ENERGY SAVINGS ASSISTANCE (ESA) PROGRAM

Impact Evaluation Program years 2015–2017

Southern California Gas Company

April 26, 2019

Table of contents

List of figures

List of tables

1 EXECUTIVE SUMMARY

1.1 Program background

The Energy Savings Assistance Program (ESA) provides no-cost energy efficiency services and no-cost direct installation energy efficiency measures to income-eligible households via ratepayer funding. ESA was developed in the early 1980s to improve the access of income-eligible households to utility conservation programs and provide relief from rising energy costs.

From 2014 through 2017, ESA served approximately 260,000 households per year.[1] Program services include energy education, an in-home energy assessment, and installation of one or more qualifying (or feasible) measures that are identified during the in-home assessment. ESA is implemented by four California investor-owned utilities (IOUs): Pacific Gas and Electric Company (PG&E), Southern California Edison (SCE), Southern California Gas Company (SCG), and the San Diego Gas and Electric Company (SDG&E).[2] The California Public Utilities Commission (CPUC) defines the program budget and maintains an oversight role.

The goals of ESA are to provide 100% of all eligible and willing customers the opportunity to participate in ESA by 2020, improve the health, safety, and comfort (non-energy benefits) of ESA customers, and produce cost-effective longer-term energy savings in income-eligible households that provide a reliable energy resource for California.[3]

Homeowners and tenants who receive electric or gas service from at least one of the four IOUs may receive ESA program services if they meet eligibility criteria in all the following categories:

- Household income must be at or below 200% of the Federal Poverty Guidelines[4]
- The building type is either a single family, multifamily, or mobile home with an active utility account/meter (including master meters) on a residential billing rate. CARE-eligible group living facilities on non-residential rates are also eligible if the structure is a single family, multifamily or mobile home if it meets ESA standards[5]
- In rental properties, the household members must obtain approval from the homeowner
- The type and frequency of previous ESA Program participation[6]
- For direct installation of feasible measures, the first two measure must meet a minimum energy savings threshold[7]

---

[1] From 2014–2017, ESA treated 1,039,720 households.

[2] In 2016, ESA program services are also provided by Southwest Gas Corporation, Liberty Utilities, Golden State Water Company/Bear Valley Electric, PacifiCorp, and Alpine Natural Gas Operating Company. Program results for these jurisdictions are not included in this impact analysis.

[3] California Public Utility Code Sections 382(e), 386(a)(3), 900, 2790, and the California Energy Efficiency Strategic Plan (CAEESP), adopted in D.10-09-047. Southern California Edison. 2014. Energy Savings Assistance (ESA) Program Plan and Budget Proposal for the 2015–2017 Program Cycle. California Public Utilities Commission (CPUC) Internal Audit Unit, Energy Savings Assistance (ESA) Program, October 2017.

[4] ESA Income Guidelines. All household members are considered when determining household income eligibility.

[5] California Alternate Rates for Energy Program (CARE).

[6] For program year 2017, the go-back rule, the three-measure minimum rule, and measure caps limiting the number of measures per household were removed. Commission Decision D.16-11-022.

[7] Additional measures are not required to meet a savings threshold

1.2 Evaluation background

The two most recent ESA evaluations (2009[8] and 2011[9]) raised concerns regarding the methodology. In both cases, a perception that savings estimates were either low or inconsistent influenced results. For the 2009 evaluation, an overly aggressive data trimming rule was identified as a driver of unexpectedly low results and changed for the final report. For the 2011 evaluation, a decision rule was applied that favored results closer to the ex ante levels, inflating results to levels well above the reported whole house results. For both evaluations, concerns regarding the methodology ultimately led to inconsistencies with the results.

The concern with the improving the evaluation method is understandable because evaluating the ESA program is challenging. The program serves a large population of households across a diverse set of housing types, namely single family (SF), multi-family (MF), and mobile homes (MH), and across all 16 of California's climate zones. An added challenge is that the expected household savings are less than 5% of total household consumption. Given the natural variability of consumption across sites and over time, these are unavoidable fine margins that test the limits of evaluation methods.

DNV GL's proposed approach is designed to offer a robust, routinized approach that provide a foundation of consistent and replicable results going forward. The approach follows standard evaluation protocols as simply and transparently as possible while maintaining the fundamental requirement of billing analysis: weather normalization and a comparison group to account for non-program related change over time. With consistent methods across the 3 evaluation years included in this evaluation (2015–2017) and a reasonable expectation that these results will be used in future evaluations, we can focus on the remaining variation across IOUs, years, housing-types, and geography as outgrowths of variation in program offerings and implementation.

1.3 Evaluation objectives

The research plan for the impact evaluation of the 2015–2017 ESA program was finalized in September 2017. The evaluation was divided into two phases. The first phase of the evaluation added the addition year of data from 2014 and used program data from 2014–2016 to set up the modeling frame work and developed phase 1 results for use in the ESA mid-cycle review in the summer of 2018. The first phase of the impact evaluation had the following priorities:

- Develop a routinized evaluation methodology.
- Produce consistent year-over-year savings estimates for electric consumption (kWh) and gas consumption (therms) for program participant.
- Produce an Excel dashboard of whole house and measure-level savings estimates.

The second phase of the project included an additional year of program data and looked to refine the modeling approach for the final results. The objectives for phase 2 of the impact evaluation were:

- Incorporate program year 2017 into the phase 1 impact evaluation analysis.
- Conduct additional modeling analysis into measure groups to refine results.
- Produce an Excel dashboard of whole house and measure-level savings estimates and a comprehensive report.

---

[8] Impact Evaluation of the 2009 California Low-Income Energy Efficiency Program Final Report, Study ID: SCE0273.01

[9] PY2011 Energy Savings Assistance Program Impact Evaluation, Final Report, Study ID: SDG0273.01. Evergreen Economics, August 30, 2013

1.4 Analysis methods

A primary objective of this evaluation was to develop a routinized methodology that supports consistent, reliable results year-over-year. The two-stage approach has a long track record in energy program evaluation and is effectively the basis for current methods developed for new pay-for-performance programs in California and beyond. The methodology is attractive for a several reasons:

- Focus on the site-level
- Full use of weather information at the monthly level with flexible site-level models
- A comparison group as a proxy for non-program-related change
- Separation of the weather-normalization process from savings estimation
- Flexibility to expand to daily or hourly data in the future

As a widely implemented residential program with a complex, multiple-measure offering, the evaluation of ESA is best served with billing (consumption data) analysis. This simple and transparent approach offers the best vehicle for a routinized methodology that will provide robust evaluation over time.

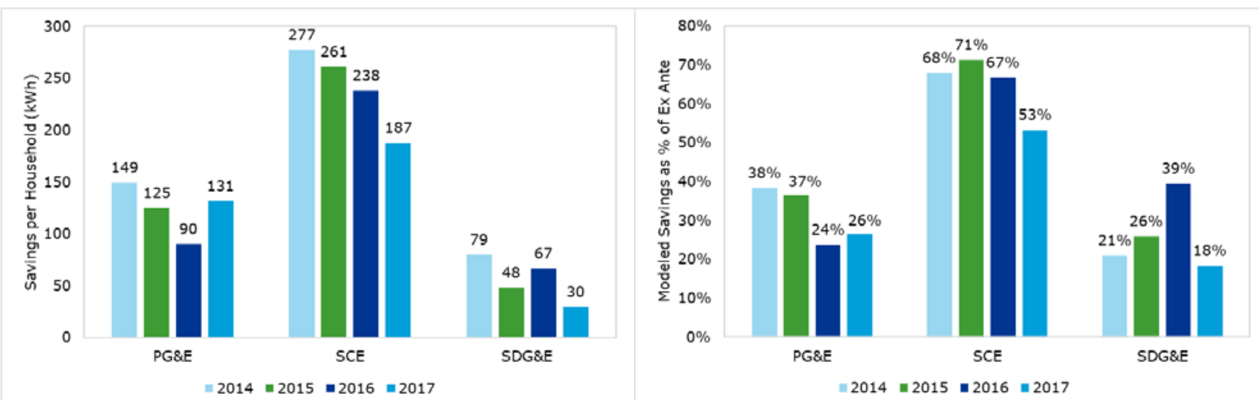
1.5 Impact results

We proved a high-level overview of the impact evaluation results here, with more in-depth discussion in Section 5. The evaluation also produced an excel results dashboard with all modeled results.

1.5.1 Electric impact estimates

Figure 1-1 provides the annual electric savings at the household level as well as savings as a percent of ex ante savings. The savings show three distinct levels of savings across the three IOUs. The savings as a percent of ex ante savings indicate that SCE's savings are substantially closer to expected savings than either PG&E or SDG&E. Despite distinctly lower savings levels, SDG&E is like PG&E in the level of achievement of expected savings. All three IOUs show a slight downward trend in annual savings. The trend remains in the savings as a percent of ex ante savings except for SDG&E.

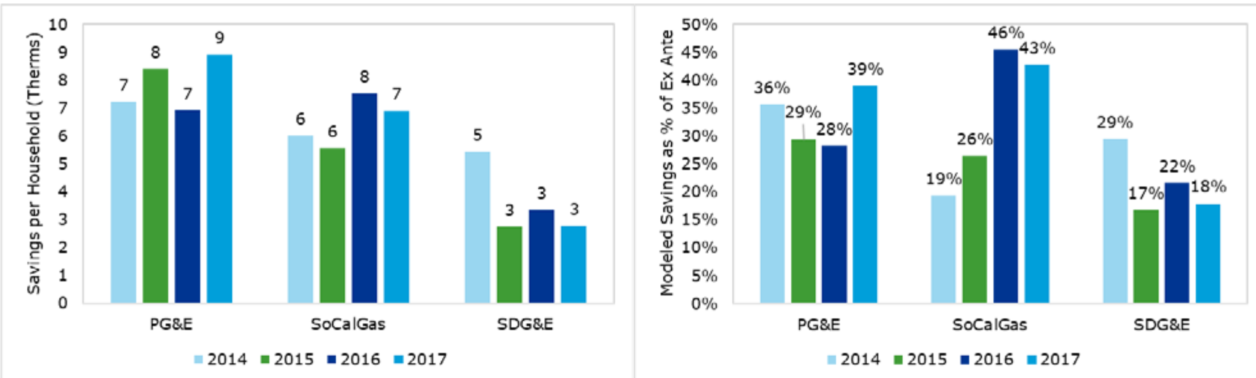
Figure 1-1. Electric savings per household and percent of ex ante savings over time



1.5.2 Gas impact estimates

Figure 1-2 provides the annual gas savings at the household level as well as savings as a percent of ex ante savings. The results show three levels of savings across the three IOUs, though the levels are not as distinct as the electric savings. The savings as a percent of ex ante savings for PG&E and SCG are at similar levels, on average, but SCG's realization rate improved dramatically in the latter 2 years. SDG&E's savings as a percent of ex ante savings are lower and may exhibit a downward trend.

Figure 1-2. Gas savings per household and percent of ex ante savings over time



1.6 Conclusions and recommendations

Key Conclusions	Key Recommendations
Ex ante savings assumptions were higher than achieved savings, with some measures leading to an increase in consumption.	ESA program planners can use the impact results to develop new ex ante savings assumptions.
ESA tracking data are organized differently across IOUs	ESA program administrators should look to improve program tracking with standardized fields, and better align program data with billing system.
The evaluation methodology produced consistent results at the household level but not at the measure level.	Future evaluations should explore other statistical analytical methods
There are limits to the answers that a billing analysis can provide for how program delivery effects savings.	Future evaluations should include a process evaluation element to better research how program delivery is linked to the impacts.

2 INTRODUCTION

2.1 Program background

The Energy Savings Assistance Program (ESA) provides no-cost energy efficiency services and no-cost direct installation energy efficiency measures to income-eligible households via ratepayer funding. From 2014 through 2017, ESA served approximately 260,000 households per year.[10] Program services include energy education, an in-home energy assessment, and installation of one or more qualifying (or feasible) measures that are identified during the in-home assessment. ESA is implemented by four California investor-owned utilities (IOUs): Pacific Gas and Electric Company (PG&E), Southern California Edison (SCE), Southern California Gas Company (SCG), and the San Diego Gas and Electric Company (SDG&E).[11] ESA was developed in the early 1980s to improve the access of income-eligible households to utility conservation programs and provide relief from rising energy costs. The CPUC defines the program budget and maintains an oversight role.

The goals of ESA are to provide 100% of all eligible and willing customers the opportunity to participate in ESA by 2020, improve the health, safety, and comfort (non-energy benefits) of ESA customers, and produce cost-effective longer-term energy savings in income-eligible households that provide a reliable energy resource for California.[12]

To maximize participation of eligible households, the IOUs refer eligible customers to each other, work with community agencies, local government, and the Department of Community Services and Development (CSD).[13] Coordination between the IOUs and governmental agencies increases the number of available measures by sharing the cost of measures offered by both programs. Private contractors selected by each utility are authorized to solicit low income households directly, determine income eligibility, and provide program services. The ESA policies and procedures manual governs ESA service providers and program activities. Customer outreach policies defined by the IOU and the policies and procedures manual covers promotional guidelines, limitations on representations made by Service Providers, customer interactions, and tracking.[14]

In areas where a customer receives natural gas and electric services from separate utilities, those utilities work together to ensure the customer receives all feasible ESA measures.

---

[10] From 2014–2017, ESA treated 1,039,720 households.

[11] In 2016, ESA program services are also provided by Southwest Gas Corporation, Liberty Utilities, Golden State Water Company/Bear Valley Electric, PacifiCorp, and Alpine Natural Gas Operating Company. Program results for these jurisdictions are not included in this impact analysis.

[12] California Public Utility Code Sections 382(e), 386(a)(3), 900, 2790, and the California Energy Efficiency Strategic Plan (CAEESP), adopted in D.10-09-047. Southern California Edison. 2014. Energy Savings Assistance (ESA) Program Plan and Budget Proposal for the 2015–2017 Program Cycle. California Public Utilities Commission (CPUC) Internal Audit Unit, Energy Savings Assistance (ESA) Program, October 2017.

[13] In 2016, Commission Decision D.16-11-022.10 authorized CSD service providers to provide ESA services.

[14] Statewide Energy Savings Assistance Program, 2017-2020 Cycle Policy and Procedures Manual, March 16, 2018

Homeowners and tenants who receive electric or gas service from at least one of the four IOUs may receive ESA program services if they meet eligibility criteria in each of the following categories:

- Household income must be at or below 200% of the Federal Poverty Guidelines[15]
- The building type is either a single family, multifamily, or mobile home with an active utility account/meter (including master meters) on a residential gas or electric rate. CARE-eligible group living facilities on non-residential rates are also eligible if the structure is a single family, multifamily or mobile home if it meets ESA standards[16]
- In rental properties the household members must obtain approval from the homeowner
- The type and frequency of previous ESA Program participation[17]
- For direct installation, the first two measure must meet a minimum energy savings threshold[18]

2.1.1 Program services, delivery, and measures

ESA services include determination of eligibility, energy education, an in-home energy assessment, and low-cost energy efficiency measures including:

- Weatherization
- Replacement domestic hot water equipment
- Heating, ventilation, and air conditioning replacement equipment
- Lighting
- Appliances
- Maintenance

A detailed list of eligible measures can be found in Table 4-1. If no feasible measures are identified in the assessment the customer is still eligible for energy education.[19]

2.1.1.1 In-home energy assessment

During the in-home visit, the program contractor determines whether the customer is income eligible. If the customer meets the eligibility criteria, the contractor completes the paperwork, the education portion of the assessment, and identifies feasible energy savings measures using utility approved forms and/or tools. The contractor then returns to the household to install the measures. Inspections must be conducted for a sample of ESA measure installations but are mandatory for projects which include attic insulation or a furnace replacement.

2.1.1.2 Education

ESA provides in-home energy education covering heating, cooling, lighting, domestic hot water consumption, large and small appliance usage, greenhouse gas emissions, water conservation, and information on other available energy and social programs.

---

[15] ESA Income Guidelines. All household members are considered when determining household income eligibility.

[16] California Alternate Rates for Energy Program (CARE).

[17] In 2016 the IOUs could provide program services to any household previously served by the program prior to 2002 (the go back rule). For program year 2017, the go-back rule, the three-measure minimum rule, and measure caps limiting the number of measures per household were removed. Commission Decision D.16-11-022.

[18] Additional measures are not required to meet a savings threshold.

[19] California Statewide Energy Savings Assistance Program, Policy and Procedures Manual. July 15, 2013; Statewide Energy Savings Assistance Program 2017-2020 Cycle, Policy and Procedures Manual, March 16, 2018. These manuals accompany the ESA Program California Installation Standards Manuals.

Specific topics must include (but are not limited to): [20]

- The general levels of usage associated with specific end uses, installed program measures and appliances
- The impacts on usage of individual energy efficiency measures offered through the ESA Program or other Programs offered to low-income customers by the utility
- Practices that diminish the savings from individual energy efficiency measures, as well as the potential cost of such practices
- Ways of decreasing usage through changes in practices
- Information on CARE, the Medical Baseline Program, Family Electric Rate
- Assistance (FERA), Low Income Home Energy Assistance Program, (LIHEAP), Community Help and Awareness of Natural Gas and Electric Services (CHANGES), and other available programs
- Appliance safety information
- Understanding utility bills and current utility rates
- Greenhouse gas emissions
- Water conservation
- CFL disposal and recycling

2.1.1.3 Measures

The ESA program installs all eligible measures that are approved for a site. Until it was abolished in 2017, a minimum energy savings threshold had to be met if only two measures are installed. Otherwise, at least three measure had to be installed. The following measures were included in the evaluation cycle reported here:

- Domestic hot water measures such water heater insulation blankets and low flow showerheads.
- Envelope improvements to improve insulation or reduce air infiltration
- Lighting
- Major appliances
- Repair or replacement of HVAC or water heating equipment

Other technologies may be considered measures for the ESA program if they promote cost-effective energy savings or reduce energy related economic burdens.[21] Measures and measure mixes are reviewed every cycle (approximately every three years), and for any mid-cycle updates.

