Appendix O. Energy Division Responses to Public Comments on the 2006-2008 Energy Efficiency Evaluation Report

A summary of the specific changes to the results based on these comments are included at the end of this document.

Source	Comment	Response
SCE	**The Draft Report's Measure-Level Approach Does Not Accurately Quantify Program And Portfolio Level Estimates.** The Draft Report is a failed attempt to address the Commission's direction for estimating the cost-effectiveness of each utility's portfolio. Both the shift from program evaluation to HIM evaluation and the misuses of measurement study and DEER data violate the Commission's direction in D.07-09-043.	Energy Division used the best available information to estimate the cost-effectiveness of each utility's portfolio. The High Impact Measure (HIM) approach was a means to organize the research focus on the technologies that had the greatest influence on the portfolio savings and provide results in such a way to inform the cost effectiveness calculators and tools that have been adopted by the Commission, and faithfully implement all applicable Decisions as outlined in the Decision Framework section of the report (3.4 Policy Direction for Updating IOU Claims). Energy Division is in full compliance with D. 07-09-043.
SCE	The result is a Draft Report and ERT full of inconsistent application of EM&V findings from a measure in one program for one utility to another measure in another program for another utility. This is done without justification of the methods used for extrapolating findings – as required by the Commission in order to calculate PEB.	Energy Division was given clear authority to extrapolate findings as necessary. However, given the HIM approach, there were very few instances where EM&V findings were extrapolated outside of the sample population. These instances are flagged in the database as "OtherEM&V", and their insignificance to the portfolio results is illustrated in Figure 26, Figure 28 and Figure 31. These figures show that these types of extrapolations were applied in 1% of utility reported kWh claims for installation rates, 2% of kWh claims for UES, and 1% of kWh claims for NTG; even fewer cases for kW and therms had this treatment. The policy justification is provided in Section 3.4 of the report and the justification from the contractors regarding the proper application of the result is provided in Appendix C.
SCE	More importantly, reviewers are left with no information on the confidence and precision level of the final portfolio results.	Energy Division has completed an analysis of the portfolio level confidence intervals using a method prescribed in the California Evaluation Framework, which was developed under contract with SCE. This has been included in section 4 in the report.

Source	Comment	Response
SCE	Many of the savings parameters supplied by the measurement studies were created at too general a level to be useful or accurate for applying to other programs and other measures. For example, the savings estimate for a 13 watt screw-in CFL should not be the same as that for a 23 watt screw-in CFL, and yet that is all that is provided in the measurement study. Finally, there are cases where savings parameters simply are not clearly delineated. The purest example of this defect is with residential duct testing and sealing. The evaluation passed through the unit energy savings (UES), but recommended a significant decrease in the installation rate of the measure.	We created delta watts values -- and HOU profiles -- for the entire population of CFLs encountered during onsite metering. The delta watts analysis did not lend itself to disaggregating by wattage, largely due to the sample size (would need fairly substantial sample size for each wattage in order to determine at such a level) and our estimation procedures. The analysis was substantially more complex than this, but essentially we had to take average CFL wattage encountered (by a factor such as room type) and compare with average non-CFL wattage encountered (at the same room type). It was not possible to estimate delta watts at the wattage level without some sort of extensive pre- and post-installation inspection of CFL installations/replacements, which is essentially impossible given the nature of the ULP (no customer information, no installation date, no way of determining who is about to replace an incandescent with a CFL, etc). The HOU profiles were created for all wattage ranges as well, as we were able to meter approximately 3,300 CFLs (not a large enough sample to sufficiently represent the many wattages/wattage ranges listed in the tracking databases. Since these are the two variables in UES estimation, we were unable to produce UES estimates at the wattage level. we didn't want to introduce arbitrary pre/post wattage pairings; rather, we let room and lamp type drive the averages DTS - The UES for duct sealing was passed through because a significant sample of pre and post testing and monitoring was not achieved. The post-only testing for duct sealing verification did achieve the planned sample and was compared to the contractor collected duct leakage percentage to determine the installation rate. The installation rates are described in detail in the Specialized Commercial report.
SCE	**Bias And Statistical Reliability Issues In The Use Of The ERT Process To Apply Study Results**	

Source	Comment	Response
	For programs or measures that were not evaluated, the ERT was based on other EM&V studies, DEER updates, and/or ex ante estimates. In the case of application of other EM&V updates to parameters, the ERT report indicates by reference the precision and confidence level of the updated parameter. This is wrong; these statistical measures of accuracy apply only to the population and program that was studied. Applying them to any other set of participants or program will always result in a greater potential for error. The ERT in many instances makes use of other ex post parameters as an off-the-shelf parameter update strategy without much careful modification to address applicability.	The DRAFT 2006-2008 Energy Efficiency Evaluation Report (Report) states that results from the 06-08 EM&V studies were extrapolated to other programs or measures in other programs only if the results were reliable (i.e., an evaluated result or value that has met statistical expectations based on the study design and professional evaluators can confidently defend and have fully documented in their evaluation reports) and the programs were comparable (e.g., similar types of customers (e.g., NAIC, SIC, size), similar quality control for measure installations, similar building type , similar operating hours, and similar climate). By referring to the achieved precision and level of confidence, the evaluator is only providing the audience with information related to the reliability of the estimate, an important factor in their decision to extrapolate the parameter. The difficulty in meeting these two primary conditions was also noted in the Report. Energy Division reviewed the application of "OthEMV" results and found that contractors were conservative in their application of results. For the Install Rate, UES, NTGR and EUL parameters, only one to two percent of the reported net energy impacts were updated using results from other EM&V studies ("OthEMV"). Specific authority for extrapolating results is presented in section 3.4.
SCE	**Mis-Categorization Of Ex Post Parameter Updates**	
	In many instances, the ERT mis-categorizes the sources of updates to the program parameters for unit energy savings (UES), net-to-gross ratios, installation rates, and effective useful lives as based on "EMV" or "Other EMV," implying that the source of data in all such cases is field-measured data for that program population. In fact, substantial UES estimates developed by ex post measurement studies were modified during the ERT process by applying DEER-modeled interactive effects that are not based on any field data or actual ex post billing analysis, but still the estimate was categorized to be EM&V-based.	SCE has made and Energy Division has responded to more specific comments (that are included in the Appendix of SCE's comments) regarding the mis-categorization of ex-post parameter updates. If the sources of updates were incorrect, and those instances were specifically identified in comments, Energy Division has made these corrections. UES estimates that included interactive effects can be identified in the database by their inclusion in the columns with a small "i" (e.g. UESkwhi) and the affect of this difference is shown in aggregate in section 4 "Comparative Interactive Effects in Evaluated Savings". UES estimates that are categorized as "EM&V" mean that some portion of those savings are based on field research conducted on 2006-2008 programs.
SCE	**The ERT Updates Are Not Consistent With Measurement Study Results**	

Source	Comment	Response
	Lighting measures in the Nonresidential Direct Installation Program (SCE2511) do not match the energy savings estimates of those in the Small Commercial study. The study provides savings estimates by measure and building type for each program. In many cases the ex ante building type was updated, but often the savings estimate matches neither the ex ante nor the updated building type, instead seeming to match with other building types for the same measure, even though those building types are not associated with the specific ERT line item.	Lighting results were applied to each measure by program and building type. The analysis aggregated records from the tracking data into sites. Each site was assigned a building type based on the EDDEERBuildingType as shown in the SPTdb. Sometimes, the records associated with a particular site were assigned different building types at the tracking record level. Only one building type was used for evaluation analysis and therefore the results were applied using that building type.
	In the Integrated School-Based Program (SCE2504), the evaluated UES included an installation rate, and yet the ERT also applies an installation rate less than 1.0, double penalizing the energy savings because attention was not paid to the meaning of the evaluated inputs.	Commenter is correct, on Green Schools/Green Campus, we did apply a realization rate on top of the installation rate SCE had already incorporated into their estimates. The corrected savings are: 306 MWh and 27 kW for Green Campus and 876 MWH and 77 kW for Green Schools. This error has been corrected in the final version.
	Savings from CO Sensors in the MAP Program (SCE2537) were completely mis-assigned. The realization rate from the study is 81%, yet the applied realization rate in the ERT is .000003. While it appeared that the savings unit associated with the UES might have been changed, in fact the unit count remained at 1. So projects that had claimed thousands of kWh	KEMA does not have the document that the commenter references, nor is it clear if the commenter is discussing kWh or kW. KEMA has reviewed Appendix C to the ERT documentation and cannot find the reference to .000003 realization rate mentioned in the comments. However, examining ERT inputs, it looks like the ERT value quoted here has extra 0s and is rounded up from .000249 in the field labeled EDUESkW. That is not a realization rate. The HVAC - Specialized Commercial Evaluation report shows that the realization rate for gross energy savings (kWh) for the four sites metered was 81% (page 193, page 195). The ERT inputs show 1 as the quantity, representing 1 unit. The ERT input sheet doesn't show CO Sensors.
	and were evaluated to have saved 81% of what was claimed were each reassigned a savings of less than one kWh. A similar problem occurred with the Demand Control Ventilation measures in the EE for Entertainment Centers Program (SCE2561), which had their savings changed from the per project value to the per ton value, but their unit count remained at one, rather than recalculating the number of tons per project.	Commenter is correct for SCE2561. The unit energy and demand savings are presented per ton, but the quantity is presented per unit. The ERT units were amended to reflect the tonnage of the unit rather than unit quantity (which is always =1). The tonnage for each record is available in the Standard Program tracking database.

Source	Comment	Response
SCE	The ERT Misapplies Ex Post Parameter Values	
SCE	**The ERT lacks Consistency In Its Treatment Of Ex Post Results**	
	In the Palm Desert Partnership Program (SCE2566), the central AC maintenance measure installation rates are not consistently applied for the same two measure records in climate zone 15, residential sector.	There were two different central AC maintenance measures performed under SCE 2566 that were named the same thing in the E3 calculator but had different names in the program tracking data. One measure was for one-time major maintenance and was evaluated using on-site verifications in SCE2566 for installation rate, measure-specific NTG surveys and UES values from the specialized commercial contract group. The other measure was for maintenance contracts. No direct measurement of savings or installation rate for maintenance contracts was conducted in SCE2566 or elsewhere. As a result, the installation rate and UES were passed through.
	Additionally, these two records have different NTG values (should be 76% instead of the ERT's 69%).	See above for explanation of why these two records were treated differently. The one-time major maintenance measure was given a measure-specific NTG value because there were enough sample points in the NTG survey to do so. The maintenance contract measure was lumped in with all of the residential measures that were not given measure-specific NTG results. These measures were surveyed as their own stratum and received a different NTG than the one-time major maintenance measure.
	UES updates for the same screw-in CFL measures vary in the ERT with reported source to be sometimes DEER, EM&V, or other EM&V.	Results were applied at the building type level. For some building types, it was determined that the results from other contract groups were not applicable to Palm Desert (poor precision for building type, building type not studied, etc.) For those remaining building types, results were mapped from DEER where possible. For the remaining measure-building type combinations, values were passed through.
SCE	**The ERT Lacks Justification In Its Treatment Of Ex Post Results**	
	For many programs with CFL measures, the ERT updates to UES for programs used the data gathered for the Residential Lighting Program (SCE2501). SCE finds this treatment very problematic as the Residential Lighting Program is a mass market upstream program unlike others that target CFLs to certain market segments through give-aways and direct install activities.	Since the ULP provided the most comprehensive information on delta watts, hours of use, and coincidence factors, these estimates were applied to other programs that incented (or gave away) CFLs for use in extrapolating the UES. Direct install or giveaway programs, however, had installation and NTG rates that could vary from ULP and therefore NTG and Irates were not extrapolated between ULP and DLP programs.

Source	Comment	Response
	For the lighting prescriptive measures in the California New Homes (SCE2505) Program, the ERT applied the same UES update based on average CFL to non-CFL wattage ratio that it did for the Residential Lighting Program measures.	Lighting savings for the Residential New Construction program were passed through, they did not use the "same UES update based on average CFL to non-CFL wattage ratio that it did for the Residential Lighting Program measures", unless that is what the utilities had filed.
	Pool Pump measures use SDG&E UES which is based on half as many hours of usage as what is expected in the Palm Desert area. There is other empirical data collected in recent studies including SCE workpapers that show that hours of operation in Palm Desert are very different than that of other mild climate regions that match SDG&E conditions.	The best available estimates for pool pump program savings come from the study in SDG&E. The SCE workpapers do not have any kind of field measurement behind them that documents the difference in operating hours between the two areas.
	There is a minor but persistent realization rate conflict between the input value and the value given at a webinar for SCE's 2517 program (97% vs. 95%). In addition, other small NTG variances between ERT and study values were observed in kWh (but apparently not kW) for SCE2517. In the Comprehensive Mobile Home Program (SCE2502) measures, the therm NTG is different (78%) than that for kWh/kW for interior and exterior lighting fixture.	It is not clear what item in the SCE2517program was assigned a realization rate of 97% or 95% or the NTG variances cited. The values that were used were consistent with the final evaluation reports as cross-referenced in Appendix C.
	For the same interior screw-in CFLs there are three variants of NTG updates for kWh, kW, and Therms (77.6%, 78.5%, and 78.05% respectively). While small, their propagation across large numbers is disconcerting.	Each fuel type NTG was weighted individually based on each survey participant's answers for the NTG battery of questions. The overall weighted averages were based on fuel type (kw, kwh, therms). Therefore the NTG for kw, kwh, and therms SHOULD be different. Applying NTG to therms (since there were no original therms savings declared) was required since there were interactive effects. Cadmus applied an average NTG across kw- and kwh-based NTG values.
SCE	**Transparency Issues In Ex Post Updates In The ERT Process**	
	Because the process of developing energy saving estimates in the impact evaluations and implementing them in the ERT was not collaborative, it is imperative that the implementation of these values be transparent so that all parties can understand how values from studies were applied to measures in the portfolio. The ERT's documentation is seriously lacking in this dimension. Often it is impossible to know where values came from, how values were mapped or the true meaning of the source coding (e.g. "EMV" or "Passthru").	Energy Division was focused on ensuring that the values that were presented in the evaluation reports were clearly and transparently passed through to the ERT final result and held 3 workshops prior to and after the release of the ERT to describe the ERT process, definitions and was available to answer questions. Appendix C is the core documentation provided to ensure this transparency along with the input files. Definitions for EMV, Other EMV and Passthru were provided in section 3.4 and contractors described the use and justification for each in Appendix C.

Source	Comment	Response
	One example of transparency concerns the ERT's treatment of SCE's Industrial EE Program (SCE209). As noted in Attachment A, the lack of transparency and replicability of the site-level analysis led to disagreements with the draft results and adjustments to the final results that raised the realization rate.	This comment refers to the "2006-2008 Evaluation Report for the Southern California Industrial and Agricultural Contract Group." ED notes that the purpose of the review of the draft evaluation reports was to identify actual errors and make corrections so that savings estimates were as accurate as possible, and that rather than a lack of transparency. In addition, there was no lack of transparency, as ED provided all site reports and requested data to SCE and its third party implementers, and met with them to discuss the results. Furthermore, the ERT documentation provided with the SCIA Industrial measures (SCE2509 and SCE2510) shows that the ERT results for the "all parameters" run match perfectly the program-level results reported in the Final Reports.
SCE	In many cases, the Draft Report's Appendix C contains no explanation of how savings values were assigned. For example, the non-HIM programs in the Specialized Commercial evaluation receive no text describing their savings mapping and only one table each showing program level, not measure level, savings.	The documentation in Appendix C from the contractors in combination with the summary tables at the beginning of Appendix C provides information to understand which parameters were updated and the source of the update. For example, the non-HIM programs in the Specialized Commercial documentation only received a EUL update (as cited in the text). The tables in the front of Appendix C confirm this by noting that for each of those programs 100% of the Installation rates, UES and NTG updates are "PassThru" of the utility claims. Measure level savings can be found in the ERT input files.
	Given all these inconsistencies and ambiguities, it is hard to verify the appropriateness of many of the inputs to the ERT. Because of the difficulty in verifying the inputs, it is impossible to know what the true portfolio-level savings values are or how much was truly evaluated.	The summary tables at the beginning of Appendix C illustrate how much of any given program or utility had a direct EM&V result. Figures 26, 28 and 31 of section 3.4 also illustrate by utility and fuel type what portion of the savings was directly evaluated.

Source	Comment	Response
SCE	The Draft Report Does Not Include Benefits From 2006-2008 Codes & Standards Activity	D.05-09-043, OP 14(f), states that "savings from pre-2006 codes and standards advocacy work shall not be counted when calculating net resource benefits ("performance basis") or cost-effectiveness associated with portfolios plans for 2006 and beyond, either on a prospective or ex-post basis. OP 14(e) of the said decision also states that, "On a forward looking basis, savings from codes and standards advocacy work undertaken in 2006 and beyond shall be counted when calculating either net resource benefits ("performance basis") or cost-effectiveness (TRC or PAC tests)." The original analysis in the codes and standards program evaluation report treated savings of a couple of standards that went into effect after 2006 but before 2009 the same as the standards effective as of Jan 1, 2006. The logic was that these were basically Tier 2 of standards developed prior to 2006 or modifications to the pre-2006 standards that resulted from mostly from IOU efforts prior to 2006. One standard went into effect in April 2006, for example, and it seemed most of the effort must have occurred prior to 2006 to get it adopted. We will refer the issue of verifying and counting savings that result from post-2006 Codes and Standards program efforts to the EMV effort of the next program cycle. Savings related to post-2005 C&S support efforts are beyond the assigned scope of the codes and standards evaluation conducted for the 2006-2008 program cycle. As noted in D.07-09-043. Section 9.3.2, the scope of the evaluation efforts was to estimate the "bonus savings" for the 2006-2008 program cycle: "All parties commenting on this issue recommend that 50% of the savings attributed to pre-2006 C&S advocacy work count towards establishing whether the MPS has been met for the 2006-2008 cycle. They also recommend excluding these savings from the calculation of PEB. We find these recommendations to be fully consistent with our determinations in D.05-09-043, as discussed above, and will adopt them. As stated in that decision, for this purpose the C&S savings are to be verified (as opposed to ex ante estimates used for planning purposes). Energy Division's EM&V contractors are in the process of verifying those savings estimates, and Energy Division will be including the verified numbers in its Annual Verification Reports."

Source	Comment	Response
SCE	The Draft Report improperly includes the program costs from SCE's Emerging Technologies program (SCE2515) in the net benefit calculation. Decision 07-09-043 specifies how the net benefits should be calculated, "with the exception of the Emerging Technologies Program and LIEE, all energy efficiency portfolio costs including associated evaluation, measurement and verification (EM&V) shall be included in the calculation of PEB."	The costs for Emerging Technologies and the associated EM&V costs for Emerging Technologies were removed in the final report.
SCE	The Draft Report systematically undercounts the avoided cost benefits associated with SCE's largest program.	
	The ERT however, without documentation, reclassifies all valid climate zones to that of "System," which does not refer to an average temperature climate zone within SCE's service territory, but instead defaults to the lowest value climate zone (i.e. climate zone 8). This has but one consequence – the systematic devaluing of several million dollars in avoided cost benefits associated with SCE's largest program. This modification is inappropriate, not grounded in logic, and furthermore, no rationale for this change was provided.	The zip code to climate zone mapping was based on the CEC look up table. There are additional (new) zip codes that were not in the CEC look up table, and therefore, they were assigned based on the CEC Google map found here: http://www.energy.ca.gov/maps/building_climate_zones.html The majority of changes to the utility assigned climate zones were due to incorrect zip codes in the tracking data. There were many cases of null, out-of-state, and impossible city/zip combinations. These zip codes were cleaned as documented in the ERT process. The clean zip codes were then mapped to the expanded CEC climate zone lookup table. This table sometimes contains multiple climate zones per zip code. Changes were only made if the utility assigned climate zone did not match any of the choices for that zip code in the CEC expanded mapping table.

