will benefit with across the board reductions not just in some areas. Thank You.	The project was designed as an impact evaluation and did not include a process evaluation. Some process work was done for the SCE program based on interim findings of WO32, but the process evaluation recommendation seems reasonable. The recommendation will be raised within the HVAC Project Coordination Group.
9	Darryl DeAngelis, Belimo	The investigating team is correct when they state that combined Economizer and RTU efficiency testing has not been performed in certified testing laboratories. I wanted to make the Commission aware that the Consortium for Energy Efficiency is currently supporting a combined approach for these two technologies as an improved energy rating. It is my understanding that AHRI has formed a committee to work on a combined test standard. This would provide a test method that and metrics	We appreciate this comment and support development of combined testing standards.

		to fully evaluate the RTU with Economizer. I am not arguing that unintended leakage has a negative effect nor that of the intended ventilation air, however what is not highlighted is the efficiency benefit of Economizer and part load efficiency improvements at lower ambient.	
10	Don Tanaka, UA	Nothing was mentioned about the evaluators, how many years of actually working with the tools and any ac certifications or education? What qualifies the evaluators to perform the evaluation? In volume I, page 20 2.2.1, "Observation Methods". Who is a "Master HVAC Tech"? What is his background in the industry	All evaluation firms are vetted by the CPUC Energy Division prior to contracting. The prime contractor has performed numerous evaluations including the previous program cycle evaluation of the RCA and duct seal programs. The Master Technician firm includes staff with over ten years' experience each and all relevant technician certifications such as NATE as well as several manufacturer specific certifications.
11	Don Tanaka, UA	In volume I, page 21, the fourth bullet point. To disassemble and removing panels and condenser fan motors to clean inside and outside of coils. This cannot be cost effective to the owner of the unit (depends on size of unit) to take apart the unit to clean it. There are other ways to do this, and how much can we save (energy) the owner by taking this apart?	The energy savings and persistence associated with following the manufacturer recommended procedure and any alternatives that take less time must be quantified and then cost effectiveness considered.
12	Don Tanaka, UA	Volume 1, page 34, second paragraph. "Most technicians did not install fan belts with proper tension or adjustment". A good technician can tell proper alignment on visual inspection for wearing of the belt or use a string or a straight edge. For proper belt tension, a simple ruler can be used to check belt tension and what is not noted in this report is checking amp draw of the motor. You can have proper belt	We evaluated this measure according to manufacturers' recommendations, which includes checking alignment with a straight edge and checking tension with a tension gage. Power draw is an

		tension but over amp the motor or have a bearing go bad.	important measurement and our inspections did not include this aspect of the fan system.
13	Don Tanaka, UA	Volume I page 43. Reconsider Diagnostic Based Refrigerant Charge Measure. "Identifying and adjusting units with low charges proves to be extremely difficult on commercial units with outdoor air intakes due to problems making accurate measurement of coil entering wet bulb temperature and general unreliably of diagnostics protocols" What about dry bulb reading, it takes two readings to be correct. That is the WB and DB to be the correct way to check a properly charged unit.	It is true that coil entering DB is important, but the statement is based on the fact that unit performance is driven more by WB. Both readings have the issue of determining appropriate sensor placement due to unmixed and temperature-stratified air on inlet side of the coil.
14	Don Tanaka, UA	Volume I page 43, Use manufacture maintenance and diagnostic protocols. "Energy savings potential exists by adjusting fan to factory specifications and following fan-belt alignment and tension guidelines" Not all units have fan belts. There are many 5 ton package units at the mini malls that have direct drive fan units. I have been in the field for 22 years and 13 years behind the desk. Being out so long and new and modern equipment are in the field, I check with the younger techs to see if direct fans are still out in the field. The techs said yes on mini malls and some computer room application.	This is true, but many systems in the programs are larger than 5 ton and have belts. There are also many direct drive small systems that this recommendation does not apply to.
15	Bob Sundberg BNB Consulting	The CQM program evaluation seemed to be missing even a mention of a vital part of implementation of Standard 180, that is, the discussion and evaluation of the agreement maintenance program consisting of the HVAC equipment/component inventory (NOT equipment evaluation but a listing of equipment to be maintained by the service provider) and the maintenance plan. Also missing is any discussion,	The impact evaluation was designed to measure energy savings resulting from maintenance activities.. The evaluation was not designed to study longer term market transformation or market

		suggestion or criticism as to whether there even was establish maintenance "performance objectives" or "condition indicators" which would guide the ongoing discussion of adequate service being provided. Without these key elements, all outlined in detail in Section 4 of the standard, right after the Section 3 Definitions, there is hardly the likelihood of any higher level of QM being continued after the utility incentives end. Far from effectively promoting "market transformation" such a practice encourages the same old practice of bribing end users with utility incentives, only to have them quit the practice after the incentives end. It would be far, far wiser to invest in QM practices which would establish the benefits, energy savings, reduced repairs/service, improved IAQ and equipment reliability, drastically reduced failures and forced replacements, far wiser to work to establish evidence of the undeniable benefits of a QM approach so that after the incentives expire, the QM practice continues. Where did the evaluators address this key element of the CQM programs?	effects, although these could be addressed in a future study.
16	Bob Sundberg BNB Consulting	CQM is a process. Its benefits and repeatability cannot be adequately captured simply in technical terms and measurements. I'll let others focus on argue the merits or demerits of the evaluator's proficiency, or lack thereof, in evaluating QM maintained systems. But, we're talking about the goal of changing behavior here. The degree on the part of building owners, or their designated representatives, to move from total neglect, run to fail or very minimalist service task list approaches to those of adequate inspection and evaluation AND taking action to authorize and remedy "out of acceptable range" operation of their HVAC equipment. I think that a more adequate evaluation approach would have to include interviews with those owner decision-makers or surveys at least to get some idea about how many, what %, to what degree they would consider OR would commit to	The long term approach presents an issue for program evaluation. The evaluation goals were stated to program staff and implementers. Even if the program allows multiple years for service actions the key actions must be taken as soon as possible to be considered first year savings. The process outlined to the evaluation team was that most issues would be identified and repaired early in the process. If it was known that most

		continuing with a QM practice. Where's the focus on that whole market transformation issue? I don't see it in the evaluation at all.	actions would not be identified and addressed until the end of the 3 year agreement then it would be recommended that savings be claimed in later years.
17	Bob Sundberg BNB Consulting	Evaluators seem to be criticizing the QM program, like fault detection and diagnosis, where there is hardly, if any agreement in the industry on what FDD means or how it should be implemented. One example being charge evaluation. Even the "master technicians" retained by the CPUC/ED as experts seem to only quote the manufacturer's original instruction for newly installed equipment correct charge procedures. Those do NOT apply for systems installed one, five, ten or more years with the impact of all the variables and past contractor practices or neglect of the system. Standard 180 does NOT go into any detail on how to correctly assess charge, for instance. Check our Table 5-22, task "n" to see. Check refrigeration temperature and ......which one? How about system pressures which correlate to proper liquid and suction temperatures. Enough said. Others will or could provide better comments on these more technical issues.	We agree there is no single set of FDD standards. We have referenced studies by others on these issues. We agree that Standard 180 does not provide specific guidance on evaluating refrigerant charge. Most important for the evaluation is that because different FDD methods produce different outcomes then the savings associated with charge adjustment may produce positive savings, negligible savings or negative savings. We disagree with the statement that manufacturers' service recommendations only apply to new units. Manufacturers' recommendations represent an independent benchmark to evaluate various FDD protocols.
18	Bob Sundberg BNB Consulting	RQI: I believe the approach to evaluation was incorrect in its comparison target for claimed savings. The evaluation team chose to compare program participants to a sample selection of non-participants. I and many others involved think that the best	This comment speaks to the difference between customer bill savings analysis and program evaluation. Program evaluation attempts to

		comparison would be comparing the newly installed system to the system voluntarily retired or one that was still operating but in a failed or failing mode and most likely guzzling energy in the process. Since such a small sample of home systems were actually evaluated, full performance of the old and new system could be subjected to exactly the same measurements, calculations and performance assessment. The home would have remained the same -- avoiding the necessity altogether of properly selecting some other home of similar size, location, design/layout, altitude, insulation, window and overhang....you get the point. Avoiding all of the variables which skew results and make them questionable at best.	quantify what the customer would have done absent the program (the counterfactual). If the customer would have otherwise chosen not to replace the system, then the previous system is the baseline for the remaining useful life of the existing system. This is the "early replacement" scenario. If the customer would have otherwise replaced the system but not hired a QI contractor, then the baseline is the installation practice for a non-QI job as estimated by the non-participant surveys. This is the "normal replacement" or "replace on burnout" scenario. Note, programs can claim early replacement savings, but documentation of the viability of the existing system must be provided. Once the existing equipment reaches the end of its remaining useful life, the normal replacement baseline is used to estimate savings from that point forward.
19	Bob Sundberg BNB Consulting	I believe the approach to evaluation was incorrect even where non-participants were selected for comparison and calculation of energy usage differences and claimed savings. From my understanding at the recent WHPA RQI Committee September meeting, it was confirmed by the PG&E and CEC representatives that non-participants were	1 – This is incorrect. Non-participants were recruited and fielded prior to the PG&E study. There were 50 sites completed. It is still incorrect to assume that the previous

		selected from a PG&E Compliance study participants. This small select sample of about two dozen homes was intended to represent the broad baseline of residential installed systems. Point 1 - too small a sample to be statistically significant. Try looking at data collected from a much larger pool like the SCE RQI program participating homes for data on systems which were removed and replaced. Best yet would be to evaluate performance of those homes PRIOR to replacement with the collaboration and cooperation of program participants. Now you could build a dbase that could be expanded and improved over time as region, climate, house type and other key variables would allow a far more valid "representative sample" to be initially compiled and continuously improved over time. Better yet, increase that "sample" size by taking advantage of additional industry program home system evaluations to add hundreds or thousands of homes to that dbase. Point 2 - I am convinced that few or no homeowners who knowingly had systems illegally installed, that is, without permits being pulled and system inspections conducted and passed for Title 24 compliance, would volunteer to have state government representative test and evaluate their systems. To me it would seem only common sense that those "illegally installed system" home owners would politely decline participation. Only or primarily home owners who were proud of their systems and contractors, who HAD gone through the Title 24 compliance process, BUT did NOT participate in the program, would volunteer. Whether the selected homes had any installation inspections for compliance could and should be determined. I suspect that many did. This could be determined since the actual sample size is so small. Is it really any surprise that the evaluators concluded that the participants and non-participants they selected had very similar results? So much for the ultra low realization rate	installed unit represents the baseline for program installations. See previous response on replace on burnout and early replacement baseline definitions. 2 - Based on the PG&E followup study we know less than half of non-participants pulled permits. The study itself may show bias towards permitted jobs, but it is not true that non-participants are all permitted. The PG&E study also shows little difference in measured aspects (duct leakage, sizing, and airflow) between permitted and non-permitted jobs. The issue of permit rates and compliance will be studied further in a separate project.

		conclusions. I'd suggest that if the sample was flawed or the wrong target, that all of the conclusions at best are suspect. At worst and likely the best answer, you have to set those conclusions entirely aside.	
20	Bob Sundberg BNB Consulting	Technical issues. Other commenter entries will provide much more detailed and thought through comparisons between how the RQI program contractors have been trained to determine capacity requirements, evaluate home loads and select equipment as well as evaluate and establish the need to correct installed system serious issues such as inadequate return airflow and duct leakage. Issues such as the proper implementation of Manual S & Manual J. Evaluators are suspected of taking measurement shortcuts, relying on invalid assumptions and software defaults, among other questionable practices. The RQI Committee has offered to discuss and even provide the training they conduct with participating contractors for program evaluators. RQI program implementers have stressed that thoroughly and consistently following the procedures (ACCA Standards 5 & 9 as guides) is very difficult and not likely to be mastered by evaluators who do not conduct these protocols frequently. Please refer to the ACCA/IHACI and Roltay Inc. (Buck Taylor) and other RQI program implementer comments for more details.	These referenced additional comments have individual responses.
21	Abram Conant, Proctor Engineering	CQM: The evaluation found no savings for the statewide CQM program and lower than expected savings for other HVAC maintenance programs. The report notes that some program modifications have occurred since 2012, but does not attempt to project the 2010-12 results to current programs. In light of these results, we suggest that a recommendation of urgency regarding evaluation of the current residential and nonresidential HVAC maintenance programs in all IOU territories would be appropriate. We further suggest the field impacts should be of primary concern and greatest urgency. While laboratory activities are instructive	There are plans for both field and laboratory efforts in the 2013-14 evaluation cycle. Given the uncertainties in field measurements, laboratory measurements are critical to providing empirical evidence of the potential benefits. Lab tests are coordinated with field tests to establish realistic test conditions. Field

		for future program improvements, the field impacts are essential to understanding if the current programs are using ratepayer funds effectively to produce measurable energy savings.	measurements are also made to establish in-situ efficiency changes within the limitations of field measurement accuracy.
22	Abram Conant, Proctor Engineering	The CQM program design assumes savings will occur through future maintenance activities. Did the EM&V study find compelling evidence to support this hypothesis? Is there evidence to suggest that future maintenance activities can be confidently projected to be substantially different from and more effective than the activities that occurred during the 2010-12 EM&V study? If so, please detail the findings and describe how confidence estimates were derived.	It was always know that service actions could be performed over several years under the CQM program. The pre-post monitoring had some limitations on duration and coordination with the programs indicated that most of the service would be performed first, followed by scheduled routine maintenance. We made it clear the intention was to get the service actions to be performed within the monitoring period and alert the team when service was performed in order to observe the process. We cannot project benefits for what may have occurred later. The CPUC expects IOU program planners will review and discuss the evaluation findings, including the issue you note, when planning for future programs.
23	Abram Conant, Proctor Engineering	CQM: We suggest modifying the report to be extremely explicit regarding which equipment types the findings are applicable to. Many readers of this report are not HVAC experts, or have only limited understanding of a certain class of equipment. Misunderstanding and misinterpretation of these	We agree that CQM is focused only on commercial rooftop systems. However smaller rooftop systems do have outside air intakes even if