Most ESA attributes are the same across the IOUs but differences in implementation, program end-uses, and measures exist.[22] Variations between programs may or may not cause differences in observed savings between utilities and over implementation cycles.

---

[20] 2017-2020 Cycle, Policy and Procedures Manual

[21] CPUC, October 2017.

[22] Utilities can vary in their contractors' assessment methods, program tracking procedures, customer outreach, and energy education.

2.2 Concerns related to prior evaluations

A concern with the 6 prior ESA evaluations going back to 2001 is the inconsistency of results. It is a primary goal of this kind of impact evaluation to reflect changes in the program implementation, measure offerings, and the location of participants. Some portion of the observed variability over the years is likely due to these factors. Another more problematic cause of variability in results is likely the application of different evaluation methodologies.

Billing analysis, the general approach used for each of these evaluations, can be performed in different ways. They either compare change in consumption over time (pre- and post-installation consumption), compare participant consumption to a comparison group's consumption, or both. All approaches use regression analyses to develop a relationship between consumption and weather to account for different weather between pre- and post-installation periods and to put results on typical weather terms. Within this broad scope, there are several choices that evaluators can make that will affect results. In addition to the basic billing analysis approach, the interpretation can also have an important effect on results, and this proved to be of importance in this instance.

Two aspects of the 2011 impact evaluation[23] raised concerns that we addressed in our methodological approach. First, the 2011 evaluation did not use a comparison group. The evaluation used a pooled, time-series approach that compares participants pre- and post-installation consumption while controlling for weather. This approach has been widely used for impact evaluation but relies on the regression structure to address non-weather changes that may occur through the analysis timeframe, such as macroeconomic effects like recessions, which had an impact in 2011. Economic and other non-weather changes can shift consumption up or down by 3% or more, which is approximately the magnitude of the expected savings of the ESA program. The lack of a comparison group could have either decreased or increased the savings estimates. Comparison groups are now generally considered an essential addition to a billing analysis to address whatever limitations may exist in approaches that do not include a comparison group.[24]

The bigger concern from the 2011 evaluation was the development of savings and ex ante estimates from the regression results. Billing analysis regression results can be estimated at the household level or the measure level. Household-level, or whole-house, savings estimates are the most accurate and reliable billing analysis estimate of savings. Measure-level results are the household-level savings distributed to measures based on the relative savings of households with different measure groupings. There are multiple challenges estimating measure-level savings in a billing analysis context.

- Two or more measures may be installed in combination at most sites making it difficult for any algorithm to separate the effects.
- Two measures may interact when installed together, producing very different savings than when they are installed on their own.
- There are many measures with extremely small savings that are simply too small to definitively identify given the natural consumption variability across sites and over time.

[23] PY2011 Energy Savings Assistance Program Impact Evaluation, Final Report. August 30, 2013.

[24] The Universal Method Project Chapter 8, "Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol" discusses the rationale behind including a comparison group.

- The IOUs recognize that some measures may appear as an increase in consumption in the billing data (a repaired AC or furnace that was not previously working at all) or a decrease (an inefficiently functioning AC or furnace adjusted for better performance) depending on the context of the situation.

Considering these extensive challenges, it is not uncommon to get some measure-level results that indicate negative savings or are otherwise not statistically significant. It is tempting to pick and choose among the individual measure-level savings, but this approach has the potential to distort overall savings. This is particularly the case for a program that has some measures that are justified based on health and safety rather than energy savings (e.g., furnace and AC repairs).

The 2011 evaluator stated the following:[25]

Energy savings values were assigned to a measure group from the billing regression models using the following algorithm:

1. If the 95 percent confidence interval of the impact estimate from the Basic Model included the ex-ante savings value, then the estimate from the Basic Model was used.
2. If the confidence interval for Basic Model estimate did not include the ex-ante value, then evaluator judgment was used to assign an impact value from among the Basic Model, Measure Model, or ex-ante values.
3. In a couple of instances, an engineering estimate was assigned when the ex-ante values appeared to be unusually high and neither of the regression models could provide a reasonable result.

The effect of this algorithm was to accept regression-estimated measure savings that were positive and closer to the ex ante values and consider alternative savings estimates for negative or smaller values. The effect of this algorithm can only increase the overall savings when summed across measures.

Table 2-1 shows that for all IOUs and both gas and electric, the savings estimates were inflated above the whole-house model result. While there is only a small increase for SCE electric savings, other results are doubled, tripled, and increased nine-fold. The whole-house model estimates will always be the most accurate estimate of overall household savings. There is no reasonable methodological justification for diverging from whole-house results based on measure-level results.

Table 2-1. 2011 ESA evaluation: gas and electric impact estimates of single family households[26]

IOU	Electric (kWh)			Gas (Therms)		
	Whole-House Savings	Final Measure-Based Model Results	% Difference	Whole-House Savings	Final Measure-Based Model Results	% Difference
PGE	36	367	919%	7.6	21.5	183%
SCE	267	279	4%			
SCG				9.5	13.4	41%
SDGE	158	279	77%	8.1	26.1	222%

[25] PY2011 Energy Savings Assistance Program Impact Evaluation, Final Report.

[26] PY2011 Energy Savings Assistance Program Impact Evaluation, Final Report, Tables 25 and 27, pp. 44 and 45.

There are 3 key lessons from the 2011 evaluation.

1. Whole-house billing analysis results can vary across methodologies. Billing analyses should include a pre- to post-installation difference in consumption of participants in the context of a pre- to post-installation difference in consumption of a matched non-participant comparison group. The combination of these two different ways of assessing program effects in a difference in difference framework will produce the most robust and consistent results regardless of other methodological details. For instance, it is possible, that the very low PG&E whole-house electric savings for the 2011 evaluation were due to the lack of a comparison group in the base regression model.
2. The whole-house savings estimates are the most accurate and reliable result. The estimation of measure-level results for the ESA program faces numerous challenges. In contrast, whole-house savings estimates give a single result that addresses potential thermodynamic interactive effects and is big enough relative to consumption, across measures, that it is more likely to be statistically significant. We understand that measure savings estimates are important for planning purposes, but measure level savings cannot be allowed to drive the savings results from the evaluation.
3. Measure-level regression results must be interpreted with care, always within the context of the whole-house result. The primary issue of the 2011 evaluation was that the effort to develop measure-level results took precedence over the basic validity of the evaluation whole house results. The measure-level results are directional and informative for identifying measures that make substantial individual contributions to savings. The measure estimates offer a template for one, mathematically-correct way to split household-level savings out to measure groups. DNV GL developed a tool that allowed the IOUs to adjust measure-level savings while maintaining an overall savings level that was consistent with the household-level results.

This evaluation seeks to put the ESA evaluation on a footing that will support accurate and robust evaluation results and provide the necessary tools and understanding to enhance the program going forward.

2.3 Evaluation objectives

The research plan for the 2015–2017 impact evaluation of the ESA program was finalized in September 2017. The evaluation was divided into two phases. The first phase of the evaluation added the addition year of data from 2014 and used program data from 2014–2016 to set up the modeling frame work and developed phase 1 results for use in the ESA mid-cycle review in the summer of 2018. The scope of the evaluation was adjusted slightly over the course of the evaluation to accommodate shifts in the timeline and to keep the project within budget; for example, the research plan included analysis to explore program redefinition for Aliso Canyon area. This analysis was reprioritized over the course of the evaluation and not included in the updated research plan. Overall the main objectives of the evaluation stayed the same and are included in this report.

The first phase of the impact evaluation had the following objectives:

- Develop a routinized evaluation methodology.
- Produce consistent year-over-year savings estimates for electric consumption (kWh) and gas consumption (therms) for program participants.
- Produce an Excel dashboard of whole house and measure-level savings estimates.

The second phase of the project included an additional year of program data and looked to refine the modeling approach for the final results. The objectives for phase 2 of the impact evaluation were:

- Incorporate program year 2017 into the phase 1 impact evaluation analysis.
- Conduct additional modeling analysis into measure groups to refine results.
- Produce an Excel dashboard of whole house and measure-level savings estimates and a comprehensive report.

The primary evaluation results in this study are household level savings estimates. For a program with a complex, multiple-measure offering and average expected savings of lower than 5% of consumption, reliable measure-level savings estimates for all measures are not feasible. This evaluation provides guidance on measure-level savings to the extent feasible within the limitations of the data. Unlike prior evaluations, we do not provide explicit measure-level savings estimates for use in planning processes.

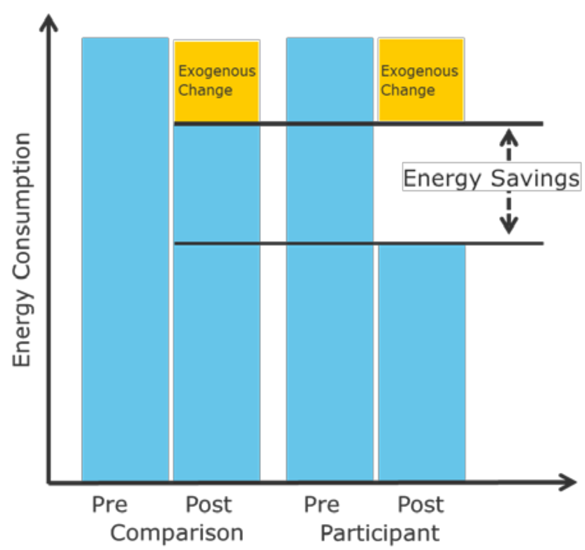
2.4 Analysis methods

A primary objective of this evaluation was to develop a routinized methodology that would support consistent, reliable results year over year. The two-stage approach has a long track record in energy program evaluation and is effectively the basis for current methods developed for new pay for performance programs in California and beyond. The methodology is attractive for a variety of reasons:

- Focus on the site-level
- Full use of weather information at the monthly level with flexible site-level models
- A comparison group as a proxy for non-program-related change
- Separation of the weather-normalization process from savings estimation
- Flexible to expansion to daily or hourly data in the future

As a widely implemented residential program with a complex, multiple-measure offering, the evaluation of ESA is best served with billing (consumption data) analysis. This simple and transparent and approach offers the best vehicle for a routinized methodology that will provide robust evaluation over time.

2.4.1 Importance of site-level approach and results



The site-level modeling approach treats each household as a unique entity. The approach requires a complete year of pre- and post-installation data so that the unique energy consumption characteristics of the household can be clearly identified. The models, applied at the household level, incorporate engineering principles to characterize the baseload, heating, and cooling dynamics of the household. These models facilitate weather normalization by putting annual consumption on typical weather terms while maintaining the unique household thermodynamic characteristics. Site-level modeling applied consistently to the participant and matched comparison groups, produces weather-normalized pre- to post-installation differences that reflect the change in consumption through the analysis period. The participants and comparison group are combined in a second-stage model that applies a simple but robust difference in difference structure. The difference in difference structure can be understood two complementary ways: Participant pre- to post- installation difference is adjusted with a proxy estimate of non-program related exogenous change from the comparison group; The difference between participant and comparison group post-period consumption is adjusted by any differences identified in the pre-installation data. The following is a simplified formula for a difference in difference

$$Energy\ Savings = (Treatment_{Pre} - Treatment_{Post}) - (Control_{Pre} - Control_{Post})$$

2.4.2 Implications of site-level approach for measure-level results

For programs like ESA with many small savings measures, billing analysis-based estimates of measure-level savings will always be challenging regardless of the methodological approach. The total expected savings as a percent of overall consumption is small compared to the natural variation across households. At the measure level the savings are extremely small.

The site-level approach provides a well-defined decomposition of consumption to baseload, heating, and cooling.[27] This decomposition supports an understanding of the effects of measures on consumption in these three areas. Also, household normalized difference in annual consumption enters a second stage regression where measure level savings are distributed to measures based on the savings levels of households with different measure bundles. The second-stage results offer a more simple and transparent distribution of savings to the measure level than other options.

2.4.3 Two-stage vs. panel model approach

A panel model approach, the most common alternative to the two-stage approach, combines the weather-normalization process and the savings estimation in a single regression across customers and months. For measure-level results, the panel models further distribute savings to measures in the same mathematical optimization. With similar data inclusion rules (12 months pre- and post-installation) and the same comparison group, panel and two-stage approaches will usually give similar results at the whole house level. The site-level approach offers model flexibility and transparency while the panel approach may offer improved precision of estimates.

2.4.4 Consistent and replicable results across years

A central part of savings estimation is addressing change in weather year-over-year and weather-normalizing consumption to typical weather year terms. The two-stage approach, with its flexible site-level models, achieves this at the individual site level. The panel approach weather-normalizes with a single model structure across the whole population. At the panel level, weather normalization is determined by a group of unique households facing different weather. A panel model of PG&E territory applies a single linear trend from a single degree day base to characterize cooling across households in both the Bay Area and hotter portions of Central Valley. This sensitivity can affect estimate of weather normalized savings at both the overall and measure level.

[27] As discussed in Section 4.2, baseload is the non-weather-correlated load remaining after heating and cooling-correlated loads have been removed.

3 DATA

We present the data and processing steps used in the evaluation in the following sections:

- Section 3.1 provides an overview of the data we used in the analysis.
- Section 3.2 summarizes the ESA program tracking data.
- Section 3.3 explores the combinations of measures installed by ESA participants to provide context for estimating measure level savings.
- Section 3.4 presents the weather data we used to normalize energy consumption.
- Section 3.5 discusses data preparation.

3.1 Data sources

DNV GL used the following four primary data sources in this evaluation:

1. *Tracking data.* The IOUs provided ESA program tracking data for the 2014 through 2017 program years. We used it to identify program participants and obtain measure-level savings estimates for each participant.

2. *Billing data.* We used monthly billing data records from January 2010 to June 2018, provided by the IOUs, to evaluate the energy savings from the program. We also used the billing data to identify non-participant customers on the CARE rate plan at any point between January 2010 and July 2018 to serve as the potential comparison group.

3. *Customer data.* The IOUs provided customer data which included information on customer location, climate zone, and housing type.

4. *Weather data.* We obtained the weather data from NOAA and CZ2010.[28]

3.2 Tracking data summary

The tracking data provided information about program participants, including individual and program identifiers, installation dates, and measure-level information. Installation dates flagged program initiation at individual sites while measure-level information listed items installed and expected energy savings (ex ante savings). We requested supplemental information from SDG&E and SCG for tracking files that did not include ex ante savings for program years 2014–2016.

Table 3-1 summarizes the tracking data used in this evaluation. We verified the total count of participants, overall and measure level expected savings and the installed number of measures, with claims that the IOUs file annually with the CPUC.[29] Information from the IOU tracking data generally conforms with data reported to the CPUC though there are areas noted where there are issues.[30]

---

[28] National Oceanic and Atmospheric Administration Hourly Weather Data; California Energy Commission Title 24. https://www.energy.ca.gov/title24/.

[29] Income Qualified Assistance Programs, http://www.cpuc.ca.gov/iqap/

[30] Some of the ex ante savings from SDG&E's tracking files did not align with CPUC tracking data sources for the years 2014–2016. Values from SDG&E's 2017 tracking data, however, were well aligned with CPUC reported information.

Table 3-1. Tracking data summary by year across IOU

IOU	Year	Number of Participants	Claimed Savings		
			kWh	Therms	kW
PG&E	2014	120,099	42,422,718	1,947,923	8,168
	2015	96,878	31,443,738	2,221,789	5,853
	2016	71,709	26,003,820	1,569,712	5,285
	2017	85,159	58,254,754	1,641,681	69,358*
SDG&E	2014	23,049	10,167,536	277,825	0**
	2015	21,423	4,075,803	197,041	487
	2016	20,340	3,796,839	190,128	93
	2017	21,862	3,444,033	208,290	414
SCE	2014	69,377	32,982,424	NA	12,543
	2015	65,287	27,965,788		4,499
	2016	63,176	27,616,052		4,443
	2017	70,808	31,651,052		4,791
SCG	2014	93,630	NA	3,041,960	NA
	2015	71,112		1,534,184	
	2016	65,576		1,175,007	
	2017	82,271		1,502,002	

*PG&E's unusually high kW claims in 2017 are due to an error in LED kW savings assumptions.

**SDG&E's 2014 tracking data reported negligible demand savings.

Across IOUs, the number of ESA participants decreased from 2014 to 2016 before increasing in 2017. Claimed savings generally declined over the same period. SDG&E's electric and gas claimed savings decreased over this period. Expected electric savings by SCE, gas savings by SCG, and electric and gas savings by PG&E decreased from 2014–2016 but increased in 2017. The increases for SCE and SCG were modest while PG&E's claimed electric savings doubled, and demand savings increased by more than 10 times from 2016 to 2017. Tracking information indicates that these increases are due to new LED based lighting installations including LED A-lamps and LED light fixtures.

3.3 Measure groups

Energy consumption data analysis reflects changes in whole-house energy use due to program intervention. Such analysis appropriately accounts for interactive effects and possible take-back (or increase in energy consumption over and above baseline use). However, consumption data analysis does not indicate the contribution of program elements, such as measures or measure groups, to the estimated whole-house energy change. Despite the challenges discussed in Section 2.2, we seek to understand measure-level program effects to help inform improvements in program design.