Source	Comment	Response
SCE	**The Draft Report Contains 538 "E3 Calculator FALSE" Errors Which Assign Zero Avoided Cost Benefits To Valid Measure Installations.** The ERT incorrectly processes SCE's input files through the E3 Calculator in order to calculate energy savings, demand reduction, cost-effectiveness, and PEB.	The comment incorrectly states that the "ERT incorrectly processes SCE's input files through the E3 Calculator". The function of the ERT is to automate the import of data from the ERT input files into the E3 calculator, running the runs, and saving the results of each run. If there is a FALSE flag generated in the E3, then it is because there is a combination of climate zone-Target Sector-Measure Electric End Use Shape not recognized by the E3. This would indicate an issue with the evaluation data and/or the E3, but not a systemic issue with the ERT. Out of the 105 programs evaluated, two had FALSE flags detected by the E3, SCE2501 and PGE2007. For SCE2501, out of the 207,463 lines evaluated there were 538 (0.23%) where the E3 generated a FALSE flag. All of these FALSE flags were due to one Target Sector-End Use Shape combination, Residential:Outdoor Lt. For PGE2007, out of 593 records evaluated, 5 (0.84%) had a FALSE flag generated by the E3 calculator. All of these FALSE flags were due to one Target Sector-End Use Shape combination, Commercial:Unknown. In summary, all of the FALSE flags generated by the E3 Calculator were due to two target sector-end use shape combinations. These have been corrected in the final version.
SCE	The Draft Report errs by completely re-assigning the building types of the customers that participated in SCE's programs. The tracking data submitted by SCE included the appropriate customer building types for those who participated in the programs. Furthermore each measure was further documented by DEER or a workpaper reference that included the appropriate building type. The Draft Report and the ERT completely ignored reality, and instead re-mapped all customer building types based on NAICS (North American Industry Classification System) codes. This process questions the validity of the NAICS codes and the mapping approach used.	A standardized building type assignment was necessary to support several items, including the assignment of EDTargetSector, supporting contract group sample designs, and informing the application of interactive effects. Because IOU building type data was frequently missing, highly vague, and non-standardized, the ED team produced standardized building types based on the DEER building type naming scheme. The process used for mapping IOU building type data to standardized DEER building types is described in detail in the SPTdb documentation. Per the documentation, NAICS codes were rarely used in this mapping process, as the IOUs reported unreliable NAICS data. A significant number of line items received program weighted-average building types, as there was insufficient data to confidently map these

Source	Comment	Response
		line items to a DEER building type.
SCE	**The Draft Report Alters The Location Of SCE's Measure Installations By Unknown And Unwarranted Parameters**	
	In SCE's measure-level reporting, it relied upon the latest zip code to climate zone mapping received from the California Energy Commission (CEC). However, the Draft Report relies on a completely different and un-documented methodology to map customer location zip codes to climate zones. The ERT instead links the customer zip code to its own lookup table (with no reference to where it is from). Since the CEC is the definitive source on this issue, the Draft Report should utilize the CEC look up table, as SCE did, and not an undocumented and un-vetted source.	The zip code to climate zone mapping was based on the CEC look up table. There are additional (new) zip codes that were not in the CEC look up table, and therefore, they were assigned based on the CEC Google map found here: http://www.energy.ca.gov/maps/building_climate_zones.html The majority of changes to the utility assigned climate zones were due to incorrect zip codes in the tracking data. There were many cases of null, out-of-state, and impossible city/zip combinations. These zip codes were cleaned as documented in the ERT process. The clean zip codes were then mapped to the expanded CEC climate zone lookup table. This table sometimes contains multiple climate zones per zip code. Changes were only made if the utility assigned climate zone did not match any of the choices for that zip code in the CEC expanded mapping table.
SCE	**The Draft Report Should Include A Disposition Of The Residential Lighting Program's Un-Installed Bulbs**	

Source	Comment	Response
	The Draft Report makes a determination that a substantial percentage of CFLs delivered upstream were either purchased and not installed or not purchased at all. SCE disagrees with the methodologies used by the measurement study to reach its conclusions. The Draft Report acknowledges that "bulbs sold at a later date may still result in future energy savings." A customer is expected to eventually install all of the purchased CFLs.	In accordance with current Commission policy, the 2006-2008 impact evaluations only credited the IOUs for impacts associated with measures installed and operable within IOU service territories by year-end 2008. As stated in D.05-04-051: "For these reasons we will require that the savings and resource benefits associated with installations completed in a given year, regardless of the year in which any given installation was funded, will be counted towards the performance basis for that program cycle. Nonetheless, we will require the IOUs to report and track both installations and commitments for each program year. This information will be useful for resource planning purposes and enable us to link program activities with a particularly funding cycle, as needed." CFLs that were purchased and not installed during 2006-2008 can be considered in future evaluations and savings associated with these installations credited to the IOUs in the program cycle that they occurred.
	While this is the logical conclusion, the Draft Report makes no estimation of or recommendation on the disposition of the "to be installed" CFLs. The wholesale elimination of the savings and benefits associated with these CFLs undercounts the effects of SCE's program. SCE recommends that since utility programs incurred the costs within the 2006-2008 program cycle, it is appropriate to consistently provide the PEB credit in the 2006-2008 true-up (consistent with where the program costs were incurred) and provide the energy savings credit in the year where those CFLs are eventually installed. Conversely, the utilities should receive energy savings credit in 2006-2008 for those CFLs that were purchased as a result of the 2004- 2005 program, but not installed until 2006. This is consistent with the Commission's stated intent to use CFLs to fill the cumulative goal gap created from CFLs dying faster than the EUL assumption used to set the goals.	See response above. As to savings from CFLs purchased during the 2004-2005 program cycle, the Commission policy at that time includes both actual installations and commitments in the determination of energy savings for 2004-2005; hence, those savings have already been accounted for in the cumulative savings determination.

Source	Comment	Response
SCE	The Draft Report is out of compliance with Commission policy on the use of the Database for Energy Efficient Resources (DEER). Specifically, the Draft Report should use DEER 2008 v2.04 to measure the 2006-2008 program cycle and the 2009 bridge funding programs. However, the Draft Report uses DEER 2008 v3.02, a version of DEER that has not yet been fully released much less fully vetted, for the determination of the interactive effects.	Energy Division did not use version DEER v. 3.02. It does not exist. In 2009, to better estimate interactive-effects, the ED DMQC Team, with assistance from the ED DEER Team, provided an interactive-effects spreadsheet for the ERT Team. Several heating and cooling system types were added to the DEER dataset, and air-conditioning and heating saturations were applied which mitigate the negative therms impact. Additionally, a couple of errors identified in the DEER 2008 analysis software tool were corrected. Energy Division has the right and responsibility to enhance interactive effects estimates in the DEER Database, and to correct errors identified.
SCE	The Draft Report is the epitome of a nontransparent document that only serves to confuse and debilitate the review process and increase the uncertainty about the accuracy of the report. The report itself is 127 pages, coupled with 16 attachments, utilizing measurement studies of over 2,200 pages and a software data tool of more than 4 million tracking records totaling 4.2 gigabytes. Given the overwhelming complexity and limited review time and guidance, SCE's comments only break the surface of the issues that are contained in the Draft Report. However, it is clear, even after a limited review period, the Draft Report and the ERT are so systematically flawed that they must not be used as a thoughtful assessment of SCE's 2006-2008 program accomplishments.	The Draft Report synthesizes into 127 pages (ES of 20) 3 years of program implementation, and evaluation across four IOUs and the present the final outcomes of X billion in ratepayer investments. The multiple attachments are data and tools provided to allow for detailed review by stakeholders and most pieces (Contractor Reports, Decision Framework and ERT) have been introduced to the public in advance of the report release. The largest and most complex portion of the data (over 4 million tracking records) was provided by the IOUs and standardized (Standardized Tracking Database) in collaboration with ED consultants over the course of a 3 year period.
SCE	The Draft Report should be corrected to accurately characterize the results of the three audits the CPUC conducted on IOU program costs throughout the 2006-2008 program cycle. In each case, SCE's costs were deemed to be accurate and reasonable. However, the Draft Report does not acknowledge this fact and instead misstates that the 2008 program year audit is not finished.	This correction has been made in the text.
SCE	The following comments dissect problematic and overarching themes, with a limited set of examples that by no means fully encapsulate the breadth of errors contained in the Draft Report and ERT. SCE wishes to present its detailed findings to the Commission staff in order to identify and correct all errors before the publishing of the Final Report.	Energy Division responds in the following paragraphs.

Source	Comment	Response
SCE	**Appliance Recycling Program (SCE2500)**	Energy Division has responded to these comments in the course of finalizing the Residential Retrofit evaluation report.
	The ERT updates for this program are based on DEER-adjusted in situ UES estimates for refrigerators. Despite the fact that the freezer recycling measure was an integral component of the program, it was not evaluated. Furthermore, despite the past history of greater program influence on freezer recycling, the NTG for this measure is incorrectly assumed to be the same as that for a recycled refrigerator. Previous evaluations have consistently demonstrated freezer recycling NTG to be higher than refrigerator NTG (70% vs. 61% in 2004-2005). The UES estimate for freezer recycling was also assumed to have the same order of realized savings as that for the refrigerator recycling measure rather than invoking the Commission-allowed option four for using prior studies where such measures were evaluated.	Option 1 (Extrapolate findings from comparable programs) as outlined in Appendix C, was selected for freezers because it was not a HIM, and refrigerators were determined to be the most comparable measure.
	There are a number of analytical issues with the UES estimates for the refrigerator recycling measures. The estimate relies on small-sample-based, limited time duration in situ metered data – an approach known to be tainted by self-selection bias and extrapolation issues. Such an approach was proven in the 2004-2005 EM&V study not to be reliable on its own given the need for careful projections of expected in situ UECs into contexts that are observed in particular participant samples, as well as others from which appliances were picked up in the program, and still others in which appliances might have been found had transfers not been prevented by the program. Analytically, there is very little evidence on how a very consequential in situ UEC regression solution was arrived at, either in terms of a progression of specifications, rules involving significance, or terms required to adequately deal with stratification (particularly in the absence of sample weights).	The selection of in situ metering for the ARP UES estimate was addressed in the responses to the comments on the Residential Retrofit High Impact Measure Evaluation Report (http://www.energydataweb.com/cpuc-cms/home.aspx) in Appendix M (Comment numbers 2.13, 2.29, 2.41, 56.07).

Source	Comment	Response
	The result is that we see inconsistent findings from the regression -- "no statistically significant difference was found between appliances in conditioned v. unconditioned spaces" (see p. 138). Yet it is reported that outdoor temperature is significant (see p.139) and appliances in warmer climate zones use more energy than those in cooler climate zones (see p. 142). Another methodological issue that SCE finds problematic is how the UEC estimation threw away available past metering data that covers the manufacture years of more than half the 2006-2008 data -- information that has always contributed so heavily to the precision of the program's evaluations for the past decade's worth of work in this area. The evaluated savings estimates ignored the 2004-2005 EM&V study's emphasis on precision, blending of methods, reliability of the DOE estimate as a regressor in estimating in situ consumption, and the capability to generalize results to populations, including tracking populations other than those studied.	The discussion regarding the space/temperature parameters was addressed in the responses to the comments on the Residential Retrofit High Impact Measure Evaluation Report in Appendix M (Comment number 56.2). The decision not to include the oldest DOE data was also addressed in Appendix M (Comment numbers 2.34 and 2.41) and in Section 11.5 of the report.
	Finally, the un-substantiated DEER-adjustment to the ex post evaluation study UES does not follow the allowed rules set for the ERT updates (Options 1 through 4 in D.07-09-043). According to the Commission's post-2005 policy rules (D.05-04-051), energy savings updates should be the purview of the load impact studies. The Appliance Recycling Program is one studied obvious case where we see the logic of an adjustment to ex post load impact results for interactive effects quickly falling apart. We note again that the overarching goal of the program is to prevent the continued operation of appliances in current or would-be transfer locations, where the latter may entail different environmental, household characteristics, or appliance uses that need to be accounted for in making any attempt to make such an adjustment. The capability of the DEER adjustment is severely weakened by the available empirical EM&V data or program tracking data to substantiate the relevancy and applicability of this adjustment to appliances in current or would-be transfer locations. Without relevant empirical evidence for the conditions of program participating refrigerators and freezers in current or would-be transfer locations, any DEER adjustment is unjustified.	As+D83 shown in Figure 9 of the Residential Retrofit High Impact Measure Report, about 21% of the recycled refrigerators that were not freeriders (and thus claimed savings) would have been used in the same location (10.6/(10.6+39.5%). For those that would have transferred, the best predictor of use is actual use in the prior location, as measured by the in situ metering in the study.
SCE	**Residential Energy Efficiency Incentive Program - Residential Lighting Program (SCE2501)**	

Source	Comment	Response
	The Residential Lighting Program received its own evaluation, coming closer to a true program evaluation than other programs. None the less, there were serious methodological problems with the study and additional problems with how the savings parameters were updated in the ERT. There were two significant problems with the UES estimation: the modeling exercise for the hours of operation and the methodology for the change in wattage. The hours of operation estimate comes from a regression analysis of a large sample of metering data collected for the impact evaluation. Several important variables were left out of the analysis for not being statistically significant, such as dwelling type, fixture type and lamp type. The lack of apparent statistical significance is likely due to the collinearity of variables that will artificially produce this result.	The regression analysis for hours of operation was addressed in Comment ID #130, pg 241-242 of the "Final Upstream Lighting Evaluation Report – 020810 – Volume 1 – FINAL.doc" ((http://www.energydataweb.com/cpuc-cms/home.aspx)
	The estimate of the change in wattage should be based on a comparison of base wattage and efficient wattage. In this case, the estimate from the study is based on the average wattage of all installed CFLs and all installed incandescents. This is not equivalent and will produce incorrect results if there has been any change over time in the relevant adoption behaviors. A survey of participants in the Integrated School-Based Program found that the average change was from a 70 watt incandescent to a 14 watt CFL, which matches well with the program assumptions, and actually measures the change in wattage. However, this is significantly different the study's assumed value.	Delta watts was addressed in Comment ID #20, pg 235 of the "Final Upstream Lighting Evaluation Report – 020810 – Volume 1 – FINAL.doc"
	Additionally, program-level savings for the Upstream Lighting Program are heavily dependent on the breakdown of residential and nonresidential lamps within the program. The study estimates this breakdown at roughly 95% residential and 5% nonresidential from a modeling exercise based on the relative prevalence of CFLs in on-site surveys of homes and businesses, but coming after the median expected life of some of the lamps in nonresidential applications, the nonresidential share is likely downwardly biased. Survey results from the study indicate between 13% and 20% of lamps were purchased for nonresidential applications, not the 5% allotted by the flawed study.	Residential vs. nonresidential was addressed in Comment ID #84, pg 300 of the "Final Upstream Lighting Evaluation Report – 020810 – Volume 1 – FINAL.doc." Additionally, the nonresidential on-sites captured burnt out lamps through interviews during onsite visits.

Source	Comment	Response
	The NTG was estimated in the study, which in this application is truly a net-of-freerider ratio (NOFR), relied on a "preponderance of evidence" approach. The idea had been to estimate the NOFR through various methods that would all coalesce around a value. During the planning stages for the study the evaluators were asked what would happen if the results did not coalesce, but no clear plan was apparent. When the analysis was complete, there were disparate results between the methods' implications.	NTG was addressed in Comment ID #39, pg 268 of the "Final Upstream Lighting Evaluation Report – 020810 – Volume 1 – FINAL.doc"
	The final recommended value was based primarily on the Revealed Preference results, but these were heavily dependent on the unsupported assumption that when a respondent reported they would have bought fewer CFLs at twice the price, they still would have bought 80% as many. Furthermore, the pricing analysis indicated that the actual effect of the program was to reduce prices to 1/3 of their original price, so the stated price comparison does not even apply to the results of the program.	These comments are addressed in Comment ID #95, pg 272 and Comment ID #98, pg 273 of the "Final Upstream Lighting Evaluation Report – 020810 – Volume 1 – FINAL.doc"
	The installation rate is final problematic input from the study. The installation rate estimate from the study is based on a modeling exercise that tries to predict the installation rate based on changing inventories in an attempt to estimate total installations within the program timeframe, even if they occurred after the first year. Unfortunately, the results almost exactly match the program-volume-weighted first-year installation rates from each of the three program years, which mean the result is the same as assuming no installation after the first year. This is clearly inappropriate and drastically reduces the true energy savings achieved by the entire three- year program.	Installation rate is addressed in Comment ID #66, pg 247-248 of the "Final Upstream Lighting Evaluation Report – 020810 – Volume 1 – FINAL.doc"

Source	Comment	Response
	The savings parameters estimated in the impact evaluation were problematic, but they were not even implemented correctly in the ERT. UES estimates for the Residential Lighting Program were inputted on an aggregate basis rather than on the more direct basis supplied in the study. That is, even though there are three different gross savings estimates for Globe, Reflector and Twister/A-lamp CFLs, a fourth value for all CFLs was used. While this should yield the correct program savings under some assumptions, changes in the savings scenario could yield incorrect results. Similarly, the savings estimates for LEDs and CFL fixtures do not match table 30 and 34 of the study, as indicated in the documentation in Appendix C. Finally, installation rates for LEDs and CFL fixtures do not match the documentation. The ERT combines installation rates from the study with realized shipment rates, also from the study. But these values do not lead to the installation rates in the ERT for LEDs and CFL fixtures.	The ERT was not initially designed to handle the type of evaluation results determined by the ULP evaluation, thus the ERT process for the ULP required an additional amount of analysis. For example, the installation rate for ULP needed to include additional adjustments for leakage, shipments vs. sales, re-appropriated residential versus nonresidential, percent sell-through, and installation rate. The entire process was reviewed with the ERT team and the CPUC to ensure quality control.
SCE	**Residential Energy Efficiency Incentive Program – Home Energy Efficiency Rebates (SCE2501)**	Energy Division has responded to these comments in the course of finalizing the Residential Retrofit evaluation report. The specific responses can be found in the final report which is posted at www.energydataweb.com
	There are two major measures in this SCE program: room AC and refrigerators. Only room AC fit the HIM description for the portfolio evaluation. The UES estimates are based on gross consumption and demand saving from a combination of 102 metered room A/C's (metered in summer 2009). A four-part model is developed to account for hours of use, compressor on- time, consumption with compressor off, and consumption with compressor on. Compressor on- time is inferred from a threshold of 250 watts. The model was used to produce annual hourly use and hours of use based on an hourly regression involving temperature and day of week inputs. The four models have very low explained variance except in the case if compressor usage. Hours of use estimated may be low (at 1,007 annually in climate zone 10), and the models supporting these annual estimates explain only 7% and 6% of the variation, indicating over 90% of the differences between individuals is still unexplained with much potential for error.	The importance of the compressor model on savings (with an R-squared of 0.72) was addressed in the responses to the comments on the Residential Retrofit High Impact Measure Evaluation Report (http://www.energydataweb.com/cpuc-cms/home.aspx) in Appendix M (Comment number 1.25).

Source	Comment	Response
	Furthermore, the basis for estimating achieved savings from the annualized operation of the 102 efficient appliances is unclear – the observed appliances' EERs are assumed to be 10.8, and the base case is an EER of 9.8. Backup in the form of only four lab-metered appliances (three non-Energy Star) that were selected based on popularity. How this work translated into any adjustments of the EER-delta-based "savings" estimated from the annualized appliance data is very unclear, if it occurred at all.	As noted on page 167 of the report the lab testing was based on a limited sample, thus the lab data were not applied to the UES estimates.
	While very low UES values were obtained, ranging from 20 kWh to 60 kWh (climate zones 6 through 10), precision on these estimates was also very low at 90/25.	The precision for the room AC estimate was addressed in the responses to the comments on the Residential Retrofit High Impact Measure Evaluation Report (http://www.energydataweb.com/cpuc-cms/home.aspx) in Appendix M (Comment number 1.27).
	The verification rates (96% for SCE) appear to confuse retention (related to operability) with first year savings. This is of course a minor point by comparison to the UES issues just mentioned, and the NTG estimation discussed below.	The use of verification rates was addressed in the responses to the comments on the Residential Retrofit High Impact Measure Evaluation Report in Appendix M (Comment numbers 1.29 and 1.36).
	The NTG estimates rely upon the CPUC's standardized NTG algorithm, which confuses ordinal ratings with probabilities of taking a particular action, ignores differences in the meaning of ratings between respondents, and ignores differences in one unit changes in an ordinal "scale" for a single respondent. For SCE, the result implies 63% free rider (which means, effectively, than an average of four ordinal influence items turned out to be 6.3), counter-intuitive to the Energy Star maximum retail share observed nationally ranging from 36% to 50% over 2006- 2008. The gap may well be related to the social desirability and/or cognitive balance/cognitive dissonance effects – ignoring for the moment that the ordinal average has very little to do with the probability of an action anyway.	These comments were addressed in the responses to the comments on the Residential Retrofit High Impact Measure Evaluation Report in Appendix M (Comment number 1.32).
SCE	**Multifamily Energy Efficiency Program – Multifamily Energy Efficiency Rebates (SCE2502)**	Energy Division has responded to these comments in the course of finalizing the Residential Retrofit evaluation report.

Source	Comment	Response
	The ERT update gives a very misleading picture of actual savings for this program because it relied on data from a very differently delivered Residential Lighting Program for UES and claimed the source to be EM&V instead of "other EM&V" and, it lumped the major measures of this program with "downstream lighting program HIM" that includes totally distinct programs (a lighting exchange program, a mobile home customer direct install program, and a multifamily landlord rebate program).	
	The UES update was based on actual field measurement data for this program for measures that represents roughly only 5% of net kWh savings for this program; the rest of the UES update representing the majority of the net savings for this program was based on the Residential Lighting Program UES parameters without any consideration of extrapolation issues involved in applying such general population-based results to the multifamily program participant population.	The claimed savings for SCE2502 Multifamily come from four measure groups, as shown in Table 177 (page 191) of the Residential Retrofit High Impact Measure Evaluation Report. The exterior CF Fixtures represented the highest savings group, and with about 80% in common areas (Table 189, page 199)were included in the direct EM&V activities
	From Pg 193 of the report: "Due to the relatively small budget and savings claims attributed to the downstream lighting programs, the intent of this evaluation was to rely heavily on the findings from ULP and to supplement that data when necessary. This evaluation did not attempt to replicate the ULP methodology."	The data were supplemental, rather than duplicative, with the ULP. For example, the evaluation focused metering on fixtures that were not covered by ULP, including common areas. The analysis, however, was consistent with the ULP.
	The confidence and precision estimates seems to be 90/5 for a small portion of the multifamily savings that received direct measurement, but the confidence and precision level for the remaining majority of the savings is unknown for this program.	The relative precision estimates for DLP Multifamily are provided in Table 196 (page 202) of the Residential Retrofit High Impact Measure Evaluation Report
	For exterior CFL fixtures installed in this program, there is a higher NTG for therms (78%) than for the kW/kWh. An example of a quality control issue: the screw-in interior lighting measures have different NTG values for kW, kWh, and therms (78.5%, 77.6%, and 78.1%, respectively). The study reports a 78% NTG.	NTG were weighted based on site-level savings and should not be consistent across parameters (kw/kwh/thm).
	Despite the direct field measurement for common area and outdoor linear fluorescents, these measures were treated as "pass through" in the ERT documentation without any explanation on why evaluation results were not used for linear fixture UES, NTG, and IR.	Linear fluorescents represented a very small portion of savings from the SCE2502 multifamily program, and thus received a combination of evaluated and pass-through values.