		findings leads policy discussions in counterproductive directions. For example, the findings regarding challenges obtaining return air wet bulb temperature measurements are already being misunderstood and misinterpreted as applicable to residential split system equipment. In reality, these findings are specific to mid to large tonnage nonresidential equipment with economizers. According to the appliance saturation surveys these comprise only 10.4% of all nonresidential systems and 0% of residential systems.	they do not have economizers. The measurement issue mentioned does apply to all units with outside air intakes. A minor note, There is a fraction of residential systems that are small rooftops, but the saturation studies did not evaluate whether fixed outside air dampers were set to closed
24	Abram Conant, Proctor Engineering	CQM: The report body discusses the following program design and implementation findings: a. "Observations of training classes indicate that the programs provide training on how to enter data into the program database, but do not provide procedures for technicians to diagnose and repair faults that lead to improved energy efficiency." (section 2.4) b. "the training did not include functional testing of economizers or repair processes" (section 2.4) c. "Field observations of technicians indicate a lack of understanding regarding how to properly diagnose faults and implement repairs to save energy, both related to program measures and other issues that affect the measure performance." (section 2.4) d. "In one program, program technicians from multiple contractors installed new economizer sensors and new controllers, but the economizers were not set up such that the controller or sensor actually changed the damper position." (Executive Summary) e. "The statewide programs did not record the amount of refrigerant charge added or removed or specific adjustments made to the system such as economizer minimum	Items a through e are quotations from the report without specific questions or comments. f) The report methodology was to evaluate the programs and savings claims. The evaluation was based on the workpapers which only attribute savings claims to a subset of the ASHRAE/ANSI/ACCA Standard 180 checklist. The evaluation did not include a process evaluation and focused on technical issues observed. The conclusions were revised to explain the actions observed that led to the estimated impacts and installation rates. The conclusions cannot make connections to program processes that were not

		position" (section 2.4) – meaning that the program design failed to capture basic information needed to validate energy savings claims f. The EM&V report does not specifically address the effectiveness of ACCA Standard 180 as the basis for energy efficiency program design, but it has been discussed in other venues (California Public Utility Commission presentation to the WHPA and guests, Nov. 7 2013) that Standard 180 is not an energy efficiency standard. The majority of activities specified in Standard 180 do not directly impact energy efficiency. For the activities that can directly impact energy efficiency, no specific procedures or specifications are provided. Interpretations of these activities may vary, and there is no enforcement mechanism specified to ensure any of the activities are completed in a manner that accomplishes energy savings. It is therefore difficult to understand why the Conclusions focus on the refrigerant charge test and do not discuss program design and implementation issues. The programs were built around a standard that was not designed to accomplish energy savings, implementers were observed to provide inadequate technical training, post installation inspections found that work was not done correctly, technicians were observed doing work incorrectly even when they knew they were being observed. Yet the conclusions state that the reasons for the finding of 0% realization rate "center on the programs providing incentives for adding or removing refrigerant charge without diagnosing other faults such as low airflow and refrigerant restrictions or non-condensables." (section 2.5.1, Conclusions, first paragraph). We recommend that the report clarify the directly observed causes for 0% realization rate, including a discussion of program design and implementation issues and their impacts. If in fact, the EM&V work actually confirmed the presence of the stated complicating	evaluated. Comment on procedures – added more emphasis that for all units the evaluation team measured data to evaluate units based on manufacturer procedures. A subset of that data also allows the calculation of comparison diagnostics such as the Title-24 superheat and subcooling metrics on all systems. In addition to the revised conclusions mentioned above a table of the frequency of faults across and within programs was added to the Appendix.

		faults for charge diagnosis (contaminants, line restrictions, etc.) please provide the method by which they were confirmed (sampling of refrigerant, etc.), how often they occurred in the sampled locations, and the distribution of their severity.	
25	Abram Conant, Proctor Engineering	CQM: The 2006-08 HVAC impact evaluation found savings that were below expectations, but above zero. The 2006-08 programs suffered from design and implementation flaws, including: a. Incentives were offered for activities with low probability of producing energy savings. For example very small duct leakage improvements, very small refrigerant charge adjustments, coil cleaning without supporting evidence that coils were sufficiently dirty to impact performance, etc. b. Inadequate, or no technical training requirements for HVAC contractors and technicians. c. Inadequate oversight and technical support to ensure HVAC technicians understood how to perform the work correctly. d. Inadequate quality assurance and quality control to ensure that contractors and technicians participating in the programs were actually performing the work correctly. e. Inadequate enforcement of quality standards. Poor performing contractors faced no consequences and were allowed to continue doing work under the programs. We suggest that a comparison to the 2006-08 programs and results would be useful. In particular it would be helpful to discuss whether the public funds expended redesigning the HVAC programs for 2010-12 and applying the redesign statewide for four years resulted in improved program designs and implementation strategies that successfully corrected the 2006-08 design and implementation problems. a. Is less ratepayer money being spent on activities that don't produce energy savings? b. Has technician training and competence improved? c. Has the quality of work being performed	A full comparison to previous evaluations and evaluations from other regions was not part of the evaluation. The CPUC expects that IOU program planners will take into account all of these study findings when planning for future HVAC programs.

		improved? d. Have robust quality assurance procedures been implemented, including the removal of contractors who fail to perform work according to specifications? This is of particular importance since previous studies have found no statistically significant difference between industry trained and certified contractors and technicians and those with no special training or certification when programs do not implement quality assurance procedures and enforce quality control standards (Northeast Energy Efficiency Partnerships 2006, Strategies to Increase Residential HVAC Efficiency in the Northeast, Appendix C, pages 55-58). "Findings of this study were consistent with recent baseline studies in the Northeast, in that that the quality of the majority of the central AC installations was inadequate. There were no statistically significant differences between installations in new and existing construction or by certified and uncertified contractors, with respect to sizing, airflow, or refrigerant charge at the 90% confidence level. The quality of duct sealing was higher among certified contractors; this difference was significant at the 99% confidence level. However, even the higher quality duct sealing was not compliant with building code, since duct tape rather than mastic was used in many cases." e. Have the delivered energy savings improved?	
26	Abram Conant, Proctor Engineering	CQM: It is very important to deal with the reality of the field, rather than a laboratory construct. A particular fault is worthy of diagnostic effort IF: a. It occurs a significant amount of the time in the field b. The severity of the fault is large enough to produce significant efficiency degradation c. The cost of diagnosing and repairing the fault is less than or equal to the value of the energy saved and reduction in peak demand. An example of these primary principles not being applied is shown in the evaluation. On package rooftop units, 60% of the units received refrigerant	Cost effectiveness is not part of the impact evaluation and therefore we cannot make statements for or against which faults are worthy of diagnostic effort. If you have data supporting the cost effectiveness of various fault diagnosis /repairs, it would be helpful for that to be provided to the IOUs.

		charge adjustments. It is frankly inconceivable that 60% of these units have sufficiently incorrect charge to warrant an adjustment. It is important to make a distinction between the installation standards for a split air conditioner and the service checks for equipment (both package and split) in the field. Service checks in the field must have a wider range of acceptability (decision to leave it alone) to be viable. This program design obviously did not have that protocol in place. It is possible, but unclear if the 60% installation rate refers to the fraction of the instances where refrigerant charge adjustments were reported where the work was found to have been done correctly, or how "correctly" was defined. We request that these findings be clarified, including the specific procedures used to determine the installation rate. We suggest that the report Conclusions and Recommendations, along with discussions in various other sections of the report regarding fault diagnostics, include data showing that recommended efforts are worthwhile. We suggest that "worthwhile" be defined as cost effective for ratepayers as evidenced by data regarding prevalence in the field, severity of efficiency impacts, and cost to diagnose and repair. Conclusions, Recommendations, and other discussions of activities that are not proven to provide cost effective energy saving benefits for ratepayers should be removed from the report.	Unproven methods are suggested for further study given the uncertainties with methods that are currently part of programs. This was also not a process evaluation, which would have assessed the program design issue raised. The section on installation rate was revised and provides additional information. We provided installation rate information specifically to inform the review of workpapers that address a specific set of issues. Fault frequencies and repair frequencies for other energy-related issues are also provided in the revised section.
27	Abram Conant, Proctor Engineering	CQM: The report Conclusions state that: "The program team reviewed all of the detailed data and observations and concluded that the underlying reasons for the result of a realization rate of zero were that the current program has various inherent challenges, many of which were likely unknown until the field and laboratory research under this study are available. The reasons center on the programs providing incentives for adding or removing refrigerant charge without diagnosing other faults such as low airflow and refrigerant restrictions or non-condensables."	Revised the conclusions to focus on the findings based on the field work and laboratory work. The misunderstanding here is that we are assuming that the measures are not capable of saving energy. The actual statement is that assessing the faults and making adjustments in the direction of energy

		(section 2.5.1, Conclusions, first paragraph). We request that above conclusion be substantiated with supporting evidence including measured data, or else that it be removed from the report Conclusions. Supporting evidence should include: a. Data conclusively proving that no energy savings can be accomplished when refrigerant charge is adjusted using the manufacturer specified charge testing procedures if the system also happens to have non-condensables or some other problem. Claims that such adjustments do not produce the exact refrigerant weight specified by the manufacturer are not particularly meaningful in the context of efficiency programs. Energy efficiency programs should be concerned about kWh, not lbs of refrigerant. The relevant question is: "Can energy savings be achieved?", not "Is the refrigerant weight exactly equal to the amount indicated on the nameplate?", or even "Is the EER the maximum it could possibly be?". Pre-Post treatment change in energy performance is the only relevant question. It is important to understand that the reports referenced in the EM&V study do not provide conclusive proof that energy savings are not possible. These reports have been extrapolated well beyond the authors' findings, and have been misinterpreted to infer meanings that are not supported by the data. The data do not show that refrigerant weight must equal the nameplate charge for any energy savings to occur. The data do not show that EER must equal the maximum possible value achieved in a laboratory for any energy savings to occur. The data do not show that energy performance improvements are not possible through refrigerant charge adjustments if another fault happens to be present. The data do not show that energy savings are not possible unless every single fault that could potentially exist is tested for and corrected, or that it would be cost effective to do so. b. Data proving that the incidence and severity of other faults was sufficiently present in the sample	savings are more challenging than previous thought. This is supported by the data and other research that it is much more difficult than previously known to diagnosis commercial unit faults. IOU workpapers assume maintenance activities will optimize system efficiency. Agree that incremental improvements in efficiency are valuable, assuming the benefits outweigh the costs. The impact evaluation attempted to measure incremental savings wherever they existed.

		to significantly impact the potential of HVAC programs to improve efficiency through refrigerant charge adjustments. The data should include the methods by which such faults were confirmed to exist, and the methods by which their severity was measured. The data should be representative of the equipment types addressed by the evaluated programs. c. Evidence that the 0% realization rate was not actually caused by program design and implementation issues.	
28	Abram Conant, Proctor Engineering	CQM: The report states that: "Approximately 92% of technicians had issues with tools or procedures. None had proper tools to evaluate economizers or outdoor air damper position. Approximately 50% did not have EPA low-loss fittings on their refrigerant hoses. Those that did have low-loss fittings often did not purge hoses of air and water vapor prior to attaching to the system. Lack of low loss fittings and failure to purge hoses causes non-condensables or contaminants to enter the system when adding refrigerant or attaching hoses. " (section 2.4) Given the number of times the team observed technicians and their use of fittings without low loss attributes, what volume of non condensables were introduced in to the systems, how was that determined and what is the measured efficiency effect of that amount of non-condensables in the systems of the size tested?	The weight of non-condensables introduced through the use of non-EPA fittings was not measured in the field. This is the subject of current lab tests. Field tests indicated the presence of non-condensables, but it was not possible to quantify the amount. Current lab tests are also quantifying the impact of non-condensables on unit capacity and efficiency.
29	Abram Conant, Proctor Engineering	RQI: Was construction date collected for these homes? If so, we request that data be presented by building vintage as there may be differences in duct leakage and airflow results for different construction dates and applicable building codes.	There are gaps in vintage data which would eliminate relevant data points. Therefore data by vintage is not presented.
30	Abram Conant, Proctor Engineering	RQI: We are concerned that the mean values of participant and non-participant installations were fed into the simulation. A number of factors (airflow for example) have non-linear effects on efficiency and the use of the mean produces estimates different from the actual effects, which will be dominated by the more extreme cases.	The analysis plan to aggregate participant and non-participant data was part of the research plan and the workpaper assumes averages not a distribution of cases.