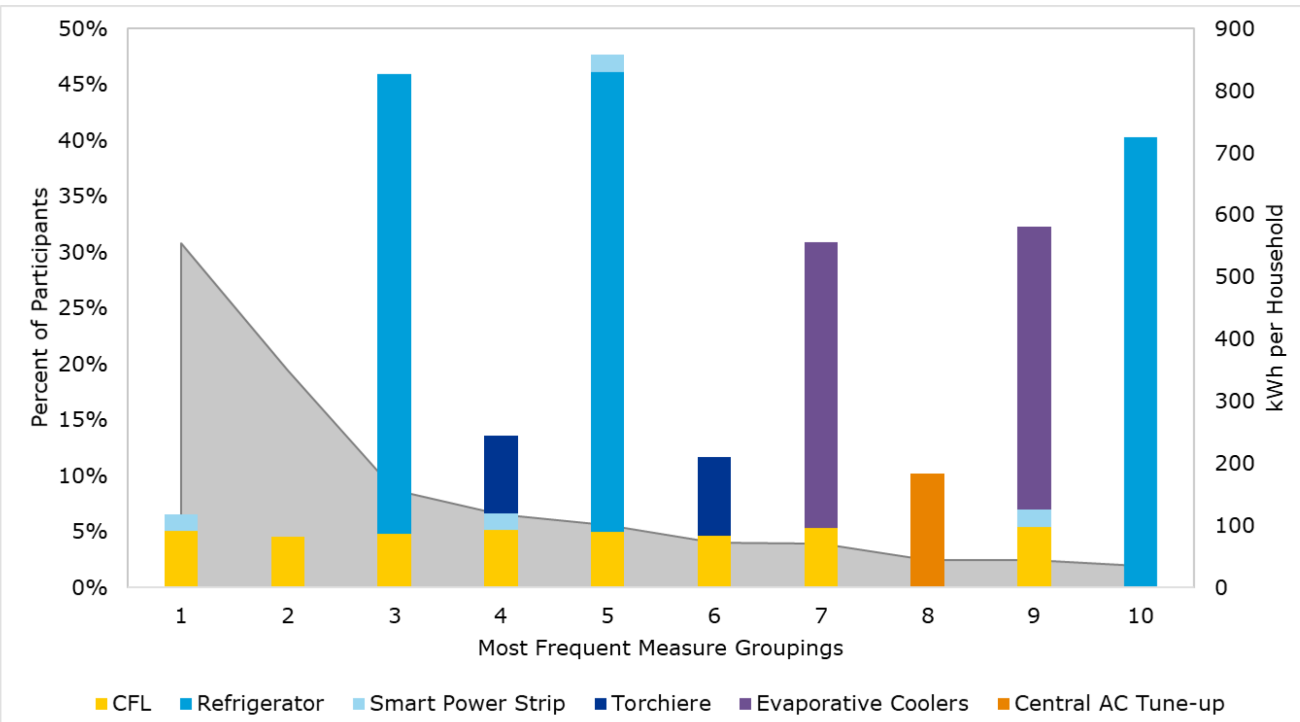
To calculate savings, we first identify the measure bundles installed by the program and then determine how they can be grouped to estimate program effects. The measure groups used to estimate savings may be composed of a single measure or multiple measures frequently installed together. When measures are consistently installed together, it is not possible to estimate reliable savings for individual measure.

In Section 3.3.1, we identify the measure bundles installed by the program to determine the measure groups used to estimate measure-level savings. A description of electric measure groups is followed by a description of gas measure grouping used to identify reasonable measure groups for estimating measure-level savings.

3.3.1 Electric measure groupings

ESA offered over 30 different electric measures which were installed in a variety of combinations which we refer to as measure groupings below. Figure 3-1 to Figure 3-3 provide a summary of the top 10 unique electric measure groupings installed in program years 2016–2017 by SCE, SDG&E, and PG&E, respectively.[31] The x-axis shows the ranking of the measure combination. Measure group 1 was the most frequent bundle. The grey plot indicates the percent of households who installed each measure grouping (left axis), while the stacked bar chart shows the average savings per household for each measure (right axis). As an example, in Figure 3-1, 31% of households installed CFLs and smart power strips with an average savings per household of 90 kWh for CFLs and 28 kWh for smart power strips.

Figure 3-1. Most frequent SCE electric measure groupings, 2016–2017



Tracking data indicate that SCE and SDG&E installed 341 and 377 unique measure bundles, respectively. By contrast, PG&E installed 3,206 unique electric measure mixes, a nearly 10-fold greater combination than existed in the programs run by the two other electric IOUs. The greater measure mix offered to PG&E's customers is reflected by the fact that the 10 most frequent measure bundles account for installations at premises of only 40% of participants. Across the 3,206 measure unique bundles there is an average of 14

[31] Measure permutation analysis of this kind was also performed for program years 2014–2016 during Phase 1 of the evaluation. The analyses from both phases informed the measure bundles used in measure-level models. We present measure permutation results from phase 2 to illustrate the process used to determine measure bundles.

customers. SCE's and SDG&E's top 10 measure bundles, on the other hand, account for installation in about 80% to nearly 90% of their participant households.

The 3 most frequent measure combinations included a CFL. As a result, lighting savings estimates could be conflated with other widely installed measures. Both SCE and SDG&E included a smart power strip in their most commonly-installed group of measures. Additionally, SDG&E and SCE had a refrigerator in their third-most common group of measures.

SCE's most common measure groupings and the percent of household that installed each group:

- CFL + smart power strips installed (31%)
- CFL only (19%)
- CFL + refrigerators (9%)

Figure 3-2 shows SDG&E's most common electric measure groupings and the percent of household that installed each group:

- CFL + LED night lights + smart power strips (29%)
- CFL + LED night lights + water heater blanket (26%)
- CFL + LED night lights + smart strip + torchiere + refrigerator (4%)

Figure 3-2. Most frequent SDG&E electric measure groupings, 2016–2017

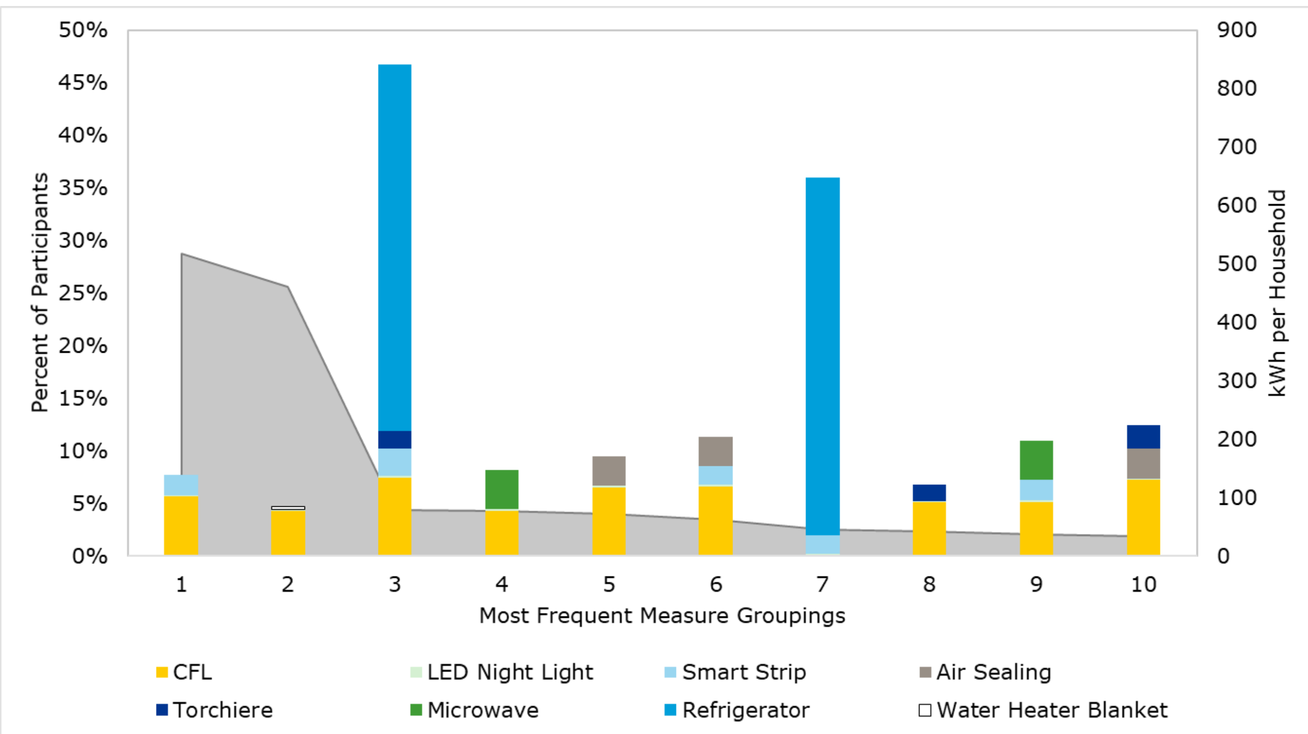
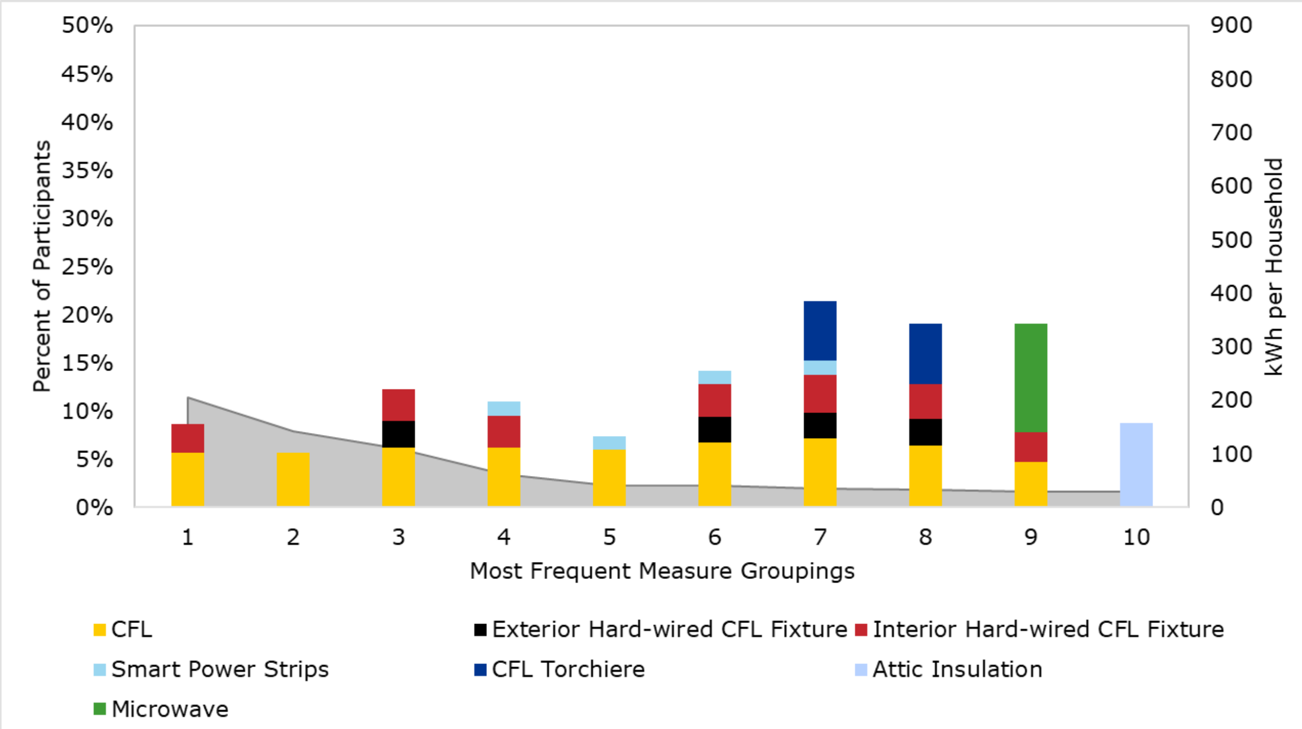


Figure 3-3 shows PG&E's most common electric measure groupings and the percent of household that installed each group:

- CFL + interior hard-wired CFL fixture (11%)
- CFL only (8%)
- CFL + exterior hard-wired CFL interior hard-wired CFL bundle (6%)

Figure 3-3. Most frequent PG&E electric measure groupings, 2016–2017



3.3.2 Gas measure groupings

ESA offered 15 different gas measures which were installed in a variety of combinations. This section provides a summary of the top 10 unique combinations of gas measures installed in program years 2016-2017 by SCG, SDG&E, and PG&E.

Nearly all the most common measure bundles include water conservation and energy savings measures such as faucet aerators and low flow showerheads. SCG and PG&E installed air sealing in two out of their top three most common measure combinations. Only SDG&E's top measure groupings included microwaves. SDG&E and PG&E installed over 300 unique gas measure bundles while SCG installed over 500. The 10 most frequent measure bundles account for 70% – 80% of participants for the gas IOUs.

Figure 3-4 shows SCG's most common measure groupings and the percent of household that installed each group:

- Air sealing + faucet aerator + low flow showerhead + thermostatic shower valve (24%)
- Air sealing + faucet aerator + low flow showerhead + thermostatic shower valve + furnace tune-up (19%)
- Faucet aerator + low flow showerhead + thermostatic shower valve (16%)

Figure 3-4. Ten most frequent SCG measure bundles, 2016–2017

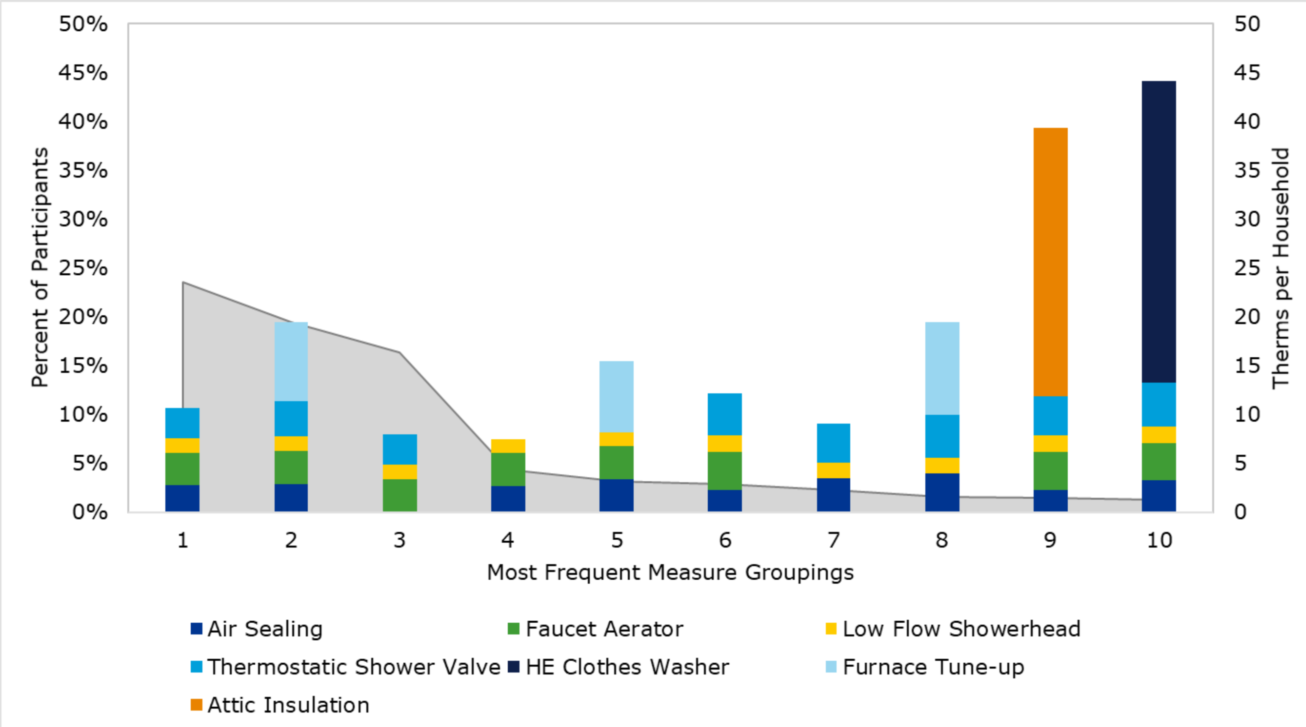


Figure 3-5 shows SDG&E's most common gas measure groupings and the percent of household that installed each group:

- Faucet aerators (43%) of households
- Microwave at (13%) of households
- Faucet aerator + microwave (8%)