Source	Comment	Response
	From the evaluation records it is not always clear if a particular measure record is a part of the study sample. It is not clear if interior CFLs included both reflector and twister screw-in bulbs, yet both have the same updates despite being very different measures.	Interior CFLs included both reflectors and twister screw-in bulbs. Hours of use and delta watts represented both of these measures.
	We also found that the indoor dwelling area linear fluorescent fixtures NTG are incorrect (81.4% instead of 77%).	Indoor dwelling area LF had incorrect NTG - 77% is correct value.
SCE		Energy Division has responded to these comments in the course of finalizing the Residential Retrofit evaluation report.
	Multifamily Energy Efficiency Program - Comprehensive Mobile Homes (SCE2502)	
	There are four major measures that were direct-installed in this program: 23 watt interior screw-in CFLs, 18 watt exterior fluorescent fixtures, AC diagnostics and tune-up, and duct testing and sealing. We find several issue areas in the ERT parameter updates, including extrapolation of results from other studies and application of the updates for these measures.	Specific issues are addressed in the following responses, this is a general comment.
	The UES estimates for lighting are based on average CFL to non-CFL wattage ratio obtained for the Residential Lighting Program evaluation from general population surveys without making any adjustment to the characteristics relevant to the mobile home target population of this program.	The mobile home population delta watts assumes the same values as the multifamily delta watts. There was not a targeted statistical sample of mobile homes for the study.
	It is not clear if correct installation rates were applied for measures from the EM&V study table. The ERT measure name is interior fluorescent fixture, which can be confused with linear fluorescent CFL fixture and fluorescent fixtures have a different NTG and IR.	Parameters were assigned to measures based on the HIM measure designation. This designation was assigned by the ED group responsible for the Standard Program Tracking Database (SPTD).
	The ERT updated the UES for AC diagnostics and tune-up based on the results from the HVAC study, which has a very different delivery model and population mix. The duct testing and sealing also got an installation rate applied from the HVAC study. Note that the ERT tool lists the source of the updated for these parameters to be “EM&V” versus “other EM&V”.	The ERT relied on both a mix of direct EM&V and some Other EMV parameters. Note that mobile homes were included in the HVAC sample.

Source	Comment	Response
	The self-reported NTG with all its methodological issues has been inconsistently applied. The therm NTG is different than that the kWh/kW NTG for interior and exterior lighting fixtures. For interior screw-in CFL for the same measure there are three variants of NTG updates for kW, kWh, and therm. (78.5%, 77.6%, and 78.1% respectively).	NTG were weighted based on site-level savings and should not be consistent across parameters (kw/kwh/thm).
SCE	**Integrated School-Based Program (SCE2504)**	
	The Integrated School-Based Program works with educational institutions to promote energy efficiency to students within schools. Originally, the evaluators in charge of this program under the Specialized Commercial contract group planned to do a full evaluation of direct and indirect program impacts. Unfortunately, "because of the reallocation of resources to HIM programs, the evaluation was limited to only evaluating direct-savings measures" (Specialized Commercial Appendices, p.144). In truth, most of this program received no true ex post evaluation. For the Green Campuses and Green Schools programs, the evaluation consisted	Commenter is correct, on GS/GC, we did apply a realization rate on top of the installation rate SCE had already incorporated into their estimates. The corrected savings are: 306 MWh and 27 kW for Green Campus and 876 MWH and 77 kW for Green Schools; this update has been made in the final results.
	of using values from the Local Government Program verification report. The LivingWise Program received data from participant surveys to update savings values. In all cases, a NTG value of 80% was "assumed."	Correct, there was no direct EM&V of NTG. These should be coded in the ERT as pass through values. This correction has been made in the final results.
	And yet, even though most values received no direct EM&V, the values in the ERT are all coded "EMV" rather than "othEMV". Furthermore, the installation rate for all measures were incorporated into the gross UES estimates in the study, but the ERT then includes them again, which will penalize the energy savings twice for non-installation. Finally, the report does not present kW reductions, yet these are included in the ERT as EM&V values.	KEMA used the kW savings supplied by the utility, and did not recalculate them in the evaluation. They are included in the Specialized Commercial report as EM&V values. These should be coded in the ERT as pass through values. This correction has been made in the final results.
SCE	**CA New Homes Program (SCE2505)**	NCCS
	Only the single-family whole house measure was directly evaluated in this program for UES. For the multifamily whole house measure, another major measure in this program, there were no updates despite the results from the single family whole house measure of nearly 400% gross realization rate and the study findings (pp. 3-47) that:	The gross realization rate was based on data collected and analyzed for single-family homes only; these findings are not transferrable to multifamily homes.

Source	Comment	Response
	"(The realization) was much greater than one for SCE. This was a result of low ex ante per-unit savings estimates across the board, as well as having a large concentration of program participants in high usage climate zones. This indicates a downward bias in a "pass through" treatment in the ERT."	Evaluation contractors only applied evaluation results that were reliable and applicable as presented in the Decision Framework, which does not create a systematic downward bias in the use of "pass through" results; when applicable results were not available the utility estimates were applied.
	The ERT update intended to use the residential retrofit lighting measures evaluation to update the CA New Homes programs. However, it did not end up using those evaluation updates due to lack of data on distribution of lamps by room type in the new homes – not because the two program populations are very distinct even by the rule set for the ERT parameter updates. This is another example of distasteful consequences of the HIM approach. In the ERT update for UES, the same therm savings (83 therms) measured for the inland and coastal regions was applied for the Desert region without any consideration of proportional relationship to the kWh and kW savings of this climate zone.	Lighting savings for the Residential New Construction program were passed through.
	The study's arbitrary approach to arriving at a final sample size for SCE when the original sample target was 45 is problematic. The evaluation team should have assessed the reasonableness of the arrived sample size for each IOU due to an arbitrary cut-off date in time. The report should have explained how the final sample impacts the reliability and validity of the findings. In the evaluation report the development of sample designs is described, the actual execution of the sample design is not always explained. It is conventional to report on the plan(s), the number of attempts to recruit, the number of refusals, the numbers of final surveys, and the numbers used in the final analysis, with explanations for how and why reality deviated from the ideal. These would, in this case, be broken out by surveyed, on-sites, metered, by utility. This is typically related to the population, with the sample weights indicated. Also, the evaluation report did not show how these varied by utility, and by whole building/systems analysis/industrial participants.	The sample size for SCE non-residential new construction (NRNC) projects was 70 and the planned and achieved sample size is shown in Table 3-11 of the NRNC volume. Table 3-12 illustrates how these varied by utility, and by whole building/systems analysis/industrial participants. The study was planned to achieve 10% relative precision at the 90% confidence interval and as shown in Table 3-15, the achieved precision for SCE exceeded the minimum at 8.4%
SCE	**Comprehensive Packaged Air Conditioning Systems Program (SCE2507)**	

Source	Comment	Response
	Savings estimates for the Comprehensive Packaged Air Conditioning Systems Program come from the Specialized Commercial contract group study. The study conducted on-site metering and other diagnostics, but did not do any billing analysis or other analysis that would derive savings estimates directly from the program population, which is the intent of ex post impact analysis. Instead, the savings estimates come from simulation modeling using characteristics measured from the program participants.	The filed and approved evaluation plan did not include billing analysis. Savings were determined using data collected from in situ metering of representative units, as described in the Specialized Commercial HVAC report. Savings were applied to the population. See Specialized Commercial Comment Responses 7 (#7 - Since the AC replacement measures did include medium-term metering and regression analysis, we assume this comment refers specifically to RCA. The RCA evaluation approach in the high impact measure plan presented the approach of monitoring units pre and post charge adjustment to determine relative changes in efficiency and inputting those efficiencies into energy models to develop 8760 energy savings. Medium term pre and post energy measurements were not feasible under the actual project timeline and the limited number of metered units that could ultimately be used in the analysis precluded robust analyses via regression. Likewise, the timeline did not allow for sufficient post data for a reasonable billing analysis. Past experiences with billing data convinced us that billing analyses can be confounded in many ways, and should only be used where a large number of sample points are available and we are certain that the measure savings are large enough to be discerned from the revenue.) and 8 (# 8 - The report text has been clarified in the Section 5 introduction as isolating the savings due to refrigerant charge adjustment, which accounted for the major proportion of IOU portfolio savings. The reported savings of attached measures, such as coil cleaning, were not evaluated in the RCA HIM evaluation.)
	The study also demonstrates some problematic characteristics with respect to other components of the analysis. The operating efficiency was chosen over the energy savings as the analysis variable in order to justify smaller samples. The analysis imposed external requirements for a site to meet the definition of installation, which seems have become necessary given the reliance on simulation results rather than statistical techniques.	See Specialized Commercial Comment Responses 7 and 9: (#7 - provided in response above) (#9) The definition of installation within some programs may be "some work was done" and the evaluation used a definition of "installed and working properly" which is the CA Protocols definition of verification. The specific criteria applied to all RCA HIMs were that the system superheat or subcooling must be within a target tolerance of five degrees for superheat or three degrees for subcooling. The criterion for all duct leakage HIMs was that after program rebated sealing the total system leakage would be 15% or less of the nominal system flow. Specific exceptions to these criteria are also explicit in the text of these sections. Systems which were within three degrees of the upper and lower limit of the target tolerance were

Source	Comment	Response
		assessed as passing if the nominal measured EER was within 15% of rated EER. For duct leakage, systems which had final measured leakage within 3% of contractor measured final leakage were considered passing.
	Like many other evaluations for the 2006-2008 portfolios, the NTGs were estimated by means of the self-report algorithm, which has significant methodological problems, not the least of which is assuming without justification that if a respondent rated the importance of the program as seven out of ten then 70% of the savings were net and 30% were free-rider.	The NTG surveys conducted for commercial participants were with vendors; NTG ratio =.96. NTG surveys were conducted with 245 residential customers. The FR score was based on the algorithm. No findings were reported that stated the FR ratio depended on respondents' rating of the importance of the program alone.
	The documentation for the ERT is not very clear for these measures. It is clear that some measures received evaluation but were passed through without a clear reason while some received evaluation results, even thought the line item appears identical in other respects.	ERT coding of pass through and EMV values has been corrected. Only the UES, Installation Rates, and NTG in the Specialized Commercial Report should be coded as EMV.
	Retro-Commissioning Program (SCE2508)	
	The Retro-Commissioning Program did receive an evaluation. The evaluators estimated kWh, kW and therm impacts, regardless of whether the utility had goals in each of those categories. They developed a factor to estimate the savings from those missing values. Appendix C states that these factors were used for therm savings estimates, but each project has	1. The documentation in Appendix C was incorrect in stating that a therms per kWh factor was used. The therms UES was an average therm savings per measure which was applied to all measures. The ERT is consistent with the methodology described in the RCx report; and Appendix C has been corrected in the final document.
SCE	the same value for the therm savings, meaning no such factor was used. As with many other programs, some values were inexplicably passed through when other similar measures did not receive a pass-through.	2. Projects which received the Pass-thru values were not included in our sample-frame, and thus not part of our study. They were excluded due to low relative savings.
SCE	**Industrial Energy Efficiency Program (SCE2509)**	

Source	Comment	Response
	Although the Draft 2006-2008 Energy Efficiency Evaluation Report fails to mention this, the major focus of SCE's Industrial program evaluation was pump-off controllers. This aside, SCE has numerous questions about the transparency of the process that determined key savings parameters. Our main concern with the novel method of determining program savings from program components, i.e., the ERT combined with HIM evaluation is based on the fact that because we had no input into its development, we have major concerns about the potential for errors and resulting mischaracterizations of program performance. The 45 tables and even more queries provided in the SCE2509 Access database simply magnify these concerns. There are many significant problems with how the evaluation conceived of the program and developed energy savings parameters.	These comments refer to the "2006-2008 Evaluation Report for the Southern California Industrial and Agricultural Contract Group" and have been addressed during the comment/answer period for that Report. Regarding the issue of transparency of the ERT process: the ERT documentation provided with the SCIA Industrial measures (SCE2509 and SCE2510) shows that the ERT results for the "all parameters" run match perfectly the program-level results reported in the Final Reports.
	• Access to Engineering Data: Global Energy Partners has presented some key questions regarding the report on this site (See comAttach_2045.doc attachment posted 1/14/10). Instead of rehashing them, SCE's main concern is that even after data requests, there is still no way to reproduce key elements of Itron's analysis. Under circumstances of collaboration and openness, this result would be less important, but given the recent evaluations that could be substantially more collaborative, the ability to understand and replicate research results are crucial for a sense of fairness and accuracy in reporting.	It should be noted that this is a comment on the evaluation report that was originally posted on ED's public comments website in connection with Itron's draft evaluation report. Global Energy Partners reviewed the results reported in site reports for specific projects included in the Southern California Industrial evaluation and made technical comments on those projects. Itron's methodology and savings calculations were provided in the site reports reviewed by GEP. Its technical comments, among others, were addressed in meetings ED held with the evaluation contractors SCE and other utilities, and utility engineers and third party implementers (including GEP), to discuss the evaluation methodologies. In addition, detailed responses were provided to GEP's comments and there was additional discussion between GEP and the evaluation contractor. GEP's comments were posted on the public website and can be viewed at www.energydataweb.com/cpuc, under the Southern California Industrial and Agricultural topic, the "SCIA Evaluation Report Comments/Responses," the SCE comments chapter.

Source	Comment	Response
	• Insight Into NTG Determination: SCE has expressed concern over the NTG methods in other contract groups. We believe that further explanation of the NTG rationale and scoring methodology is warranted not simply because it is a problematic measure but precisely because it is a problematic measure that requires program level insight and research collaboration. Since we are unable to verify the process and algorithm we are left only with the hope that inadvertent errors or mistaken assumptions have not corrupted the estimates.	ED posted for review on www.energydataweb.com/cpuc/, and SCE had and continues to have access to, the appendices to Iron's evaluation report for the Southern California Industrial and Agriculture evaluation, which include Appendix D-1 (Nonresidential Net-to-Gross Methodology), Appendix D-2 (Nonresidential Net-to-Gross Survey Instruments), Appendix D-3 (Detailed Site-Specific Net-to-Gross Results), and Appendix D-4 (Onsite Data Collection Forms). In addition, ED posted on the same website, under the SCIA topic, a paper titled "Response to Overarching IOU Concerns Regarding the Estimation of the Net-To-Gross Ratio Using the Nonresidential Self-Report Approach." In addition, ED provided responses to comments posted by SCE's third party implementer, Neat, at www.energydataweb.com/cpuc, under the Southern California Industrial and Agricultural topic, the "SCIA Evaluation Report Comments/Responses," the SCE comments chapter.
	• Standard Industry Practice (SIP): The report both suggests SCE should defer to SIP and simultaneously suggests the need to define SIP. We believe the null hypothesis should remain the existing customer baseline (in situ) unless compelling reasons dictate otherwise. As PG&E notes, "Industry standards and baselines used by the evaluation are subjective and questionable." Had there been more collaboration in this research effort, the errors and uncertainties introduced by these seemingly arbitrary decisions could at least have been understood as a cost of research business given the complexities and time pressures involved. Instead, we are simply left to wonder if these determinations reflect the real conditions our customers are facing.	This is an issue relating to the evaluation methodology that was raised in comments on Iron's draft evaluation report. We note that in the evaluation, Itron found instances where "compelling reasons" dictated against using in situ equipment as the baselines. These results were provided to SCE in the site reports, and were addressed in meetings with the utilities as well as in the technical comments that were filed in connection with the draft report and to which we responded. To the extent SCE suggests that baseline determinations were "arbitrary" and that it was "left wondering if these determinations reflect the real conditions our customers are facing," we note that baseline determinations were made after on-site evaluation visits, and refer to the site reports and comment responses for detailed discussion of the bases for baseline determinations.

Source	Comment	Response
	• Program Influence: Finally, SCE agrees with ACEEE's comments on PG&E's Industrial program. In particular: "Relationships built over many years with multiple people within a company often lead to consideration of energy efficiency investments as a result of this sustained interaction. ACEEE cautions that the Itron methodology may have discounted this type of historical relationship, and at times may have relied on the input from staff members at various industrial companies that may not have a working knowledge of the historical interactions..." and continues, "Working to change the manner in which companies think about energy is a long-term process, and one that can take years to influence. ACEEE's experience shows that many projects that have been "decided upon" in a given year actually rely on groundwork laid by various interactions in many years prior. ACEEE does not believe that such groundwork is accurately reflected in this methodology."	This is a copy of a comment that was filed on the draft evaluation report. We nevertheless note that program influence is investigated in depth in connection with the net to gross surveys and interviews. Detailed comments and responses on program influence can be found at www.energydataweb.com/cpuc, under the Southern California Industrial and Agricultural topic, the "SCIA Evaluation Report Comments/Responses," the SCE comments chapter, in the responses to the group of comments made by Nexant.
	Aside from the nuts and bolts of the algorithm and the problems of reaching the correct decision maker as opposed to the facility handyman, there is the problem of focusing only on the rebate and measure as determinates of program influence and ignoring the information value a utility recommendation brings. In other words, the dollar value of a utility rebate for a POC project may often not be what matters to nonresidential customers. The fact that an independent and trusted organization with a much more indirect financial benefit to this technology or process is recommending this technology may be the only reason an installation/upgrade decision is made. By focusing on whether they would have done it without the rebate substitutes an accountant's view of business transactions for an economist's. This seems to be a symptom of treating energy efficiency as a set of measures, rather than as programs that are each designed to achieve specific goals.	Continuation of comment, please see response above.

Source	Comment	Response
	We agree with Nexant's comments on PG&E's Industrial Program on this point: "Rebate is not the most important influence for the industrial segment. The evaluation seems to place rebate as the most important factors to determining the NTG. The reality is that the influence of an energy efficiency (EE) program for industrial segment is broader than just the availability of the incentives. The IOU programs reach out to the customers through multiple channels including mass marketing, education and training, influencing customers' decision makers (top- down approach), handholding the customers through the project development phases (bottom-up approach), providing energy audits to help customers identify projects, and providing technical assistance for project design....Instead, the evaluation should credit the project as long as the IOU program provides sufficient influence to overcome the market barriers—that is, in the absence of the program that sufficient market barriers would not have been removed and the project savings would not have been realize. Sufficient reduction of market barriers can sometimes be as simple as pointing out an obvious deficiency to a customer whom is not aware of the deficiency, and yet in the absence of the sponsoring program, the resultant savings would not have occurred."	Continuation of comment, please see response above.

Source	Comment	Response
	In response to a data request, Itron provided incomplete data (some raw data, but no analysis); GEP was not able to derive Itron's ex post savings values. The data included show zero values, however GEP cannot validate that Itron took the correct steps to properly condition the data. Itron metered only four wells in their evaluation; GEP metered five wells (out of 20) to determine an average motor load factor to apply to all pumps (even those metered). Yet Itron comments on Page 5: "The applicability of the 72% based on motor load factor gathered from five pumps and applied to the remaining 15 is uncertain." The GEP methodology is more rigorous (5 wells metered out of 20) than Itron's (4 wells metered out of 20). Itron states that the wells were equipped with POCs before the installation of the POCs in the project and discounted the annual operating hours from 8,585 hr/yr (based on 98% availability) to 4,400 hr/yr because of this. Review of the DOGGR site and DOGGR drilling permits shows, with the exception of three wells, all wells were new drills from early to mid 2006. The POCs were installed in late 2006, and final application submitted in Dec 2006. Based on this timeline, the fact that these were new wells precludes them from having existing POCs. Consequently, the Itron conclusions on that account are without any merit.	Please note that this is a comment on evaluation of a specific measure -- pump off controllers -- that was previously posted on the public documents website for the evaluation report and to which we responded. Itron made changes to its calculations and report as a result of this comment, as shown on the response [www.energydataweb.com/cpuc, under the Southern California Industrial and Agricultural topic, the "SCIA Evaluation Report Comments/Responses," the SCE comments chapter].
	Itron determined a Baseline Adjustment Factor (BAF) of 78.4% for POCs based on the assumption that wells without POCs operate with partially filled pumps and consume less power than wells with POCs that operate with completely filled pumps. The BAF was determined from testing 28 wells in two fields in PG&E service territory (see Itron B001 and B007).	This comment was copied from the comments posted on Itron's evaluation report. The comment and ED response can be viewed at www.energydataweb.com/cpuc, under the Southern California Industrial and Agricultural topic, the "SCIA Evaluation Report Comments/Responses," the SCE comments chapter.
	Based on the following two significant factors, we proposed the Baseline Adjustment Factor be removed from the Realization Rate analysis for POCs:	Continuation of comment, please see response above.

Source	Comment	Response
	First, the method of deriving the adjustment factor did not provide a reasonably valid result. Itron did not provide the data for the wells metered to obtain the BAF; however, based on analysis of data provided for project B109, the interpretation of the metered data for this project was flawed, resulting in extremely low average field motor loadings. Periods of non-operation were not adequately removed from the power kW average. Including these 0 kW values from the non-operation periods significantly skews the average downward. In one case, Well 36B 6- 6R, the pump went out of service for a period of time. When it restarted the voltage value on channel 2 dropped out, most likely due to a disconnected lead wire. The metered data following the voltage drop out was included in the average, which significantly skewed the average downward. The metered data results for well 5C 5-4A and the current draw suggest a motor that is much smaller than the reported 25HP. The current draw is representative of a 10HP motor. At the time of the metering was a record taken of motor face plate data to validate the 25 HP? This calls into question whether Itron validated the motor HP.	Continuation of comment, please see response above.
	Second, the broad utilization of the adjustment factor to other producers and oil fields throughout California is not reasonable. The results from 28 wells were used to determine the BAF – the average of the 28 wells in the test sample was 78.4%. The results on an individual well ranged from a low of 32.7% to a high of 162.5%. This factor was applied to all wells even though a review of DOGGR shows over 60,000 wells in California. The sample size is grossly inadequate and is not statistically valid. In addition, the BAF provides no consideration for how the producer designs the pumping unit for the application, or differences in the oil field, i.e. depth of wells, fluid conditions, and production rates. An example is that the wells evaluated to determine the BAF use steam injection for enhanced oil recovery, something that coastal oil producers in SCE service territory do not do. For vapor recovery units, we were unable to evaluate the calculations, as the requested data was not provided.	Continuation of comment, please see response above.