31	Abram Conant, Proctor Engineering	RQI: It appears clear that there was insufficient quality assurance for the program, since on the average the participant systems' airflow fell 62 CFM per ton below the 400 CFM per ton upon which the program design was based. Had the 400 CFM per ton or better been achieved, the airflow savings should have achieved a higher value than specified in the Workpaper.	No response necessary
32	Abram Conant, Proctor Engineering	RQI: It is dangerous to assume that sizing changes alone will produce energy savings. These are interactive systems and the resistance to flow, surface area, insulation value, and location of the ducts interacts with sizing and airflow. Furthermore, SEER 16 or better systems were near half of the installed systems. Some or many of these systems may have been two speed machines. If they were, oversizing is likely to produce energy savings since the machine will spend more time in its more efficient low speed mode.	The workpaper assumes one oversizing estimate for all installed efficiencies. Our analysis was not setup parametrically to simulate savings for each case. This may be an additional analysis that can be performed in the future.
33	Abram Conant, Proctor Engineering	Upstream: Please clarify the apparent contradiction between findings of the Market Baseline Characterization Study and the Upstream HVAC EM&V Study. The Baseline Characterization Market Effects Study of Investor-Owned Utility Residential and Small Commercial HVAC Quality Installation and Quality Improvement Programs in California found that "40% of single-phase air cooled units and 56% of air cooled three-phase packaged and split equipment sold in 2011 and 2012 meet a Tier 1 or better performance standard.". How do these findings relate to the Upstream HVAC program? If these percentages are truly market baselines, then questions regarding the finding of 0.8 net to gross ratio for the Upstream Program arise. Alternatively if these percentages include a disproportionate number of units within the Upstream Program or other programs that are incentivizing higher efficiency equipment, then the market baseline estimates may assume efficiencies that are too high and not representative of the general market.	We are uncertain on the number of program participants present in the market baseline characterization, but the interview data leveraged may be overlapping.
34	Abram Conant,	Lab: We understand that the laboratory results are	In the results presented,

	Proctor Engineering	still being compiled and we look forward to reviewing them. Please include more information regarding possible causes for performance differences between various system types. For example, the insensitivity of the larger tonnage two circuit units to refrigerant overcharge. Other than general equipment class, is there any other characteristic of these systems that contributes to the finding that these units are sensitive to refrigerant undercharge but are very insensitive to refrigerant overcharge? Is there a receiver or some other refrigerant management component present that is not typically used in smaller systems?	the underlying cause for overcharge insensitivity is the metering device used in the system. The metering device for the particular manufacturer operates similar to a thermostatic expansion valve (TXV) which also mitigates the impact of overcharge on system efficiency.
35	Chandler von Schrader, National Program Manager, ENERGY STAR Verified HVAC Installation	ENERGY STAR currently partners with Southern California Edison (SCE) and San Diego Gas & Electric (SDG&E) to deliver the ENERGY STAR HVAC Quality Installation (ESQI) program. The ESQI program is designed to achieve energy savings by verifying that contractors are installing replacement HVAC equipment in accordance with guidelines based on the Air Conditioning Contractors of America's (ACCA) HVAC Quality Installation Specification (ANSI/ACCA 5 QI-2010) and their verification protocols (ANSI/ACCA QIVP 9). SCE has made tremendous progress in their program's delivery; it is the flagship program for ENERGY STAR's QI program and serves as an important resource for ESQI to share best practices across the country. Successful energy efficiency programs in California often lead the nation to achieve greater energy savings. However, if a California program is perceived as not being successful or cost-effective, it can have a chilling effect on other utilities across the country that might be considering a similar program. As the manager of ENERGY STAR's existing homes' HVAC programs, I've been invited to join the Western HVAC Performance Alliance in order to gain insight on how best to leverage and grow our ENERGY STAR brand inside California's HVAC	The pre and post energy consumption is not equivalent to the savings from quality installation since the baseline for unit efficiency is energy code, not the existing system. See discussion of customer bill savings vs. program evaluation above. A primary conclusion from the study is that some aspects of the program could further exceed energy code, but much more importantly the baseline homes indicate many cases where performance is not as poor as assumed. The larger sample would allow for more analysis to determine the frequency in which there are poor installations and the frequency of installation that already achieve code minimum performance.

		programs. Our recent WHPA calls have focused on Work Order 32. I wish to echo comments I've heard by members of the working group concerned that this study's methodology could have been more rigorous, and as a result the study's findings may be of limited validity. The report's authors seem to agree, as they propose to expand the non-participant sample size "given the relatively small size of the sample in this study." Unfortunately, many readers of the study may jump to the seemingly precise realization rates presented in Table 5 and conclude that HVAC QI programs are not cost-effective, while not realizing that 1) the baseline assumptions were built on a small – and likely unrepresentative - sample size and 2) a more accurate baseline study might significantly alter the realization rates shown in Table 5. I suggest that the realization rates be removed or at least be bounded by an appropriate margin of error based on the appropriate statistical methodology. I certainly recognize the importance of quantifying the impact of residential HVAC replacement and maintenance programs through evaluation, monitoring, and verification activities. Establishing a meaningful and transparent baseline condition to account for savings going forward is challenging. Yet I encourage the CPUC to consider taking a hard look at the actual energy usage (pre and post utility history) of SCE's 8,000 plus HVAC quality installations as a backstop to this and any future Work Order studies. Here's a novel thought: – let's make energy savings decisions by looking at metered data!	For the results in the executive summary the upper and lower bound and sample sizes have been added to the results tables.
36	ACCA & IHACI	CQM: In January of 2014, ACCA and ASHRAE sent a joint letter in regards to the KEMA study's portion of the final report on commercial quality maintenance (CQM) Standard 180 practices. In that letter, ACCA and ASHRAE stated that: "… the Study draws conclusions that are either generalizations or that cannot be verified. Furthermore, important aspects of a complete	No response.

		analysis were omitted." ACCA and ASHRAE recommend further analysis of applications in the field before CPUC undertook any actions based on the report. See the attached letter for the CQM inputs provided in January 2014.	
37	ACCA & IHACI	RQI: Poor Sample Selection: The report seeks to compare installations of program participants and nonparticipants. However, Section 3.2.1 (page 46) describes an inherent flaw, "This program is subject to self-selection effects at two different levels, and at each level it is more likely that the self selection effects will lead to increasing bias in the savings results." (Emphasis added) It later admits that "… the results should be interpreted as most likely being an upper bound" for comparison of program participants to upper tier contractors with very similar installation practices. Therefore these findings, which seek to compare the installations of QI program participants and "typical" nonprogram participants should be rejected.	The report includes this quote and other statements to clearly outline the limitations of the study. In addition, in the report the sample sizes and confidence intervals have been added with additional text on the validity and statistical significance of the findings. The direction of potential bias is acknowledged, but the study recruited customers and got a variety of contractors across tiers and followup study indicated a majority were not permitted.
38	ACCA & IHACI	RQI: Non-representative Baseline Group: It is widely acknowledged that less than 10% of the residential market has permits pulled when existing HVAC equipment is replaced. Hence, utilizing permitted homes (which entails that Title 24 requirements are followed) as the comparative baseline group to utility program participants (entails satisfaction of Title 24 requirements) is misleading and greatly underrepresents the performance degradation of non-compliant installations in California. Furthermore, since the requirements of Title 24 and the ANSI/ACCA 5 QI Standard are well-aligned (i.e., Title 24 embodies the QI requirements), it should be expected that the participant group and the baseline group would have smaller performance differences than if the participant group was compared to a cohort sample drawn from typical	The wide acknowledgement of 10% is not based on a definitive study on permitting or the efficacy of non-permitted installations. Please see the PG&E Permit study which followed this study and shows that the non-permitted installations are also not as inefficient as assumed in workpapers

		California installations (i.e., the non-permitted, other 90%).	
39	ACCA & IHACI	RQI: Incorrect Procedures: Section 3.3.2 (page 50) describes equipment sizing procedures, "… nonparticipants… were assigned the default SHR in the Manual J software used in the analysis." Yet, use of sensible heat ratio (SHR) is not an appropriate method for equipment selection. Manual S procedures only use the SHR to select a preliminary target airflow for the heating / cooling system. The use of the SHR for any other aspect of equipment selection is in violation of Manual S. Furthermore, the verification process for the underlying load calculations were inconsistently undertaken between the program participants and the non-program participants. Non-participant's equipment sizing was based on assumptions or field measurements/observations. Hence, the Manual J load calculations were inconsistently followed and the equipment sizing comparisons were flawed; thereby, invalidating the comparative load sizing results. Therefore, the resulting findings fail to properly evaluate equipment sizing and should be dismissed.	It is incorrect that Manual J processes were inconsistent between participants and non-participants. It is correct that Manual S was part of program participant analysis and not part of non-participants. This key difference is acknowledged in the report
40	ACCA & IHACI	RQI: Comparison of Different Metrics: Section 3.2.3.7 (page 47) describes the data collection protocol to measure duct leakage – the metric is duct leakage to the outdoors (LTO). However, Section 3.3.1.1 (page 49) describes the Title 24 duct leakage tolerances – the metric is Total Leakage. Section 3.3.3 (page 52) states, "Both participant and non-participant duct leakage could be compared with Title 24 requirements and the work paper assumptions on duct leakage." Section 3.3.3 Table 19 (page 52), seems to compare measured LTO metric to the Title 24 Total Leakage metric. The comparison of disconnected metrics should be discarded.	Total leakage and leakage to outside was measured for all sites. We only compare like metrics. This is incorrect that TL is ever compared to LTO. This was clarified in the revised report.
41	ACCA & IHACI	RQI: Failure to Meet the Studies Purpose: Section 3.2 (page 45) states, "The team conducted site visits	It is important to first note that the overall average

		at 50 program participant and 50 non-program participant sites." Section 3.3.3 (page 52) states, "Five non-participants were in climate zones where duct testing and sealing were not required; these sites were excluded from the duct leakage findings." Section 3.3.3 Table 19 (page 52), notes only 36 nonparticipants duct systems that were tested for leakage, while Table 20 (page 53) notes only 32 duct systems were tested. This indicates that an additional 14 - 18 duct systems were not tested, or not reported. A valuable opportunity to evaluate the difference between the installations of participants and California-typical non-participating contractors was lost, or not reported. Because the study failed to meet an important objective, the findings need re-evaluation.	leakage findings do not change when reporting the results for the climate zones without duct leakage requirements. The remaining difference is due to additional exclusions due to testing issues. Note that five participants and nine non-participants were excluded due to QC. Leakage to outside was a secondary test and the difference in sample size between tables was due to unusable blower door conditions making the LTO test invalid. The data collected has been added as an Appendix, but report comparison table is unchanged.
	ACCA & IHACI	RQI: It is noted that Volume 2 (Appendices) of the Draft Report has extensive supporting data for all sections, except the Residential Quality Installation Section. Given the glaring faults in the report, the supporting data might have help third parties to duplicate the results or to better review the validity of the often conflicting, inconsistent, or flawed sections.	Added Appendix with RQI data
42	ACCA & IHACI	Based on the comments above, and the absence of collaborating information, ACCA and IHACI recommend that: • the WO32 report not be released in its current form; • report flaws be addressed and once the subsequent update is available, that	The reporting process in place allows a round of review and comment. Edits to the report based on these and other comments have been undertaken.

		adequate time be provided for dissemination, review, and vetting; and • in the interim, the WO32 report information and purported findings not be used for decision making.	
43	Donald Prather, ACCA	The California Public Utility Commission Energy Division programs are driven by incentives. Those incentives are designed to save energy during peak demand under cooling loads. WO 32 is designed to establishing maintenance and installation methods that will result in higher efficiency levels on peak demand days. This methodology ignores the consumers need to have savings based on monthly billing averages for heating and cooling. Until the utilities have the authority to look at what is the best plan for the home and building owners: peak demand energy saving measures will not be widely embraced and implemented by the market.	Noted, but the evaluation did show issues in achieving annual energy savings as well as peak demand reductions The CPUC expects the IOUs, as program administrators, to research, develop and propose new HVAC programs or modifications to existing ones to the CPUC for approval.
44	Donald Prather, ACCA	HVAC Impact Evaluation DRAFT Report WO 32 HVAC- Volume 2: Appendices 8/11/14 A.1: Performance Changes and Savings Estimates from Pre- and Post-loggers appear to be flawed. Since there is no airflow measurement taken after the service by the data loggers, the formula must use a constant value for CFM/SVsupply-I. Since part of the maintenance is changing filters one would expect the difference between enthalpy temperatures to decrease thus skewing the scores. A better approach would have been to use the fan law to evaluate the airflow based on the supply duct static pressure and the measured CFM. Thus, results on tables 11, 12, and 13 do not reflect the actual operational efficiency based on actual CFM at the times when the data was measured.	Airflow was measured pre and post. Clarification added to the report.
45	Donald Prather, ACCA	Table 24 on Page 61: It appears that energy savings for heat pumps were based solely on peak demand for cooling. From a consumer perspective that may not be the actual annual savings and benefit received. Thus, ignoring annualized energy usage and focusing on peak cooling demand probably does not reflect the annual cost for the consumer.	Since the CA IOUs are "summer peaking," the benefit to the IOUs occurs during the summer when system demand is highest.