Figure 3-5. Ten most frequent SDG&E gas measure bundles, 2016–2017

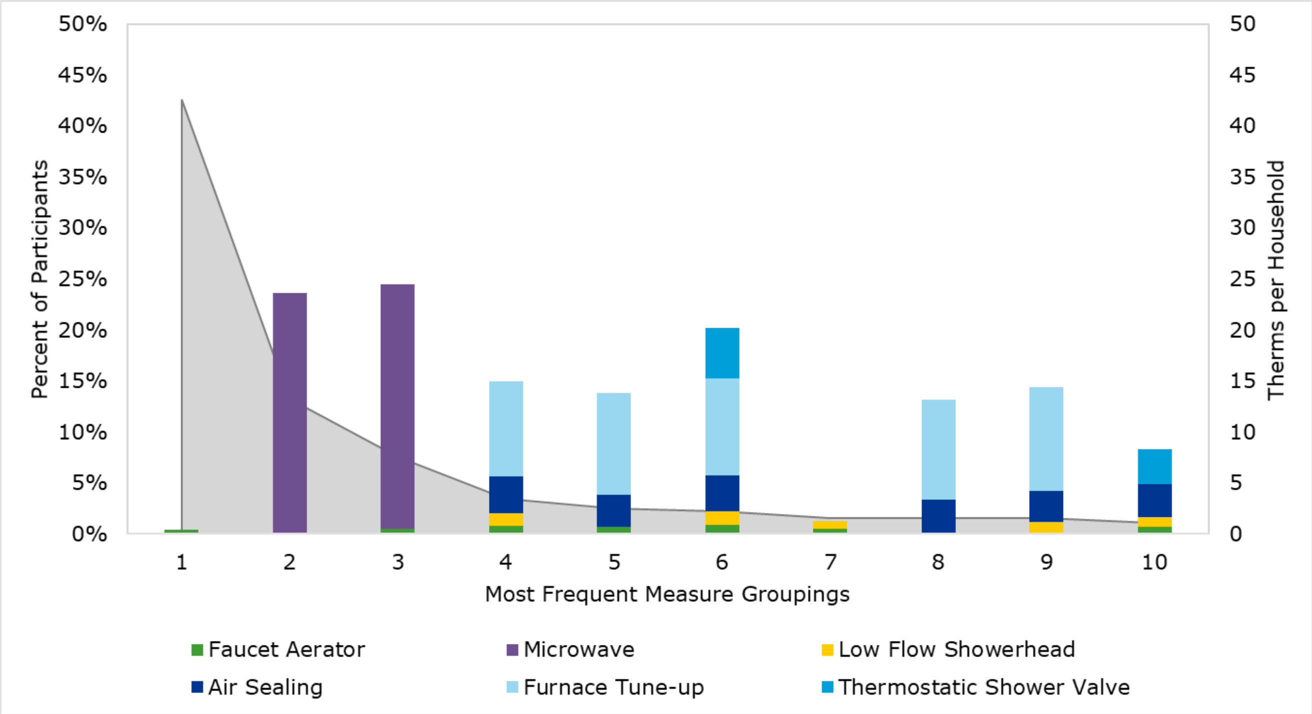
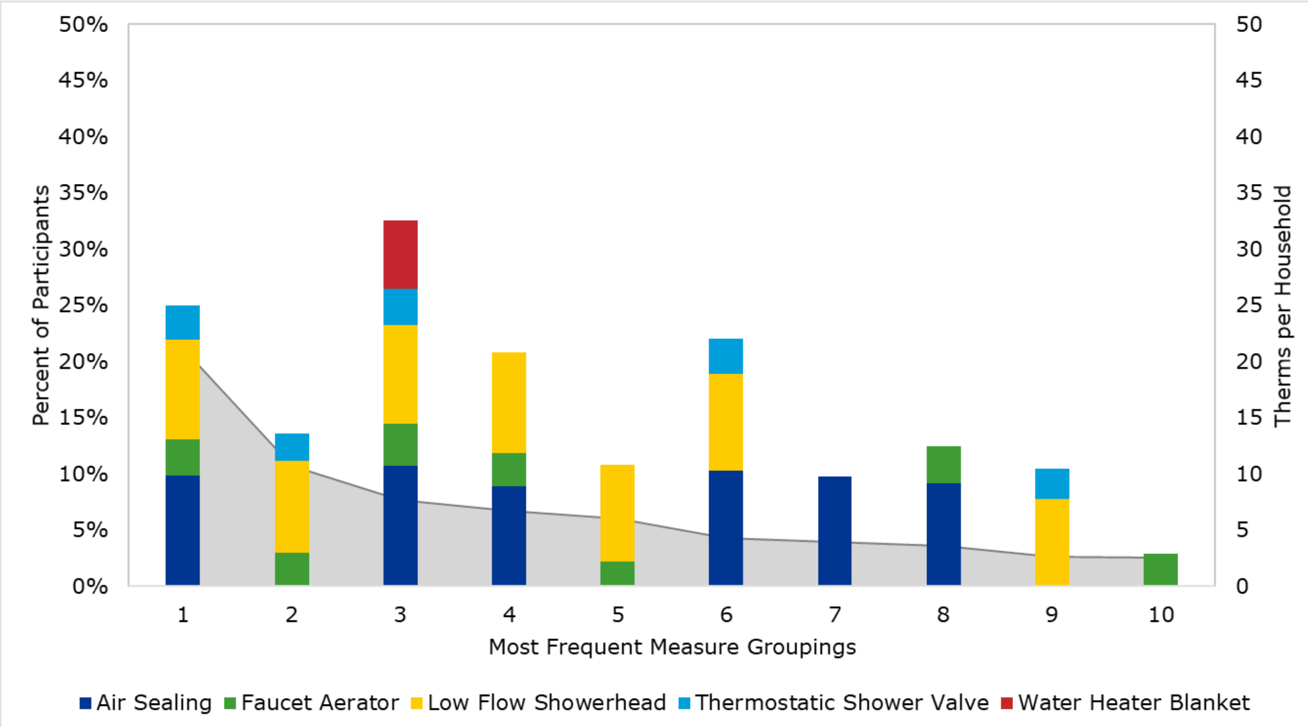


Figure 3-6 shows PG&E's most common gas measure groupings and the percent of household that installed each group:

- Air sealing + faucet aerator + low flow showerhead + thermostatic shower valve (21%)
- Faucet aerator + low flow showerhead + thermostatic shower valve (11%)
- Air sealing + faucet aerator + low flow showerhead + thermostatic shower valve + water heater blanket at (8%)

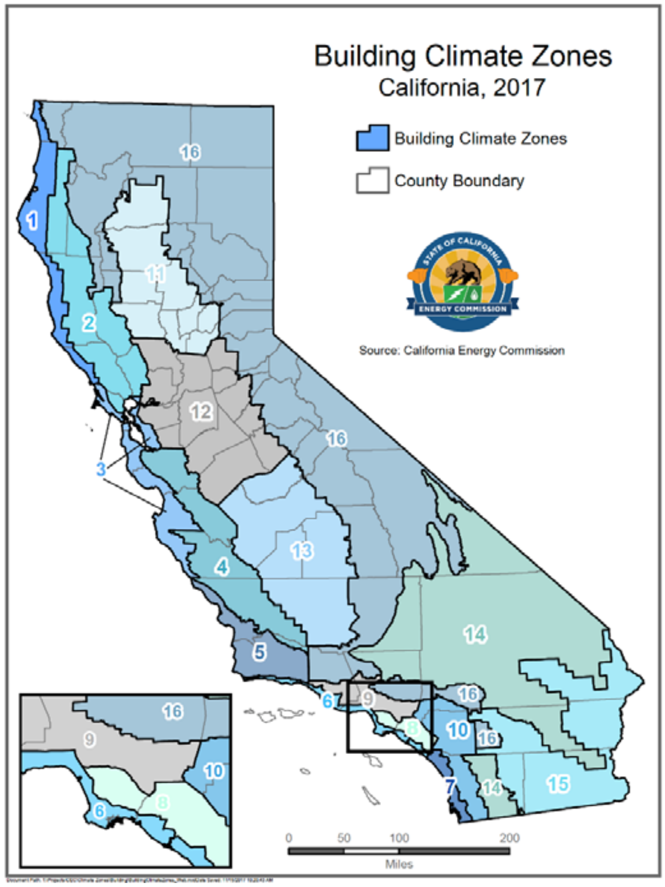
Figure 3-6. Most frequent PG&E gas measure bundles, 2016–2017



3.4 Weather data

Observed and typical meteorological year (TMY) data are important inputs for addressing changing weather conditions and their effect on energy consumption. There are 86 NOAA weather stations across California that provide historical weather observations and for which TMY series were developed (CZ2010). The 86 weather stations are mapped to the Title 24 Climate Zones, displayed in Figure 3-7.

Figure 3-7. California Title 24 climate zones (2017)

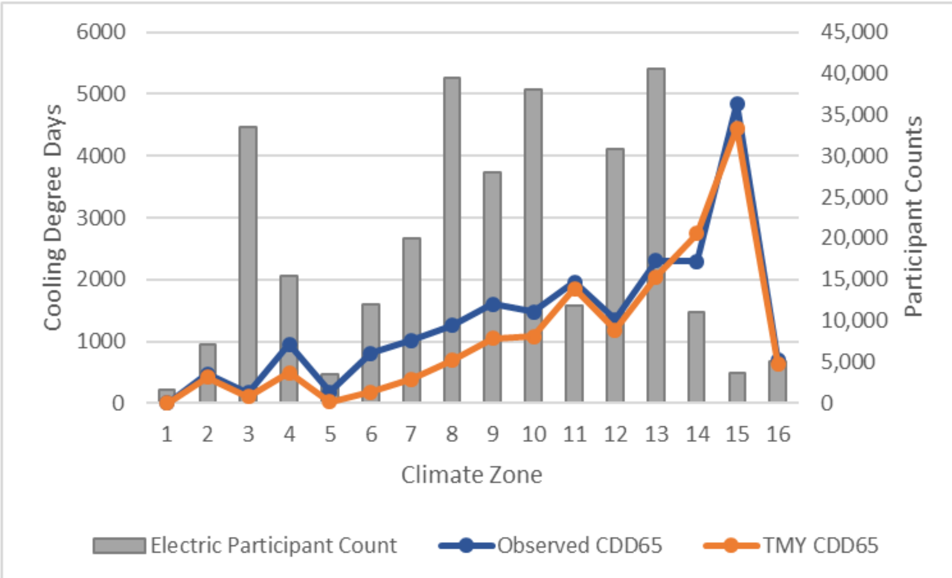


Weather data enter the analysis as degree days, which are values above or below some reference point or degree day base. Reference or base points indicate temperatures at which individual households switch to using cooling in the summer (for example, 70°F) and heating in the winter (for example, 60°F). Cooling degree days (CDD) are degrees above the base temperature (temp minus base) or zero. If daily average temperature is 80°F, for example, CDD would have a value of 10 for a reference temperature of 70°F. Heating degree days (HDD) are degrees below the base temperature (base minus temp), expressed positively, or zero.

For summarizing and comparing weather data over the timeframe of this analysis, we used HDD and CDD with base 65 (denoted by HDD65 and CDD65). For the sake of summary and comparison, we summed degree days to the annual level. We also aggregated weather data from the 86 weather stations to the 16 CA climate zones.

Figure 3-8 provides average annual observed and TMY CDD65 for each climate zone. The CDD65 values generally increase as the zones move southward and inland. The gray bars provide the number of participants across the analysis period from each climate zone. Most participants are in climate zones with observed CDD65 of at least 1000. The higher CDD65 will be correlated with greater cooling consumption.

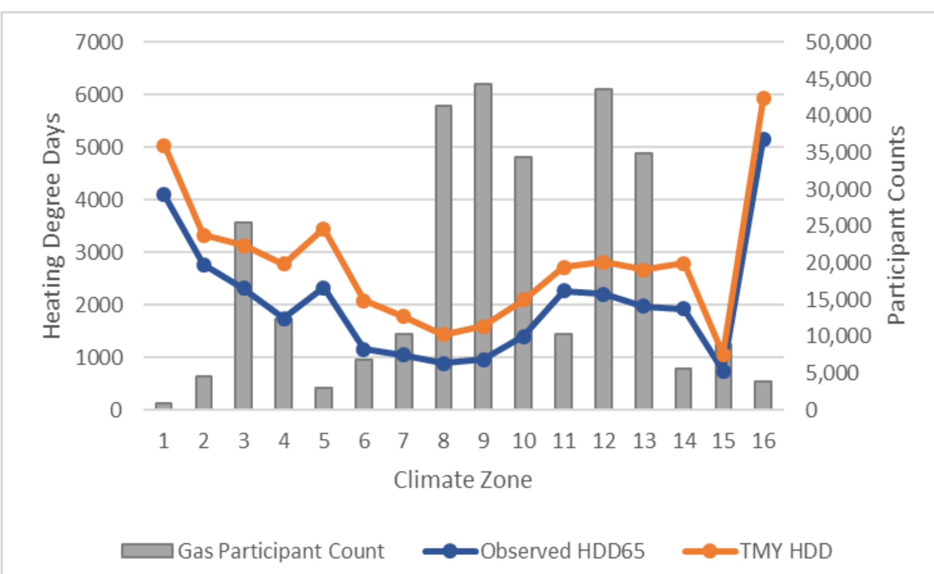
Figure 3-8. Observed CDD and TMY, by climate zone



Currently, CZ2010 TMYs reflect typical weather from 1980 to 2010. In many areas, the weather during the analysis time period was warmer than weather that undergirds the TMY data. In some climate zones, CDD values were greater by more than 500-degree days. Coastal climate zones in southern California observed CDD65 that look more like those of their inland neighbors. The difference between observed and TMY is much less dramatic in the Bay Area (zone 3) and the inland climate zones (zone 11 and up).

Figure 3-9 provides the parallel figure for observed and TMY HDD65. In keeping with the generally higher temperatures during the analysis period compared to TMY, observed HDD65 values are substantially lower than the TMY HDD65. The reduction in HDD, averaging over 700 HDD across all weather zones, is consistent across all climate zones across the state. The lower observed HDD65 will be correlated with lower heating consumption.

Figure 3-9. Observed HDD and TMY, by climate zone



3.5 Billing data

This analysis is conducted using monthly billing data for all IOUs and all years. In the planning phase of the analysis, the IOUs provided information regarding the availability of AMI data as the basis for this evaluation. While AMI data were reasonably complete in the early evaluation years for some IOUs, for others, AMI data were not fully available even at the end of the analysis timeframe. In consultation with the IOUs, the study team decided that the lack of complete AMI data in some years for some IOUs made it impractical to pursue AMI data as the basis for this evaluation.

AMI data coverage should be sufficient to allow its use for future evaluations. Daily consumption data derived from AMI data can be seamlessly integrated into the routinized modeling process used in this evaluation. The additional amount of data (daily consumption values for 365 days compared to 12 monthly average daily values) make it much easier to characterize a home's heating and cooling characteristics and, as a result, the weather-normalization process will be much more accurate.

3.5.1 Billing data screening

In this section, we provide an outline of the monthly energy consumption data preparation steps. Table 3-2 to Table 3-7 present the counts of ESA households used in the evaluation for each IOU by fuel and over time. The tables indicate participant counts with billing data, the number of participants with quality billing data, participants that had 12 months of pre and post data, and the final number of participants included in the analysis. The final participants included in the analysis have weather data and savings for the analysis fuel.

First, customer tracking data counts differ from customers with billing data for two reasons. The first is related to billing data availability and the second is related to tracking data preparation. The goal of the tracking data prep was to identify which customers had participation patterns that would allow us to conduct a billing analysis. For all IOUs, we applied the following criteria to prepare the tracking data:

1. Participation frequency. We kept only participant records with three or fewer unique installation dates. Most participants with more than three installation dates had dates that were too close together to include in the analysis as the effect of intervention with adequate pre and post participation data would not be available. Using three dates allowed us to include participants who had participated multiple times over a long period of time or had multiple installation dates that were closely clustered.

2. Participation window. For participants with three or fewer installation dates, we included participants whose latest installation date is either less than three months or more than 12 months apart from the two other possible installation dates. While assessing the distance between installation dates, we assigned program start and end dates. For participants with one installation date, we defined the program start date and the program end date as the installation date. For participants with more than one installation date within a three-month period, we defined the program start date as the earliest installation date and the program end date as the latest installation date. When installation dates were more than 12 months apart, we used the most recent installation date for the analysis and assigned the installation date as the program start and end date.

3. Participation year. Finally, we included participants who were in ESA in 2014 or later.

As a second step in data screening, we reviewed quality of the consumption data. We removed records with billing period abnormalities such as duplicate or overlapping read dates, billing periods less than 15 days or greater than 40 days, and records where the period end date preceded the start date. We also removed records if consumption was greater than the difference between the standardized maximum consumption and the average consumption.

The availability of 12 months of pre- and post-period billing data was the biggest factor in determining the final counts of households used in the analysis. Twelve months of pre- and post-program intervention data is required for a robust analysis of the program effect because it allows us to account for the effect of the program in all seasons. From 2014 to 2016, for each fuel and IOU, the household data attrition rate ranges from 27% to 57%. The 2017 attrition data rate is higher than 2014 to 2016 because 12 months of post data was unavailable, and the evaluation covered only the first half of 2017.

Table 3-2. PG&E electric billing data attrition

Billing Data Attrition	2014	2015	2016	2017
Customers in tracking data	120,015	96,775	71,665	85,084
Customers with billing data	94,483	76,219	57,893	67,844
Customers with quality billing data	94,431	76,180	57,847	67,812
Customers with 12 months pre and post billing data	55,585	48,601	34,034	13,940
Customers in final analysis dataset*	51,334	45,118	31,517	12,638
Percent remaining	43%	47%	44%	15%

*These customers have weather data and savings for the analysis fuel.

Table 3-3. SCE – electric billing data attrition

Billing Data Attrition	2014	2015	2016	2017
Customers in tracking data	69,377	65,287	63,176	70,808
Customers with billing data	56,688	51,679	47,688	57,578
Customers with quality billing data	56,673	51,667	47,675	57,578
Customers with 12 months pre and post billing data	37,540	35,947	35,240	24,261
Customers in final analysis dataset*	36,691	35,403	34,777	24,062
Percent remaining	53%	54%	55%	34%

*These customers have weather data and savings for the analysis fuel.

Table 3-4. SDG&E – electric billing data attrition

Billing Data Attrition	2014	2015	2016	2017
Customers in tracking data	23,049	21,423	20,340	21,862
Customers with billing data	17,247	15,662	15,294	18,622
Customers with quality billing data	17,228	15,644	15,270	18,609
Customers with 12 months pre and post billing data	9,941	8,918	8,731	3,985
Customers in final analysis dataset*	9,532	8,458	8,485	3,712
Percent remaining	41%	39%	42%	17%

*These customers have weather data and savings for the analysis fuel.

Table 3-5. PG&E – gas billing data attrition

Billing Data Attrition	2014	2015	2016	2017
Customers in tracking data	106,746	86,640	62,702	72,604
Customers with billing data	86,867	69,963	50,209	69,756
Customers with quality billing data	86,726	69,825	49,816	69,078
Customers with 12 months pre and post billing data	51,286	41,855	28,804	12,543
Customers in final analysis dataset*	46,623	38,417	25,867	11,077
Percent remaining	44%	44%	41%	15%

*These customers have weather data and savings for the analysis fuel.

Table 3-6. SCG –billing data attrition

Billing Data Attrition	2014	2015	2016	2017
Customers in tracking data	93,630	71,112	65,576	82,271
Customers with billing data	76,411	63,759	57,815	73,869
Customers with quality billing data	76,401	63,747	57,810	73,866
Customers with 12 months pre and post billing data	56,923	28,022	43,187	30,397
Customers in final analysis dataset*	55,340	27,409	42,318	29,852
Percent remaining	59%	39%	65%	36%

*These customers have weather data and savings for the analysis fuel.

Table 3-7. SDG&E – gas billing data attrition

Billing Data Attrition	2014	2015	2016	2017
Customers in tracking data	12,851	12,336	11,894	13,332
Customers with billing data	9,603	8,838	8,799	11,695
Customers with quality billing data	9,583	8,825	8,774	11,667
Customers with 12 months pre and post billing data	5,677	5,305	4,849	2,247
Customers in final analysis dataset*	4,338	3,842	3,833	1,805
Percent remaining	34%	31%	32%	14%

*These customers have weather data and savings for the analysis fuel.