Source	Comment	Response
	Energy Intensity was used to evaluate the energy savings and the post-installation production increased significantly. Itron revised the energy savings using pre- installation production stating the use of a "protocol." Itron committed to provide the "protocol" in a webinar, but GEP has not received any documentation from Itron to support their methodology. The SCE SPC Manual clearly states "In general, these measures will be based on post-production..." (in Section 1, see page 1-10 under 1.4.5 Increased Load/Production Measures).	Continuation of comment, please see response above.
	For variable speed drives, pump flow varies. Both Itron and GEP assume the specific energy usage is constant (kW/Barrel).	This comment was copied from the comments posted on the draft evaluation report. The comment and ED response can be viewed at www.energydataweb.com/cpuc, under the Southern California Industrial and Agricultural topic, the "SCIA Evaluation Report Comments/Responses," the SCE comments chapter.
	• The GEP methodology uses flow rate from the most recent individual well tests (performed 2-3 times per month) to determine kW/BBL.	Continuation of comment, please see response above.
	• Itron methodology uses a 29 month average extracted from the DOGGR data (Min of 51 BPD, Max of 346 BPD with an average of 202 BPD).	Continuation of comment, please see response above.
	The Itron methodology used kW reading taken at some point during the 29 months and the applied the average 202 BPD from the 29 month average. Theoretically the well flow rate during metering period could have occurred on the 51 BPD or the 346 BPD period but it is unlikely that the average flow occurred during this metering period. Using a flow rate that occurred during the metering period would significantly change the results. The GEP methodology more accurately represents the kW/Barrel. Of the 10 Smart Wells Itron evaluated, they found 4 wells to have negative energy savings.	Continuation of comment, please see response above.

Source	Comment	Response
	Their analysis of these under-performing wells also shows higher oil production in the base case. The Itron methodology leads to the conclusion that 40% of the time the application of advance drilling techniques by the producers to complete a SMART well will result in reduced oil production, and increased power consumption. This is not a reasonable conclusion, as SMART Well completion techniques involves a body of work and research conducted by an entire industry of geologist and petroleum engineers and represent significant investment by the producers.	Continuation of comment, please see response above.
	SMART well production and baseline well production rates are based on a geologist estimate and specify a quantity of water shutoff. Actual production values are gathered in the post production inspection and the baseline production rates are adjusted based on initial estimates. This adjustment is necessary as the gross values can often vary from the initial estimates.	Continuation of comment, please see response above.
	Itron's methodology for new wells with estimated baseline relied on the same initial estimates used by GEP, however in the verified savings the pre and post values are mixed. The Itron assumption is that the geologist was targeting specific % water cut. The case of project C017 shows the energy estimates from GEP vs. Itron. The ex ante savings reported by Itron do not match the ex ante savings reported by GEP in many cases; based on the submitted data the discrepancy cannot be rectified.	Continuation of comment, please see response above.
SCE		
	Agricultural Energy Efficiency Program (SCE2510)	
	Pump testing is a key component of SCE's Agriculture program; however, it did not receive and impact evaluation. As Itron's report noted, "The parameter examined for SCE2510 measures is the Net-to-Gross Ratio (NTGR) only." The report states that "Following CPUC's "Requirements for Evaluating High Impact Measures" directive of July 21, 2008, the SCIA contract group was instructed to finalize impact evaluation work for the SCE Industrial Measures and Agricultural Measures (note, not programs but measures). The report continues:	Two issues have been confused in this comment: - SCE2510 Pump Testing program was evaluated, please refer to Section 5, Pump Testing, in the "2006-2008 Evaluation Report for the Southern California Industrial and Agricultural Contract Group." - Evaluation of the SCE2510 Agricultural program was indeed limited to the NTGR parameter as a result of CPUC's "Requirements for Evaluating High Impact Measures" directive. SCE's ex ante gross impacts were accepted by the evaluation and the ERT process, and only NTGR results were applied to the SCE2510 Agricultural program.

Source	Comment	Response
	"This report contains impact results for SCE2509 only. Efforts to complete impact evaluation for the 10 sites drawn for program SCE2510 were stopped when the Q4, 2008 extract was received from SCE. That extract, containing 1,133 records and a very diverse set of measures, led to the conclusion that any results based on a sample of 10 points wouldn't be robust enough to support a realization rate result. (p. 6-8)"	Continuation of comment, please see response above.
	While SCE provided the evaluators over a 1,000 records for the program, they were directed, due to the unilateral and unjustified change in evaluation methodology, to stop impact work on this important program leaving only 10 sample points – rightly ruled out for reliable savings estimation.	Continuation of comment, please see response above.
SCE	**Nonresidential Direct Installation Program (SCE2511)**	
	The Nonresidential Direct Installation Program did not receive a program evaluation; instead its measures are split between various evaluations including the Commercial Facilities and Small Facilities contract groups.	The majority of the savings in the Nonresidential Direct Installation Program were evaluated through a combination of high impact measure studies.
	The measures in the Commercial Facilities study suffered from extremely large confidence intervals, largely due to very small sample sizes caused by recruitment problems.	As discussed in previous public comments, the large confidence interval is primarily due to a mathematical anomaly. Since the ex post energy savings is such a small fraction of the ex ante savings (two orders of magnitude smaller), the error bounds is calculated using the ex ante savings (on the order of a hundred) but it is applied to the ex post savings on the order of 0.5 to 3. While the resulting confidence interval seems large in absolute terms, it is completely insignificant. Consider, for example, if our estimate were off by the maximum error, the greatest savings would still be an order of magnitude smaller the ex ante savings.
	Additionally, many study participants were not even administered the net-to-gross battery, and instead were assumed to be full freeriders.	Results from the NTG surveys were appropriately extrapolated to the sample population based on the sample design.

Source	Comment	Response
	The study does not detail the kW reductions from Strip Curtains or Door Gaskets, only providing kWh savings.	As documented in the Commercial Facilities report, the results of Strip Curtains and Door Gaskets were based upon an hourly analysis of the savings, thus; kW impacts are an integral part of our results. The kW results are presented in the HIM Appendices to the Commercial Facilities contract report as well as the ERT write-up include these data. To quote the ERT write-up: "Demand Savings for Strip Curtains and Gaskets: The demand savings for door gaskets are determined with a kWh to kW multiplier. The multiplier is the ratio of the average energy savings during the peak nine-hour period (2 PM to 5 PM on weekdays during heat wave) to the annual energy savings. The ratio is 0.000167 for door gaskets. For strip curtains, the ratio is multiplied by a "demand coincidence factor" which represents the likelihood that the refrigeration system would be removing the infiltration heat load during the peak nine hours. On average, this likelihood is 15%. The kWh to kW multiplier for strip curtains is 0.000025."
	In the ERT, an unexplained multiplier, not found in the study, is used to create kW savings estimate.	The "unexplained multiplier" is explained in the Commercial Facilities report in the HIM appendix. This multiplier does not "create(s) kW savings estimates," but rather it is the ratio of the kW impacts to the kWh impacts.
	For Door Gaskets, the study does not provide savings by climate zone.	The savings for gaskets were only reported by climate zone for SCE. SCE was the only utility that reported by climate zone, and also because the SCE realization rates for door gaskets for SCE were much higher than those for PG&E and SDG&E. Therefore, the small differences in savings that result from climate zone were reported in SCE territory only. However, please note that the large confidence interval casts the relatively small differences that occur from climate zone to climate zone as moot. For PG&E and SDG&E, the savings were not reported by climate zone because the ex-ante savings were climate zone independent, because the variation in the ex-post savings due to geography were insignificant compared to the overall error bound of the ex-post savings estimates, and because the variations had insignificant effects on the realization rates. In essence, the question could be cast as whether the savings are 2.90 or 2.97 kWh/ft, though the confidence interval is approximately 0 to 4.5 kWh/ft and the ex-ante is 105 kWh/ft (note: for SCE the ex-ante is closer to 15 kWh/ft)

Source	Comment	Response
	The ERT, on the other hand, includes such a table in the documentation in Appendix C, and those values, which are not included in the study, are used.	Climate zone information was required for the ERT process. Therefore, the climate zone variations were included in the ERT provide an expansion upon the results found in the report but do not conflict with the summary findings presented in the report.[ADM]
	There are different values for cooler and freezer savings for door gaskets, which seem to be randomly assigned to measures.	Savings are not "randomly assigned to measures." The assignment is a direct result of the data provided to us in the IOU program tracking databases. For example, many records showed "cooler" as the measure but claimed "freezer" as the energy savings. In such cases we used the energy savings claims as the more reliable data.
	Some measures are ambiguous (Glass Door Cooler/Freezer Gaskets) and receive one value or the other, but even the measures that explicitly state whether it is a cooler or a freezer receive a mix of both savings levels.	The confusion, if any, is a result of the data provided to us in the IOU program tracking databases. For example, many records showed "cooler" as the measure but claimed "freezer" as the energy savings. In such cases we used the energy savings claims as the more reliable data. SCE reported savings by case temperature (e.g. freezer, cooler) and by climate zone. As such, we created a lookup table that recreated the climate zone and case temperature from the SCE ex-ante per-unit savings. We chose to reconstruct the measure characteristics from the claimed per-unit savings because in our experience, the per-unit savings in program tracking data are usually more reliable than most other fields. We have checked our records and for SCE2511, if we had used the program reported case temperature and climate zone rather than the ex-ante savings, the total program level savings would be about 5% lower. We reported the higher number.
	The NTG value for Strip Curtains matches the NRDI-specific value in the table on page 136 of Appendix C, but this program breakdown does not exist in the study and the text preceding the table states that “No program has enough statistics to warrant a program-specific NTG. All programs are assigned the statewide average NTGR for door gaskets.”	The breakdown can be found in the "Commercial Facilities" contract report where all results for strip curtain and door gasket impacts can be found. Consistent with the ERT methodology, it is perfectly valid to use a measure-weighted NTG value for a specific measure and sum these up across all of the program's measures to achieve valid program-wide NTG impact values, just as it is valid to use a single NTG value to apply across all associated HIM results.

Source	Comment	Response
	The lighting measures for the Nonresidential Direct Install program seem to indicate systematic errors in the ERT. Many of the measures appear to have had their building type updated. Even so, many of the measures updated UES simply do not match either the ex ante or the updated building type. In one particularly egregious section, a set of measures with the exact same measure name, ex ante target sector and ED target sector and load shape had no fewer than seven different UES values. This is in no way supported by the evaluation, which specifies UES by building type (target sector).	Lighting results were applied to each measure by program and building type. The analysis aggregated records from the tracking data into sites. Each site was assigned a building type based on the EDDEERBuildingType as shown in the SPTdb. Sometimes, the records associated with a particular site were assigned different building types at the tracking record level. Only one building type was used for analysis and therefore the results were applied using that building type.
SCE	**Savings By Design Program (SCE2512)**	
	The ERT update is based on the evaluated realization rate to the SBD projects. For some projects it is not clear if the sample included projects involved refrigeration and refrigerated warehouses, and hence jeopardizes the appropriateness of applying the same realization rate to those projects.	Updated report: The refrigeration measure listed under the systems approach represent any refrigeration savings that were calculated from the building simulation tool. All refrigerated warehouses were considered industrial measures and used engineering calculations not the building simulation tool to calculate savings. The savings from these measures were included in the industrial measure savings.
	With regard to sample design, the evaluation used two phases of sample designs, but just as in some other studies, never made it clear how the sample designs were actually executed. It is conventional to report on the implemented sampling plan, the number of attempts to recruit, the number of refusals, the numbers of final surveys, and the numbers used in the final analysis, with explanations for how and why reality deviated from the ideal.	Section 3.1.2 [of the New Construction Evaluation Report] We have added to the sample design section of the report a summary that outlines the final sample design, number of calls places, number of sites dropped due to inability to recruit, reasons for dropping sites, along with the stratum weights by utility for electric and gas savings. Samples were designed at the utility level and not by approach, measure, or building type.
	In the evaluation study it was noted that the basic therm savings estimation techniques used for the IOU ex ante estimates need to be fundamentally re-examined. The study noted that "the lack of relationship is so poor that error bounds and relative precision have essentially no meaning." As a result no relative precision was provided for the therm savings. During the draft report commenting phase, because of the sampling issue, the Joint utilities had recommended that the therm model, and the therm portion of the results, not be accepted as reliable and used in the ERT process. The ERT nevertheless seems to have used the therm results from the evaluation study.	Section 3.4.3 [of the New Construction Evaluation Report] With over 70% of the total program tracking gas savings sampled we feel confident that our sample is more than adequate to represent the gas savings for the NRNC program. Weights were calculated separately for electric and gas savings so that each site would accurately represent its contribution to the population. The lack of correlation between tracking and evaluated savings for simulation building models meant that we could not use the ratio model that was used to analyze the electric savings. The precision and error bounds for the Mean Per Unit (MPU) analysis that was used on gas savings measures the variability of site savings from the weighed mean which means that if there is a large amount of variation in savings across the sample then the precision will be high and the error bound will be big. This does not mean that the

Source	Comment	Response
		estimates are inaccurate but simply that there is a great deal of variation (from -800,000 therms to 3,000,000 therms in the case of gas savings). The ratio model is able to deal with the larger variation by calculating at realization rate between tracking and evaluated savings, in that way even with a range of 0 to 5,000,000 MWh, the precision measures the variability of each site to the realization rate and not the weighted mean of the sample.
SCE	**Business Incentives & Services Program – Express Efficiency (SCE2517)**	
	Express Efficiency, a sub-program of Business Incentives & Services Program, did not receive a unified program evaluation. While no direct reason was given why this radical approach was necessary, it can be explained again by a misguided focus on measures and not programs evident in the report itself:	Express Efficiency was a program that consisted of several program elements and measures that also were prevalent in other programs. The combined evaluations covered over 80% of the energy savings claims for installation rates, unit energy savings, and net to gross ratios (see Appendix C; SCE2517 citations), with the remainder of the savings claims being passed through. The high impact measure approach was intended to achieve these high levels of coverage of energy savings with available resources, which may not have been possible with program evaluation approaches.
	"The major objectives of the impact evaluation are to estimate the energy and demand impacts produced by the HIMs and non-HIMs, to conduct research to inform the Commission's energy efficiency policy and program planning needs, and to provide feedback to program administrators and implementers in order to improve programs." (p. 1)	[See above]
	Finally, in checking source data for the ERT, SCE notes that the NTG ratio for strip curtains matches the documentation in Appendix C of the ERT, but is slightly different than the value in the Small Commercial study.	SBW applied to strip curtains an NTGR of 0.41. This matches the values noted in Appendix C by Small Commercial (p. 73) and Commercial Facilities (p. 135).
SCE	**Business Incentives & Services Program – Standard Performance Contract**	
	(SCE2517)	

Source	Comment	Response
	The same misguided focus on measures versus programs is supremely evident in SBW's Major Commercial Evaluation:	This comment was copied from the comments posted on the draft evaluation report. The comment and ED response can be viewed at www.energydataweb.com/cpuc, under the Major Commercial topic, the "Major Commercial Evaluation Report Comments/Responses," the SCE comments chapter.
	"The primary goal of the full impact evaluation was to assess the gross and net program-specific energy and demand impacts for high impact measures (HIM) and non-residential programs in the Major Commercial contract group." (p. 4)	Continuation of comment, see response above.
	While SBW did not announce, as other evaluators did, how the HIM switch hurt sample achievement, a total of only 18 sample points was achieved for an impact evaluation of SCE's major commercial program? As noted in our comments on the study, an impact evaluator team member had this to say regarding the robustness of the study results: "The projects become case studies, but you can't say anything at the program level." SCE agree. The ERT is similarly hard- pressed to justify program-level conclusions.	Continuation of comment, see response above.
SCE	**Business Incentives & Services Program – Nonresidential Audits (SCE2517)**	
	The Draft Report should contain an acknowledgement that, in actuality, the research analysis design was flawed – it assumed certain data would be available, and regardless of whatever information was found in the tracking systems, the data were forced to the design rather than altering or abandoning the design. The researchers relied on a proxy for site-specific energy savings. These numbers were then treated as if they were truly interval data, used to develop stratification, and ultimately compared to an on-site, measure-specific engineering estimate to generate a realization rate calculated to three significant digits — with a relative precision of 1.26.	This comment was copied from the comments posted on the draft evaluation report. The comment and ED response can be viewed at www.energydataweb.com/cpuc, under the Major Commercial topic, the "Major Commercial Evaluation Report Comments/Responses," the SCE comments chapter.

Source	Comment	Response
	In short, this is false precision. For the Nonresidential Audit part of the Business Incentive & Services Program, the quantitative analysis should be discarded because the starting point is unreliable and leads to "increased uncertainty around the program level savings estimate" (p 55). The application of some highly imprecise numbers to any quantification of the realization rates is not an appropriate analytic technique.	Continuation of comment, see response above.
SCE	**California Community Colleges Partnership Program (SCE2526)**	
	The Local Government Partnership program evaluation study claimed that for the CCC program "all project line items had UES and NTGRs applied based on direct EM&V studies. No results were applied from "Other EM&V" studies and no DEER updates were made to EUL's, as all projects were "custom" and not able to be directly mapped to a DEER measure." CFL giveaway UES updates are set as EMV-based, but the LGP study did not evaluate this measure, hence this update is not based on direct measurement. Rather a realization rate from customer projects was applied to the ex ante savings for this measure, which is totally inappropriate for this kind of non-custom measure.	While we concede that CFL giveaways are not custom measures, there were a number of factors that prevented the rigorous evaluation of this measure: 1.) CFL giveaways represented a marginal portion of CCC Energy Efficiency Partnership's savings. Given that the evaluation effort was already resource constrained (the impact evaluation sample frame targeted 90% confidence with 20% precision), the decision was made to focus on projects and measures that represented a larger portion of partnership savings. 2.) At the time realization rates were being calculated for the CCC Energy Efficiency Partnership's savings, findings from studies that would have been more applicable to CFL giveaways were not available.
	Similarly, there are other measures that get a realization rate adjustment based on two CCC partnership sample points regardless of their relevancy from a "custom project" perspective to all projects in the CCC program participant population.	The evaluation of the CCC Energy Efficiency Partnership was designated as Protocol Guided Direct with an Enhanced Level of Rigor. And as noted earlier, the impact evaluation sample frame targeted 90% confidence with 20% precision with individual samples drawn for each IOU-Fuel type permutation. More specifically, a cluster sampling approach was used where campuses were randomly drawn based on total expected savings of all projects at the campus. In the second stage, individual projects within the chosen campuses were selected to receive M&V activities. Therefore, even though a realization rate adjustment may have been based solely on two CCC Partnership sample points, those points represent a statistically significant portion of program savings. Overall, the impact evaluation sample represented 38% of gross ex-ante kWh savings and 37% of gross ex-ante Therm savings for custom retrofit projects in the 2006 – 2008 Program cycle.
SCE	County of Los Angeles Partnership Program (SCE2528)	

Source	Comment	Response
	Most of the customized projects in this partnership did not get directly evaluated. There are some cases of NTG updates in the ERT that are different for kW, kWh, and therms for the same measure. For most of the retro-commissioning measures in this partnership program the UES parameter is updated using the realization rate from the retro-commissioning program that was based on project-specific M&V plans, without any regard to the measure type or measure mix in the partnership program. There is no discussion or justification provided for the relevancy of the realization rate to this partnership program. In addition, the ERT mis- categorized the UES update to be EM&V rather than "other EMV."	RCx measures in this partnership program were part of the RCx HIM study population, and received RCx results. As such, "EM&V" is the correct category.
SCE	**UC-CSU Partnership Program (SCE2530)**	
	A large portion of the evaluation is based on project-specific M&V plans. The UES estimation did not report on the extent of measurement error that exists in such a type of analysis, which is also associated with individual evaluators and/or individual projects. participating project.	The evaluation objectives of the UC/CSU Energy Efficiency Partnership was to determine the impacts of all custom retrofit projects on gross annual program energy and peak demand, while accounting for interactions among them. As such, UES savings were calculated according to the following equation: UES = ∑ (IOUPrgTrkNTG * EDFilledExAnteGrSav(Fuel) * EvaluatedRR(Fuel,IOU) ,) A thorough discussion of potential measurement error and steps taken to mitigate their effects is provided in Section 6.4.1 (Key Uncertainty Sources and Mitigation Methods) of the Local Govt. Partnerships Impact Evaluation Report . The evaluation team identified sources of measurement bias and error and attempted to correct for them in the project-specific M&V plans.