46	Donald Prather, ACCA	3.2.3.4 Refrigerant Charge on Page 47 does not state if Superheat and Subcooling were recorded. Pressures alone do not provide the information needed to evaluate system charge.	Superheat and subcooling were measured for all units. Statement in report clarified
47	Donald Prather, ACCA	Section 3 States in the 2nd bullet point on page 45 airflow should be 400 CFM. However, it is not noted what the design CFM should have been based on Manual S equipment sizing guidance (Manual S equipment selection is a code requirement).	Manual S was not a requirement of the 2008 code which pertained to the installations in this report. We are stating that the workpaper assumed 400 CFM in the measure case
48	Donald Prather, ACCA	Section 3 page 45 States in the 1st bullet point that units are typically oversized by 20%. It is not clear if this is over the Sizing limits in Manual S or over the sizing based on load calculations.	Agreed. The workpaper makes this assumption without documentation on whether that is based on Manual J or Manual S.
49	Donald Prather, ACCA	Page 42 2nd paragraph mentions decommissioning economizers as an option. Locking them in the minimum outside air position may provide the best result for obtaining the optimal peak load efficiency for facilities where they see no value in the additional maintenance required by economizers (needed to save energy during off-peak operation hours).	The recommendation referenced says replace economizers, which means replacement with a modern working economizer and not removing or using the non-operating economizer as a fixed outside air damper. De-commissioning is not an energy efficiency measure for programs, the statement is not intended to suggest decommissioning as a measure.
50	Donald Prather, ACCA	Section 2.5.2 page 41notes that contractors see economizer set up as retro-commissioning not maintenance is partially true. Making sure the dampers operate, and open and go to a marked minimum position would be considered as maintenance and technicians would have the tools for doing that as maintenance. However, measuring airflow and marking the minimum damper positions	Agreed that TAB technicians are needed since many units need minimum positions identified. Additionally, technicians should be able to verify the economizer responds to favorable

		requires additional tools, and the higher skill set. In the field that work is generally done by TAB certified technicians or senior start up technicians during equipment commissioning.	outdoor conditions through a cold spray test or similar test procedure.
51	Donald Prather, ACCA	Page 34 3rd paragraph notes that maintenance technicians do not have the tools to set economizer airflows. This relates back to the recommendation for Page 11 bullet 3: Upgrades in economizer controls that include temperature and humidity measurement increase the building owner's savings.	Agree with the statement that upgraded controls when setup correctly may aid technicians with this task
52	Donald Prather, ACCA	Section 2.3 page 27: Performance Changes and Savings Estimates from Pre- and Post-loggers appears to be flawed. Since there is no airflow measurement taken after the service by the data loggers, the formula must use a constant value for CFM/SVsupply-I. Since part of the maintenance is changing filters one would expect the difference between enthalpy temperatures to decrease thus skewing the scores. A better approach would have been to use the fan law to evaluate the airflow based on the supply duct static pressure and the measured CFM. Thus, results on tables 11, 12, and 13 do not reflect the actual operational efficiency based on actual CFM at the times when the data was measured.	Airflow was measured pre and post. Clarification added to the report.
53	Donald Prather	Table 8 on Page 19 does not appear to be justified for the number of pieces of equipment in each type of program. The CQM Package vs the Individual Measures is based on gross program numbers. Not savings per ton, Btuh, or per square foot or some other method for correlating what the numbers reflect based on the equipment. Thus, the claims are meaningless for comparison other than as a popularity contest.	The total savings by program are still relevant given the large differences. Clarification on total tonnage added, but total savings are most relevant for the comparison.
54	Donald Prather, ACCA	On page 12 all three bullet points: The idea is correct, however, the percentages listed are totally unfounded for field comparisons and the 1 finger and two finger damper setting will almost always result in exceeding ASHRAE 62.1 requirements and drive up the peak load cooling cost.	The percentages are based on laboratory measurement and are not estimated. The second statement is true that the rule of thumb methods do not achieve ASHRAE

			standards
55	Donald Prather, ACCA	Page 11 bullet 3 states: "The test results demonstrate that reducing minimum damper position can be a more reliable measure to improve cooling efficiency." Based on my experience in TAB in the field, dampers in the closed position will often provide up to 200% of the ASHRAE 62.1 required outside air. Upgrades in economizer controls that include temperature and humidity measurement increase the building owner's savings. To increase peak demand savings the only metric that will save on the load is to set the minimum outside air properly. Thus, from a utility peak load perspective, incentivizing upgrades on economizers that measure the operating airflow through the economizers offers the maximum energy saving potential.	This statement supports the economizer replacement/upgrade recommendation
56	Donald Prather, ACCA	From the recommendations I am in agreement with the first bullet point on page 9. Since all of the homes compared were Title 24 or title 24 plus QI the common practice bar was probably too high. Thus, the savings that could be achieved if equipment in all homes had permits pulled would be an interesting study. The one question that remains unanswered is: how many of the title 24 only homes were done by contractors that implement the QI standard for all of their installations? Many ACCA members have opted to do all jobs to the QI Standard to eliminate call backs and increase customer satisfaction.	There is not clear information on installing contractor for all non-participants. In many cases the installations were unpermitted based on the PG&E followup study. CALMAC ID: PGE0349.01
57	Donald Prather, ACCA	It was noted on Page 7 in the 1st paragraph that the older DEER calculation method was used based on data collected based on a previous QI work paper. That data and the work paper do not appear to be available for review. It appears that energy savings for heat pumps were based solely on peak demand for cooling. From a home/building owner's perspective that may not be the actual annual savings and benefit received. Thus, ignoring annualized energy usage and focusing on peak cooling demand probably does not reflect the annual cost for the consumer.	Annual cooling savings was included, and the heating savings was not addressed in the either the workpaper or the evaluation analysis. IOU workpapers and documentation of the DEER prototypes are available on the DEER website www.deeresources.com.

			Customer bill savings and program evaluation have separate objectives. See response on program evaluation objectives above in response number 18.
58	Donald Prather, ACCA	Page 5 the 1st paragraph appears to support retro-commissioning before implementing an ongoing maintenance plan. This is a nonstarter due to the expense involved. However, requiring system documentation and installation to meet the QI Standard's requirements on new equipment and using that data as a reference point for maintenance plans would be a viable approach for transforming the market in the future and for providing the information needed to do the maintenance plan design.	The utility programs could consider overcoming this cost barrier, since much of the energy savings result from repairing faults that affect unit efficiency through a retrocommissioning activity. The statement on using maintenance after QI for a long term MT strategy is noted
59	Buck Taylor, Roltay	LAB "The optimal charge was established using manufacturer charging charts where target suction temperature (ST) is a function of suction pressure and outdoor air temperature as shown in Figure 22 (Carrier 2005). The measured ST must be within +/- 5°F of the target ST." Can evaluation team better explain this? The target ST is also a function of mass flow across the coil – what AF setting was the "Laboratory Optimized" charge determined to be, and what was the barometric pressure?	The charging chart conditions are 400 CMF/ton at .25" static for near sea-level conditions.
60	Buck Taylor, Roltay	LAB "The out-of-box verification "A" test with sealed cabinet and 10" pulley was 13% less efficient than the manufacturer-rated EER. The out-of-box factory and optimal charge tests with no seal, 7 inch pulley, and A49 V-belt were 18 to 27% less efficient than the manufacturer-rated EER." Why was the pulley size changed to affect airflow? Is this something an HVAC contractor would do during normal set-up or maintenance?	No this procedure is done in AHRI lab testing to establish static pressure conditions specified in the AHRI standards. It would not generally be done in the field.
61	Buck Taylor, Roltay	LAB "The actual outdoor air leakage with closed dampers was 462 CFM or seven times greater than the 67.1 CFM rated leakage (Table 37)." Can you elaborate on the method to determine leakage	Airflow measured with and without sealing off the economizer. Evaluating HRV was not in

		here? Given the extremely high efficiency hit economizers seem to cause, would it make sense to explore the use of HRV ventilation for some of the cooling zones?	scope, but could be a measure for future IOU program consideration.
62	Buck Taylor, Roltay	LAB "Economizers save energy when the outdoor air is cooler than the indoor cooling set-point and they replace fixed outside air dampers to provide fresh air to satisfy ventilation standards in CA Title 24 and ASHRAE 62." What is the potential savings for the use of economizers versus the penalty?	We did not evaluate potential savings. A thorough examination should be made that considers an efficiency estimate of "free cooling" that acknowledges the small amount of cooling provided relative to fan power consumption when outdoor conditions are close to the economizer changeover setpoint.
63	Buck Taylor, Roltay	LAB "These standards are written in terms of CFM/person and CFM/building square feet, requiring a calculation specific for each building. HVAC technicians rarely know the required building ventilation, and often err on the side of providing too much outside air." Technicians may set them up this way initially, but they are also savvy and practical – usually economizers are closed and disabled immediately after the first cooling, comfort of bill complaint.	This acknowledges a primary issue that could be addressed through programs.
64	Buck Taylor, Roltay	LAB Sect 5.2 & 5.3, Tables 36 and 38: Why isn't there any data showing the factory charge in these series of tests? No technician is going to mess with the charge until he/she is paid to do so (or by a utility program); these tests introduce too many variables (although they are good bits of information).	Additional laboratory report is needed to show more performance at factory charge
65	Buck Taylor, Roltay	LAB Sect 5.2 & 5.3, Tables 36 and 38: If the intention of changing pulleys is to simulate changes in airflow, it would seem to me that you are getting false fan power information – was any attempt made to alter the airflow by increasing flow resistance to mimic an actual application?	Note that certain adjustments are made to pass rating tests and measurements with appropriate static pressures found in the field are part of ongoing tests

66	Buck Taylor, Roltay	RQI "When time permitted, after measuring the air flow in cooling mode, the team switched the unit to either fan only or heating mode depending on the effect on residents' comfort, allowed it to stabilize, and measured airflow again." Why was airflow measured in fan only or heating mode? How was this (above) used in the airflow analysis? Was airflow volume corrected for temperature and air density (altitude)? Without static pressure data we can only speculate whether "poor" airflow results are due to restricted ducts, poor coil choices or improper fan settings. Was Total External Static pressure data collected during the airflow tests?	Pressure data was collected and will be shown in the new RQI Appendix. Additional measurements were taken to assess whether cooling airflow was set differently in other modes, meaning there would be potential to easily increase flow through fan settings. Temperatures were recorded but mass flows were not calculated.
67	Buck Taylor, Roltay	RQI "The refrigerant liquid line and suction line temperatures and pressures were measured with the unit in cooling mode." What instrumentation was used to do this with?	This data was not used or reported. Similar high grade instrumentation as was used in 2006-08 evaluation, including industrial grade Crystal pressure gauges and platinum thermistor temp sensors CALMAC ID: CPU0028.01;
68	Buck Taylor, Roltay	RQI "For split systems, the cooling mode power and power factor for the entire condensing unit and the entire air-handler unit were measured." What instrumentation was used to do this with?	Fluke 345 meters
69	Buck Taylor, Roltay	RQI "Next, the Duct Blaster fan was then attached to the untaped return register and used to pressurize the conditioned space to 25 Pa above ambient pressure. Then the Duct Blaster fan was used to pressurize the duct system to 0 Pa relative to the conditioned space." Did you use the Duct Blaster to pressurize the building in addition to the ducts?	A blower door pressurized the building and ductblaster pressurized the ducts
70	Buck Taylor, Roltay	RQI "…there was no consistent information about what scope of installation work was done or how equipment was selected for non-participant sites. It was unknown if the installations included other changes to duct systems or if furnaces were	Data on fan motor was not explicit gathered but is part of Furnace information based on make and model number.

		included in air conditioning replacements." Was any data collected regarding the size, nominal efficiency rating and refrigerant type of the previous system? Was Fan Type (ie. PSC vs ECM) data collected? Was Fan Speed Setting (Tap or dip switch) data collected?	Fan speed setting types were also collected. All evaluation data is being made available after the final reports as part of the data warehouse effort.
71	Buck Taylor, Roltay	RQI "HVAC replacements within California must comply with relevant sections of Title 24, Part 6, which define the baseline regulatory codes for the program" Can you please confirm or provide a clearer citation for this – the 2008 Title 24 RCM we have refers to HVAC work under section 8.4 (page 8-20).	Footnote reference to Title 24 Residential Compliance Manual added.
72	Buck Taylor, Roltay	RQI Sect 3.3.1.1 pg 46. Title 24 requires a Load Calculation on all replacements. This is not mentioned in the requirements section. How many of the Non-participants that were permitted complied with this requirement?	Evidence of the load calculation was not available
73	Buck Taylor, Roltay	RQI "Residential-sized units are typically available in half-ton increments, so the analysis looked at distributions of units sized within half ton increments higher or lower than calculated." Except there are typically no 4.5 ton units. Additionally, most multi-speed systems (high-efficiency) are only manufactured in 1-ton increments.	Added clarification that 4.5 ton is not available
74	Buck Taylor, Roltay	RQI "The QI programs require the use of both Manual J and Manual S30 for equipment sizing; therefore participants' project folders contained a calculation of the sensible heat ratio (SHR) in accordance with Manual S. The SHR was used by participants' contractors to inform equipment selection." Manual S does not inform or create a SHR. Equipment capacity is a function of outdoor ambient, elevation, and return duct loads (airflow, temperature and grains of moisture). The AHRI capacity rating is performed at 95° outdoor, 80°F dB/67°F wB entering coil with 400 to 450 CFM and no additional duct loads that would modify the entering air density. The Manual S procedure is about calculating the return air entering coil conditions which are not	Manual S was not a requirement of the 2008 code which pertained to the installations in this report.

		only different than AHRI ratings, but may be different than the interior design or space conditions due to return duct location (attics) and leakage. Since the Manual S procedure is an obvious difference from code, why was this ignored in the study?	
75	Buck Taylor, Roltay	RQI "The data in Table 17 show oversized and undersized units in both the participant and non-participant samples." The QIV program requires units at design conditions to fall within 115% of load for A/C and 125% of load for Heat Pumps. How do the units depicted in Table 17 compare to the QIV requirements?	15% of any of the loads will be within a half-ton
76	Buck Taylor, Roltay	RQI "The evaluation team applied Manual S site-specific SHRs provided by the participants' contractors." Please better describe this process by providing an example. What document did the contractor site-specific SHR come from? How was this used to determine oversizing?	Each program participant had a project file including calculations and other documentation. The data source was the program project file. This altered the target value to determine if the program participant unit was installed within one half ton of the Manual S result.
77	Buck Taylor, Roltay	RQI "Since non-participants did not use Manual S calculations, they were assigned the default SHR in the Manual J software used in the analysis." How does this method compare to the actual Manual S procedure used by participant contractors? How do the Manual S capacity calculations compare to the Manual J load calculations for participating sites? If Manual S was used on non-participants, would it result in a different outcome? How do the Manual S capacities provided by participant contractors compare to the AHRI rated total capacity? Did the evaluation team utilize the coil interpolation tool in the RHVAC software to perform any Manual S calculations?	Participant Manual S and Manual J results are shown in the report. Comparison of Manual S to AHRI are exactly what is shown in the final comparison figure. We did not perform independent Manual S calculation using RHVAC