4 ANALYSIS METHODS

ESA is a comprehensive whole-house energy efficiency retrofit program. Thus, methods suitable for analyzing energy use changes in the entire home are required to isolate the effect of multiple interventions. A two-stage modeling approach with a comparison group is considered the best practice approach for estimating whole house savings.[32] The rationale for this method is discussed in Section 4.1.

The first stage involves site-level models that weather normalize household energy consumption. The site-level model is presented in Section 4.2. In the second stage, weather normalized household energy consumption is used in a pre-post analysis. The difference-in-difference models used in the second stage produce savings per household. A detailed discussion of this method is provided in Section 4.3.

Comparison groups are used in in the second-stage models to account for changes that are not caused by weather or the program such as economic conditions or changes in the number of people in a household. Comparison groups can be constructed in a variety of ways. In this evaluation, we construct a matched comparison group that provides a reasonable means of controlling for non-program related energy consumption trends. This method is discussed in Section 4.4.

ESA must install all feasible measures to improve the energy efficiency as well as the health, comfort, and safety of a home. In the face of the variety of measure mixes installed in ESA homes our approach to estimate such savings involve bundling measures into groups. The measure groups used to calculate measure level savings estimates as well as the estimation model are provided in Section 4.5.

Whole-house interventions affect both the rate at which households use energy (consumption measured in kWh for electricity, therms for gas) and energy use during a specified period (demand measured in kW). In this study, factors for converting kWh savings to demand (kW) savings are provided in Section 4.6.

As noted earlier, the evaluation was completed in two phases. In Phase 1 we reported draft results. Based on stakeholder discussions, three additional tasks were completed in Phase 2. Two measure bundles were split up (shell measures, other water heater measures), interaction terms between evaporative coolers and room and central ACs were added, and a decision rule to limit the effects of poor model fits at the site level was added. The additional analytical tasks are discussed in Section 4.7.

4.1 Choice of methodology

The evaluation approach DNV GL chose reflects 3 primary goals. First, the approach met the needs of the initial impact evaluation scope of work, which requested a routinized evaluation process that provides consistent savings estimates over time both at the whole-house and measure-levels. Second, the appropriate method needed to evaluate energy changes after a whole-house energy efficiency intervention. Finally, the approach needed consistency with accepted evaluation methodologies for analyzing the effect of intervention of the sort ESA involves.

[32] Ken Agnew and Mimi Goldberg. (2017) Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol. The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures. NREL/SR-7A40-68564, NREL; The Princeton Scorekeeping Method (PRISM®). Fels, Margaret. (1986). PRISM: An introduction. *Energy & Buildings* 9, 5–18.

These goals are related, and the method chosen to address them reflects this overlap. As a start, the methodology is consistent with the approach laid out in the Uniform Methods Project (UMP) Chapter 8 modeling approach, which provides whole-house savings estimation protocols for retrofit projects like ESA.[33] It is also consistent with the general approach of the International Performance Measurement and Verification Protocol (IPMVP) Option C for Whole Facility that addresses evaluation conditions applicable to whole-house retrofit interventions. The modeling approach is also closely related to all other forms of program analysis that use energy consumption data including time-series, cross-section approaches. Finally, it is also consistent with CalTRACK, the recent effort to develop agreed upon steps for the site-level modeling portion of the analysis.[34]

The approach we are using in this evaluation estimates savings from multi-measure or retrofit projects using consumption data from utility billing records. The method is suitable for analyzing the effect of the intervention at the program or program segment level but not for individual participating sites. It uses consumption data from relatively homogenous sites to provide savings estimates that are applicable to residential populations at the program level.

The analytical framework lends itself well for a developing a routine and replicable process that ESA can use to obtain savings estimates that are reliable and consistent over time. The whole-house savings estimates reflect variation in program implementation year by year, but aid in the comparison of results across years as well across programs.

4.2 Stage one: site-level modeling

The first stage site-level model correlates daily energy consumption with heating degree days (HDD) and cooling degree days (CDD). Based on PRISM,[35] this model is used to estimate each household's response to (1) outdoor temperatures, (2) the temperature points (base or balance points) that trigger cooling and heating, and (3) weather-adjusted consumption that reflects typical weather for each site. The outcome of this process is weather normalized energy consumption.

[33] Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol. The Uniform Methods Project.

[34] CalTRACK, http://www.caltrack.org/

[35] The Princeton Scorekeeping Method (PRISM®).

The site-level model is given by:

$$E_{im} = \mu_i + \beta_H H_{im}(\tau_H) + \beta_C C_{im}(\tau_C) + \varepsilon_{im}$$

Where:

E_{im}	Average electric (or gas) consumption per day for participant i during period m
μ_i	Base load usage (intercept) for participant i
$H_{im}(\tau_H)$	Heating degree-days (HDD) at the heating base temperature τ_H
$C_{im}(\tau_C)$	Cooling degree-days (CDD) at the cooling base temperature $\boldsymbol{\tau_C}$ (not included in gas models)
β_H	Heating coefficient determined by the regression
β_C	Cooling coefficient determined by the regression (not included in gas model)
τ_H	Heating base temperatures, determined by choice of the optimal regression
τ_C	Cooling base temperatures, determined by choice of the optimal regression
ε_{im}	Regression residual

Consumption is estimated over a range of 64°F to 80°F for cooling and 50°F to 70°F for heating to identify the temperature base points for each site (household); statistical tests identify the optimal set of base points. The outcome of the site-level model is parameters that indicate the level of baseload (consumption not correlated with either HDD or CDD) and the relationship between heating and cooling consumption and HDD and CDD, respectively.

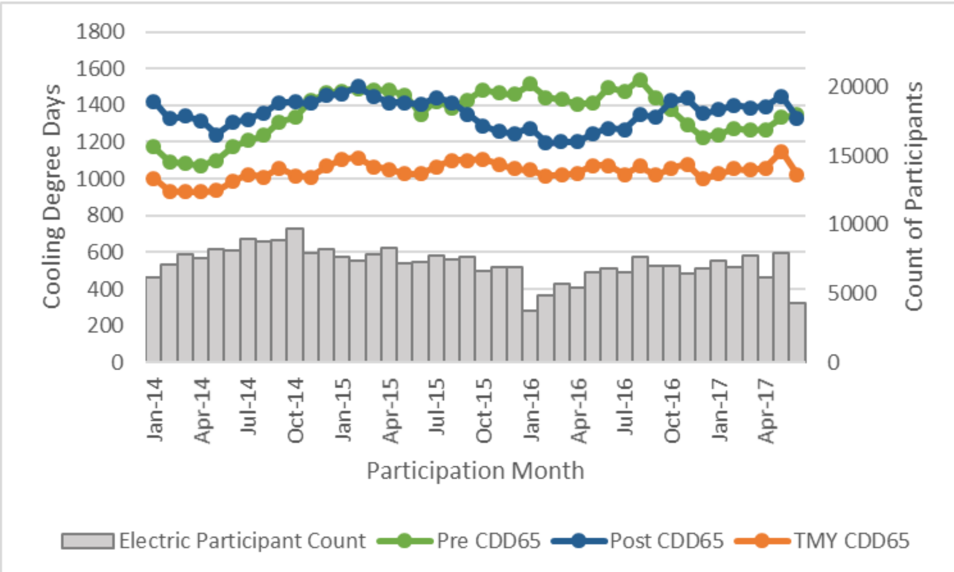
Model parameter estimates for each site allow the prediction of consumption under any weather condition. For evaluation purposes, all consumption is put on a typical weather basis called normalized annual consumption (NAC). NAC for the pre- and post-installation periods are calculated for each site and analysis time frame by combining the estimated coefficients $\hat{\beta}_H$, and $\hat{\beta}_C$ with the annual typical meteorological year (TMY) degree days H_0 and C_0 calculated at the site-specific degree-day base(s), $\hat{\tau}_H$ and $\hat{\tau}_C$. Normalized annual consumption is given by:

$$NAC_i = (365 \times \hat{\mu}_i) + \hat{\beta}_H H_0 + \hat{\beta}_C C_0$$

Individual household level regression models are estimated using observed weather data from the NOAA sites. Associated TMY data are used to weather normalize annual consumption using the estimated model parameters. The process serves two purposes; first, putting pre- and post-installation consumption on the same weather basis so that change in weather is not conflated with program effect, and, second, choosing a weather basis that represents a reasonable expectation of future weather for the ex ante projections.

Figure 4-1 summarizes how weather affected the analysis over time. All households are modeled on 12 months pre- and post-installation data. A substantial change in weather from the pre- to post- installation period increases the importance of the weather-normalization process to avoid conflating weather-related change (up or down) with program-related savings. The figure plots pre- and post-installation CDD separately, by participation month cohort. For example, the first month cohort that participated in January 2014, faced relatively cooler weather in the year prior to participation (2013) compared to the weather in the first year after participant (2014). Without weather-normalization and a comparison group, cooling related savings could be reduced or even removed altogether because of increased cooling occurring in the post-installation period for this cohort. The opposite effect, a potential misguided increase in savings, would have been a concern for much of 2016.

Figure 4-1 Observed CDD and TMY, by participation month



This figure illustrates how the pre-post weather differential varies over time. Prior ESA impact evaluations highlighted the challenge of variation in results, year over year. This figure illustrates one of the likely drivers of year to year variation. The figure indicates that if weather were not appropriately addressed, there was a potential for substantial upward trend in cooling-related savings from 2014 through 2016. This would occur because of a clear downward weather-related bias of results in 2014 and a clear upward bias of results in 2016.

The billing analysis methodology for this evaluation is designed to address this challenge in two ways. The household-level regressions model the interaction between consumption and degree days for each household and allow pre- and post-installation consumption to be adjusted to a common weather basis. The weather-normalization process removes the differential effect of weather as effectively as is possible with, in this case, monthly data.[36]

Figure 4-1 also illustrates the effects of normalizing evaluation results with CZ2010 TMY data. The lower TMY CDD values reduce cooling consumption and cooling savings by approximately 40%. This is the case for areas where the observed CDD values are higher than the TMY CDD where observed cooling load and savings correlated with cooling load are reduced (Figure 3-8).

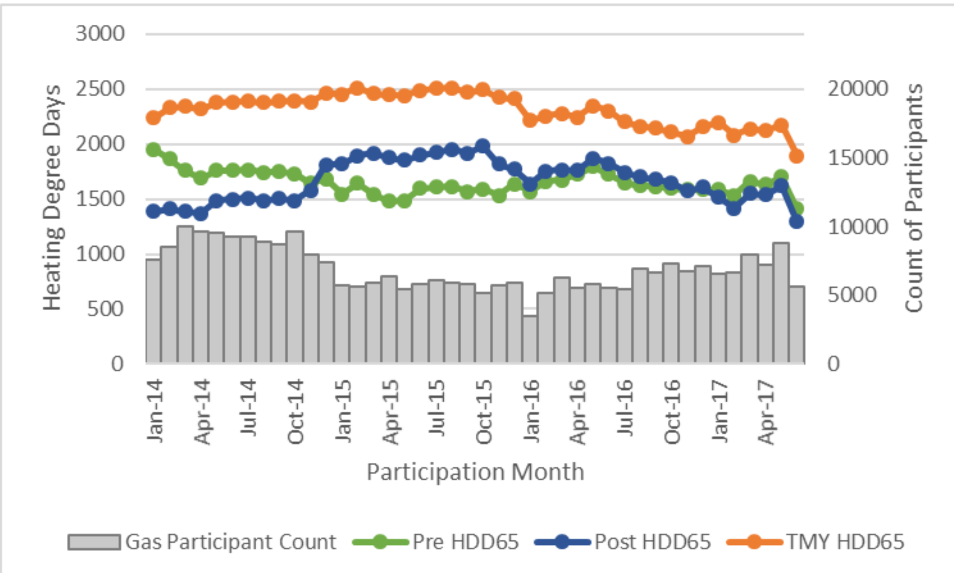
Figure 4-2 shows the pre- to post-installation HDD differentials by participation month. HDD differentials are even greater than CDD differentials on a percentage basis. For the January 2014 cohort, the post-

---

[36] The inclusion of the comparison group in the second stage regression should address any remaining weather-related differential not controlled for directly by the household models. We discuss comparison groups in Section 4.4. The comparison group households also control for other kinds of non-weather-related changes that may be occurring between pre- and post-installation periods. Some billing analysis approaches do not weather normalize individual households' energy use or do not include comparison groups. These approaches risk not fully addressing weather differentials and producing biased and variable savings estimates.

installation period HDD is 29% lower than the pre-installation period implying a similar reduction in heating consumption unrelated to the program installations. For 2015 participants, an uncontrolled weather differential would increase consumption and savings by 20%. The figure also illustrates the effects of normalizing with CZ2010 TMY data. The higher TMY HDD values increase heating consumption and savings. Observed heating load and savings correlated with heating load are increased by an average of 48% as they are put on a TMY basis.

Figure 4-2 Observed HDD and TMY, by participation month



4.3 Stage two: difference-in-difference model

Normalized annual consumption from site-level models form the basis for the second-stage of the analysis. A model based on the pre- and post-difference in NAC for participant households and a matched comparison group is estimated using a difference-in-difference modelling approach. This model is given by:

$$\Delta \text{NAC}_i = \alpha_i + \beta T_i + \varepsilon_i$$

In this model, i subscripts a household and T is a treatment indicator that is 1 for ESA households and 0 for the matched comparison homes. The effect of the program is captured by the coefficient estimate of the term associated with the treatment indicator, $\hat{\beta}$.

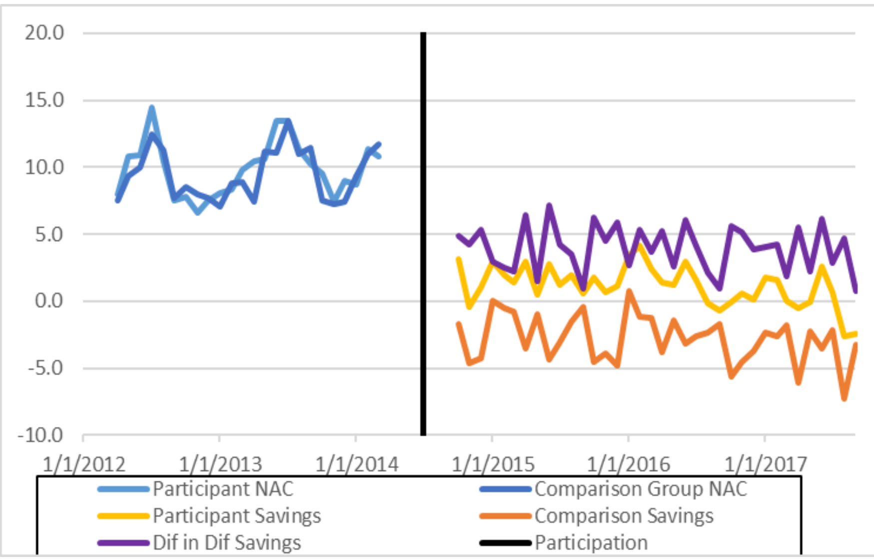
Pre- and post-program periods are based on a definition of a blackout period for each participant. Based on the CalTRACK recommendation and the IOU-provided tracking data, DNV GL defined a blackout period that reflects installation month(s) reported in the tracking data. Typically, the tracking data indicates a single installation date, though some sites have two or three installation months indicated.[37] These installation dates are used to define the blackout period. According to CalTRACK, an intervention period is a "time between the end of the baseline period and the beginning of the reporting period in which a project is being

[37] For each IOU, 99% of projects involved no more than 3 installation months. Projects involving more than three installation months were dropped from the analysis.

installed." It advises the use of "the earliest intervention date as project start date and the latest date as the project completion date."[38]

Figure 4-3 illustrates the second-stage difference-in-difference framework by estimating changes in consumption following participation in ESA. Two households with similar energy use (left panel) have changes in such use that is different in the post-intervention period (right panel). The difference or reduction in energy use in the pre- to post-intervention period is greater for the ESA treated household indicated by the light orange line. The difference in this pre-post difference between the participant and comparison groups (purple line) measures program savings.

Figure 4-3. Second-stage difference-in-difference



4.4 Comparison group development

The goal of any energy efficiency evaluation is to estimate change in energy use due to a program, while accounting for the effect of other changes in consumption, such as weather, income, and household characteristics. Weather normalization accounts for the effect weather has on consumption changes. After weather normalizing consumption, there remain two other possible explanations for pre-post differences: program-related savings or exogenous changes (non-program, non-weather changes in consumption). Exogenous changes may be driven by economic or other factors, but they occur across all customers not just program participants. For instance, if customers are coming out of a period of economic recession, an average two to three percent increase in consumption may occur across all customers. If this increase is not addressed, it will directly undermine true savings.

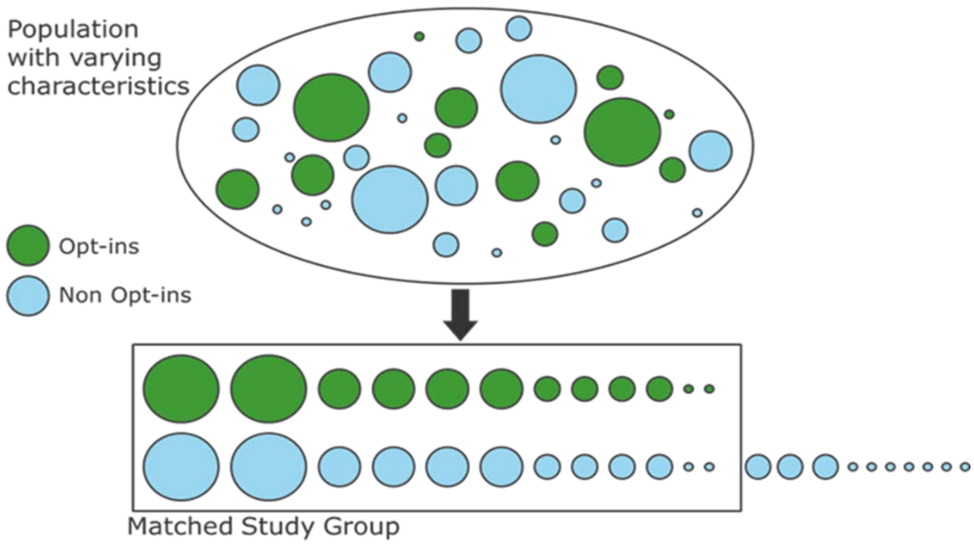
We control for the effect of these types of exogenous changes by using a comparison group. The comparison group is like the participant group except for program participation. Data from a comparison group is used, alongside participants' data in the difference-in-difference model to estimate program impact.