Source	Comment	Response
	Any realization rate adjustments to UES to the program projects need to account for measurement errors, and difference in measure mix in the program	The evaluation of the UC/CSU Energy Efficiency Partnership was designated as Protocol Guided Direct with an Enhanced Level of Rigor. As such, the impact evaluation sample frame targeted 90% confidence with 20% precision and individual samples were drawn for each IOU-Fuel type permutation. More specifically, a cluster sampling approach was used where campuses were randomly drawn based on total expected savings of all projects at the campus. In the second stage, individual projects within the chosen campuses were selected to receive M&V activities. In other words, the sample was not stratified by measure mix because of the breadth of custom technologies offered, and the lack of documentation available at the sample design phase. The Stratified Ratio Estimation approach, consistent with the California Evaluation Framework, was used to calculate program level realization rates and relative precision estimates by fuel type. Overall, the impact evaluation sample represented 32% of gross ex-ante kWh savings and 54% of gross ex-ante Therm savings for custom retrofit projects in the 2006 – 2008 Program cycle.
SCE	**MAP Energy Efficiency Program (SCE2537)**	
	The ERT inputs for the MAP program are quite problematic. The evaluation found gross realization rates of 81% and 110% for kWh and kW, respectively, for CO sensors. But the ERT contains a small UES that is equal for each line item that is significantly different individual energy savings that were claimed. The Turbocor Chiller measure was evaluated, but the evaluators decided to accept the claimed value because of lack of information. These measures were coded "PassThru" which does not match the methodology used elsewhere where values passed-through by the evaluators were coded as "EMV." Appendix C has absolutely no explanation of how values were changed. It only has a table with program level savings and no description.	For Turbocor chillers, the evaluation included case studies, as described in the Specialized Commercial/HVAC Evaluation Report. We collected trend data from seven sites and monitored performance at three of these. Resulting data were used to create actual performance curves for each site, which were compared to performance data from other compressor manufacturers. Ex ante savings and NTG assumptions were passed through.
SCE	**Energy Efficiency Program for Entertainment Centers (SCE2561)**	

Source	Comment	Response
	The Demand Control Ventilation measures in the program were evaluated. It appears that the savings estimates in the ERT were changed from per project to per ton. Again, Appendix C has absolutely no explanation of how values were changed. It only has a table with program level savings and no description.	Commenter is correct for SCE2561. The unit energy and demand savings are presented per ton, but the quantity is presented per unit. The ERT units should be amended to reflect the tonnage of the unit rather than unit quantity (which is always =1). The tonnage for each record is available in the Standard Program tracking database. We corrected the input sheet for this program and re-ran it through the ERT.
SCE	**Palm Desert Partnership Program (SCE2566)**	
	The linear fluorescent fixtures installation rate is applied from the SCE2511 Direct install program at 93%, while for the same fixtures there are "pass through" values applied to other line items.	Results were applied at the building type level. For some building types, it was determined that the results from SCE2511 were not applicable to Palm Desert (poor precision for building type, building type not studied, etc.) For these building types that could not be mapped to a good number, the results were passed through.
	There is neither consistency in application nor, justification of why the Direct Install verification rate for fluorescent is as equally applicable to the Palm Desert Partnership measures that have different set of third party contractors doing the installations.	The Direct Install verification rate for the linear fluorescents in the rest of the portfolio is the best available estimate of what the verification rate is likely to be in an average program installation. There is no evidence that the Palm Desert program would have significantly higher installation rates than the average program in the portfolio.
	The duct sealing measure installation rate is used from the Specialized Commercial report (51%) for SCE2507, with precision levels at 22%, which makes it even more questionable to use this installation rate for the Partnership program that is delivered very differently than the measures in the comprehensive AC program.	The duct sealing measure installation rate from the Specialized Commercial report represents the best available estimate of installation rate for an average duct sealing measure in SCE's portfolio. There is no evidence to suggest that the results in the Palm Desert program would be significantly higher than the average for the rest of the portfolio. With a precision level of 22%, the 90% confidence interval would be between 40 and 62%. The upper end of this range is still much lower than 100%.

Source	Comment	Response
	Early retirement CAC was only given up-to-code savings, not full savings per table 8-8 of the evaluation report. The installation rate is incorrectly used as "quantity of tons installed" at 3. 73 tons/installation rather than converting it into proportion of installed units for a direct comparison with the ex ante estimates.	Every single AC early retirement reported in the program tracking database also had an energy efficient AC installation reported separately. Since both an early retirement and an energy efficient installation were reported for each actual AC early retirement, there were two options: 1. Associate all savings associated with each installation to the early retirement measure and give 0 savings to the energy efficient AC installations reported in conjunction with an early retirement. 2. Split out the savings associated with existing up to code minimum and code minimum up to high efficiency. We elected to use option 2. Both options will give the same total savings for each actual installation event, regardless of whether we assign the savings only to the early retirement line items or split them between the early retirement and high efficiency AC line items. The 3.73 tons/installation was necessary to correct for errors in the way that the installations were reported. The units for early retirements and central AC installations were stated to be tons. In earlier versions of the program tracking database (e.g. Q208), quantities were actually being reported in terms of tons. In the final program tracking database (Q408), quantities had inexplicably been changed to number of units installed, with the same ex ante savings per ton values used, and units still claimed to be tons. The actual installation rate associated with the measure was 100% on the basis of number of units installed, and 373% on the basis of tons installed, to reflect that we saw an average 3.73 tons per unit installed. In order to maintain the integrity of the comparison in ex ante and ex post savings per ton, an installation rate of 373% was used to account for the units error in the tracking database.

Source	Comment	Response
	CAC maintenance services get two different sets of UES- one from EM&V and one simply a pass through. The installation rates are also not consistent for the same climate zone. For example, CZ 15 in the residential sector, some measures received a pass through, and others 10%. Additionally, these two records have different NTG values, which should be 76% and not the 69% in the ERT.	This comment is essentially the same as an earlier one. The response is copied from above. There were two different central AC maintenance measures performed under SCE 2566 that were named the same thing in the E3 calculator but had different names in the program tracking data. One measure was for one-time major maintenance and was evaluated using on-site verifications in SCE2566 for installation rate, measure-specific NTG surveys and UES values from the specialized commercial contract group. The other measure was for maintenance contracts. No direct measurement of savings or installation rate for maintenance contracts was conducted in SCE2566 or elsewhere. As a result, the installation rate and UES were passed through. The one-time major maintenance measure was given a measure-specific NTG value because there were enough sample points in the NTG survey to do so. The maintenance contract measure was lumped in with all of the residential measures that were not given measure-specific NTG results. These measures were surveyed as their own stratum and received a different NTG than the one-time major maintenance measure.
	Room AC received an installation update from the Residential Retrofit HIM results, again not accounting for the differences in programs that deliver this measure.	The Room AC installation rate from the Residential Retrofit contract group represents the best available estimate of installation rate for an average SCE program delivering this measure. There is no evidence to support using a significantly higher installation rate in Palm Desert.
	The baseline gasket conditions can be widely different between Palm Desert and PG&E's high technology and large customer segment programs that provided the samples for the field measurements data for this measure. Door Gasket UES does not match the ADM reported UES kwh/linear square feet in Table 5-1 of the report. Also, kW UES values are not found in the report. No interactive effects were applied to commercial refrigeration measures in this case. We also did not find installation rate in the ADM study for door gasket yet the ERT indicated "other EMV" as the source of the 1.0 IR.	Table 5-1 gives an average result across all installations. A higher UES for freezer applications in Climate Zone 15 was applied to Palm Desert, obtained from ADM. The kW UES values are in the HIM appendix of the Commercial Facilities Report. ADM assigned an installation rate of 1, and it was categorized as "OthEMV".
	The reflector CFL has wrong the NTG (67% versus 69%).	This is not true, a 67% NTG was not applied to ANY measures in Palm Desert.

Source	Comment	Response
	The In Home survey NTG is not found in the EMV report contrary to the reported source for this update in the ERT documentation.	This same comment was made in a previous section. The response is copied here. The in-home survey measure was lumped in with all of the residential measures that were not given measure-specific NTG results. This included all residential measures that were not residential major central AC maintenance, energy efficient AC, or AC early retirement. These remaining residential measures were surveyed as their own stratum and received a NTG value weighted for the mix of measures in this stratum, which was then applied to all residential measures in that stratum.
	The Pool Pump measure uses the SDG&E UES, which is based on half as many hours of usage as expected in the Palm Desert area. There are other empirical data collected in recent studies, including SCE workpapers that show hours of operation in Palm Desert to be very different than those of other mild climate regions like SDG&E's. Also, the installation rate is based on SDG&E's program that postulates voluntary change to off-peak usage, which in the Palm Desert program is delivered very differently.	The best available estimates for pool pump program savings come from the study in SDG&E. The SCE workpapers do not have any kind of field measurement behind them that documents the difference in operating hours between the two areas.
	The source of installation rate is not documented for screw-in CFL in the ERT documentation. UES for the same screw-in CFL varies sometimes DEER, other EM&V, or EM&V.	Part of this comment was addressed earlier. That response is copied here. Results were applied at the building type level. For some building types, it was determined that the results from other contract groups were not applicable to Palm Desert (poor precision for building type, building type not studied, etc.) For those remaining building types, results were mapped from DEER where possible. For the remaining measure-building type combinations, values were passed through. The source of installation rate is documented for screw-in CFL in the ERT documentation wherever an updated value was applied. Many (especially residential) were passed through.
	The Draft Report Systematically Reduces SCE's Program Avoided Cost Benefits	Please see the responses in the main comments; SCE/18
	The Draft Report Contains 538 "E3 Calculator FALSE" Errors Which Assign Zero Avoided Cost Benefits To Valid Measure Installations	
	The Draft Report Alters The Building Types Of SCE's Measure Installations By Unknown And Unwarranted Methodologies	
SCE	**(THESE ARE ALL LISTED IN THE MAIN COMMENTS)**	
SCE	**ERT Quality Control Errors**	

Source	Comment	Response
	There are a number of quality control errors found in the ERT. The tool itself has multiple built-in quality control checks that are shown to fail in multiple instances.	There are two things to keep in mind about the quality control (QC) functions of the ERT. First, the QC checks are intended as a preliminary check on the input data and a tool to help identify potential issues. They have nothing to do with the actual results of the ERT or the processing of the data files. It is not required to run the QC checks. Second, the QC checks can produce false positives depending on the scenario. Because there are twelve scenarios, the QC checks may detect a potential problem for one scenario that may not apply to the other scenarios. For example, one of the QC checks will check to see if there is a Gas Profile and Gas Sector if Therms are reported. Because SCE has no Gas Profiles or Gas Sectors, evaluated records with Therms will result as a false positive, but only for one scenario. We acknowledge that the QC checks can be improved by making it more clear which scenarios apply and which data source the issue is with (for example, no gas sector in the IOU Claim file). However, the comment "With such a large majority of quality control issues, the ERT and Draft Report results are shown to be extremely flawed" is incorrect because the QC checks err on the side of caution, and have no bearing on the actual processing of the data files through the E3 calculators, but were intended for the QC on the input data.
	The Draft Report uses the incorrect data field from the 2006-2008 program cycle tracking data. In this case, the "Estimated Incentive" field was populated by the ERT instead of the actual incentive value located in the "Calculated Incentive" field.	This comment seems to be a comparison between the EDFilledIncentivePaid ($4,897,719.37) in the SPT and the IOU E3 Claim incentive ($4,935,742.57). SBW double checked when the IOU E3 claim incentive does not match the ED Filled incentive. This occurred on 6 occasions for SCE programs because there were multiple incentive fields. In all occasions, ED selected the incentive value that was closest to the IOU E3 value. The different in impact ranged from .773% to .0024%. No changes are warranted. The numbers used in the report are correct.

Source	Comment	Response
	Another example of a programmatic error found in the ERT database would be the misrepresentation of the ex ante net energy savings.	This mismatch occurs between option 0 and option1. Response. Option 1 of ERT uses the values from the IOU submitted E3 calculators and reruns the E3 calculator. If there is a discrepancy between option 0 and option 1 it is either due to a misreading of the submitted E3 file values, or the submitted E3 calculator is an incorrect version, a customized E3, or the values have been manually altered. A review of IOU submitted E3 calculators where there is an option 0 – 1 discrepancy found no instances where the values were misread from the submitted E3 claim file. The review seems to indicate that results in the submitted claim E3 files do not match the results if the input values are manually re-input into a clean E3 calculator. In other words, any discrepancies between the submitted files and option 1 results are due to errors in the submitted claim E3 files. It is impossible to tell where the errors are in the submitted claim E3 calculators because they are 'exported', and therefore do not have the original formulas intact.
SCE	**Misapplication Of DEER**	
	The process for evaluations laid out specific guidelines for the evaluation of SCE's portfolio. Included in this was specific guidance to use DEER 2008 v2.04. The analysis within the evaluation report did not consistently follow that directive. In the process of modifying stated policy, the evaluation report introduces other changes which raise numerous questions including lack of transparency, lack of consistency, increased uncertainty, and quality of data issues.	Energy Division followed specific guidelines for the evaluations of all utility portfolios and has provided detail regarding how the evaluations were conducted and how the results were applied to develop the final report. Specific issues are addressed in the following responses.

Source	Comment	Response
	Not only was the utilization of DEER 2008 v2.04 process not consistently followed, but a new version of DEER, which has not been formally released or fully vetted, was used in multiple cases to apply interactive effects. The newer non codified version of DEER uses a different version of the DOE2.1 software, contains building code updates that occurred in 2010, and contains adjustments for numerous modeling assumptions. Some of the modeling assumptions found in DEER v3.02 differ (which has almost the same factors as Appendix B of the Draft Report) from those in DEER 2008 v2.04; the most important of which is weighting each measure value by prototype building/HVAC type using RASS and/or CEUS which impacts both the direct and the interactive effects. Note that the RASS and CEUS data used in this weighing is several years old.	Energy Division did not use version DEER v. 3.02. It does not exist. In 2009, to better estimate interactive-effects, the ED DMQC Team, with assistance from the ED DEER Team, provided an interactive-effects spreadsheet for the ERT Team. Several heating and cooling system types were added to the DEER dataset, and air-conditioning and heating saturations were applied which mitigate the negative therms impact. Additionally, a couple of errors identified in the DEER 2008 analysis software tool were corrected. Energy Division has the right and responsibility to enhance interactive effects estimates in the DEER Database, and to correct errors identified.
	• The interactive effects worksheet/book is using the DEER v3.02 approach to calculating the interactive effects. This introduces a weighted average of cooling/heating system types that were not present in DEER v2.04 which modifies the interactive effects from the DEER database used for this period. Further description and examples of interactive effects will be addressed in the next section.	We knew that HVAC interactive effects data were important, and made a conscious decision to use the HVAC interactive effects data from the latest DEER modeling for consistency with the 2010-2012 planning. The new DEER models also have additional HVAC system types not present in the earlier versions.
	• The interactive effects from the Draft Report Appendix B tool had five measure types, two for Residential Measures and three for Non-Residential measures. Based on these measure types, the interactive effects were not consistently applied to the measures. In the example below it can be seen that an Exit Sign Measure that, according to Appendix B (DEER 2008 v3.02), should have had interactive effects applied to it, but did not.	• Interactive effects were often inconsistently applied across different measure types. In some instances, custom measures, regardless of savings impact, received the same therm benefits value from interactive effects, in others the values changed by measure. Therm impacts for SPC component of SCE2517 appear to have a processing error. Non-interactive therm savings were incorrectly assigned to measures. The result is an over prediction of the SCE net therm savings by 0.19%. A similar problem was observed with SDGE3025, resulting in an over prediction of SDGE net therm savings by 0.4%. Energy Division has corrected this error in the final results.

Source	Comment	Response
	- In general, it is unclear how the interactive effects values were obtained for Misc. Commercial. There is no Misc. Commercial building type in the Draft Report Appendix B tool (DEER 2008 v3.02). In the example below, however, it can be seen that some of the misc. commercial building types seem to have been linked to the Assembly building interactive effects. The chart is trying to portray that for an identical measure the interactive effect for assembly is coming up on every other measure line while the other interactive effects' origin cannot be determined.	Misc. is not a DEER building type. Misc/unknown HVAC interactive effects were defined as a weighted average of known building types, which varies by program.
	As with previous examples, DEER 2008 v3.02 interactive effects were misapplied to the residential lighting measures.	PG&E. HVAC interactive effects were incorrectly applied to some exterior lighting measures in programs PGE2036 and 2080, resulting in an over-estimation of the kW and kWh savings by 38 kW and 62,242 kWh, and an under-estimation of the therm savings by 7,164 therms, representing 0.01%, 0.02% and -0.3% of the total PG&E ex-post net kW, kWh and therm savings respectively. Holiday lights were incorrectly assigned to the Exterior Lighting Other measure group. HVAC interactive effects were appropriately applied to that measure. SCE Therm impacts were incorrectly applied to exterior lighting measures in programs SCE2517, resulting in an under-estimation of the therm impacts by 554 therms, representing 0.004% of the total SCE ex-post net therm impacts. SDG&E Outdoor CFL fixture measure group assigned to project in SDGE3025. Not clear from tracking data if these are outdoor or indoor fixtures. Assuming the fixtures are indoor; HVAC interactive effects are applied correctly. Outdoor CFL fixture and outdoor screw lighting measure groups incorrectly assigned to some interior CFLs in Program SDGE3035. HVAC interactive effects were appropriately applied. [Energy Division has made these corrections in the final results]

Source	Comment	Response
	By introducing a new version of DEER for this report, the analysis is also introducing a new un-verified data set that has not been adjusted for the most recent impact data, which is the primary purpose of this evaluation report.	Energy Division did not use version DEER v. 3.02. It does not exist. In 2009, to better estimate interactive-effects, the ED DMQC Team, with assistance from the ED DEER Team, provided an interactive-effects spreadsheet for the ERT Team. Several heating and cooling system types were added to the DEER dataset, and air-conditioning and heating saturations were applied which mitigate the negative therms impact. Additionally, a couple of errors identified in the DEER 2008 analysis software tool were corrected. Energy Division has the right and responsibility to enhance interactive effects estimates in the DEER Database, and to correct errors identified.
SCE	**Interactive Effects**	
	• As indicated in the report, no part of the 2006-2008 impact evaluations specifically addressed this lack of interactive effects data, even though this is a known gap in the data set.	
	• The data set for customized projects does not clearly indentify how the interactive effects were developed when they were applied. Except in few cases where simulations models were used, and the interactive effects were implicitly dealt with (as explained in the report), the balance of the custom projects would not have Interactive effects that could be directly derived from DEER since their savings values use different assumptions than DEER.	Full building models were not conducted on all the custom lighting projects. HVAC interactive effects multipliers are a secondary effect, and using factors derived from building prototypes is adequate even if some of the assumptions in the prototypes are different from any particular site. An uncertainty analysis would show the uncertainty in the HVAC interactions are a minor contributor to the overall uncertainty in the estimate.
	• As indicated previously, in some cases, the different versions of DEER ascribes multiple sets of interactive factors which lead to inconsistencies in the application of these values from the data sets. In other cases, the source of the interactive factors was unclear.	The specific versions of the DEER runs used for the HVAC interactive effects calculations were specified by ED.

Source	Comment	Response
	• Interactive effects were not properly applied across all measure types. External Lighting measures, which are known to not have any HVAC interactive effects, received interactive effects credit, which is a blatant error. In the example below, interactive effects were applied to an exterior lighting measure. This is an erroneous application of interactive effects as exterior lighting has no effect on the HVAC systems of the building.	PG&E. HVAC interactive effects were incorrectly applied to some exterior lighting measures in programs PGE2036 and 2080, resulting in an over-estimation of the kW and kWh savings by 38 kW and 62,242 kWh, and an under-estimation of the therm savings by 7,164 therms, representing 0.01%, 0.02% and -0.3% of the total PG&E ex-post net kW, kWh and therm savings respectively. Holiday lights were incorrectly assigned to the Exterior Lighting Other measure group. HVAC interactive effects were appropriately applied to that measure. SCE Therm impacts were incorrectly applied to exterior lighting measures in programs SCE2517, resulting in an under-estimation of the therm impacts by 554 therms, representing 0.004% of the total SCE ex-post net therm impacts. SDG&E Outdoor CFL fixture measure group assigned to project in SDGE3025. Not clear from tracking data if these are outdoor or indoor fixtures. Assuming the fixtures are indoor; HVAC interactive effects are applied correctly.. Outdoor CFL fixture and outdoor screw lighting measure groups incorrectly assigned to some interior CFLs in Program SDGE3035. HVAC interactive effects were appropriately applied. [Energy Division has made these corrections in the final results.]
	Interactive effects were often inconsistently applied across different measure types. In some instances, custom measures, regardless of savings impact, received the same therm benefits value from interactive effects, in others the values changed by measure.	A line item in the tracking system represents a project, which often consists of multiple measures. Therm impacts come from two sources: non-interactive M&V and interactive effects. The non-interactive and interactive therms were distributed differently. The non-interactive average UES therms were extrapolated to all electric measures. The M&V found an average therms savings per sampled electric measure. Every measure in the population received that value. In addition, interactive therms were assigned to projects according to the kWh share of lighting to the total kWh savings.
	Another example of Linear Fluorescents in the Palm Desert Partnership where the “end use” is from DEER v2.04, but the “whole building” values do not match DEER v2.04.	Contractors followed guidance given by ED to use the HVAC interactive effects multipliers while using the end-use savings for DEER measures from DEER v2.04. Differences in the HVAC interactive effects multipliers account for the differences in the Whole Building savings. Note, the commercial building HVAC system type selections were expanded. The HVAC system types were weighted, and the weighted

Source	Comment	Response
		average multipliers were used to calculate the HVAC interactive effects.
	The same trend occurs for CFL lighting in the Palm Desert Partnership, for example, where it appears that the Draft Report's values proposed for a 23 Watt indoor CFL "end use" do match DEER v2.04, but the "whole building" values do not match. It is unknown where the "whole building" values were derived from since they do not match DEER v3.02.	*See response above.*
	In the Nonresidential Direct Install Program, it is unclear what source of interactive effects were utilized since they do not match any of the likely versions of DEER.	The HVAC interactive effects apparent in the ERT match the DEER derived values exactly for all cases checked. Spot checks comprised all linear fluorescent lighting measures in small retail and small office buildings across all climate zones.
	• In the Misc. Commercial building type, the therms interactive effects for the same measure type and climate zone are not consistent. This goes against the logic in DEER v3.02 that has a specific value based on building type and climate zone.	Misc. is not a DEER building type. Misc/unknown HVAC interactive effects were defined as a weighted average of known building types, which varies by program.
	For Non-Residential audits, the interactive effects used could not be derived from DEER 2008 v3.02.	Audits are a combination of HVAC interactive and non-interactive measures. HVAC interactive effects are applied to a portion of the tracking line item, so it is reasonable that the HVAC interactive effects do not resemble the DEER values, which are for a single measure. The factors derived from the ERT do not vary by climate zone or building type, because the realization rates are calculated relative to the ex-ante values which also do not vary by climate zone or building type.
	• In many cases, including SCE's Residential Lighting Program, the interactive effects were applied from the wrong climate zone, illustrating quality controls issues with the processing of this data.	Upstream residential screw-in CFL component of SCE2501 used a system wide rather than CZ specific number for HVAC interactive effects, since locations of customers is not known. The HVAC interactive effects for these lamps are described in the next section below. These lamps make up 88% of the kW savings, 80% of the kWh savings and 96% of the therm impacts. The downstream residential screw in CFL installations report the climate zone. The HVAC interactive effects assigned to these measures match the HVAC interactive effect factors.