		If yes, how did they compensate for the entering air conditions?	
78	Buck Taylor, Roltay	RQI "The QI workpapers assumed that non-participant units provide airflow of 350 CFM per ton." Can you please define the CFM per ton metric? Is this "per nominal ton", per AHRI rating divided by 12,000 or per Manual S capacity divided by 12,000 BTUH? Can you expand Table 21 to illustrate the above 3 definitions of CFM per ton applied to the sample data?	Cooling tons established per AHRI rating The New RQI Appendix can show all data needed to make the calculation
79	Buck Taylor, Roltay	RQI "There were participant systems with low airflow as well and improving the airflow of those systems would increase program savings." Were tap speeds or dip switch settings noted for any of these tests? The tap speed setting should inform the evaluator to the ballpark CFM. This should then be compared to the designed CFM from the contractors Manual S results. How do participant sites actual CFM compare to the design CFM?	Speed settings were captured We did not compare measured to design CFM
80	Buck Taylor, Roltay	RQI "Table 22 shows the mean power consumption per ton was lower for participants (0.486 W/CFM) than for non-participants 0.569 W/CFM)" What was the distribution of fan types between the respective populations? What were the static pressures relative to the relative populations? What was the change in fan power (pre versus post) and power factor by fan type for the various populations? What affect would altitude have on power draw for an ECM motor? What affect would station pressure and humidity have on the measured CFM for and ECM motor? What affect does a high-efficiency evaporator coil have on ECM fan power consumption versus a standard 13 SEER coil?	Fan type was not an explicit variable so this analysis was not conducted. Recorded pressures added to new RQI Appendix. There is no pre-post information. Sites were in low altitude locations We did not evaluate the variation in fan power based on coil efficiency
81	Buck Taylor, Roltay	RQI "Simulation…. The results were then input into the same single-family prototypes used in the program workpaper…" Can you elaborate on what this simulation is? Is this a spreadsheet analysis or	The IOUs used an eQuest DEER prototype. The same model was used and only the inputs addressed by

		something else? Do all the IOU's use the same assumptions and/or prototypes?	the workpaper were altered. The DEER prototype models are documented on the DEER website: www. Deeresources.com
82	Buck Taylor, Roltay	RQI "For non-participant sites, peak cooling loads were calculated with Rhvac Version 9.33" Can the evaluation team please confirm this version number? The highest (latest) version publically available from Elitesoft and in current use by RQI contractors is Version 9.01.129.	Corrected - 9.01.129 was used
83	Buck Taylor, Roltay	RQI "The analysis team reviewed each of the participant models to confirm that the input values were consistent with the data we collected at that site." Were any differences found, and if so, how were they treated?	The data was consistent as stated in the sentence.
84	Buck Taylor, Roltay	RQI "Ambient conditions were modeled by selecting a reference city based on geographic proximity to the site or, in the absence of a nearby reference city, on the most climatically similar reference city available." How does this method compare to the individual IOU implementation teams methods for design weather data? Were there any differences between the various IOU implementation teams in their design methodology including this parameter? Title 24 requires the use of the design data in Table J2. The QIV standard requires compliance with all code specified requirements. The J2 table data can be several degrees different and altitude may be dramatically different. Can you explain how this may affect your load calculations?	This is the same method, but not all communities are listed in the table. Sensitivity of load calculations to differences in temperature and altitude was not evaluated.
85	Buck Taylor, Roltay	RQI "Within The General Project Data Section, modelers selected duct roughness factors based on the observed predominant duct material (e.g., aluminum, fiberglass, steel), as well as minimum and maximum air velocity based on the observed predominant duct type (rigid or flexible)" What impact does this have on the load calculations?	A sensitivity analysis was not performed
86	Buck Taylor,	RQI "The winter (heating) and summer (cooling)	We did not do a

	Roltay	thermostat set-points were entered from the observed heating and cooling schedule.36 Where schedule information was not available, default-modeling setpoints were 68°F for winter and 80°F for summer." What interior design temperatures are required by Title 24 versus the QIV Standard? Were any deviations in the QIV standard interior design values observed among participating sites? What impact(s) on the load calculation will any differences in this design parameter have?	comparative analysis on actual onsite values and Title 24
87	Buck Taylor, Roltay	RQI "The difference between the air entering the rooms and the inside design temperature of the system is the winter or summer leaving coil-room delta T value (in degrees). The inputs for these modeled values were taken from the difference in supply and return temperature logger readings during the cooling and heating mode operations. In two cases the supply and return temperature information were unavailable for the heating or cooling modes, the default values of 70°F for winter and 20°F for summer were used." What is this information used for in the load calculations or software in general? If the purpose of a load calculation is to design a system prior to its installation, why did you use logged data for the model? Are the default values listed here the same in all the design software suites?	For the ex post evaluation we focus on measured in-situ conditions, as these conditions provide the best estimate the impact of design choices on energy consumption. We are not certain of defaults in all software available as this effort was not part of the scope of the evaluation.
88	Buck Taylor, Roltay	RQI "The building shell infiltration data were entered into the Rhvac single-point blower door method section" What method did the participant sites use? Was infiltration testing performed on participant sites, and if so, were the participant sites load calculation modified to reflect the actual measured data? How did the participant contractors values compare to the evaluation teams measurements?	Infiltration data was recorded and variation between program files and measured value was within the measurement uncertainty of the blower door
89	Buck Taylor, Roltay	RQI "When TrueFlow data on system air flow in cooling or heating modes were available, the modeler changed the default system air type from "auto" to "fixed" and entered the average recorded airflow rate for the cooling or heating mode. When	Measured airflow provides a key metric for the assumed workpaper savings. The airflow element is not directly

		TrueFlow data were not available, the default 'auto' system air type was used, and Rhvac calculated the required airflow rates." Please explain what the "auto" setting provides for airflow compared to the "fixed" setting? How does the evaluation team's method compare to the participating contractors or IOU Implementation methods? If the purpose of the load calculation and system design procedure is to prescribe design airflow (the goal or target CFM), can you explain the justification for using measured airflow? How can your model describe a delta if the goal is the result? How does the measured airflow compare to the initial Manual J design CFM?	measured for participants using the same method so we focused on comparison of like methods between participants and non-participants
90	Buck Taylor, Roltay	RQI "For ductwork characteristics, modelers selected the inclusion of sensible and latent cooling duct gains." If this setting is not selected, how does it affect the load calculation? How does this compare to the methods utilized in the QIV program?	For sizing calculations this was done consistently with program calculations. In addition, this is consistent with program modeling using the DEER prototypes. Sensitivity analysis was not performed
91	Buck Taylor, Roltay	RQI "The duct leakage rate was calculated as the quotient of the flow rate (in CFM) result from the DuctBlaster duct leakage to the outside at 25 Pa and the Rhvac calculated total duct surface area to produce an overall ductwork leakage rate in CFM/ft2" How does this compare to the method used by QIV participants? How does this compare to the Title 24 leakage metric of 15% of Total System Fan Airflow?	Duct leakage for Title 24 is separate in the report, this is just focused on and is consistent with load calculations done by the program
92	Buck Taylor, Roltay	RQI "If the attic R-value was unknown, the modelers selected insulation values based on the vintage of the home and the historical Title 24 attic insulation R-value code requirements40, which are presented in the table below." How many sites used the assumed values versus actual values for attic insulation?	This was collected but is not in the new RQI Appendix data table. All building shell information will be compiled in the data warehouse
93	Buck Taylor, Roltay	RQI "The following window characteristics were used for modeling all window fenestration except	Consistent with program

		skylights: internal shading blinds 45 degree light color and 50% drawn, insect screens outside with 50% coverage, ground coefficient = 0.23 (green grass), and no overhangs or projection included in models.” How does this compare to the method or specifications used by QIV participants?	
94	Buck Taylor, Roltay	RQI “For all modeled sites, internal equipment loads of 2,400 Btuh sensible heat and 350 Btuh latent heat42 were used.” How does this compare to the method or specifications used by QIV participants?	The evaluation reviewed program assumptions in models which did include site variation as well as variation in detail based on which of the two approved software packages were used. The site-collected data was used to develop assumptions for non-participants that were consistent with the level of detail of assumptions in program models
95	Buck Taylor, Roltay	RQI “Internal occupant loads were accounted for by entering the reported year-round occupancy into Rhvac…” How does this compare to the method or specifications used by QIV participants?	Unclear on how program collected occupancy data
96	Buck Taylor, Roltay	RQI “Load-calculation models were created for non-participants, while existing load-calculation models for participants were reviewed for consistency.” Are the load calculation procedures used for non-participants consistent with the methods and specifications used by the participant population? Were any discrepancies found in the participant population calculations?	Yes although all non-participants used one software and participants could use two options. We chose the option that allowed inclusion of all available collected data
97	Buck Taylor, Roltay	RQI “Energy savings for participants relative to non-participants was then determined on the basis of proper system sizing, correct airflow, and reduced duct leakage.” The savings for any projects is driven by the changes in overall system performance. The IOU’s are forced to utilize existing DEER measures to estimate savings. The report does not describe any deltas for the respective populations, nor does it provide any insight into the relative mix of	Permitting is part of the PG&E follow-up study, and will be further studied in an upcoming project. CALMAC ID: PGE0349.01 NP contractor generally unknown but in many cases would be impossible due to location of NPs as

		permitted versus non-permitted projects and assumes the evaluation method correctly estimated loads, target airflows and system sizing. Were any of the non-participant projects installed by QIV trained contractors outside the program? Since Title 24 does not require sizing, wouldn't the evaluation team need to perform the same Manual S procedure to determine if the selected default SHR method described in the paper provides the same results? What affect does the difference in R-22 versus R-410A have on system performance at higher (than AHRI rating) outdoor temperatures? If there is a difference in system performance due to refrigerant properties at higher (or lower) outdoor temperatures, wouldn't this compound the sizing issue?	they were outside QIV qualified contractor service territories. We acknowledge the Manual S calculation difference, but are not performing an additional analysis. R-22 is not used in any of the sampled projects. Updated DEER models that include new performance curves for R-410A systems were not used in the IOU workpapers or this analysis. The impact of the new curves was not specifically studied, but agree they should be used in future studies.
98	Buck Taylor, Roltay	RQI "Participant units were found in single-family homes, duplexes, and apartment and condominium complexes throughout the SCE service area." What was the mix of the various housing types described?	Building type added to Appendix table
99	Buck Taylor, Roltay	RQI "Since the workpaper assumed non-compliance, it essentially presumed a standard practice baseline, not a code baseline. If it was assumed that participants would have met code requirements in the absence of the program it would have been possible to estimate free-ridership." Part of the purpose of this study was to determine a baseline. Please indicate the number of non-participants that meet the Load Calculation requirement in the code? Of the non-participants, how many in this study paid a permitting fee? Wouldn't you be able to determine the Free-ridership issue by defining the baseline case practice for load calculations? Since code doesn't require Manual S, wouldn't this make free-ridership "0" for this element?	See PG&E Permit Study for number with permit documentation on Load calculation – CALMAC ID: PGE0349.01