We use propensity score matching to construct a matched-comparison group. This approach is based on propensity scores that measure the probability that a household can be assigned to a program given its

[38] http://docs.caltrack.org/en/latest/methods.html#section-2-data-management

characteristics. An alternative way to think of propensity scores is as a metric that summarizes several dimensions of household characteristics into a single metric that can be used to group similar households. Propensity scores are used to match program participants with similar non-participant households. As Figure 4-4 illustrates, matching is a process of funneling population members with varying characteristics into a matched set that share similar traits. Further details on the matching methodology and results are provided in Appendix A.

Figure 4-4. Propensity score matching process



For this evaluation, matched-comparison households for participants are identified using information on all households' consumption levels and patterns. Household and other external characteristics that drive consumption habits are embedded in such data, which provide a readily available source of information that can be used to identify similar households. Matching is done using monthly billing data prior to any program start months for candidate participant groups and comparators. We would have preferred to also conduct matching within housing-type groups, but consistent indicators of housing type were not available for the eligible comparison group population. In this case, the embedded customer information was used to make the best feasible match.[39]

We use a pool of general CARE population customers that are not in ESA as candidate comparators. This kind of matching is made practical because the pool of eligible CARE customers is substantially large. In general, there were approximately 20 to 30 CARE comparison candidates for each ESA participant. ESA households are drawn from a population whose characteristics is like those on CARE, which makes CARE households an ideal comparison pool.

We match households using pre-participation consumption. Further, we stratify households into climate regions so that homes with similar consumption patterns from the same climate region are matched. For the analysis, we choose one matched-comparison household for each ESA program participant. The quality of

[39] In the specific instance of manufactured homes with master metered gas, an additional limitation occurred. The comparison group process requires common support. To the extent that master metered gas data are outside of the range of eligible comparison group customers, they will drop out of the analysis.

matches is tested using procedures discussed in the paper "Model Based Matching and Other Benefits of High Frequency Interval Data."[40]

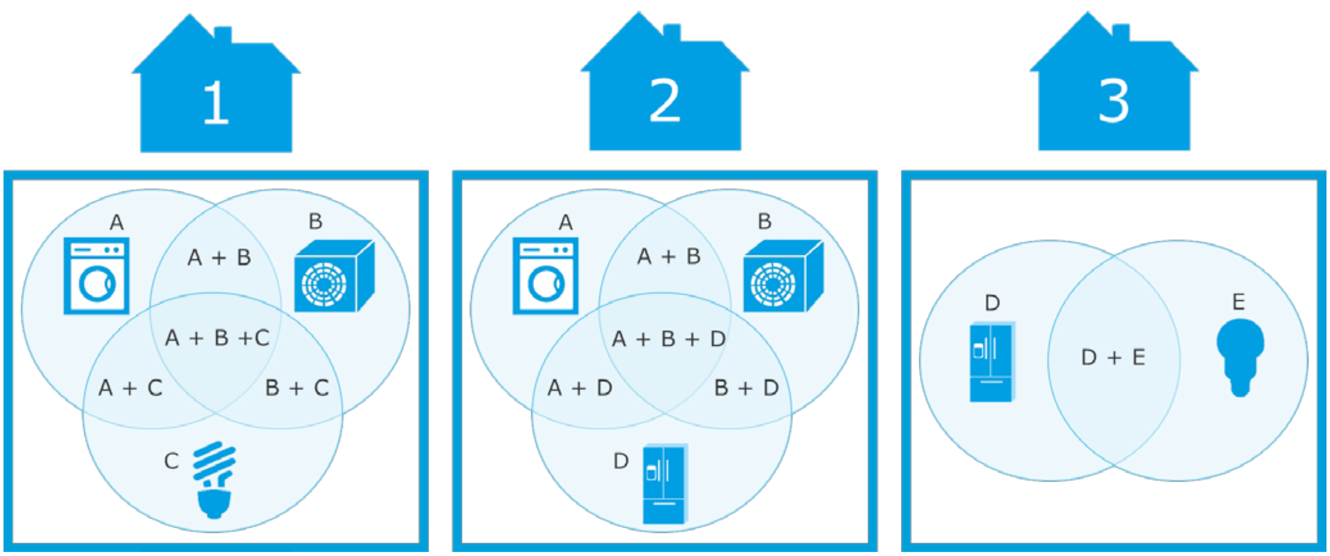
4.5 Measure-level models

The primary objective of this evaluation is producing sound and consistent site-level energy savings estimates of ESA efforts. Measure-level savings estimates can be useful to program administrators tracking the effectiveness of their program implementation over time and planning for future iterations of the program.

For the ESA program, with many unique measure bundles installed across the program population and climate region, estimating individual measure level savings is challenging. First, there is substantial variation in measure savings across populations and climate regions because of different usage patterns based on weather and housing type. In addition, measure-level savings will vary by measure bundle. The interactive effects that necessitate the use of site-level billing analysis also necessitate measure-bundle- savings to capture the interactive effects. The site-level savings estimates in a given year will reflect the program population from that IOU and year, with respect to geography and measure bundles that were offered. The savings will reflect the average savings across measure bundles.

Even with bundling, measure- or measure-bundle-level savings estimates are informative but have limitations due to the interactive effects mentioned above. The different measure scenarios presented in Figure 4-5 illustrates these challenges. Each house in the figure has a different measure bundle. Each house gets a more efficient appliance, cooling or lighting system; for example, the House 1 installed efficient cooling, washing machine and lighting measures.

Figure 4-5. Measure bundles and program savings



Estimated energy use reduction reflects the combined effect of the three measures and their interactions (the union of the circles), and not the sum of their individual effects (the sum of each individual circle).

[40] "Model Based Matching and Other Benefits of High Frequency Interval Data" Getachew et.al IEPEC 2017, Baltimore, Maryland

Further, as different measure mixes are installed across a given service territory both whole-house and measure-level energy savings reflect the interactive effects of a whole range of different measures mixes installed to effect change. Estimating the measure level effect of energy use under this scenario is challenging and requires measure groupings based on the frequency of measure combinations installed and the extent of their overlap in generating energy efficiency.

We consider these realities when developing measure bundles used to model measure-specific savings. Over the years, the ESA program has installed 33 electric and 16 gas measures. Table 4-1 provides a list of these measures. It also provides the measure groups or bundles that DNV GL, in consultation with the Commission's Energy Division and IOU staff, developed based on the different installed measures and their possible interactive effects. We estimate measure-level models based on the 20 measure-categories listed in the table.

Table 4-1. Measure bundles used for modeling

Measure Bundle	Electric Measures	Gas Measures
Air Sealing	Air Sealing/Envelope	Air Sealing/Envelope
Central AC	Central A/C Replacement	
Duct Testing & Sealing	Duct Testing and Sealing	Duct Testing and Sealing
Evaporative Coolers	Evaporative Coolers Replacement and Installation	
Furnace Repair/ Replacements		FAU Standing Pilot Light Conversion and Furnace Repair/Replacement
HE Clothes Washers	High Efficiency Clothes Washers	High Efficiency Clothes Washers
Heat Pump	Heat Pump Replacement	
Insulation	Attic Insulation	Attic Insulation
Lighting	Compact Fluorescent Lights (CFL), Exterior Hardwired CFL fixtures, Exterior Hardwired LED fixtures, Interior Hardwired CFL Fixtures, Interior Hardwired LED fixtures, LED A-Lamps, LED Diffuse Bulbs (60W Replacement), LED Night Lights, LED Reflector Bulbs, Torchieres - CFL, Torchieres - LED, and Vacancy Sensors	
Maintenance - ACs	Central A/C Tune-up	
Maintenance - Furnaces		Furnace Clean and Tune-up
Microwaves	Microwaves	Microwaves
Other Hot Water	Faucet Aerators, Low Flow Showerheads, Thermostatic Shower Valves, and Tub Diverter/ Tub Spouts	Faucet Aerators, Low Flow Showerheads, Thermostatic Shower Valves, and Thermostatic Tub Spouts
Pool Pumps	Pool Pumps	
Refrigerators	Refrigerators	
Room AC Replacements	Room A/C Replacement	
Smart Fan Delays	Energy Efficient Fan Control	Energy Efficient Fan Control
Smart Power Strips	Smart Power Strips - Tier 2 and Smart Power Strips - Tier 1	
Tank and Pipe Insulation	Water Heater Blankets and Water Heater Pipe Insulation	Water Heater Blankets and Water Heater Pipe Insulation
Water Heater Repair/ Replacements		Water Heater Repair/Replacement

Measure-level savings for the bundles developed are estimated based on the following model specification:

$$\Delta NAC_{jm} = \alpha + \Sigma_k \gamma_k T_{jk} + \varepsilon_{jm}$$

where:

ΔNAC_j	=	Change in NAC for customer *j* for post period m
T_{jk}	=	An indicator variable that is 1 if customer *j* is a participant installing measure bundle k, 0 otherwise
α, γ_k	=	Coefficients determined by the regression
ε_j	=	Regression residual

Savings for measure bundle k is given by:

$S_k = \gamma_k$ = Savings for measure bundle k

This regression specification is applied to each year of program data. The model gives results consistent with single-year whole-house models. This modeling framework makes the routine application of the approach to additional years of data possible and straightforward while maintaining analytical consistency.

4.6 Demand impacts

The impact evaluation methodology described above does not provide methods to estimate demand impacts (kW) for electric measures. The IOUs use peak demand impacts (kW) as part of the cost effectiveness testing for program delivery. It is our recommendation that ESA program administrators continue to use energy-to-demand conversion factors to estimate kW savings. We recommend using the conversion factors from the tracking data for existing measures and from the DEER database for any future measures added to the program. The conversion factors we recommend for current kW saving estimates are taken from the 2017 ESA program tracking data and are provided in Appendix C; DNV GL cross-checked these values with the 2011 impact evaluation when inconsistencies were found.

In the process of developing these conversion factors, we observed that PG&E is using inflated kW savings for new LED measures in 2017.

4.7 Additional analytical tasks

The analysis of Phase 1 results highlighted undertakings that could be more useful for the IOUs or make findings more robust. Our analysis of the contribution of measures to overall total savings indicated that the other water heating and shell measure bundles were each responsible for substantial portions of savings. Both bundles were comprised of measures with divergent expected savings. The other water heating bundle include water heater tank insulation blankets and pipe insulation along with low-flow showerheads. We separated the insulation-oriented measures because they are a purely physical, non-behavioral measure. Similarly, the shell measure-bundle combined attic insulation with air-sealing. The latter is much more frequently performed but insulation was expected to produce greater savings. Splitting these measure bundles provided more targeted savings estimates for these measures and improved the quality of the overall model because they better control for the variation in savings.

SCE was interested in understanding the interactive effects of evaporative coolers with room and central air conditioners. By adding an indicator variable for households where both measures were installed, the interactive effect is made explicit. This change provides a better estimate of savings for the room and central ACs.

We also investigated options for assessing the quality of first stage models and then addressing the potential bias issues therein. First, it is important to understand that the consistent treatment of the comparison group limits the negative effects of first stage modeling issues on the results. The root causes of poor models should be evenly distributed across the participant and comparison groups and thus the issues in one group counteract the possible effects of the other. We focused on a concern rooted in the limitation of billing analysis on monthly data. Sometimes there is enough of a heating or cooling signal to include a heating and/or cooling slope in the model, but the overall model does not fit the data well. This problem is particularly a concern when the poorly determined slope is steep. Under this condition, the process of weather normalization can dramatically change modeled consumption levels. We created a decision rule that flagged households where modeled consumption increased by more than 50% over actual consumption and the model fit was not strong (an R-square < 0.8). In these cases, we reverted to a baseload only model, rather than removing the household. The baseload only model is effectively actual load and removes the possibility of weather-normalizing these sites. This adjustment affected less than 3% of household for any of the IOUs or fuels. This process corrected implausible and anomalous increases in normalized annual consumption relative to observed consumption reducing the underlying variation in the models.

5 IMPACT ESTIMATES

This section provides electric and gas savings estimates at the household, measure, and program level. Section 5.1 provides estimates of electric savings per household and compares expected and ex ante savings. Additional details of savings per household by housing type and climate zone are presented to shed light on the possible sources and variations of these savings across IOUs. Measure-level results broken down by whole-house level estimates savings per household are also presented. Analogous discussion for gas savings is presented in Section 5.2. Aggregate program savings are presented in Section 5.3.

5.1 Electric impact estimates

5.1.1 Whole house

Annual electric whole house savings estimates for the three electric IOUs are provided in Table 5-1. The table provides the count of participants with electric savings in each year used in the analysis. Values for 2017 reflect ESA program activity for the first half of the year only. Like the trends noted at the overall program level in Section 3.2, the analysis data indicates that SCE and SDG&E decreased the size of their programs by 5–10% from 2014 to 2016. PG&E's program participation showed a significant drop of 40% from 2014 to 2016.

Table 5-1. Electric modeled and ex ante savings per household over time

IOU	Year	N	Average Savings (kWh)		Savings as % of Ex Ante
			Modeled	Ex Ante	
PG&E	2014	51,334	149	390	38%
	2015	45,118	125	341	37%
	2016	31,517	90	381	24%
	2017	12,638	131	495	26%
SCE	2014	36,691	277	408	68%
	2015	35,403	261	366	71%
	2016	34,777	238	356	67%
	2017	24,062	187	352	53%
SDG&E	2014	9,532	79	379	21%
	2015	8,458	48	184	26%
	2016	8,485	67	169	39%
	2017	3,712	30	162	18%

The whole-house electric results are relatively stable for all IOUs between 2014 and 2017. Savings estimates are statistically significant, except for the SDG&E 2017 estimate.[41] The three IOUs deliver savings at three distinct levels, as seen in Figure 5-1 and Figure 5-2. SCE's savings steadily decreased from 277 kWh to 187 kWh between 2014 and 2017. PG&E's savings decreased from 149 kWh to 90 kWh through 2016 with a

[41] The 2017 electric savings estimate is not statistically significant.

rebound in 2017. SDG&E's savings decreased from 79 kWh to 30 kWh with greater variability across the years. Figure 5-2 provides modeled electric savings as a percent of average household consumption.

Figure 5-1. Electric savings per household and percent of ex ante over time

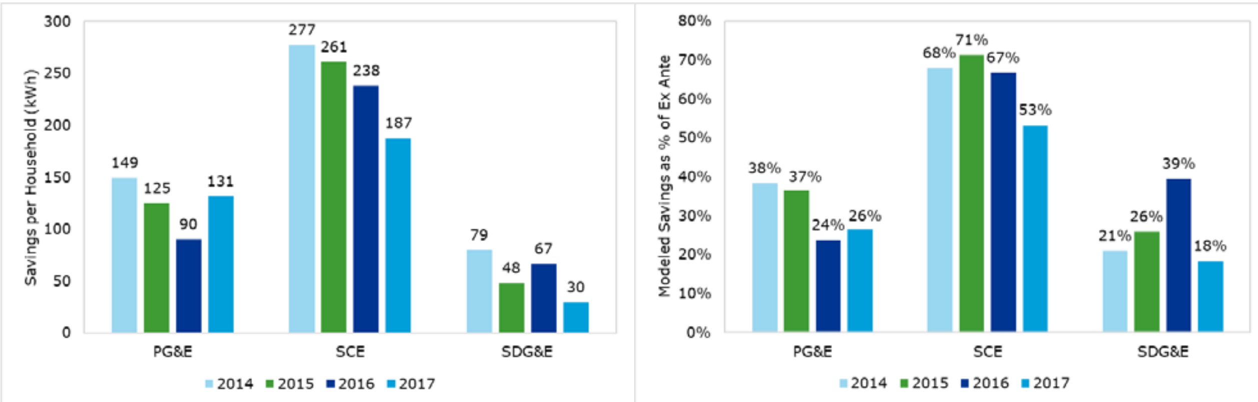


Figure 5-2. Electric savings as a percent of household consumption

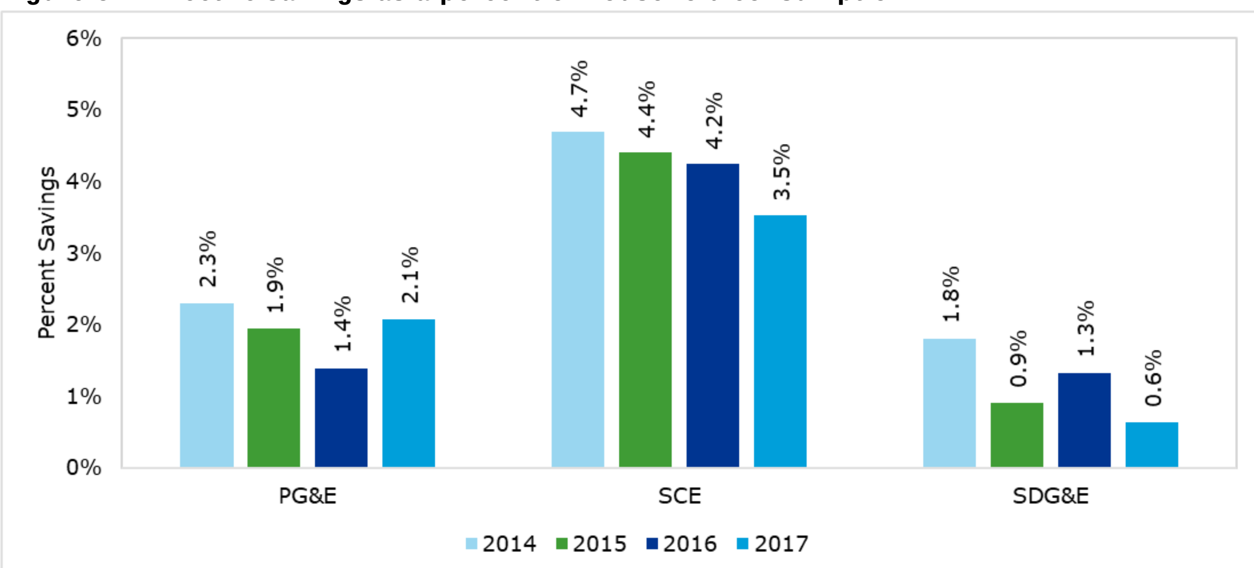


Table 5-1 also provides the expected savings from the tracking system as well as the savings as a percent of expected savings (also illustrated in Figure 5-1). PG&E and SCE expected savings average about 400 kWh per household over the four years. On the other hand, SDG&E was expected to save half as much on average over the four years. In 2014, SDG&E's expected savings were about the same as the expected average savings of PG&E and SCE, but the expected savings fell 50% over the remaining three years. The consistent patterns across Figure 5-1 and Figure 5-2**.** indicate that changes across years are primarily explained by changes in savings magnitude as opposed to changes in either expected savings or average household consumption.