Source	Comment	Response
	• For some of the measures in Palm Desert Partnership Program, it appears that the interactive effects were applied correctly to kW and kWh; however for the therms value, the interactive effects values are for climate zone 16 not 15. This creates a larger negative therms value.	Entries were spot checked across a variety of building and measure types, and the HVAC interactive effects for kW, kWh and therms are correct.

Source	Comment	Response
	It appears that all of the interactive effects for the Residential sector in SCE's Residential Upstream Lighting program are being applied incorrectly.	The HVAC interactive effects multipliers were derived weighted averages across HVAC types and building types. The HVAC interactive effects in the ERT match the weighted HVAC interactive effects calculations exactly when the weights used by the contractor are applied. See tables below. SCE System type weights by building type Building Type GasPac HP ElecHeat GasFurn DX/Other Unconditioned Single Family Residential 67% 4% 0% 19% 2% 8% Multi-Family Residential 38% 8% 18% 27% 0% 8% Double Wide Mobile Home 80% 16% 0% 0% 0% 4% Weighted HVAC interactive effects by building type Building Type (kWh/kWh) (kW/kW) (therms/kWh) Single Family Residential 1.0775 1.4345 -0.0210 Multi-Family Residential 0.9852 1.2252 -0.0127 Double Wide Mobile Home 1.1970 1.7317 -0.0198 Building Type Weights Building Type Weight Single Family Residential 0.682 Multi-Family Residential 0.269 Double Wide Mobile Home 0.049 Weighted upstream CFL HVAC HVAC interactive effects Building Type (kWh/kWh) (kW/kW) (therms/kWh) All 1.0586 1.3929 -0.0187 Non-interactive and Interactive UES Values from ERT EDUESkWh EDUESkW EDUESkWhi EDUESkWi EDUESThermsi 31.07 0.00287 32.89 0.00399 -0.580 HVAC interactive effects calculated from ERT UES (kWh/kWh) (kW/kW) (therms/kWh) 1.0586 1.3929 -0.0187 Note, values in the table above match the weighted values exactly

Source	Comment	Response
		PG&E System type weights by building type Building Type GasPac HP ElecHeat GasFurn DX/Other Unconditioned Single Family Residential 49% 4% 2% 30% 7% 9% Multi-Family Residential 34% 7% 21% 31% 0% 7% Double Wide Mobile Home 17% 38% 0% 18% 17% 11% Weighted HVAC interactive effects by building type Building Type (kWh/kWh) (kW/kW) (therms/kWh) Single Family Residential 1.0323 1.3824 -0.0239 Multi-Family Residential 0.9151 1.2032 -0.0172 Double Wide Mobile Home 0.9920 1.6221 -0.0107 Building Type Weights Building Type Weight Single Family Residential 0.683 Multi-Family Residential 0.271 Double Wide Mobile Home 0.046 Weighted upstream CFL HVAC HVAC interactive effects Building Type (kWh/kWh) (kW/kW) (therms/kWh) All 0.9987 1.3447 -0.0214 Non-interactive and Interactive UES Values from ERT EDUESkWh EDUESkW EDUESkWhi EDUESkWi EDUESThermsi 30.72 0.00284 30.68 0.00381 -0.659 HVAC interactive effects calculated from ERT UES (kWh/kWh) (kW/kW) (therms/kWh) 0.9987 1.3447 -0.0214 Note, values in the table above match the weighted values exactly SDGE System type weights by building type Building Type GasPac HP ElecHeat GasFurn DX/Other Unconditioned Single Family Residential 59% 2% 2% 25% 4% 8%

Source	Comment	Response
		Multi-Family Residential 32% 33% 11% 17% 2% 5% Double Wide Mobile Home 79% 0% 0% 21% 0% 0% Weighted HVAC interactive effects by building type Building Type (kWh/kWh) (kW/kW) (therms/kWh) Single Family Residential 1.0541 1.3260 -0.0185 Multi-Family Residential 0.9732 1.2878 -0.0091 Double Wide Mobile Home 1.2007 1.5718 -0.0208 Building Type Weights Building Type Weight Single Family Residential 0.667 Multi-Family Residential 0.289 Double Wide Mobile Home 0.044 Weighted upstream CFL HVAC interactive effects Building Type (kWh/kWh) (kW/kW) (therms/kWh) All 1.0371 1.3257 -0.0159 Non-interactive and Interactive UES Values from ERT EDUESkWh EDUESkW EDUESkWhi EDUESkWi EDUESThermsi 24.31 0.00284 25.21 0.00377 -0.386 HVAC interactive effects calculated from ERT UES (kWh/kWh) (kW/kW) (therms/kWh) 1.0371 1.3257 -0.0159 Note, values in the table above match the weighted values exactly
SCE	**Misapplication Of Data**	
	The systematic errors in applying UES values include:	

Source	Comment	Response
	• Indeterminable adjustments were made to the same measure found using the same climate zone and building type which resulted in different UES values. This occurred in the Non-Residential Audit and Direct Install programs. Conceptually, these adjustments should be producing consistent UES values.	SBW used realization rates not UES values for the Non-Residential Audit (NRA). NRA measures were not updated based on DEER - climate zone and building type were not relevant - they were updated based on M&V.
	• In the Non-Residential Direct Install program, savings were applied inconsistently. ED's kWh/kWh (i) and kW/kW (i) values are different even though the measure, building type, and climate zone are the same as shown in the table below. Only an example is shown here. The same issue is encountered for all climate zones. Extreme outliers are shown in bold red.	Interactive effects were applied by EDDEERBuildingType, not by EDTargetSector, using the file provided by the DEER team (Interactive Effects_100218.xls).
	Unlike the previous example where no difference was applied, it appears that the differences between coolers and freezers where applied to this measure where they should not have been.	The issue stems from the data provided in the IOU program tracking databases. For example, many records showed "cooler" as the measure but claimed "freezer" as the energy savings. Energy savings claims were used as the more reliable data. Please see our previous responses for more details. The savings were applied in a fashion that was consistent with the most reliable field in the program tracking data - the ex-ante savings.
	• In the Business Incentives and Services (SPC and Nonresidential Audits) Program, savings were applied inconsistently for the same measure. For Nonresidential Onsite Audits for large, medium, small and very small buildings, the factors used to reduce the savings cannot be determined. However, the example below shows no pattern for EDUES kW/kWh changes for a deemed measure, having the same climate zone and same target sectors.	SBW did not use deemed UES values for SPC or NRA. The impacts of the interactive effects were determined by applying the ED-supplied interactive factors to those portions of the custom and indirect measures where they applied. After applying the interactive factors on a case by case basis to the studied measures, new realization rates and UES values were recalculated for the entire program elements. These new rates were then applied across the board to the sample populations. The resulting ratios in interactive vs. non-interactive values would not be recognizable as DEER interactive factors. This work is shown in our interactive site-specific workbooks.

Source	Comment	Response
	In some programs, UES and interactive effect changes were applied to some climate zones, but not others. Measures that are assigned to Education Primary School and Education Secondary School for climate zones other than 8 are given a kW interactive effect of 1. This infers that only climate zone 8 has any school operation during the DEER peak period of the three consecutive hottest days. This does not make intuitive sense that only schools in climate zone 8 would run a summer school program or schools outside of climate zone 8 would be completely shut down. In the screen shot below it can be seen that in DEER v3.02, only climate zone eight has a non 1 interactive effect for kW.	The DEER peak period occurs during summer break for primary and secondary schools in all SCE climate zones except CZ 8. The school occupancy schedules have not changed for quite some time.
	For Home Energy Efficiency Energy Star Room AC, half of the climate zones were passed through in the Draft Report and the other half received a large change in savings values.	The correct approach was to apply the update to all climate zones and has been corrected in the final report.
	• Some measures that would have been assigned to DEER do not have a clear source. This again indicates a lack of consistency in the data processing.	The specific issues are addressed in the following comments.
	• In the Non-Residential Direct Install Program (SCE2511) it is unclear what source of interactive effects were utilized since they do not match any of the likely versions of DEER. There are also unclear translations for building type and climate zones. For example, the Screw-in Compact Fluorescent Lamp, 14-26 watts. SCE's ex ante values are from DEER 2005, 18 W CFL. The SCE2511 input sheet's kWh/kWh (i) and kW/kW (i) values are different even though the measure, building type, and climate zone are the same.	Interactive effects were applied by EDDEERBuildingType, not by EDTargetSector, using the file provided by the DEER team (Interactive Effects_100218.xls).
	Another example is the Screw-in Compact Fluorescent Lamp, >27 wattsSCE values are from DEER 2005, 28 W CFL. The Draft Report's kWh/kWh (i) and kW/kW (i) values are different even though the measure, building type, and climate zone are the same.	Interactive effects were applied by EDDEERBuildingType, not by EDTargetSector, using the file provided by the DEER team (Interactive Effects_100218.xls).

Source	Comment	Response
	In the Non-Residential Direct Install Program, Gasket savings were not adjusted by cooler/freezer measures, even though SCE reported them this way and the report Indicates that analysis was done to that effect. The Draft Report does not differentiate between coolers and freezers regarding kWh and kW. This leads to a large reduction in kWh & kW as they are very different measures.	The gasket savings were adjusted by ratio of cooler to freezer measures. For SCE 2511, the program tracking data show that about 20% of the gaskets were installed on freezers and 80% of gaskets were installed on coolers. The typical savings for gaskets on freezers for that program was 3.6 kWh/ft, and the typical savings for gaskets on coolers for that program was 0.5 kWh/ft. Our average was 1.13 kWh/ft, which makes sense according to the previous numbers. However, in some cases we were not able to determine the store-type or case-type from the tracking data. In these cases, we did not feel justified in applying our realization rates and we deferred to the SCE ex-ante savings applied a "PASS THROUGH". As a result, the actual average per-unit savings was almost twice the 1.13 kWh/ft.
	• Similar to the other Residential Upstream Lighting measures, a blanket value for kW and kWh was applied across multiple measures/wattages; however in this case the measures are both interior and exterior CFL measures which have very different operating hours. While it is understood that the impact study analysis only presented high level results, mixing of interior and exterior is problematic since the savings, Load shapes and demand reduction values are significantly different for these two applications. In the target sector, the Draft Report labels measures Misc. Commercial for the same measure mentioned above, a blanket kW of 0.01638271 and kWh of 98.9017007 was applied to the measures.	Fixture UES values were weighted by the shipping distribution of such measures to produce one single weighted UES for all fixtures for the ERT.
	• In the table below is another example from SCE's Industrial program. This measure has the same measure name, same RunID, same climate zone, yet the reduction to the kW and kWh is inconsistent.	As explained in the "2006-2008 Evaluation Report for the Southern California Industrial and Agricultural Contract Group", SCE's Industrial Program SCE2509 was evaluated using the Realization Rate approach and five electric strata. This evaluation approach applies strata-specific results to each record, not measure-specific results.

Source	Comment	Response
	ED should revise the Report's Executive Summary to provide appropriate context and a balanced presentation of its findings. The executive summary sets the stage for the evaluation results by describing how many cars have been taken off the road or stressing that "two-thirds of these savings would not have occurred without program intervention." However, the Report should also provide information about how success of the programs is to be measured in the context of CPUC-assigned energy savings goals and the shareholder incentive mechanism. At a minimum Table 23[1] on Comparative of Program Cycle 2006-2008 Evaluated Results to Goals (summary results) should be utilized in the executive summary	Energy Division believes that the discussion of goals is sufficiently complex to deserve its own section and that is in section 4. The presentation in the executive summary provides the historical context and a complete discussion is included in the body of the report.
	These summary results should be accompanied in the executive summary by a discussion of utility energy savings targets and how they were established.	A brief discussion is provided, and the more detailed information is provided in to Section 4.
DRA	Summary results should be accompanied in the executive summary by a discussion of the shareholder incentive mechanism's required level of achievement for incentives in order for the "performance against goal" percentages to be meaningful.	The shareholder incentive findings are provided in the Energy Division Scenario Report
DRA	The Report should be revised to be appropriately balanced in its discussion of cost-effectiveness. in actuality, ratepayers break even: they invested their dollar in EE and they got their dollar back – then they netted an additional 17 cents.[2] Accordingly, the Report should not state that the net benefit is the entire $1.17.	Energy Division has clarified the discussion in the final report.
DRA	The Report should add a discussion of incremental measure costs that were not included in this calculation.	The discussion of incremental measure costs, and the fact that they are not included in this calculation is already present in section 3.4.
DRA	The Report should also clarify that shareholder incentive payments to-date ($144 million) and any future shareholder bonuses have not been included in the PEB.	The shareholder incentive findings are provided in the Energy Division Scenario Report
DRA	The Report should be revised to remove unsupported findings such as market effects. The Report accentuates presumptions of speculative energy savings through statements such as "even though accounting of costs and benefits has been done strictly according to Commission-adopted rules and practices, in reality utility programs are likely providing additional long-term societal benefits that are not captured in this analysis.	Energy Division conducted several non-resource studies, market effects studies and several indirect impact studies that provide evidence to support the statement that utility programs are likely providing additional long term societal benefits than those included in the cost benefit calculations. These findings are highlighted in section 2 (summary of evaluation results). In addition, the Commission policy to exclude spill over impacts, which were not quantified in these

Source	Comment	Response
		evaluations, also supports this statement. Energy Division has provided more specificity to support this claim in the Final report
DRA	**ED should ensure that all tables and charts are appropriately updated.** On April 22, 2010, ED provided via email updates to the Report for pages 95-97 for Tables 23 and 24. Table 23 showed a greater discrepancy between Reported and Evaluated energy savings. Yet Table 3, for example, on Reported and Evaluated Net Savings as a Percentage of Savings Goals since 2002 which serves to illustrate the growing discrepancy was not updated commensurately.	Energy Division has corrected all tables and charts in the final document.
DRA	In order to be consistent in its recommendations, ED should retract its recommendation that the findings of its Report should be segregated from shareholder performance earnings. ED states that in order to "remove the disincentives to making productive use of the information generated by the EM&V work and to encourage the pursuit of all Commission energy efficiency policy goals" the report should not be used in the shareholder incentive mechanism process. DRA supports ED's objective to utilize its findings to update savings estimates and to improve EE programs. However, given that it is the utilities own inflated savings assumptions that are being leveraged to create untenable energy savings projections in order to promote maximum shareholder bonuses (not real energy savings), it is impossible to segregate the updates to energy savings estimates from the shareholder incentive mechanism. ED's Report provides critical data which should be used across the board to provide accountability for EE program design, impact on procurement, and shareholder bonuses.	Energy Division continues to hold this position, with further detailed explanation available in the April 1, 2009 whitepaper entitled: "Proposed Energy Efficiency Risk Reward Incentive Mechanism and EM&V Activities" which can be found at: http://www.cpuc.ca.gov/NR/rdonlyres/A51D61E2-DF03-4D9B-BFDB-221109638165/0/ProposedEnergyEfficiencyRiskRewardIncentiveMechandEM_VActivities.pdf
SCG/S	**SoCalGas Seeks Clarification on Tables 11 and 29.**	

Source	Comment	Response
DGE	It is not clear to us how the CPUC gets to the "0.30%" value from the 22,212,713,417 Therm total they say they've used along with any of the Therm values they've provided in the corresponding row of Table 11. We do not agree with the base gross number of 22,212,713,417. The base of 22,212,713,417 therms should be replaced by 16,022,250,000 which is the sum of 537,337 MDth for Yr2006, 537,493 MDth for Yr2007and 527,395 MDth for Yr2008, converted to Therms (multiplied by 10,000).	Energy Division has worked with SoCal Gas to correct this error and it is reflected in the updates to the report. An additional error in the calculation was also uncovered in reviewing this comment for the other utilities. In the Draft Report Energy Division divided the gross cumulative savings achieved by the end of 2008 by the total cumulative sales in 2006, 2007 and 2008. However, the savings achieved at the end of 2008 should have been compared to the consumption in 2008 since these were savings that were "on the grid" at that time. This means that roughly 1% per year savings v. consumption were achieved in 2006, 2007 and 2008, and by the end of 2008 roughly 3% of consumption was matched with energy efficiency available on the grid. These corrections we'll be made in the final report.
SCG/SDGE	**The Verification Report Errs in Using the 2006-2008 Load Impact Evaluations**	Energy Division has provided responses to all of these comments in the course of reviewing the individual contractor reports.
	SDG&E and SoCalGas believe that relying on the 2006-2008 load impact evaluations to update the program achievements due to numerous deficiencies and errors in the reports.	The evaluations that were completed of the 2006-2008 programs represent the best available data on program performance based on field measurement.
	Errors Common to Many of the Load Impact Evaluations	
	Sampling Issues	
	Small sample sizes	Sample size comments on the specific reports have been addressed. Sample sizes were a function of the time and resources that were available for the studies. Professional evaluators have consistently defended their results as sufficiently robust to make inferences to the population. Because the results are based on field analysis they are a more accurate representation of the savings than the ex-ante estimates, studies are not designed to reject a null hypothesis, but to estimate the savings based on field measurement.
	o Results in inappropriately large error bounds, that often include the ex ante estimates. Thus, the null hypothesis (IOU achieves claimed savings) cannot be rejected.	
	o Difficult or impossible to make inferences to the population.	
	o Sample sizes are often less than the protocols require.	
	o Measurement error is much more significant in situations in which the sample size is small since each observation is weighted more heavily.	

Source	Comment	Response
	Nonrandom sample selection	Specific sample selection issues have been addressed in comments on the evaluation contractor reports.
	o Samples based solely on those individuals that were willing to participate.	
	o Biased results not representative of the overall population so inferences to populations are not appropriate.	
	Out-of sample predictions	Energy Division and our contractors have been given clear authority to apply results from these studies across the portfolio to make a more accurate prediction of the savings achieved. In the 2006-2008 studies however this was done for less than 1% of the portfolio savings (see section 3.4), and only after careful consideration of the applicability and reliability of the estimates available.
	o The sample is representative of the population given a specific program design, time frame, etc.	
	o Inappropriate to make inferences to populations outside the specific program (e.g., use information on CFLs to make inferences about LEDs) or the specific time frame (e.g., use information from the 2004 – 2006 program cycle to make inferences about 2006 – 2008 program cycle) without accounting for these differences in the statistical analysis.	
	Metering/Monitoring Issues	Evaluation contractors took measurement error into consideration in designing their studies. This is a core element of evaluation design and every effort was made to minimize measurement error within the bounds of the resources and time that was available for the study.
	Site-specific measurements are subject to significant measurement error.	

Source	Comment	Response
	Dependent on when the measurements are taken (e.g., during the economic downturn or during the most benign period of the year).	As is standard practice in energy efficiency evaluation, Energy Division directed its evaluation contractors to calculate energy savings based on actual observed post-installation conditions for the 2006-2008 energy efficiency evaluation. Consistent with prior evaluation and reporting protocols dating back to the 1990s, evaluation-based savings estimates are based on actual conditions, not conditions re-normalized to represent future unobserved economic conditions. Similarly no efforts were made in previous evaluations conducted during very high-growth economic periods to adjust savings downward. Evaluators were instructed to follow the guidelines related to changes in pre- and post-installation industrial production provided in Appendix J of the CADMAC evaluation protocols from the late 1990s. These guidelines indicate that post-installation production levels are to be used when estimating savings unless the change in pre-and post installation production is itself program induced.
	Dependent on the length of time the metering/monitoring was conducted (e.g., 4 – 6 weeks)	
	Dependent on who conducted the metering/monitoring and ultimately interpreted the results.	
	The effect of measurement error is exacerbated in situations in which the sample sizes are small.	
	Site-specific measurements are subject to the Hawthorne effect.	
	Spot measurements cannot capture fluctuations in variables of interest (e.g., related to day, week, month, seasonality or variation in either the energy efficiency application or the local environment. Thus, spot readings cannot be used reliably for extrapolation.	
	Extrapolation from logger results is problematic because they are operational for insufficient periods of time. This creates two obvious problems.	