100	Buck Taylor, Roltay	RQI "Manual S allows up to 140% sizing relative to Manual J which leads to participant systems that don't meet the workpaper assumption that all participant systems are sized to the load." Manual S allows up to 115% for cooling only and 125% for heat pumps and 140% for heating only – at design conditions. Given the evaluators used the AHRI rating total capacity to determine sizing, what affect would using the correct standard (Manual S) have on the results? What percentage of heating systems falls above the 140% threshold due to fan sizing requirements (for cooling)?	We compared Manual J to installed AHRI and then Man S for participants to AHRI. We did not summarize findings for heating
101	Buck Taylor, Roltay	RQI "Non-participants did not systematically oversize, but rather there was a wide range of sizing relative to load calculations." Since no sizing was performed by non-participants we would expect the sizing to load ratios to form a typical population bell curve, did the Evaluation Team determine the size of the system previously installed? Were any building improvements made between system change-outs? If not, is it assumed that the new system is the same tonnage as the old?	Previous system size is irrelevant to the savings calculation. This information is recommended to be collected in the report to document early replacement claims
102	Buck Taylor, Roltay	RQI "There is no data provided for reviewers to perform a proper review and engineering checks on to validate or properly react to the statements made in the report." The Non-participant houses did not get the same level of engineering that was applied in the program, how can comparisons be properly made across the different populations? Please list the training classes and experience of your modelers in the use of the RHVAC software used in the analysis? Did your modelers have access to any EPIC-ACCA certified instructors for support during the modeling efforts? Did any evaluation staff attend any of the various IOU Implementation providers QIV training prior to modeling efforts? If yes, please list dates and details of training. Did the evaluators compare program QC data to	This is begins with a quote of another comment. See previous responses. The program participants had documentation generated by the program that was not present at non-participants. The calculations and data collection for EM&V were consistent for both groups and the evaluation is based on that consistency. The report shows that the Manual S process for participants improves correct sizing. Since non-

		their test data? Please provide copies of load files, system information and site surveys (we will need location information as well) along with any testing of systems.	participants did not have the program documentation requirements their sizing remained relative to Manual J. The evaluation firms were vetted by the CPUC for qualifications related to the evaluation. Personnel information cannot be provided in the public response format. Evaluation staff have been trained on the data collection and analysis software used in the evaluation. The specific courses mentioned were not attended. The program files were compared to data, but QC data was not provided, program data was after QC when provided via request. The data request is noted and all data is provided to the CPUC after report finalization.
103	Buck Taylor, Roltay	CQM Are psychrometric calculations being adjusted for atmospheric pressure?	Yes
104	Buck Taylor, Roltay	CQM Was the Outdoor Air Fraction calculation adjusted to account for fan heat? Please provide an example or illustrative example of how this calculation was performed.	Yes, more detailed site files added to the Appendix
105	Buck Taylor, Roltay	CQM Power factor is an important consideration - if fan speeds were changed or static pressure change due to other measures then power factor may have changed. Input voltage also affects sync speed and	This is unclear, static pressure measured and true RMS power meters were used to measure fan

		slip: was any of this accounted for in air mass calculations?	power, which include consideration of power factor in the metering.
106	Buck Taylor, Roltay	CQM Was voltage data analyzed on 3-phase systems for voltage imbalances, or otherwise low-voltage? Voltage can affect fan and compressor speeds – were any adjustments made for changes in voltage if there were differences between tests? Was logged power measured on one leg, or all legs? Were spot measurements taken with hand-equipment (Fluke) to compare legs and against data suite?	True RMS watt transducers used with power measured on all legs. See Appendix A for site instrumentation documentation. These values were verified with spot true kW measurements.
107	Buck Taylor, Roltay	CQM "Figure 1: Packaged Unit Airside Metering Schematic" Can you explain what the item labeled "Relief Hood" is – what function does this serve? Did the tested units have fixed ventilation intakes with hoods?	Units had economizers and fixed ventilation, the schematic may not represent all units and is intended to be illustrative of measurement points not unit layout
108	Buck Taylor, Roltay	CQM "The metering suite included nine (9) temperature/relative humidity (temp/RH) sensors. In a typical installation, the evaluation team installed two (2) sensors in the return plenum just upstream of the unit's curb mount; one (1) underneath the economizer hood in an area shielded…" I commend the evaluation team for taking this approach; however was any effort given to determine mass flow for each or any of those sensor locations? Are reported airflow measurements adjusted for temperature and altitude or just raw output?	Enthalpy calculations included compensation for altitude and air density. Volumetric flow rates measured for supply flow rate only; no volumetric measurements made at OA intakes or returns. Air density calculations considered temperature and altitude.
109	Buck Taylor, Roltay	CQM "Units were not weighted by program savings, but rather the specific results were turned into an average savings per ton and then extrapolated to the program tonnage claimed." Any chance the analysis shows a pattern between those that showed positive versus negative savings? IE – overcharged versus undercharged, economizer adjustments, airflow adjustments, and coil cleanings?	There is no measure level analysis given the numerous combinations. Revised analysis includes an engineering review for each unit; one could make some generalizations about the types of faults or repairs made to the best and worst

			performing sites, but cross-sectional analyses were not performed.
110	Buck Taylor, Roltay	CQM "Lab testing indicated that manufacturer refrigerant charge protocols provided the best set of diagnostics for all possible faults, but still provided false diagnostics in some cases." Please clarify: does this mean weighing in charge clears other faults (other than charge) or that when the system was analyzed afterwards that other faults were now more easily assessed?	No – this means the manufacturer FDD protocols were most effective in identifying faults set up in the lab. They however were not 100% effective in identifying all fault conditions.
111	Buck Taylor, Roltay	CQM "Field staff and master technicians only assessed installation rates for charge adjustments and economizers given the challenge of assessing coil cleaning months after service." How does the Master Technician determine the effectiveness of coil cleaning? What objective data is used to do so? Wouldn't this data also be important months after the cleaning to determine a fouling rate to help guide the use of coil cleaning as a measure in the future?	We did not report the installation rate on coil cleaning due to the timing of assessments. Activity is planned for 2013-14
112	Buck Taylor, Roltay	CQM "The evaluation team found that a majority of economizers (75%) were not functional after servicing, and that a number of other issues were not addressed by the maintenance, which likely led to misdiagnosis or misadjustment of refrigerant charge." How does the presence of an economizer play a role in refrigerant charge analysis? This statement seems misleading.	This connection is based on lab data. The reason is that FDD targets such as superheat require coil entering conditions, but when these targets are developed it is on systems in labs with no outside air. Once an economizer or outdoor air intake is added to a system, the coil entering conditions are difficult to measure in the field, confounding the results of the FDD tests.
113	Buck Taylor, Roltay	CQM Are the CQM programs utilizing formal FDD methods, or does this statement refer to limited evaluations such as Charge Diagnostics? Can you provide a table of what FDD methods are being applied in the respective programs?	FDD methods can vary for 2010-12 statewide programs. This is clarified and the methods used by third-party /local are

			clarified in the text.
114	Buck Taylor, Roltay	CQM There is no mention of maintenance rates or previous program participation of equipment prior to measures or monitoring. What affect does prior service have on the program? What data can you provide that shows the various faults your team and master technician identified in the field and data logging sets? Was evaluator field data compared to technician reported data in respective IOU tracking systems?	A new summary table provides pre-maintenance issues. Yes the field data was compared where data was provided. There are gaps in the IOU data sets. Existence of stickers indicating prior program participation was noted in the site reports.
115	Darryl DeAngelis, Belimo	The following statement is apparently made without taking in consideration economizer savings, only the waste potential; "If economizer dampers are closed when units are decommissioned, laboratory tests indicate cooling efficiency can be improved by 5% to 140% depending" A non working economizer may waste energy, the resolution is to fix the problem. As noted in a previous comment, the net energy savings need to be identified. The operation of the supply fan (thermostat on Auto or On) will also have a big effect on IAQ and energy waste. The damper only leaks when there is a pressure differential. A damper, if the control and mechanical are properly functioning, will only be at a correct or excessive minimum airflow position when the fan is operating. Future study efforts need to focus on these items and the Net Energy use, not just wasted energy in Mechanical Cooling or Heating mode.	Agree that savings must account for all aspects affected by the program. We re-analyzed the data to look at ventilation and cooling and compressor heating from heat pumps with changes to the report. We did not record gas data.
116	Darryl DeAngelis, Belimo	Certainly not verifying mechanical functionality is a deficiency. Proper training is required coupled with new control with FDD. The FDD will help identify the problems, but is not a complete solutions. Training and technology must be hand in hand. Our FDD will identify if a damper doesn't fully open, it will also identify if supply air is not where is should be. This can assist with identifying a problem. Giving up and decommissioning economizer is wrong approach towards energy savings.	Agreed that decommissioning is not a measure. Replacement of the component is recommended. See comment response #49

117	Darryl DeAngelis, Belimo	One of the largest challenges with a new RTU or existing is to properly set the OA ventilation rate. Such a effort is not QM, it is TAB and Commissioning and requires special tools, skills, and equipment. Without properly applied and tested airflow measuring equipment, every adjustment is a rule of thumb or rough calculation assumption. Even with good airflow measuring systems, there are still accuracy tolerances may provide excess air. The CPUC should be careful with assuming what the ventilation should be. Do not loose focus on building IAQ chasing after energy savings.	Distinction to be made. See previous comment and response #50 and 55 on TAB versus QM.
118	Darryl DeAngelis, Belimo	Note that this study was done prior to the industry introducing and implementation new economizer solutions. Belimo developed the ZIP Economizer to assist the industry in solving many of the problems noted in this report. Belimo would be happy to work with CPUC and IOUs to identify how this new solution (built in tools) can assist in correcting some of the problems, such as setup, commissioning, diagnosis, position. It also should be noted that the IOUs are making an effort to change the CQM program to using new controllers. Although there needs to be improvement in technician training to ensure sustainability as this report has pointed out if a new control is installed without ensuring mechanical functionality, then it isn't worth the effort. Manufacturer certified training should be a pre-requisite.	Noted
119	Darryl DeAngelis, Belimo	Is there any measured data and calculations showing efficiency improvements and degradation related to the following statement "In some cases, there were negative savings due to increased loads from opening economizer dampers, which diminished any efficiency improvement benefits" This report focuses so much on economizers as energy wasters, and they may be if not working correctly, it doesn't emphasize enough the benefits. Making many people suggestion eliminating them altogether or decommissioning them as opposed to ensuring they work. For this report to be effective, the potential energy savings for a properly	The goal is not decommissioning. The goal is to get economizers functioning and understanding optimal setup. Clarification added that we are not suggesting decommissioning. See response #49.

		functioning economizer should be highlighted and only then the net loss of energy for broken economizer or assumed over ventilation.	
120	Darryl DeAngelis, Belimo	Future testing should be performed with actuators that provide position feedback. I have provided the WO32 team with a sample actuator for use in the Plano lab. Note: the new ZIP Economizer Control uses this feedback to provide position. Additionally it scales actuator to damper to always work 0-100% operating range for clearer position identity.	The results of this should be presented by the 2013-14 lab testing study as the ZIP Economizer Control was provided after the WO32 funded lab activities
121	Darryl DeAngelis, Belimo	The statement "closed dampers that deliver 15% outdoor air through damper and other sources of leakage. This amount of leakage, 15%, satisfies the minimum outdoor air requirements for most building occupancies per ASHRAE 62.1 and Title 24." is assuming RTU supply fans are on 100% of occupied time. Often owners run fan in auto position. In this case the space may be under ventilated with dampers closed. Recommend attention to operation modes and IAQ measurements in future studies before determination to close all dampers assuming that the correct ventilation is maintained.	This statement can be clarified. It is correct that fans cycling on with compressors may not meet OA requirements solely through leakage. Agree that outside air rates need to be evaluated and the leakage must be taken into account such that it is known and part of calculations
122	Darryl DeAngelis, Belimo	Certainly a replacement will ensure new parts and expected quality factory assembly, however it will not improve leakage and doesn't solve the problem of service technician lack of knowledge to properly maintain economizer. Without proper training, the cycle will repeat. Additionally, manufactures of economizers aren't currently improving designs for old RTUs. Only if an RTU is currently manufactured would there be potential to obtain 10 CFM leakage design.	The recommendation text is revised to note that factors such as unit age should be considered in determining whether to repair or replace
123	Bob Baker, ASHRAE; EC Co-Chair	Maintenance is a Process; not an Event – Most of the content of the Evaluation Report focuses on discrete technician visits to a facility (one-time events). In developing Standard 180, the consensus body determined early on that Maintenance is a process that is dependent on a close relationship between the facility owner/manager, technicians performing the maintenance and service tasks and others (such as utilities) who contribute to the	The long term approach presents a challenge for program evaluation. The evaluation goals were stated to program staff and implementers. Even if the program allows multiple years for service actions the key actions

		maintenance process and this relationship matures over time. Only through periodic visits to and inspections of HVAC system components and the intervening discussions, reports and authorizations with and from facility management do technicians (service providers) gain the knowledge about and understanding of an HVAC system that leads to greater efficiency, enhanced sustained performance and an optimal indoor environment. In addition, decision makers over time develop the trust and confidence in service providers that leads to the needed authorizations to do what is needed to best enhance operational performance. By concentrating on individual 'service' visits, the evaluators have failed to develop any way to evaluate the success of the long term approach that is the objective of Standard 180 and related utility programs.	must be taken as soon as possible to be considered first year savings. The process outlined to the evaluation team was that most issues would be identified and repaired early in the process. If it was known that most actions would not be identified and addressed until the end of the 3 year agreement then it would be recommended that savings be claimed in later years.
124	Bob Baker ASHRAE; EC Co-Chair	Reliance on Manufacturer Maintenance Instructions – The report suggests that the best source of maintenance information is manufacturer literature. The Standard 180 Project Committee concluded that manufacturer literature fails to be a useful basis for HVAC System maintenance for several reasons: • Manufacturer literature can be difficult to obtain – Owner manuals and other manufacturer provided service guidance is almost never available to service technicians. Such documentation is kept at a location distant to the equipment, may have been lost and, where available, such archives are not well organized. Attempts to access the correct literature over the internet can be frustrating due to the many different models and configurations of equipment. Similar models from the same manufacturer may have very different performance and operational specifications so correct data for one model may not apply to another. • Manufacturer generated service information cannot incorporate field conditions – The essential reality of an HVAC system is that it is a SYSTEM and as	We agree with the difficulty in obtaining the documentation, but do not believe that should deter an incentive program. We also agree that manufacturer instructions are not based on field conditions. It should be noted that manufacturer troubleshooting guides do provide some methods that may work under many field conditions and others that may require incentive programs to develop new methods. Recent lab tests indicate manufacturer protocols perform more reliably than generic protocols. The lab test conditions were established to