On average, SCE achieved around two-thirds of its expected savings, PG&E realized a third, and SDG&E a quarter of the expected savings over the four years program (Figure 5-1). Realized savings relative to expected savings vary with SCE's values ranging from about 50% to 70%. PG&E's savings relative to expected savings are below 38%. SDG&E's savings rate are close to PG&E's, despite lower savings, because SDG&E's expected savings dropped in the last three years.

The results presented here differ from the Phase 1 results because of additional modeling. The analysis described in Section 4.7 tested modifications to the original models to improve the quality and applicability of the results. The additional steps included changes to site-level model inclusion rules that were expected to an effect on whole house savings results. For PG&E and SCE, the changes caused savings to increase less than 5% in three of the four years. SDG&E, with its smaller program population and lower savings estimates, saw the largest change resulting in decreased savings. Unexpectedly, SDG&E 2015 savings changed from statistically significant prior to the changes to non-significant after. This change is caused by a substantially lower point estimate indicating that outliers may have artificially inflated the original results.

Looking at savings by housing type offers clues for the average electric whole-house savings estimates. The residential population that the ESA program targets reside in single family, multi-family, and manufactured homes. Figure 5-3 illustrates the mix of housing types targeted by each IOU based on the data used in the analysis.[42] Single family houses constitute most residences receiving measure installations by the program for all IOUs. However, the share of single-family housing in the program is higher for PG&E and SCE at three-quarters of the total. The two IOUs also have similar housing type shares for the remaining two categories. However, SDG&E has a much larger share of multifamily participants.

Figure 5-3. Percent of participants by housing type in the electric analysis dataset

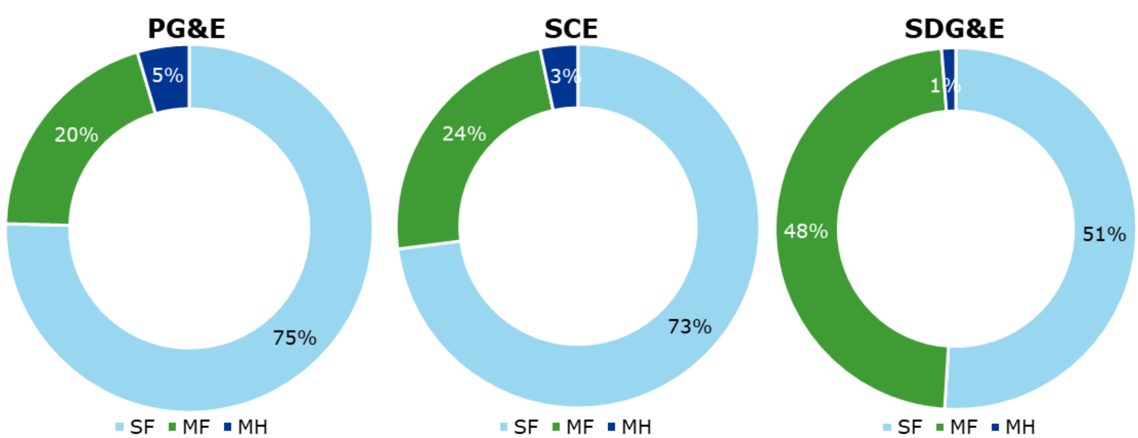
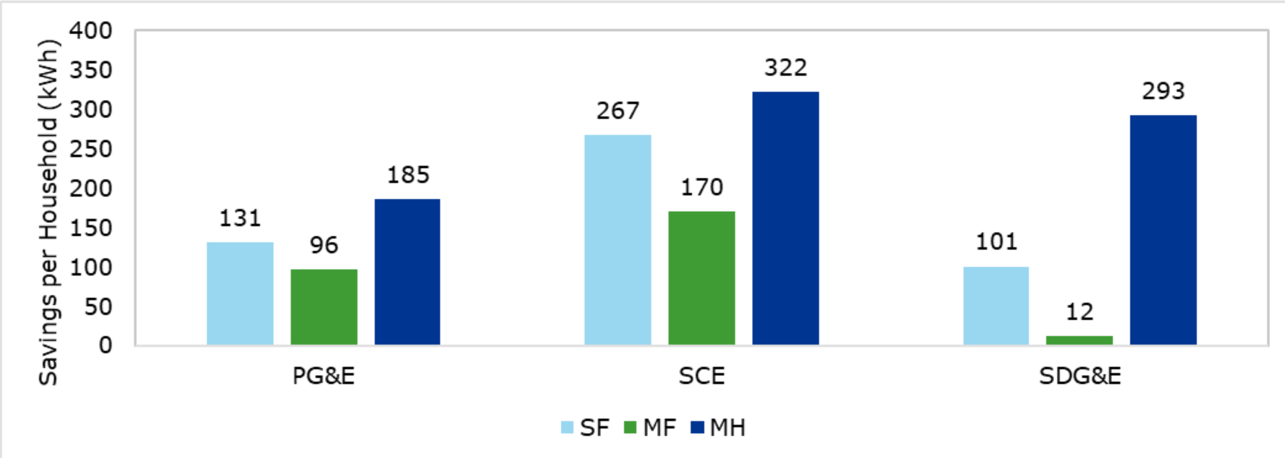


Figure 5-4 provides averages electric savings estimates by housing type for each IOU. The levels of savings reflect the distinct patterns noted earlier, but also indicate electric savings variation by housing type. All three IOUs achieve their highest savings in manufactured homes followed by single family homes. While savings per household in multifamily homes is the lowest among the three housing types, electric savings per household for SDG&E's multifamily participants is particularly low. This may drive lower overall savings noted in Table 5-1 and Figure 5-1.

---

[42] Overall program data indicates that the breakdown of housing type aligns with the analysis sample.

Figure 5-4. Electric savings by housing type



In addition to savings by housing type in the ESA program, an examination of the baseload, heating and cooling components of savings can be suggestive of what the measures installed by the program are achieving. Such an examination revealed a consistent finding across the years by IOU. Figure 5-5 provides an example from 2016. For PG&E, savings per household are primarily due to baseload savings with limited amount of cooling and virtually no heating load savings. This is consistent with the mix of measures installed at the utilities. PG&E primarily installs measures that would provide baseload savings (consistent savers at the measure level include refrigerators and lighting) with limited installation of measures that would show up as cooling savings (insulation, AC maintenance). In contrast, SCE installs AC and evaporative cooler measures that demonstrate clear cooling related savings. SDG&E installs baseload measures and appears to get electric heating savings from the weather correlated measures such as insulation, air sealing and duct sealing.

Consistent with the measure mixes, the savings have a locational dimension that explain the type of electric savings. SCE's service territory covers some of the hotter inland climate zones that probably facilitate cooling load reduction. SDG&E's ESA population is in more mild coastal areas and has a substantially higher incidence of electric heating. The analogous figures for the remaining years are presented in Appendix B.

Figure 5-5. Electric savings components, 2016

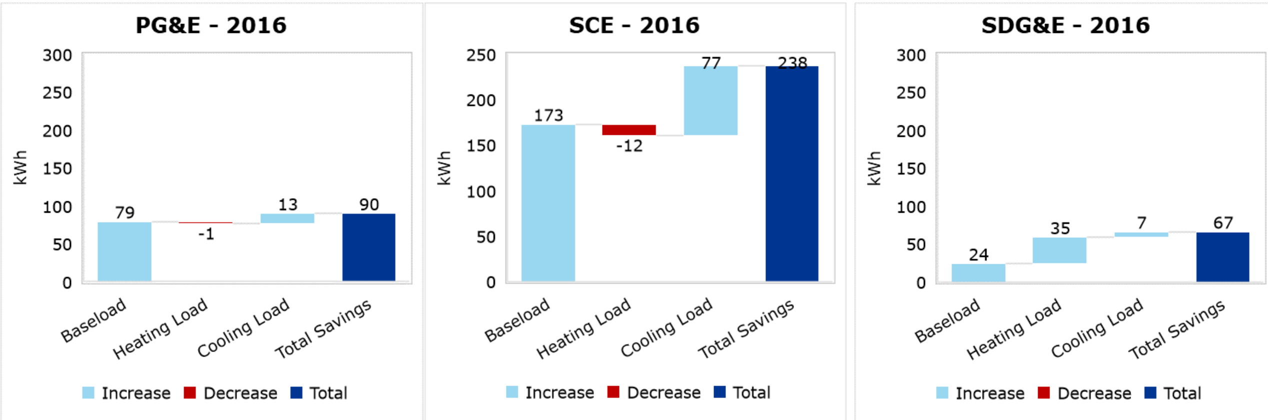
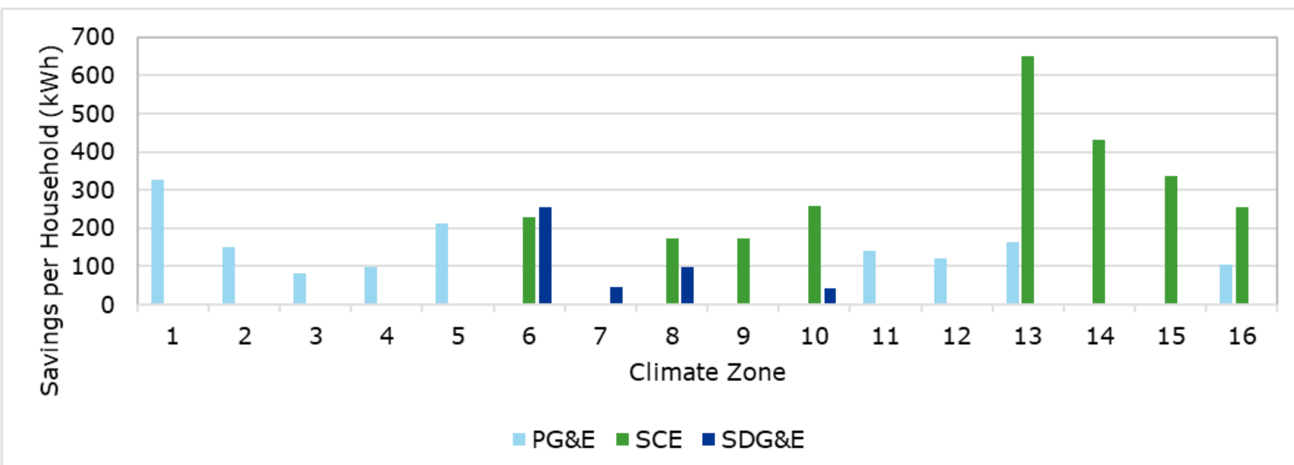


Figure 5-6 presents the pattern of electric savings estimates per household by climate zone. PG&E serves customers both in the mild climate zones of 1-5 and inland climate zones of 11-13 and 16. Despite the inland climate zones having higher temperatures, the reductions in energy use per household are not higher than in the mild regions. This is consistent with PG&E decreased focus on cooling-related measures. SCE's focus on cooling-related measures is evident in the high savings in climate zones 13-16. SDG&E's program is concentrated in the milder climate zones.

Figure 5-6. Electric savings by climate zone



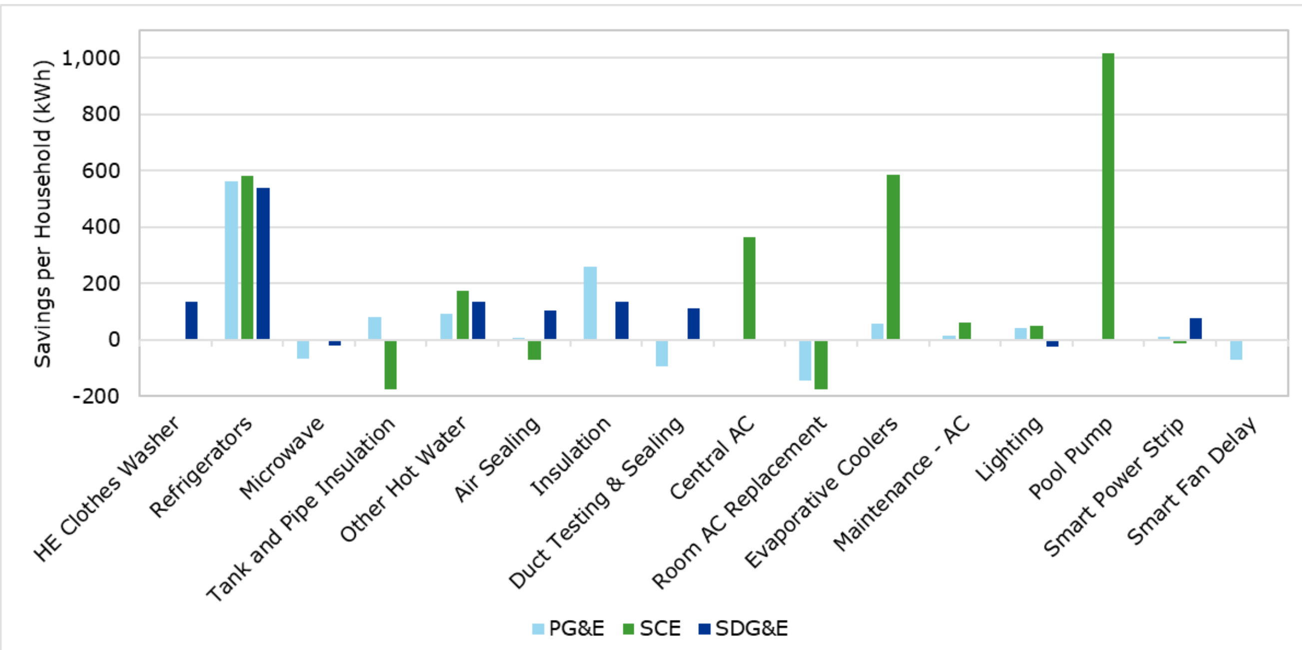
Note: Only estimates with 300 participants or more are reported here.

5.1.2 Measure level

Estimated measure-level electric savings per household can be grouped into three general categories. First is a group of measures that provide high savings of more than 300 kWh per household including refrigerators, central AC, evaporative coolers, and pool pumps (Figure 5-7). Figure 5-8 also indicates realization rates for this group of measures, with the exception of pool pumps, is high indicating that they achieve and even exceed their expected savings. Second is group of measures that provide moderate savings of between 100 and 300 kWh per households including high efficiency clothes washers, other hot water measure and insulation. Third, there are a group of measures that provide moderately low savings per household of less than 100 kWh that include lighting, smart power strips and AC maintenance.

SCE installs all the large savers, including refrigerators, central AC, evaporative coolers, and pool pumps. Its program also offers moderate saving measures, a combination that could explain why it experiences notable baseload and cooling load reductions. PG&E installs some of the large energy saving measures including refrigerators and evaporative coolers, but PG&E installs half as many evaporative coolers as SCE and they only yield an average reduction of less than 100 kWh per household in PG&E territory. PG&E also installs moderate energy saving measures such as lighting. The combination of lighting and refrigerators and hot water measures likely drives the predominance of baseload savings for PG&E ESA participants. The only high savers that SDG&E installs are refrigerators and households in its ESA programs are also estimated to get moderate electricity savings from smart power strips, insulation, and air and duct sealing measures. The higher prevalence of electric heat in SDG&E territory combined with the air and duct sealing measures explains SDG&E's heating load savings.

Figure 5-7. Electric measure-level savings per household



Note: Measure groups with fewer than 100 participants are not displayed.

At the measure level, any conclusions drawn are nuanced because of the difficulty of getting clear signals for many measures that have small expected savings and are installed widely across diverse measure bundles. When focusing on measures with high precision, many of the key high savings measures appear to have ex antes that are too high (Figure 5-8). Refrigerators (all IOUs), pool pumps (SCE) and lighting (PG&E, SCE) are all important and consistent performers but not at the level of ex ante expectations. SDG&E's refrigerator ex ante values are lower than the other IOUs' for all years and they report the only annual savings greater than ex ante expectations. On the other hand, some other measures appear to perform better than expected. SCE cooling measures, CAC, heat pump and evaporative coolers, all perform substantially better than ex ante. The PG&E insulation measure performs at or above its ex ante value in most years. As an example of where measure-level results can be misleading, SDG&E air sealing measure appears to be highly statistically significant and producing savings at three times the level of expectation. In this case, these results appear to be an outgrowth of collinearity with lighting, another widely installed measure. SDG&E lighting savings are small or even negative and the lower savings results line up with the clearly inflated air sealing results.