Source	Comment	Response
	A short installation/operation period (e.g., two weeks) will be unlikely to capture the fluctuations in use patterns that occur over the year that correspond to variations in weather, daylight hours, macroeconomic conditions, etc. In essence, an analysis based on one to two weeks of logging data is exactly akin to basing the analysis on spot readings, which cannot be used reliably for extrapolation were defined by the study authors to be woefully insufficient.	
	Appropriate sampling design must consider both the number of sites and the time period over which the loggers are installed/operational (e.g., achieving a 10 percent precision with a 90 percent requires in excess of 29 weeks of logger information).	
	Significant measurement error and imprecision in the measured/monitored variables of interest (e.g., estimated hours of use) undermines the estimation of ex post savings.	
	Program evaluators make biased interpretations/decisions (e.g., one-sided trimming in the evaluation of Small Commercial).	
	Compliance Modeling Issues	Specific issues with the use of compliance modeling in the evaluation studies have been addressed in Energy Division's responses to comments on those studies. Energy Division and its contractors employed modeling practices that were consistent with professional practice.
	Modeling is based on metered/monitored data that is measured imprecisely.	
	Models perform poorly compared to actual usage (e.g., see Residential New Construction).	
	Model results are incomplete (e.g., the treatment of interactive effects is undefined).	
	Model results are inappropriately used (due to small sample size, large error bounds, etc.) to aggregate up to produce population savings estimates.	
	Econometric Estimation Issues	Specific issues with the use of econometric estimation methods in the evaluation studies have been addressed in Energy Division's responses to comments on those studies. Energy Division and its contractors

Source	Comment	Response
		employed econometric estimation practices that were consistent with professional practice.
	Econometric analysis is severely outdated	
	Methods used are not scientifically credible. A litany of problems would include the following.	
	Lack of theoretical justification for modeling design (i.e., *ad hoc*).	
	Data sets are inappropriately small.	
	Hypothesis testing is conducted inappropriately (e.g., insignificant variables are eliminated from estimated equations or step-wise regressions are used).	
	Empirical results are not subject to robustness testing or sensitivity analysis. Such testing would include, but not be limited to, determining the relative importance of alternative sets of independent variable sets, alternative functional forms, alternative estimation methods, measurement error, outliers, and influential observations, etc.	
	Econometric results suggest model misspecification (e.g., see Commercial Steam traps) such as omitted variable bias.	
	Econometric modeling is based on incomplete knowledge of the relevant literature or an incomplete understanding of the methods involved (e.g., conjoint analysis or revealed preference modeling).	
	Data sets created for 2006 – 2008 program evaluations are inappropriately combined with data from previous evaluations without justification or an understanding of the relative impact on ultimate results.	
	Econometric results produce inconsistent and counter-intuitive results that are offered without explanation.	

Source	Comment	Response
	Self-report approach (SRA) Issues	The self report approach implemented in the evaluations to estimate net to gross ratios is a widely-used and well-established means of measuring attribution and has in fact been implemented on numerous occasions by the IOUS. Energy Division's "Net to Gross Working Group" was convened early in the evaluation process to ensure consistency in survey methods and design and scoring algorithms. Additionally, Energy Division technical advisors drafted net-to-gross supporting documents that provide detailed explanation of the use of the self report approach in these evaluations and address questions of potential bias.
	The method is plagued by numerous inherent biases. A list would include the following.	
	Self-report bias in which the respondent attempts to please the surveyor or to create the appearance of "socially acceptable behavior." This bias is especially relevant when being questioned regarding socially acceptable activities (e.g., quitting smoking, recycling, adopting energy efficiency, etc.).	
	Starting point bias in which the final respondent answer is closely tied to the suggested starting point (see Codes & Standards evaluation).	
	Non-random selection bias, in which respondents self-select into the survey or are only those individuals willing to participate.	
	Decision-maker bias, in which the survey is conducted with a single individual but decisions are not made in this manner (rather the process includes many disparate influences) or the appropriate decision maker cannot be identified (e.g., no longer works for the entity).	
	Program-influence bias, in which the NTGR for non-residential applications seems to be limited at upper end due to the "program influence" question. Specifically, it seems that this question inappropriately anchors the respondent to the 50 – 50 attribution between the "program" and "other factors."	
	Program evaluator bias, in which the evaluators make post-survey interpretations/adjustments regarding respondents' answers that are either *ad hoc* or inappropriate (e.g., converting "Don't Know" into a numerical value).	

Source	Comment	Response
	Survey implementation bias, in which surveys are conducted by overly-experienced survey personnel that achieve a pre-determined result.	
	Reality bias, in which SRA results are inconsistent with actual data (e.g., sales information, prior lack of action by participants, stocking practices, etc.).	
	There is a general lack of supporting evidence to calibrate the SRA findings.	
	Compliance with M&V Protocols	The M&V protocols were adopted as guidance, but did not preclude adaptation by CPUC staff where necessary. "While these Protocols are the key guiding documents for the program evaluation efforts, the CPUC and the Joint Staff reserve the right to utilize additional methodologies or approach if they better meet the CPUC's evaluation objectives and when it serves to provide reliable evaluation results using the most cost-efficient approaches available." p1 California Evaluation Protocols.
	Protocols are often not followed or not followed to completion (e.g., billing analysis in Small Commercial or the econometric analysis in Residential Retrofit).	
	Failure to think through the evaluation process prior to beginning the evaluation effort or “unforeseen” problems undermine the analysis.	
	Less rigorous analysis completed.	
	Evaluation effort produces less than expected at a cost per result much higher than expected.	
	Lack of supporting evidence reduces the validity of the work presented.	
	Evaluators create *ad hoc* measures of program success that are outside the M&V protocols (e.g., “vulnerability” and “leakage” in the Upstream Lighting evaluation).	
	Reporting Issues	Energy Division did its best to completely present the methods, results and conclusions for each utility (including any variability in the results) in the body of the reports and supplement the details with technical appendices. D408

Source	Comment	Response
	There is a general lack of information provided that is necessary to assess the validity of the study (e.g., econometric estimation missing information includes, but is not limited to, R-square vales, number of observations, standard errors, etc.	
	There is unexplained variation across IOUs on such variables as hours of use, NTG ratios, etc.	
	Specific IOUs seem to perform consistently worse than others without explanation.	
	NTG ratios that vary widely (see HVAC HIM and Specialized Commercial Evaluation) without explanation.	
	There is insufficient evaluation of the likely effect of sampling error, measurement error, modeling error, etc.	
	Inappropriate conclusions are drawn (e.g., inferences from small samples, inferences from data measured with error, inferences that use out-of-sample predictions, etc.).	
	In appropriate attribution (e.g., in every case in which ex post savings differs from ex ante savings it is assumed that the ex ante figures are incorrect yet the evaluation, whether it be data collection, modeling, estimation, etc., is fraught with errors and is likely the reason for the difference).	
	Evaluation results are inappropriately used in policy making.	
	Recommendations for improving performance, especially as it pertains to inter-IOU performance, are often missing.	

Source	Comment	Response
PGE	Substantively, the EM&V reports that support the Draft Evaluation Report contain various methodological errors that call into question the validity of its conclusions. These errors are documented in detail in PG&E's comments to the draft EM&V evaluations, which are attached as Attachment A to these comments. The majority of the draft EM&V evaluations reduced program savings based largely on Net-to-Gross analyses that often employed arbitrary adjustments and produced unverifiable results. In addition, evaluators often applied incorrect baselines, applied inappropriate assumptions with respect to in-service rates and commercial/residential usage for Compact Fluorescent Light bulbs (CFLs), failed to collect data in accordance with the evaluation protocols, and made numerous other errors in methodology with respect to individual program evaluations as noted in PG&E's comments in Attachment A.	Energy Division has provided responses to all of these comments in the course of reviewing the individual contractor reports.
PGE	Procedurally, the most notable weakness of the Draft Evaluation Report is that its conclusions have not been properly vetted, as they have not been subject to an adequate public review process. Specifically, given the breadth of the evaluation, as noted by Energy Division, a prohibitively short period was provided for review and comment on the individual program EM&V evaluations whose conclusions are adopted in the Draft Evaluation Report.	The Draft and Final Report is the aggregation of results that were reported in the final versions of the individual evaluation reports. Energy Division and their contractors conducted webinars on the draft reports, provided responses to all comments provided on those reports and made several adjustments to the final reports. In addition Energy Division hosted a 2 hour workshop on the April 15 Draft Report to answer questions about the conclusions and findings, and there were no substantive questions or comments made at that workshop (it was therefore concluded after one hour); a month was given for comments on this report and the majority of comments represent a re-submittal of comments on the evaluation contractor reports which have already been addressed.
PGE	In its comments to the draft EM&V evaluations, PG&E commented on serious methodological weaknesses in the reports that cast doubts regarding the veracity of the conclusions presented.3 Notwithstanding PG&E's comments in this regard, the vast majority of the findings, based on these flawed or unsupported methodologies, now form the basis for the conclusions presented in the Draft Evaluation Report. For these reasons, the Draft Evaluation Report should not be used for program planning purposes or for calculating incentive earnings.	Energy Division and their contractors provided responses to all comments provided on those reports and made several adjustments to the final reports. These results represent the best available data regarding the portfolio performance from 2006-2008.

Source	Comment	Response
PGE	For virtually each of the contract groups evaluated, the Draft Evaluation Report findings note significant reductions to savings attributable to free ridership. The Commission acknowledges that "studies that evaluate and measure net-to-gross ratios are inherently difficult. They typically involve surveying customers more than a year after adopting their energy efficiency measure to determine what motivated the customer to pursue that particular measure. PG&E agrees with the Commission, and respectfully asserts that the NTG studies used in the final performance evaluations issued by ED in the Draft Report are unreliable. In reviewing the final performance evaluations, it appears that many of the NTG ratios were based on inadequate sample size, insufficient response levels, and/or a 1 1/2 to 3 year delay in surveying customers regarding their motivation for participating in energy efficiency programs.	The self-report approach implemented in the evaluations to estimate Net-to-Gross Ratios is a widely-used and well-established means of measuring attribution and has in fact been implemented on numerous occasions by the IOUs. Energy Division's "Net-to-Gross Working Group" was convened early in the evaluation process to ensure consistency in survey methods and design and scoring algorithms. Additionally, Energy Division technical advisors drafted Net-to-Gross supporting documents that provide detailed explanation of the use of the self-report approach in these evaluations and address questions of potential bias.
PGE	In the state of California, Energy Efficiency Evaluation Protocols, in the section on Sampling and Uncertainty Protocols, Table 20, page 166, the basic rigor for net impacts evaluation requires a minimum of 300 sites or a census, whichever is smaller. By way of example, a census for the PGE 2005 group would be 200 sites (or the total number of firms that includes the 200 sites). It appears that the CPUC EM&V consultant based the NTG ratio on 18 firms (28 sites), which is clearly inadequate and does not meet the requirements of the Evaluation Protocols.	The M&V protocols were adopted as guidance, but did not preclude adaptation by CPUC staff where necessary. "While these Protocols are the key guiding documents for the program evaluation efforts, the CPUC and the Joint Staff reserve the right to utilize additional methodologies or approach if they better meet the CPUC's evaluation objectives and when it serves to provide reliable evaluation results using the most cost-efficient approaches available." p1 California Evaluation Protocols.
PGE	In addition to the inherent flaws in the self-reporting methodology, the general concept of applying Net-to-Gross ratios to estimate free ridership is based on a faulty premise—that evaluators can tease out one, single reason why an individual or business chooses to install a particular Energy Efficiency measure, especially in given current societal trends toward "green" practices.	The self-report approach implemented in the evaluations to estimate the net to gross ratios is a widely-used and well-established means of measuring attribution and has been implemented on numerous occasions by the IOUs. Energy Division's Net to Gross Working Group was convened early in the evaluation process to ensure consistency in survey methods and design and scoring algorithms. Additionally, Energy Division technical advisors drafted net-to-gross supporting documents that provide detailed explanation of the use of the self-report approach in these evaluations and address questions of potential bias.

Source	Comment	Response
PGE	Finally, the net-to-gross analyses in these evaluations fail to address spillover. This limitation is unwarranted, as many programs have significant market effects over and above the direct impacts on participants. Worse, program baselines often consider these spillover effects. Failing to account for them in the NTG analysis essentially penalizes the utilities twice when they are not applied to program savings.	Current Commission policy is to not count spillover in 2006-2008. See Finding of Fact 27 of D. 05-04-051. The speculative nature of any attempts to quantify spillover effects significantly reduces their applicability as an analytical tool at this time. Moreover, discounting the accounting of free-ridership through "spillover," as PG&E proposes, would make it particularly difficult to attribute indirect program benefits to education and information programs, without double-counting those benefits. "
PGE	Since energy savings are baseline energy usage less the energy usage with the efficiency measure, assumptions affecting baseline energy use can have a dramatic effect on savings. Some evaluations made multiple incorrect assumptions about project baselines resulting in significant reductions to gross savings. In some instances, project savings were reduced to zero.	The Draft and Final Report are based on the energy savings that were found in the energy efficiency evaluation studies, which relied on current evaluation practice, professional review, and in field data to determine the appropriate baseline.
PGE	The Draft Report incorporates CFL residential In-Service Rates (ISRs) that are modeled in the Upstream Lighting Program (ULP) Evaluation Report that are unreasonably low8. Evaluators substituted an untested modeling approach in place of the approach called for in the evaluation plan, because the original approach did not produce meaningful results. Given the uncertainties in the calculation method used (outlined below), the evaluators should have relied upon the ex ante ISRs, which were based upon publicly-vetted estimates published in DEER 2005.	The Draft and Final Report incorporate the values that were finalized in the Upstream Lighting Program Evaluation Report, and these comments have been addressed in finalizing that report. The in-service rates modeled in this analysis reflected the best available data regarding lighting that was installed and operating in the 2006-2008 period.
PGE	The Draft Report updated assumptions regarding the use of upstream CFLs from 90% residential / 10% nonresidential to 94% residential / 6% nonresidential. This adjustment dramatically reduced total kW and kWh savings from upstream CFLs given higher peak and total use intensity for nonresidential CFLs. Moreover, this adjustment was made absent compelling analytical support. The new split is based on CFL on-site surveys and estimated install rates (see Section 3.1.5 in ULP final report). Given the uncertainties discussed in the final ULP report, PG&E believes more research is needed to ascertain a more reliable estimated percentage split between Residential/Nonresidential percentages for upstream lighting and small commercial programs.	The Draft and Final Report incorporate the values that were finalized in the Upstream Lighting Program Evaluation Report. These comments have been addressed in finalizing that report, and the residential/non-residential split reflects the best available data.

Source	Comment	Response
PGE	The ERT methodology and application itself contains numerous flaws. These are discussed at length in the Joint Utilities' Comments on the on the Assigned Commissioner's Ruling Providing Energy Division Report and Soliciting Comments on Scenario Runs, filed in R.09-01-019 on May, 18, 2009. In summary, the ERT is flawed in the following ways:	Responses to specific issues are addressed in the following paragraphs.
	The ERT Does Not Include The Benefits Associated With 2006-2008 Codes and Standards Activity;	D.05-09-043, OP 14(f), states that "savings from pre-2006 codes and standards advocacy work shall not be counted when calculating net resource benefits ("performance basis") or cost-effectiveness associated with portfolios plans for 2006 and beyond, either on a prospective or ex-post basis. OP 14(e) of the said decision also states that, "On a forward looking basis, savings from codes and standards advocacy work undertaken in 2006 and beyond shall be counted when calculating either net resource benefits ("performance basis") or cost-effectiveness (TRC or PAC tests)." The original analysis in the codes and standards program evaluation report treated savings of a couple of standards that went into effect after 2006 but before 2009 the same as the standards effective as of Jan 1, 2006. The logic was that these were basically Tier 2 of standards developed prior to 2006 or modifications to the pre-2006 standards that resulted from mostly from IOU efforts prior to 2006. One standard went into effect in April 2006, for example, and it seemed most of the effort must have occurred prior to 2006 to get it adopted. We will refer the issue of verifying and counting savings that result from post-2006 Codes and Standards program efforts to the EMV effort of the next program cycle. Savings related to post-2005 C&S support efforts are beyond the assigned scope of the codes and standards evaluation conducted for the 2006-2008 program cycle. As noted in D.07-09-043. Section 9.3.2, the scope of the evaluation efforts was to estimate the "bonus savings" for the 2006-2008 program cycle: "All parties commenting on this issue recommend that 50% of the savings attributed to pre-2006 C&S advocacy work count towards establishing whether the MPS has been met for the 2006-2008 cycle. They also recommend excluding these savings from the calculation of PEB. We find these recommendations to be fully consistent with our determinations in D.05-09-043, as discussed above, and will adopt them. As stated in that decision, for this purpose the C&S savings are to be verified (as opposed to ex ante estimates used

Source	Comment	Response
		for planning purposes). Energy Division's EM&V contractors are in the process of verifying those savings estimates, and Energy Division will be including the verified numbers in its Annual Verification Reports."
	The ERT Contains "E3 Calculator FALSE" Errors Which Assign Zero Avoided Cost Benefits To Valid Measure Installations.	This error has been corrected in the final version of the ERT.
	The ERT Alters The Known Location Of Installed Measures By Unknown And Unwarranted Parameters.	The zip code to climate zone mapping was based on the CEC look up table. There are additional (new) zip codes that were not in the CEC look up table, and therefore, they were assigned based on the CEC Google map found here: http://www.energy.ca.gov/maps/building_climate_zones.html The majority of changes to the utility assigned climate zones were due to incorrect zip codes in the tracking data. There were many cases of null, out-of-state, and impossible city/zip combinations. These zip codes were cleaned as documented in the ERT process. The clean zip codes were then mapped to the expanded CEC climate zone lookup table. This table sometimes contains multiple climate zones per zip code. Changes were only made if the utility assigned climate zone did not match any of the choices for that zip code in the CEC expanded mapping table.
	The ERT contains values from non-transparent sources and contains errors that are summarized in Appendix X:	
	(1)EDUESi - Appendix C does not provide the details how UES values with interactive effects (EDUESkWi, EDUESkWhi, and EDUESThermsi) are derived based on revised UES figures used in the ERT.	The application of interactive effects for UES values are included in Appendix B that details the guidance provided to contractors in applying interactive effects. If interactive effects were part of the evaluation study they are embedded in the UES and realization rates.

Source	Comment	Response
	(2) OthEMV - Throughout the ERT, the source of some values were referenced to "OthEMV". Appendix C does not provide the details on what the other EM&V sources are for these values.	Energy Division has addressed specific comments regarding the application of "OthEMV", and reminds readers that the instances of these applications constitute less than 1% of the portfolio energy savings.
	(3) PGE2000 - RCA: There is no reference in Appendix C under Residential Retrofit where the updated values in the ERT came from. PG&E was able to find the EDIRate and EDNTGR from Table 5-44 of the HVAC High Impact Measure study, but the source of the EDUES figures could not be determined. The overall Gross Realization Rates were not used to determine the UES.	Table 5-44 in the original report is total savings for the Res RCA programs. PGE 2000 UES and extrapolation can be found in Table 5-39 of the Specialized Report. All these UES tables for Res RCA are in the documentation (5-39 to 5-43).
	(4) PGE2004, 2042 - Fab: There is no detail in Appendix C that describes how the sampling Strata results in the EM&V report maps to the measure level detail in the ERT.	As mentioned in the "2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group", all industrial programs in PG&E territory were evaluated together. Five electric strata and five gas strata were defined. Realization rates and NTGRs were estimated by strata and for three groups of measures: pump-off controllers (POC), non-POC electric measures and gas measures. RRs and NTGRs were then applied by measure group (POC, non-POC, gas) and stratum to each of the records in the tracking database. Tables 4-17 through 4-19 show the estimated impacts for the industrial programs in PG&E's territory. As shown in the ERT documentation, the "all parameters" ERT run matches perfectly the Final Report results for the PG&E Fab-Industrial contract group. Any measure-level reporting within the ERT after the evaluation parameters (RR, NTGR) were applied would be solely a result of measure grouping.
	(5) PGE2068 - Gross RR for RCA from Table 5-63 (Page 89) was .45. What is the source of the figures in the ERT that shows much smaller realization rates?	It is not clear that there is a lower realization rate reported. The gross realization rate for the RCA measures in the PGE2068 program are shown in Table 5-63 pg.89 of the Specialized Commercial Report and referenced in Appendix C. Other measures in PGE2068 had pass through savings and the total net realization rate for 2068 is much higher at 0.79 as shown in Appendix A page 6. The total gross realization rate for all RCA measures in all programs residential and non-residential was 0.47 kWh and 0.42 kW.