		such is the sum of the various components rather than being those components. A given package unit may have very different operating parameters depending on its interaction with other system components such as controls, the air distribution system and ambient conditions where it is installed. Attempts to conform to 'laboratory' conditions as provided in manufacturer literature ignores the influence of all these factors. • For these and several other reasons, the Standard 180 Project Committee (which included several manufacturer representatives) rejected manufacturer provided service instructions as a basis for our useful content to be included in the standard. Correct refrigerant charge, for example, is highly dependent on the specific instillation of a unit and ambient conditions. The 'factory charge' is a starting point and often has to be adjusted to achieve the appropriate performance for the conditions on site. It would be a serious error for the commission to mandate (or even encourage) exclusive use of manufacturer protocols and other service information.	replicate realistic field conditions. An update of the lab test results will be provided as part of the 2013-4 evaluation activities. What maintenance instructions would the Standard 180 Project Committee recommend? If manufacturer instructions are not used then technicians and programs are open to using any method they deem suitable to meet the criteria of the standard. For energy efficiency programs there needs to be specific methods. Other studies (Braun) show issues using other available charge diagnostic procedures and other procedures for faults have been studied less. The report does not recommend exclusive use of any specific method but offers a recommendation that manufacturer protocols performed most accurately, but not without some issues, within the scope of the evaluation. The evaluation team encourages and applauds efforts of development and demonstration of methods that better focus

			on energy efficiency improvements.
125	Bob Baker ASHRAE; EC Co-Chair	Sample size – In every case in the report, the observations and conclusions were based on sample populations that were so small that any statistical reliability is impossible. Although I understand the challenges that the evaluators faced in securing both large populations to study and economic limitations that restricted the scope of the investigations, I am greatly concerned about the failure to acknowledge this inherent lack of reliability when stating conclusions and proposing recommendations. The samples are simply too small to serve as a basis for changing policy direction and this should have been acknowledged in the report.	The issue of sampling and coordination of activities was a mutual issue of both the study and the program. The report does not explain the results of the supplemental observations conducted which show similar issues for top contractors. Overall better coordination and cooperation would have mitigated the primary issues, thus we have reported based on the sample available. The confidence interval is provided, along with a discussion of potential biases in the sample. The small samples will be better acknowledged and supplemental samples explained. It is also acknowledged that a more representative sample with pre-post monitoring was the goal which was not achieved despite concentrated efforts.
126	Dale Gustavson, Better Buildings Incorporated	Commendation: Energy Division is to be commended for its inclusion of so many stakeholders over the months required for this evaluation. Actions of note include distributing the Interim WO32 Reports to the full WHPA member list, co-hosting the 3-hour WO 32 Interim Reports Webinar(s) with the WHPA, custom tailored	No Response Required

		reports/disclosure to WHPA committees and working groups by ED's consultants, WHPA Working Group attendance by ED consultants even when they are not presenting, distribution of the WO 32 DRAFT Reports to the full WHPA Member list, inviting WHPA members to attend WO 32 Reports Public Workshop, and setting up a system for WHPA members and other stakeholders to record comments on the Reports. It is clear that ED has bent over backwards to been inclusive, most particularly in the last six months following the release of the WO 32 Interim Reports.	
127	Dale Gustavson, Better Buildings Incorporated	Recommendation: Future evaluation plans would benefit by including more HVAC industry input into the Project Coordination Group (PCG). The 2010-2102 evaluation plan suffered in ways that could be avoided if HVAC industry SMEs were engaged as guests in PCG meeting/dialogue environments rather than simply been solicited for comments on proposed evaluation plans by email. I will comment further on some if the deficiencies in the evaluation plan below, but here simply make the observation that despite decades of energy efficiency and policy expertise participating in the PCG as currently configured, PCG expertise is not as strong as it could be in the business models/processes/procedures one the contractor/customer aspects of the CQM programs. Many of the difficulties/challenges cited by the evaluators themselves could have been overcome by having had industry input up front. NOTE: This is not a recommendation that the PCG operate under WHPA rules of governance. PCGs have their own rules of engagement.	Efforts have been made to ensure the industry is aware of the release of 13-14 research plans to provide comments on research before it is conducted. CPUC staff have reviewed these comments and replies and note this suggestion. Some PCGs have formed Program Advisory Groups (PAGs) that consist of non-IOU stakeholders. This may be appropriate for the HVAC work as well.
128	Dale Gustavson, Better Buildings Incorporated	Recommendation: Have ED staff or consultants more comprehensively capture oral questions and comments from stakeholders during webinars and workshops and then reply to the questions and comments in writing. While sending replies to individual spoken remarks might not be possible, those replies could be forwarded to the WHPA support team for posting and/or distribution. This recommendation simply builds on another outreach	A summary of stakeholder comments at the September 8 Workshop was compiled, but many more detailed comments were received in writing. We have taken into account comments made at that meeting in our

		effort for which ED is also to be complimented. When the WO 32 Interim Reports were distributed to the WHPA member list, at ED's direction the WHPA support team requested that members to send written comments and questions to the WHPA. WHPA support team organized them into a single document and forwarded them to ED. ED and its consultants provided written responses. Expanding this to include the capturing of more oral comments is important. WHPA volunteers have donated over $3 Million worth of time and expertise to helping implement the California Energy Efficiency Strategic Plan (CEESP). When they sit in two to four to even full-day workshops and make provide constructive criticism or answers to direct questions, those comments should be considered and responded to. This would further expand the engagement of business leaders and subject matter experts in California energy efficiency efforts, a cornerstone objective of Decision 07-10-032.	final report draft. We do not have a summary of all the effort in the meetings mentioned so we cannot respond to comments outside of those provided in WO32 specific venues.
129	Dale Gustavson, Better Buildings Incorporated	Recommendation: Plan ahead to allow more time for written comments and questions especially by non-intervener stakeholders such as those who make up the WHPA membership. To ED's credit, in response to a written request from several stakeholder groups, the comment period was extended. Going forward, it would be better to build in a longer comment period at the outset. Expecting the IOUs and SMEs WHPA member organizations to reply in a matter of days to a report that took a year to write seems to fly in the face of ED's know desire to engage more fully with HVAC industry leaders/institutions.	CPUC staff have noted this comment. A balance is required on report review timelines and delivering more timely results
130	Dale Gustavson, Better Buildings Incorporated	Recommendation: Heed criticism, believe statements of fact, seriously consider recommendations, admit when mistakes are made, and cite important contributions by stakeholders, so that dialogue around EM&V literally becomes a two way street. Though ED's process of transparency and taking input have no doubt had a positive impact of the evaluation plan, substance of	The feedback will affect the analysis and plan of the 2013-14 efforts more so than the 2010-12 efforts since primary data collection between the interim and draft report was minimal

		the report, and the next round of CQM evaluations, those changes/improvements/course directions are not as obvious on the outside looking in because stakeholder input/contributions are not cited. Example: ED and ED's evaluation consultants were told by unimpeachable experts in the November 2013 WO 32 Interim Reports webinar that the consultants were misinterpreting the definition of "Inventory" in Section 4 of Standard 180. This is important because if the definition is misunderstood, the evaluation's observations/conclusions will be wrong. A more constructive process would be to acknowledge that the SME's were correct, that the consultants would adjust their evaluation to align with the facts, and thank the experts (in this case members of ASHRAE and ACCA who wrote Standard 180) in writing. Leaders of the WHPA seem to be asking, "If the WHPA is the official statewide HVAC advisory group, but nobody takes our input/advice, why should we waste our time trying to help." Of the $3 Million of volunteer in-kind consulting contributed by WHPA members since May 2009, $1 Million of effort was put forth in the last 12 months. What a waste if it is dismissed as uninformed, biased, or self-serving.	All comments on public drafts of both plans and reports will be acknowledged Comments and clarifications of the intent of the inventory were very helpful in improving our understanding of the intent of Standard 180. We also appreciate the effort of the WHPA subcommittee in recommending clarifying language to help improve the Standard going forward.
131	Dale Gustavson, Better Buildings Incorporated	Recommendation: Have ED's evaluators and researchers communicate their recommendations for improvements to standards, or training, or products, to the organizations who have the power/responsibility to make the recommended changes, and do so honoring those organizations' rules. For example, the ASHRAE Standard 180 Committee operates under a complex set of ANSI rules. Participation in the development of ANSI accredited standards must be broad-based and open. The watchwords are "transparency" and "consensus." Standard 180 was developed and is being updated for the second time by organizations, not individuals, and must adhere to the ANSI rules mentioned above. Thus comments made in as recently of May 2014 made by an ED evaluator during a WHPA Working Group that the definition	The elements outlined on the venues for providing feedback are noted and appreciated. The evaluation did not intend to evaluate industry standards but rather the programs. Since the programs were viewed by others as initial implementation of new standards then a burden was placed on the impact evaluation to do a process evaluation as well. A full list of the various

		of "Inventory" in the standard "…needs to change because the there is no point in doing maintenance if the fundamental concept of performance baseline lost." Nobody acquainted with Standard 180 or the CQM programs has lost the fundamental concept of performance baselines. Not one. The writers of Standard 180 and contractors implementing it in the field simply believe the word "Inventory" does NOT mean "diagnose all the systems to determine how to bring them to baseline." The inventory is a list of HVAC equipment/systems/parts without which a Standard 180-based maintenance plan can be developed. It is by designing implementing of the maintenance plan for the equipment documented in the Inventory that systems are brought to baseline over the timeframe acceptable/affordable to customers given their budgets and priorities, and that other energy savings measures are added. Maintenance plans can and often do include retro-commissioning activity. Maintenance plans can also include planned replacement (early retirement) of old, inefficient systems. You would not call a replacement "maintenance," but the maintenance plan can stipulate am early retirement. All that being said, if ED's evaluation team still believes the definition of "inventory" and its placement in the Standard need to change, it should respectfully submit comments to the ASHRAE Standard 180 Committee, or work with/through the WHPA CQM Committee or its "Section 4 Working Group."	elements controlled by various industry organization would be an excellent starting point for future studies to avoid generalizations made in the draft report. The 2013-14 studies will work towards identifying responsible groups and requesting additional information as necessary. We feel energy efficiency programs need specific procedures which may not align with the open requirement of the ANSI standard. Specific and transparent program procedures can piloted, evaluated, and refined. These specifics should fit within the overall standard.
132	Dale Gustavson, Better Buildings Incorporated	Commendation: One of ED's consulting team has joined or is planning to join the ASHRAE Standard 180 Committee. This ensures that lessons learned by ED's evaluation team will be offered for consideration to the ASHRAE Committee under the rules of the Society and ANSI processes required for this standard. This is a very constructive development, for which the consultant and ED should be highly commended. Collaborating with ASHRAE and the other HVAC industry institutions that donate their time to develop and maintain	No response required

		Standard 180 is an excellent way for members of ED's team to not only improve the Standard for use in California, but for use throughout the world in places such as NASA's headquarters and in Hong Kong where it has already been adopted.	
133	Dale Gustavson, Better Buildings Incorporated	Recommendation: The IOUs should give consideration to making their comments/responses to ED available to a broader range of stakeholders during the comment period. The WHPA could easily serve as conduit for the dissemination of IOU comments. The IOUs are well within their rights by responding to/commenting on the report directly to ED and, in essence, doing so privately. However, HVAC industry stakeholders would benefit greatly by having access to those IOU comments as they were preparing their own. The IOUs sharing their comments might well lead to having HVAC organizations affirm the IOU comments that they themselves had not yet considered. It is also likely that some comments submitted by the IOUs have also been submitted by HVAC industry organizations. This duplication could be avoided if the comments were shared. This is clearly an IOU decision to make.	CPUC staff have noted this comment and will bring it to the attention of and discuss with IOU staff.
134	Dale Gustavson, Better Buildings Incorporated	Recommendation: Future rounds of evaluations need more clearly disaggregate the narrative so that stakeholders can understand it, learn from it, and improve not only the programs, but practice in the private sector. The organization of Volume 1 and Volume 2 makes it nearly impossible to comment constructively and learn from the evaluation because it is too difficult to determine which programs are being discussed in the narrative. While many of the tables broke out data by program, the narrative did not. For instance I don't know which program's economizers weren't working a year after repairs. I cannot learn from this observation. This is just one of many examples.	New tables added to differentiate where possible. The variation in programs is noted and was not foreseen in planning as we assumed a statewide uniform approach
135	Dale Gustavson, Better Buildings Incorporated	Recommendation: When it is known that programs have already been improved, reports should note that changes have been made and that those changes should be applauded. For instance, the	Changes noted in these and IOU comments have been added into appropriate sections. But

		report recommends that advanced digital economizer controls (ADEC) upgrades replace economizer repairs because the new ADEC systems are far superior technology. This would be an excellent recommendation if it were made in 2010. However, ADEC was included in PG&E's program from the beginning, added to SCE's in 2013, and will be or has been added to SDG&E's programs in 2014 when the evaluation was being drafted. The evaluators know that this is the case, having used pictures of what are purported to be evidence of incorrect installations of the Honeywell JADE ADEC.	the information on 2010-12 has to be reported followed by the known changes.
136	Dale Gustavson, Better Buildings Incorporated	Recommendation: The evaluators, especially those making field observations must be cautious about becoming overly confident that they know everything and that only their opinions are correct. This will be addressed further in my responses to the Research Plan, but one example of mistaken overconfidence is worth documenting in these comments. In the WO 32 Reports Interim Report and in presentations made at the ACEEE Summer Study. As mentioned in #2, above, members of the evaluation team provided pictures of JADE ADEC installations claiming that the absence of two wires in the photo evidenced incorrect installations that rendered the ADEC systems useless and thus not able to save energy. But according to the manufacturer, the wires did not need to be connected for the JADE to be operational. One IOU program implementation team in particular was well trained by Honeywell's local JADE distributor as were contractors participating in the program. At the WO 32 Report public workshop the evaluators stated they learned about how to install the JADE ADEC from manufacturer's instructions. The program implementer and participating contractors had hands on field training by the distributor on top of the manufacturer's instructions. The implementer stands by the quality of the installations and that the economizers are operating correctly.	This is noted. Updated accurate information is required to evaluate the cause of measure performance. The units without the wired connections were not functioning at the time of site visit. The lack of functionality was verified through a functional test; the lack of the jumper was identified as the probable cause. If the instruction manual documentation is not correct then a foundation of evaluating the measures independent of the program is difficult
137	Dale Gustavson,	Recommendation: Future CQM Program evaluation	The Standard and the