Source	Comment	Response
	(5) PGE2001 - Ag & Food: The realization rates from Table 4 of the final report do not match the realization rates on Table 10 of Appendix C. The ERT also had negative 45.78 for EDUESTherms when there were no ex-ante gas savings. The ERT incorrectly applied the retrofit values for new construction projects.	Table 10 of Appendix C in the draft ED report is consistent with Table 51 (page 74) of the PG&E Ag & Food report. The realization rates in these tables reflect the base realization rates that do not take into account dual fuel impacts that were omitted from ex ante calculations. The realization rates reported in Table 4 of the PG&E Ag & Food report include the omitted dual fuel impacts. This difference is explained at the bottom of page xi of the PG&E Ag & Food report. The -45.78 value for EDUESTherms properly takes into account gas impacts for electric projects where gas savings were omitted from the ex ante calculations. The PG&E Ag & Food Program included both retrofit and new construction projects. Commercial new construction projects were assessed separately, as part of the New Construction evaluation. The other new construction projects (those with sector identified as AGR or IND in the PG&E tracking data) were evaluated as part of the Ag & Food evaluation. Hence the Ag & Food evaluation addressed a combination of retrofit and new construction projects, and the various values developed in Ag & Food evaluation reflect a combination of retrofit and new construction projects. These combined values were correctly applied through the ERT.
	(6) PGE2004 - Fab: There are errors in how certain measures were updated by retrofit values vs. new construction values (i.e., some measures that were retrofit were updated with new construction values and some measure that were new construction were updated with retrofit values).	This comment refers to the "2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group" and has been addressed during the comment/answer period for that Report. PG&E did not separate new construction projects from retrofit projects in their reporting of industrial program activity, and neither did the PG&E Fab contract group. As specified in the comment/answer period for the 2006-2008 reports, the PG&E Fab contract group evaluated all PG&E Industrial programs together and estimated "industrial" measure results (not results for "new construction" and "retrofit" measures.) Thus, the PG&E Fab contract group could only apply "industrial" results to each record (measure), not "new construction" or "retrofit" results. Also as specified in the 2006-2008 Report, the New Construction and Codes and Standards contract group, defined (with PG&E's help) and evaluated "commercial" new construction records across programs and contract groups. 58 of these "commercial new construction" records

Source	Comment	Response
		came from program PGE2004. As shown in the ERT documentation, NCCS results were applied to these 58 records, not industrial results.
	(7) PGE2070 - Retro-Commissioning results were incorrectly applied to all measures. Some measures were not retro-commissioning.	Assigning the RCx study results to all measures in PGE2070 was an error on SBW's part. The results should only have been applied to 5 measures identified as RCx and included in our study population. The remaining 6 measures should either have been assigned PassThru, or assigned a realization rate from other study results. [This has been corrected in the final report.]
PGE	The ERT also contains errors where the values in the EM&V study results were not used in the ERT or the values were applied to the wrong measures:	
	(1) EDEUL - For those measures where the DEER EULs are based on the hours of operation, like Residential CFLs, the ERT does not adjust the EULs when the EDUESs are based. on lower hours of operation resulting from the EM&V results. Less hours of operation results in longer EULs.	1) There were also open-ended ranges such as "CFL >30W" that we had to account for. 2) We created delta watts values -- and HOU profiles -- for the entire population of CFLs encountered during onsite metering. The delta watts analysis did not lend itself to disaggregating by wattage, largely due to the sample size (would need fairly substantial sample size for each wattage in order to determine at such a level) and our estimation procedures. The analysis was substantially more complex than this, but essentially we had to take average CFL wattage encountered (by a factor such as room type) and compare with average non-CFL wattage encountered (at the same room type). It was not possible to estimate delta watts at the wattage level without some sort of extensive pre- and post-installation inspection of CFL installations/replacements, which is essentially impossible given the nature of the ULP (no customer information, no installation date, no way of determining who is about to replace an incandescent with a CFL, etc). The HOU profiles

Source	Comment	Response
		were created for all wattage ranges as well, as we were able to meter approximately 3,300 CFLs (not a large enough sample to sufficiently represent the many wattages / wattage ranges listed in the tracking databases. Since these are the two variables in UES estimation, we were unable to produce UES estimates at the wattage level. 3) Additionally, we didn't want to introduce arbitrary pre/post wattage pairings; rather, we let room and lamp type drive the averages.
	(2) PGE2000 - Exterior Upstream Fixtures: In Appendix C, the referenced source for EDUES figures is Table 30, pages 63 - 64. The UES figures on the referenced table do not match the figures used in the ERT. For example, Table 30 has 59.14 as the Ex-Post UES kWh/yr and the ERT has 38.54.	Fixture UES values were weighted by the shipping distribution of such measures to produce one single weighted UES for all fixtures for the ERT
	(3) PGE2000 - Torchiere Fixtures: In Appendix C, the referenced source for EDUES figures is Table 30, pages 63 - 64. The UES figures on the referenced table do not match the figures used in the ERT. For example, Table 30 has 59.7 as the Ex-Post UES kWh/yr and the ERT has 38.54.	Fixture UES values were weighted by the shipping distribution of such measures to produce one single weighted UES for all fixtures for the ERT
	(4) PGE2000 - LED Night Lights: In Appendix C, the referenced source for EDUES figures is Table 34, pages 68 - 69. The UES figures on the referenced table do not match the figures used in the ERT. For example, Table 34 has 23.84 as the Ex-Post UES kWh/yr and the ERT has 3.68.	Fixture UES values were weighted by the shipping distribution of such measures to produce one single weighted UES for all fixtures for the ERT

Source	Comment	Response
	(5) PGE2001 - Ag & Food: The realization rates from Table 4 of the final report do not match the realization rates on Table 10 of Appendix C. The ERT also had negative 45.78 for EDUESTherms when there were no ex-ante gas savings. The ERT incorrectly applied the retrofit values for new construction projects.	Table 10 of Appendix C in the draft ED report is consistent with Table 51 (page 74) of the PG&E Ag & Food report. The realization rates in these tables reflect the base realization rates that do not take into account dual fuel impacts that were omitted from ex ante calculations. The realization rates reported in Table 4 of the PG&E Ag & Food report include the omitted dual fuel impacts. This difference is explained at the bottom of page xi of the PG&E Ag & Food report. The -45.78 value for EDUESTherms properly takes into account gas impacts for electric projects where gas savings were omitted from the ex ante calculations. The PG&E Ag & Food Program included both retrofit and new construction projects. Commercial new construction projects were assessed separately, as part of the New Construction evaluation. The other new construction projects (those with sector identified as AGR or IND in the PG&E tracking data) were evaluated as part of the Ag & Food evaluation. Hence the Ag & Food evaluation addressed a combination of retrofit and new construction projects, and the various values developed in Ag & Food evaluation reflect a combination of retrofit and new construction projects. These combined values were correctly applied through the ERT.
	(6) PGE2004 - Fab: There are errors in how certain measures were updated by retrofit values vs. new construction values (i.e., some measures that were retrofit were updated with new construction values and some measure that were new construction were updated with retrofit values).	This comment refers to the "2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group" and has been addressed during the comment/answer period for that Report. PG&E did not separate new construction projects from retrofit projects in their reporting of industrial program activity, and neither did the PG&E Fab contract group. As specified in the comment/answer period for the 2006-2008 reports, the PG&E Fab contract group evaluated all PG&E Industrial programs together and estimated "industrial" measure results (not results for "new construction" and "retrofit" measures.) Thus, the PG&E Fab contract group could only apply "industrial" results to each record (measure), not "new construction" or "retrofit" results. Also as specified in the 2006-2008 Report, the New Construction and Codes and Standards contract group, defined (with PG&E's help) and evaluated "commercial" new construction records across programs and contract groups. 58 of these "commercial new construction" records

Source	Comment	Response
		came from program PGE2004. As shown in the ERT documentation, NCCS results were applied to these 58 records, not industrial results.
	(7) PGE2070 - Retro-Commissioning results were incorrectly applied to all measures. Some measures were not retro-commissioning.	Assigning the RCx study results to all measures in PGE2070 was an error on SBW's part. The results should only have been applied to 5 measures identified as RCx and included in our study population. The remaining 6 measures should either have been assigned PassThru, or assigned a realization rate from other study results. [This has been corrected in the final report.]
PGE	THE DRAFT EVALUATION REPORT HAS NOT BEEN SUBJECTED TO ANY MEANINGFUL PUBLIC REVIEW PROCESS AND ITS CONCLUSIONS REMAIN UNVETTED	Energy Division hosted a 2 hour workshop on the ERT contents and another 2 hour workshop for discussing the report. Energy Division provided one month for comments and has been available to answer questions. In December 2009 Energy Division hosted two workshops on the ERT and one workshop on the Decision Framework which was the methodology for applying evaluated results to the reported claims and the policy justification for that approach.

Source	Comment	Response
PGE	Such an opportunity for a free exchange among the stakeholders was not present here. The review of the conclusions in the draft EM&V reports was directly hampered by (1) an unmanageably short review and comment period spanning the year-end holidays, which failed to provide the opportunity for any sort of comprehensive review of the reports' conclusions given the breadth of the evaluation; (2) the lack of data made available in a timely manner during the designated review period with which the parties could assess the veracity of the conclusions presented in the reports; and (3), the lack of meaningful responses to critical comments addressing faulty methodologies and other errors, many of which persist in the Draft Evaluation Report without sufficient explanation and that continue to have a significant effect on the savings estimates.	Energy Division was obligated to produce a Performance Basis Report in the spring of 2010 to inform the shareholder incentive mechanism which was to be completed in 2010. This dictated a very short period for review on both sides given that the implementation of the programs was not complete (and represented the majority of savings) until the end of 2008. Evaluation field work was still going on until the fall of 2009 to capture seasonal peaks and fully account for the savings.
PGE	The Designated Period for Review and Comment was Prohibitively Brief and Did Not Allow any Meaningful Opportunity for Comprehensive Review of the Draft EM&V Reports.	Deadlines for the review of the initial reports were set to comply with timelines established by Commission decision, specifically the completion of a draft report in the spring of 2010.
PGE	In addition to the prohibitively short time provided to review and comment on 11 Draft EM&V Reports, ED further curtailed the IOUs' resources available to assist with the review during this time. On December 17, 2009, Energy Division sent a letter to all consultants who assisted the CPUC with preparation of any 2006-08 EE program impact evaluations in any capacity whatsoever, stating that pursuant to the consultants' contracts with the CPUC, ED considered it to be a conflict of interest for any of those consultants or their subcontractors to consult for a utility regarding the review of any impact evaluation.	Contractors and their subcontractors on the 2006-2008 evaluation were bound by the conflict of interest clauses in the contracts. The letter was sent to carry out our responsibility to enforce the terms of the contract.

Source	Comment	Response
PGE	Because of the inherent challenges in conducting a full-scale review of a three-year EE portfolio of programs in only five weeks, let alone over the year-end holidays, PG&E served a data request on ED on November 12, 2009, prior to the release of the Draft EM&V Reports, in which PG&E clarified the categories of data it would need along with the reports themselves to be able to vet their conclusions in such a limited time frame. On December 2, 2009, Energy Division responded to PG&E's data request by voicemail and declined to provide the requested information. Energy Division informed PG&E that ED's own work finalizing the draft reports takes precedence over PG&E's request, that ED really didn't have the time to provide PG&E with the information that it was requesting ahead of the reports being posted, and recommended that PG&E simply wait until after it reviewed each individual, posted report to submit any data requests.	PG&E's blanket request was made prior to the completion of the draft reports. Energy Division had not completed its review of the drafts and believed it improper to release data and incomplete reports prior to the public review process.
PGE	ED often provided incomplete information or otherwise provided technical data in an unworkable format, or too late to conduct a meaningful review. In situations where PG&E did receive the requested information, it was provided mere days before comments were due to be posted, thereby eliminating the possibility of any sort of comprehensive and robust review. In the case of the Retro-Commissioning Impact Evaluation, PG&E requested site detailed data which was received on 12/22/09. However, PG&E was unable to locate and sort out all the files, leading it to believe the data set was incomplete at the time comments were due. This issue persisted beyond the posting of the final EM&V reports as well. For example, on March 26, 2010, PG&E served a data request to ED pertaining to its HVAC program. Among other things, the data request referenced a change in NTG from 9% in the Draft Report to 63% in the Final Report and requested an explanation for the 600% increase that appeared in that program alone. PG&E has yet to receive a response to that data request. The lack of timely access to critical data foreclosed the opportunity for the sort of "give and take" contemplated by the Commission in D.07-09-043	Energy Division did its best to provide meaningful and timely responses to data requests made by PG&E and other stakeholders. Time constraints imposed on the process are attributable to the Commission imposed timelines for informing the Risk Reward Incentive Mechanism which required the completion of the report in the spring of 2010, to allow for timely payment of incentives.

Source	Comment	Response
PGE	Energy Division Failed to Address Certain Critical Comments PG&E Submitted in a Reasonable Manner. Although the short review period, disqualification of consultants, and lack of timely access to critical data precluded PG&E from conducting comprehensive reviews of the conclusions presented in the Draft EM&V Reports (a point PG&E noted in virtually every set of comments to the evaluations) PG&E put forth its best efforts to identify "big picture" issues in the evaluations and address them both through individual comments and narrative-style summaries.19 Energy Division did respond in some form to many of the comments PG&E submitted. However, ED did not respond in a reasonable manner to others, in which PG&E raised critical issues that had significant effects on overall portfolio savings. While PG&E did receive a response to the comments, the response does not address the concern in a reasonable manner because it does not provide the explanation sought for these seemingly-arbitrary reductions in NTGR for this particular program. ED's recommendations on these issues were subsequently incorporated into the Draft Evaluation Report. This demonstrates the true fallibility of the "public review process" ED refers to throughout the Draft Evaluation Report.	Energy Division did its best to provide meaningful and timely responses to comments filed by PG&E and other stakeholders. Time constraints imposed on the process are attributable to the Commission imposed timelines for informing the Risk Reward Incentive Mechanism which required the completion of the report in the spring of 2010, to allow for timely payment of incentives.
PGE	The Draft Evaluation Report fails to provide a credible or reliable analysis of the savings associated with the utilities' 2006-08 EE programs. As such, while the Draft Evaluation Report may be a potential resource in developing protocols for future impact evaluations, it fails to advance the Commission's EM&V goals and should not be used for program planning purposes or for calculating incentive earnings.	Energy Division believes that the Draft and Final Evaluation Reports do provide a credible and reliable analysis of the savings associated with the utilities' 2006-2008 programs based on current Commission guidelines. The studies that were conducted include important recommendations for improving the programs as well as improving the estimates of savings in future cycles.

Summary of comments resulting in updates to ERT runs and action taken

#	Comment	Response	Solution/Action
1	In the Integrated School-Based Program (SCE2504), the evaluated UES included an installation rate, and yet the ERT also applies an installation rate less than 1.0, double penalizing the energy savings because attention was not paid to the meaning of the evaluated inputs.	Commenter is correct, on GS/GC, we did apply a realization rate on top of the installation rate SCE had already incorporated into their estimates. The corrected savings are: 306 MWh and 27 kW for Green Campus and 876 MWH and 77 kW for Green Schools	Corrected Irate in Input File and re-ran Specialized ERT
2	..and were evaluated to have saved 81% of what was claimed were each reassigned a savings of less than one kWh. A similar problem occurred with the Demand Control Ventilation measures in the EE for Entertainment Centers Program (SCE2561), which had their savings changed from the per project value to the per ton value, but their unit count remained at one, rather than recalculating the number of tons per project.	Commenter is correct for SCE2561. The unit energy and demand savings are presented per ton, but the quantity is presented per unit. The ERT units should be amended to reflect the tonnage of the unit rather than unit quantity (which is always =1). The tonnage for each record is available in the Standard Program tracking database.	Corrected EDQuantity and ERT Units
3	The Draft Report improperly includes the program costs from SCE's Emerging Technologies program (SCE2515) in the net benefit calculation. Decision 07-09-043 specifies how the net benefits should be calculated, "with the exception of the Emerging Technologies Program and LIEE, all energy efficiency portfolio costs including associated evaluation, measurement and verification (EM&V) shall be included in the calculation of PEB."	Commenter is correct for all utilities.	Removed SCE2515, SCG3506, and SDGE3011 from IOU_E3_Cost_Q42008 table and re-ran Option 1 in ERT. PGE's ET program was never included in the first place, so it did not have to be removed.
4	The Draft Report Contains 538 "E3 Calculator FALSE" Errors Which Assign Zero Avoided Cost Benefits To Valid Measure Installations. The ERT incorrectly processes SCE's input files through the E3 Calculator in order to calculate energy savings, demand reduction, cost-effectiveness, and PEB.	The comment incorrectly states that the "ERT incorrectly processes SCE's input files through the E3 Calculator". The function of the ERT is to automate the import of data from the ERT input files into the E3 calculator, running the runs, and saving the results of each run. If there is a FALSE flag generated in the E3, then it is because	Changed the input file to conform to acceptable combinations so no falses occur. Re-ran ERTs for both SCE and PGE program IDs listed.

#	Comment	Response	Solution/Action
		there is a combination of climate zone-Target Sector-Measure Electric End Use Shape not recognized by the E3. This would indicate an issue with the evaluation data and/or the E3, but not a systemic issue with the ERT. Out of the 105 programs evaluated, two had FALSE flags detected by the E3, SCE2501 and PGE2007. For SCE2501, out of the 207,463 lines evaluated there were 538 (0.23%) where the E3 generated a FALSE flag. All of these FALSE flags were due to one Target Sector-End Use Shape combination, Residential:Outdoor Lt. For PGE2007, out of 593 records evaluated, 5 (0.84%) had a FALSE flag generated by the E3 calculator. All of these FALSE flags were due to one Target Sector-End Use Shape combination, Commercial:Unknown. In summary, all of the FALSE flags generated by the E3 Calculator were due to two target sector-end use shape combinations.	
5	The Integrated School-Based Program works with educational institutions to promote energy efficiency to students within schools. Originally, the evaluators in charge of this program under the Specialized Commercial contract group planned to do a full evaluation of direct and indirect program impacts.	Commenter is correct, on GS/GC, we did apply a realization rate on top of the installation rate SCE had already incorporated into their estimates. The corrected savings are: 306 MWh and 27 kW for Green Campus and 876 MWH and 77 kW for Green Schools	Corrected Irate and re-ran Specialized ERT

#	Comment	Response	Solution/Action
	Unfortunately, "because of the reallocation of resources to HIM programs, the evaluation was limited to only evaluating direct-savings measures" (Specialized Commercial Appendices, p.144). In truth, most of this program received no true ex post evaluation. For the Green Campuses and Green Schools programs, the evaluation consisted of		
6	…using values from the Local Government Program verification report. The LivingWise Program received data from participant surveys to update savings values. In all cases, a NTG value of 80% was "assumed."	Correct, there was no direct EM&V of NTG. These should be coded in the ERT as pass through values.	Updated the input files.
7	And yet, even though most values received no direct EM&V, the values in the ERT are all coded "EMV" rather than "othEMV". Furthermore, the installation rate for all measures were incorporated into the gross UES estimates in the study, but the ERT then includes them again, which will penalize the energy savings twice for non-installation. Finally, the report does not present kW reductions, yet these are included in the ERT as EM&V values.	KEMA used the kW savings supplied by the utility, and did not recalculate them in the evaluation. They are included in the Specialized Commercial report as EM&V values. These should be coded in the ERT as pass through values.	Update the input files.
8	As with previous examples, DEER 2008 v3.02 interactive effects were misapplied to the residential lighting measures.	PG&E. HVAC interactive effects were incorrectly applied to some exterior lighting measures in programs PGE2036 and 2080, resulting in an over-estimation of the kW and kWh savings by 38 kW and 62,242 kWh, and an under-estimation of the therm savings by 7,164 therms, representing 0.01%, 0.02% and -0.3% of the total PG&E ex-post net kW, kWh and therm savings respectively. Holiday lights were incorrectly assigned to the Exterior Lighting Other measure group. HVAC interactive effects were appropriately applied to that measure.	For PGE, created new input sheet and re-ran, zeroing out any interactive effects for exterior lighting measures. Re-assigned holiday lights to interior lighting. For SCE, created new input sheets and re-ran, zeroing out any interactive effects for exterior lighting measures.

#	Comment	Response	Solution/Action
		SCE Therm impacts were incorrectly applied to exterior lighting measures in programs SCE2517, resulting in an under-estimation of the therm impacts by 554 therms, representing 0.004% of the total SCE ex-post net therm impacts. SDG&E Outdoor CFL fixture measure group assigned to project in SDGE3025. Not clear from tracking data if these are outdoor or indoor fixtures. Assuming the fixtures are indoor; HVAC interactive effects are applied correctly.. Outdoor CFL fixture and outdoor screw lighting measure groups incorrectly assigned to some interior CFLs in Program SDGE3035. HVAC interactive effects were appropriately applied.	For SDGE, re-assigned measure group from outdoor to interior lighting.
9	Non-resource programs were not included for PGE	This is correct. PGE was missing costs for C&S, Education and Training, Marketing and Outreach, canceled programs and EM&V expenditures. SCE was missing costs for C&S and EM&V expenditures. SDGE and SCG were missing costs for C&S. For all utilities, EMV expenditures should remove EMV costs from Emerging Technology programs.	Adjusted IOU_E3_Cost_Q408 table in ERT and re-ran Option 1 so ERT uses proper dollar amounts.
10	Savings from CO Sensors in the MAP Program (SCE2537) were completely mis-assigned. The realization rate from the study is 81%, yet the applied realization rate in the ERT is .000003. While it appeared that the savings unit associated with the UES might have been changed, in fact the unit count remained at 1. So projects that had claimed thousands of kWh	We don't have the document that the commenter references, nor is it clear if the commenter is discussing kWh or kW. We've reviewed Appendix C to the ERT documentation and cannot find the reference to .000003 realization rate mentioned in the comments. However, examining ERT inputs, it looks like the ERT value quoted here has extra 0s and is rounded up from .000249 in the field labeled EDUESkW. That is not a realization rate. The HVAC - Specialized Commercial Evaluation report shows that the realization rate for gross energy savings (kWh) for the four sites metered was 81% (page 193, page 195). The ERT inputs show 1 as the quantity, representing 1 unit. The ERT input	Corrected EDQuantity and ERT Units

#	Comment	Response	Solution/Action
		sheet doesn't show CO Sensors.	
11	The ERT inputs for the MAP program are quite problematic. The evaluation found gross realization rates of 81% and 110% for kWh and kW, respectively, for CO sensors. But the ERT contains a small UES that is equal for each line item that is significantly different individual energy savings that were claimed. The Turbocor Chiller measure was evaluated, but the evaluators decided to accept the claimed value because of lack of information. These measures were coded "PassThru" which does not match the methodology used elsewhere where values passed-through by the evaluators were coded as "EMV." Appendix C has absolutely no explanation of how values were changed. It only has a table with program level savings and no description.	For Turbocor chillers, the evaluation included case studies, as described in the Specialized Commercial/HVAC Evaluation Report. We collected trend data from seven sites and monitored performance at three of these. Resulting data were used to create actual performance curves for each site, which were compared to performance data from other compressor manufacturers. Ex ante savings and NTG assumptions were passed through.	Corrected EDQuantity and ERT Units
12	The Demand Control Ventilation measures in the program were evaluated. It appears that the savings estimates in the ERT were changed from per project to per ton. Again, Appendix C has absolutely no explanation of how values were changed. It only has a table with program level savings and no description.	Commenter is correct for SCE2561. The unit energy and demand savings are presented per ton, but the quantity is presented per unit. The ERT units should be amended to reflect the tonnage of the unit rather than unit quantity (which is always =1). The tonnage for each record is available in the Standard Program tracking database.	Corrected EDQuantity and ERT Units