	Better Buildings Incorporated	would be more constructive and instructive if ED and its teams of consultants became more fully acquainted with Standard 180—its original 2008 version, its 2012 update and the update activity recently commenced. In particular the ED team needs to become more acquainted with maintenance plans as they are envisioned by Section 4 of Standard 180, with maintenance contracts in general—particularly under what circumstances contractors are authorized to correct faults while on site and what energy impact measures must receive prior customer approval—and generally with how contractors and their customers transact business in the real world. Further, the ED team must learn to ask the right questions of the right people. Given the length and complexities of this recommendation, I will only comment on the last in the above list and address the other facets of this recommendation below. The report references interviews with technicians about whether economizer repairs or upgrades were part of the program or not. Apparently some didn't think they were. However, this particular question is more properly addressed to contractors, not technicians. A maintenance technician is charged with reporting economizer malfunctions. In most maintenance contracts, even a Standard 180-based contract, a complete diagnosis of the economizer, repairs, and/or replacements would require the contractor to provide pricing to the customer and require the customer to approve the work and when that work should occur. It is true that contractors do not always quote this work. They certainly should be doing so in the context of the statewide CQM program. A customer cannot approve an expense until provided a quote. When presented with quotes for more in depth diagnosis of economizer malfunctions, repairs or replacements, some customer still say "No" or "Not just yet." The benefit to ratepayers of leveraging the maintenance industry delivery channel, however, is that the maintenance relationship	workpaper assumptions may not be fully aligned based on this recommendation. The 13-14 direction to move towards individual measure workpapers mitigates this issue of assuming all possible repairs are performed on all units where the workpaper savings are claimed. We suggest our recommendation that enacting the standard as a non-resource market transformation effort seems more appropriate than assuming benefits when revisions to the standard is still being disseminated to contractors and technicians as well as evaluators.

		between the customer and contractor does not end. Good contractors, contractors who have a grasp of Standard 180 and the CQM programs and who are committed to promulgating efficiency will persist in those recommendations. Eventually customers say "Yes" as the result of finally grasping the value of the work being quoted and/or as budgets allow. But the entire process described above is certainly not the purview of field technicians. Tech's thinking economizers aren't a part of the program could be a totally irrelevant observation. NOTE: Were my comments being directed to contractors, I would be critical of them leaving their techs out of the loop, but these comments are not directed at participating contractors.	
138	Dale Gustavson, Better Buildings Incorporated	Recommendation: In the interest of California ratepayers, ED should seriously consider ignoring the field observations and perhaps even the metering aspects of this report and simply view it as one step of many in the right direction. The IOUs should NOT be credited with less than zero savings. I leave it to the professional evaluators and engineers to debate the more technical aspects, but make the above recommendation in all seriousness. There is no way the IOUs should be credited with less than zero savings based upon this report. Though the evaluation team will differ with the following, it is upon simple logic the recommendation is based. 1) The evaluators apparently did not review the Standard 180-described maintenance plans that are required by the program, nor the actual maintenance contracts that govern what contractors can and cannot do without prior authorization. Thus they could not accurately determine "where they were in the process." NOTE: According to Standard 180, the maintenance plan should include an equipment inventory, performance objectives, condition indicators, task tables (primarily 5-12 and 5-22 for the program), and task frequencies. To not review these	1 – All data available for M&V units was requested and this information was not provide proactively to the evaluation team. In some cases participants provided maintenance agreements. More importantly it was described to the team that the pre-post monitoring would capture all measures. In cases where additional measures required approval we returned multiple times to sites but were rarely alerted to when those visits occurred. The M&V process cannot wait for the full three years and efforts to accelerate measures to fit them within the M&V period would have been

		documents and the maintenance contracts of which they should be make a part is a serious oversight by the evaluators. This was observed in comments to ED after the WO 32 Interim Report was issued. The original authors describe Section 4 to be "the heart of the standard." Not reviewing these documents relegates them to being of secondary importance when the writers of Standard 180 clearly though otherwise. 2) If the evaluators were unaware of what aspect of the maintenance plan was being observed, the conclusions they draw cannot be relied upon. Certainly the evaluators/observers could be and likely were right about some of what they concluded, but it is also true that they could be totally wrong and often. As a result, many of the very instructive contributions the report might have made, will remain in question, and rightfully so. Example: Evaluation team members made thoughtful, convincing presentations of their findings at the ACEEE Summer Study. The photos were revealing, even disturbing. But what did the photos actually tell us? One picture was of an obviously filthy, energy efficiency killing heat exchanger. The presenters represented, as does the report, that the technician left the premises with the dirty coils still dirty and that therefore, the quality of work by technicians is terrible. By extension—though not explicitly stated in the ACEEE presentations or the report—there is the suggestion that contractors and customers are collecting incentives for work that is not being done. Perhaps surprisingly given the intensity of this criticism, I couldn't agree more that if contractors and customers are collecting incentives for work not being done—especially if they are reporting the work as completed—they should be kicked out of the program and return their incentives. Further and obviously, QA controls would also need beefing up. However, gaming not what is occurring here. Coil cleaning is not normally done "on the spot" when discovered in	captured by the evaluation. 2 – See above. The M&V goals were clearly stated to the programs in coordination meetings, but efforts were not made allow observations of all activities. The team was ready and willing to return to sites multiple times to observe all measures. The primary limitation of the M&V was the coordination with programs more so than the evaluation of whether the plans were adhered to. 3 – As noted in 1 and 2 in some cases the programs were not designed in a way that would allow effective evaluation. An "evaluability assessment" should be included in program designs so that effective evaluation studies can be planned and executed. Programs would not have been penalized for accelerating service on M&V units or allowing observations on units noted to have completed being brought to baseline after initial pre-maintenance observation. There is an obvious need to improve cooperation with program evaluators and design

		maintenance visits. Normally, if techs visually identify dirty coils, the coil cleaning is scheduled for the future—perhaps the next scheduled maintenance. Quite often the cost of coil cleaning (as is also the case for economizer work) must be approved by the customer because it is not in the maintenance budget. When the evaluators conclude that poor work is being done and that techs are leaving rooftops replete with kWh-gobbling systems unaddressed, it is likely they are wrong and their ultimate conclusions are also rendered mistaken. The efficiency measures need to be addressed pursuant to the maintenance plan and the contract that governs which measures can be corrected on the spot and what measures require additional approvals. 3) The SCE program tracks when RTUs are brought to baseline. Many of the RTUs in the report and among the field observations had not yet been brought to baseline. Some still have yet to be. But the point is that they are being and will be if that is what the maintenance plan calls for…. And it should. The evaluation plan needs to accommodate business models that, according to CBECs soon to be updated data, dominate the commercial sector. 4) I hesitate to include a rumor in comments to the CPUC, but it seems serious enough to do so. It was brought to my attention that some of the RTUs metered for the study may have been installed in RTUs that were not running. Thus, the baseline energy use would be zero. Once the RTUs were made operable the energy consumption would skyrocket. If data of this kind was not eliminated from the sample, it unfairly skews the results in the direction of small or zero savings. One could argue that this rare sort of event might occur could/should have been factored into the CQM workpapers, but this data does not belong in the evaluation, especially given the questionably small sample.	programs with evaluation in mind in the future. 4 – If a unit was not running and had no viable pre-maintenance data it was not included in the analysis
139	Dale Gustavson,	The Energy Information Agency considers	Per recommendation we

	Better Buildings Incorporated	maintenance to be an energy efficiency measure. California's plan to use the maintenance delivery channel to also increase retro-commissioning, energy efficiency measures focused on refrigeration cycle tuning, airflow improvements, controls upgrades, and early RTU retirement is sound. It is a sophisticated strategy rooted in market transformation. It is not a "direct install, widget program." The evaluation plan needs to match the sophistication of the program. I conclude with the vision for the HVAC sector as it is recorded in the California Energy Efficiency Strategic plan. Before doing so, it is important to note that Energy Division management personnel have joined members of the WHPA to change one word in what is recorded in the CEESP. The word is "industry." To keep it simple WHPA and some ED PPT presentations and documents replace the word "industry" with the word "market" which is what I plan to do below. This is totally in alignment with CPUC Decision 07-10-032 that concludes that changes in both HVAC industry practice AND utility programs must take place. The $37 Billion per year HVAC industry is not all that must change. Other market actors including regulators, their evaluators, and IOUs need to change as well. We all need to get better at our jobs. This will take innovation and unparalleled collaboration, cooperation, and communication. The residential and small commercial heating, ventilation, and air conditioning (HVAC) market will be transformed to ensure that technology, equipment, installation, and maintenance are of the highest quality to promote energy efficiency and peak load reduction in California's climate.	feel that the larger market transformation strategy would be best assessed as a non-resource effort allowing focused M&V on specific actions that lead to energy efficiency improvements and take into account adjustments for other non-energy benefits that may increase loads. Also agree with comment that unparalleled levels of cooperation are needed.
140	Don Tanaka, UA	Page B-1 B.1 Question I have is, is SCE statewide (Southern California Edison)?	Yes
141	Don Tanaka, UA	Page B-2 Par.4 If circuit 2 has a failed compressor (12th line), how	The assessment of failure is based on the readings,

		can superheat readings and other readings be obtained when the compressor has failed. It's broken.	the compressor is drawing power
142	Don Tanaka, UA	Also there is no mention on most of the comments from the evaluator if the system is a VAV or constant volume system which makes a big difference on taking readings.	Noted, where available this has been added to Appendix site information
143	Don Tanaka, UA	Page B-3 Fifth line from the top The Trane model number according to Trane alphabet may be missing. As it's written now, it's showing a 50HTZ unit and not a 60 HTZ which makes a big difference in the readings.	All model numbers have been reviewed per this comment
144	Don Tanaka, UA	Also no mention of bringing the compressor to full load amps to try to simulate design conditions and to check the charging chart to design conditions for correct charge.	Measurements were made with data loggers but spot measurements could not control ambient conditions. Diagnostic measurements made with unit in stable full cooling operation, but specific test conditions may differ from design conditions. Charging charts account for off-design conditions.
145	Don Tanaka, UA	B-5 Paragraph 7, fifth line Starts off as "Notched BX42 fan belt was improperly". The fan belt may have been to tight but no mention of checking amp draw of motor. This can be an important factor of the motor and bearings.	These measurements are made but tensioning should be independent assessment. Agree motor amp draw is also an important diagnostic.
146	Don Tanaka, UA	B-13 2nd paragraph, 7th line starting from "used discharge pressure instead of liquid". A very good observation, taking pressures on the liquid line instead of discharge line is a better way to get subcool liquid readings.	This is key as many assumptions are made on difference being 15 degrees but this varies often and manufacturers provide diagnostics based on the ports on their system (often discharge)
147	Don Tanaka, UA	B-19 Top paragraph, 10th line, using the word "superheat" The word superheat is used on R-410A, the word	Noted with footnote, but summary table will indicate superheat

		should be "Temperature glide" for this Refrigerant.	
148	Don Tanaka, UA	B-25 Seems like most techs are not checking for wear and checking voltage drop across the contacts.	Voltage was checked by master technician
149	Don Tanaka, UA	) B-26 Paragraph 4, 2nd line stating "compressor contactor was pitted" Did the tech check for voltage drop across the contacts to verify that the contacts are worn and needs to be replaced.?	Voltage was checked by master technician
150	Don Tanaka, UA	C-1 REFRIGERANT TUBE MEASUREMENT INSTRUMENT TEST Fourth line, " mounted at 10:00 or 2:00 o'clock position" is incorrect. Please see my Attachment on Sporlan Thermostatic Expansion Valves dated June 2011/Bulletin 10-11 for the correct position of the sensing bulb. It should be 8 or 4 o'clock Position.	This is based on which position came closest to the measurement from inside the line in the laboratory. Thanks for noting the discrepancy; could be an error in the product literature.
151	Don Tanaka, UA	C 12 Table 13, Fan Belt Tension and Alignment Measurement Instrument Test Should also have a column on amp draw of blower motor. Again, you can have the correct tension but over amp the motor or destroy the fan bearing.	Noted for lab test reporting, This may not be fully addressed in this report but will be included in 13-14 lab report

SAFER, SMARTER, GREENER

THIS IS DNV GL

Driven by our purpose of safeguarding life, property and the environment, DNV GL enables organizations to advance the safety and sustainability of their business. We provide classification and technical assurance along with software and independent expert advisory services to the maritime, oil & gas and energy industries. We also provide certification services to customers across a wide range of industries.

Combining leading technical and operational expertise, risk methodology and in-depth industry knowledge, we empower our customers' decisions and actions with trust and confidence. As a company, we continuously invest in research and collaborative innovation to provide customers and society with operational and technological foresight. With our origins stretching back to 1864, our reach today is global. Operating in more than 100 countries, our 16,000 professionals are dedicated to helping customers make the world safer, smarter and greener.

In the Energy industry

DNV GL delivers world-renowned testing and advisory services to the energy value chain including renewables and energy efficiency. Our expertise spans onshore and offshore wind power, solar, conventional generation, transmission and distribution, smart grids, and sustainable energy use, as well as energy markets and regulations. Our 3,000 energy experts support clients around the globe in delivering a safe, reliable, efficient, and sustainable energy supply.

For more information on DNV GL, visit www.dnvgl.com.