Figure 5-8. Electric measure-level savings as percent of ex ante across IOUs

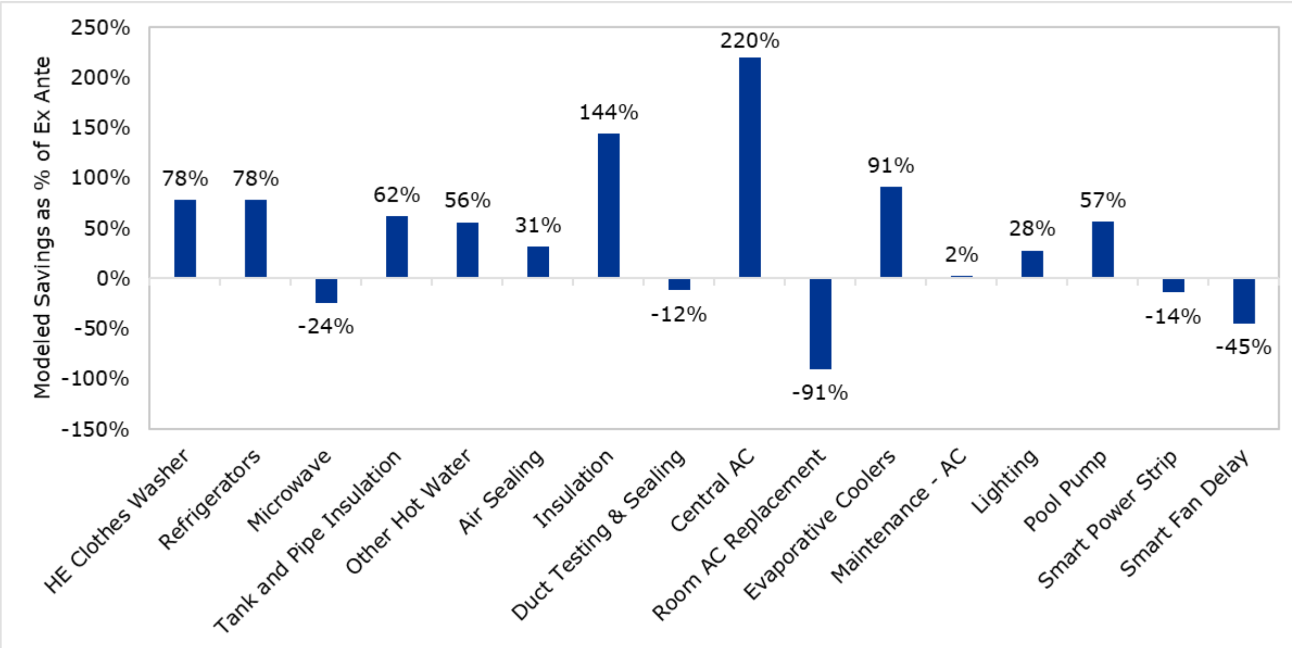
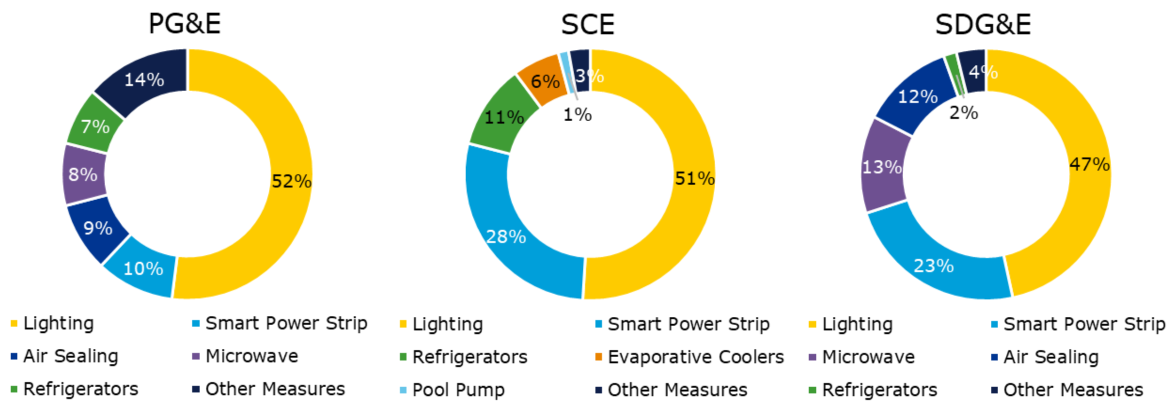


Figure 5-9 provides the 5 most frequently installed electric measures by the IOUs. In addition to the measure-level savings just discussed, these commonly installed measures yield clues for the patterns of savings seen among ESA households in each IOU's service territory. PG&E and SDG&E installed that same set of top five measures. With the exception of refrigerators, these measures provide modest savings. PG&E installed a higher percentage of refrigerators and lighting than SDG&E while installing fewer microwaves, a negative saver. These differences could explain PG&E higher savings levels than SDG&E. In contrast, three of SCE's 5 most frequently installed measures were high savings measures - refrigerators, evaporative coolers, and pool pumps. In addition, SCE did not install any microwaves. Fully 18% of the measures installed by SCE produce substantial savings as indicated by the measure-level savings estimates. The frequency of the installation of high saving measures explains why SCE's site-level sayings are the highest among the IOUs.

Figure 5-9. Top 5 most frequent electric measures by IOU



5.1.3 Results of additional measure-level analysis

The additional analysis task prioritized two areas that affected the measure-level results. In prior results, air-sealing and insulation were combined in a single measure bundle referred to as enclosure. These two measures are shell-related and quite different in their installation pattern as well as their expected savings. Approximately 15% of PG&E households claimed electric savings for air sealing as a measure, whereas insulation was installed in less than 5% of homes. A quarter of SDG&E participants received air sealing for electric savings but just a few hundred homes were insulated. Expected savings are at least double for insulation compared to air sealing.

Splitting the enclosure measure bundle makes it clear that insulation is a strong measure for both SDG&E and PG&E. Insulation savings are strong relative to ex ante savings. Air sealing alone does not perform well for PG&E, generally providing negative savings, but does much better for SDG&E. There is evidence that the air sealing savings for SDG&E are related to the greater prevalence of electric heating among SDG&E participants.

The other measure that was split was domestic hot water other. The measure includes pipe and tank insulation measures along with various water-flow related measures. Like the enclosure measure, the tank and pipe insulation measures have different expected savings and are installed much less frequently. In this case, the other hot water measure bundle has higher expected savings than the pipe and tank insulation measure because many households receive multiple water-flow related measures. PG&E is the primary installer of pipe and tank insulation. The savings for the tank and pipe insulation measures and water-flow measures stayed positive and at similar savings levels across all years.

A separate additional analysis task included interaction terms for combined installations of evaporative coolers with either central ACs or room ACs. SCE installs the largest number of these measures. The inclusion of the interaction term increases the savings of evaporative coolers and CAC measures. The interactions terms are primarily negative indicating that when evaporative coolers are installed along with either of the other measures, the combined savings is less than the sum of the two measures individually. Room ACs savings remain negative but are less negative with the interaction term.

5.2 Gas whole-house impact estimates

5.2.1 Whole house

Figure 5-10 presents the average estimated savings per household and the percent of ex ante savings by year. Values for 2017 reflect only the first half of the year. Savings estimates for PG&E, SCG, and SDG&E in 2014 are statistically significant, while estimates for SDG&E in 2015 through 2017 are not statistically significant. The whole-house gas results for PG&E and SCG are relatively stable across years while SDG&E's gas savings decreased substantially from 2014 levels and then remained steady from 2015 to 2017. SCG savings ranged from 6 therms to 8 therms. PG&E's savings ranged from 7 to 9 therms. SDG&E's savings ranged from 3 therms to 5 therms. Figure 5-11 provides modeled gas savings as a percent of average household consumption.

Figure 5-10. Gas savings per household and percent of ex ante over time

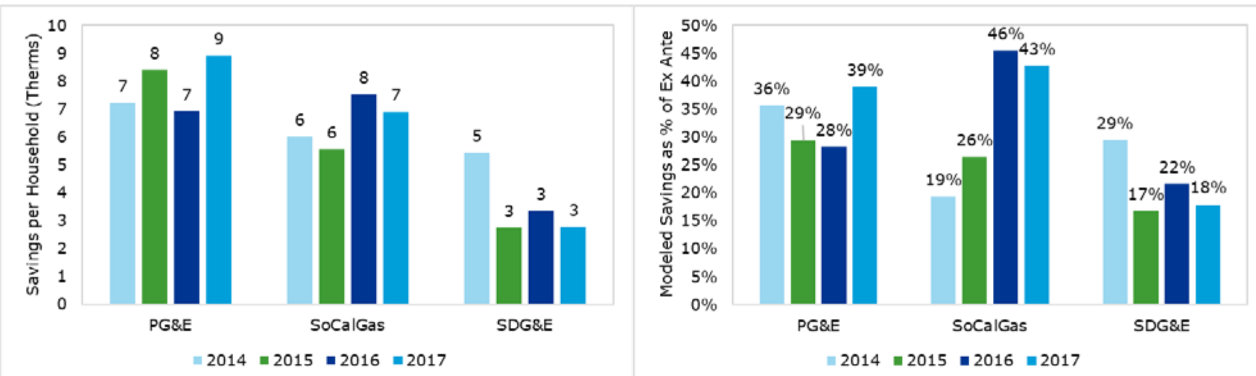


Figure 5-11. Gas savings as a percent of household consumption

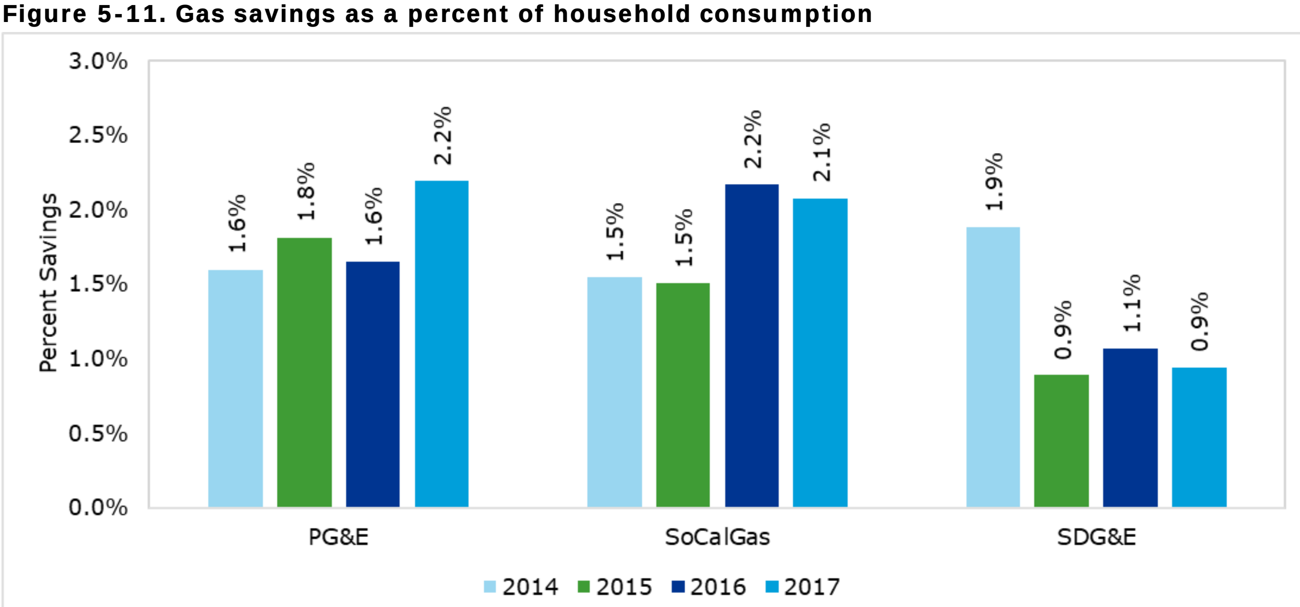


Table 5-2 presents the count of participants with gas savings used in the analysis, the savings per household, the ex ante savings, and the modeled percent of ex ante savings. PG&E's expected savings (ex ante) increased in later years while SCG's expected savings dropped by almost half. SDG&E's expected savings decrease moderately. PG&E's savings as a percent of expected savings fluctuated around 30%. SCG's steady savings drive improved realization rates with the decrease in expected savings over time. SCG's savings as a percent of expected savings increased from around 20% in 2014 and 2015 to over 40% in 2016 and 2017. SDG&E's reduction in estimated savings is more precipitous than the decrease in expected savings driving lower realization rates. SDG&E's savings as a percent of expected savings decreased from nearly 30% in 2014 to approximately 20% from 2015 to 2017.

Table 5-2. Gas modeled and ex ante savings per household over time

IOU	Year	N	Average Savings (Therms)		Savings as % of Ex-Ante
			Modeled	Ex-Ante	
PG&E	2014	46,623	7	20	36%
	2015	38,417	8	29	29%
	2016	25,867	7	24	28%
	2017	11,077	9	23	39%
SCG	2014	55,340	6	31	19%
	2015	27,409	6	21	26%
	2016	42,318	8	16	46%
	2017	29,852	7	16	43%
SDG&E	2014	4,338	5	18	29%
	2015	3,842	3	16	17%
	2016	3,833	3	15	22%
	2017	1,805	3	15	18%

The results presented here differ from the Phase 1 results because of a final in-depth modeling task. These analysis, discussed in Section 4.7, tested multiple modifications to the original models with the intent of improving the quality and applicability of the results. The additional steps included changes to site-level model inclusion decision rules that were expected to have some effect on whole house results. The modeling changes did not affect the results for PG&E. SCG's estimates for 2015 and 2016 increased by 1 therm each, while the estimates for 2014 and 2017 did not change. SDG&E, with its smaller program populations and smaller savings estimates, changed more than the other IOUs and savings mostly decreased. For SDG&E, savings estimates for 2014 through 2016 decreased by 1 therm, while their 2017 estimate increased by 1 therm. Additionally, the 2015 estimate for SDG&E flipped from statistically significant prior to the changes to non-significant after. The consistent patterns across Figure 5-10 and Figure 5-11 indicate that changes across years are primarily explained by changes in savings magnitude as opposed to changes in either expected savings or average household consumption.

Like electric savings, examination of savings by housing type may shed light as to the patterns observed at the whole-house level. Figure 5-12 indicates that gas savings per household are similar across housing types. Multifamily homes contribute the most to gas savings followed closely by savings among single family homes. Manufactured homes have small or no savings.

Figure 5-12. Gas savings by housing type

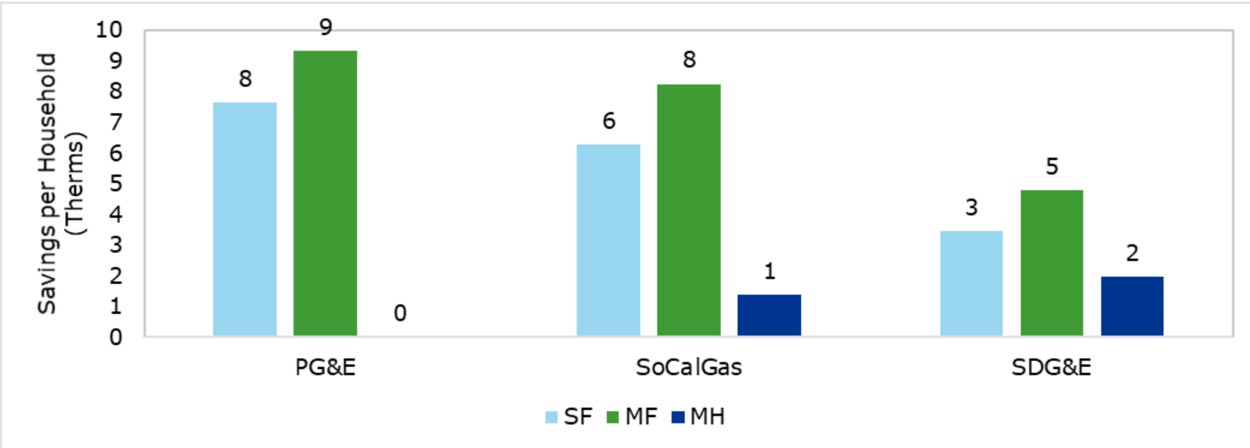
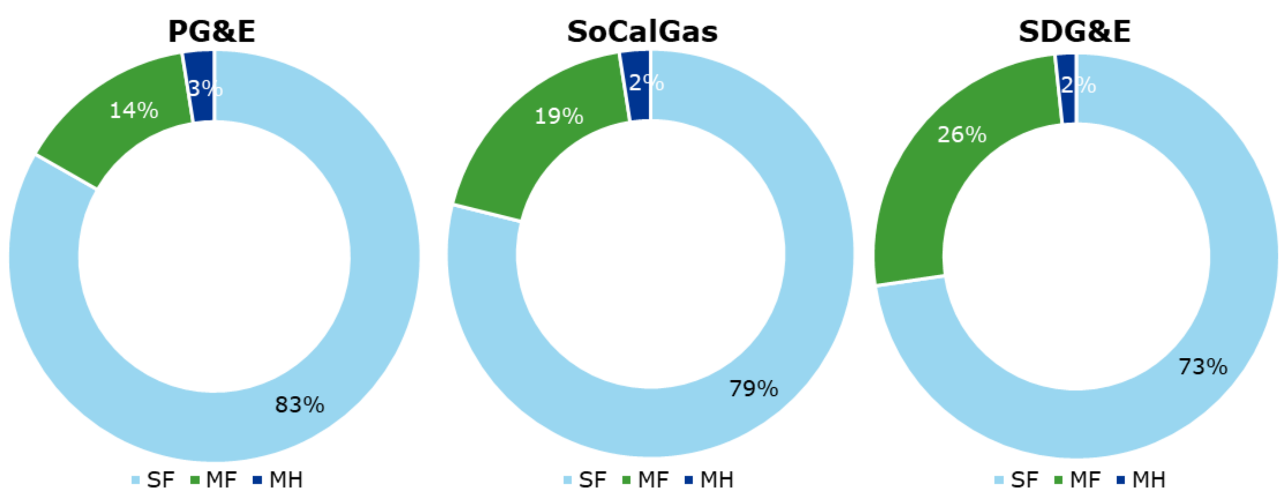


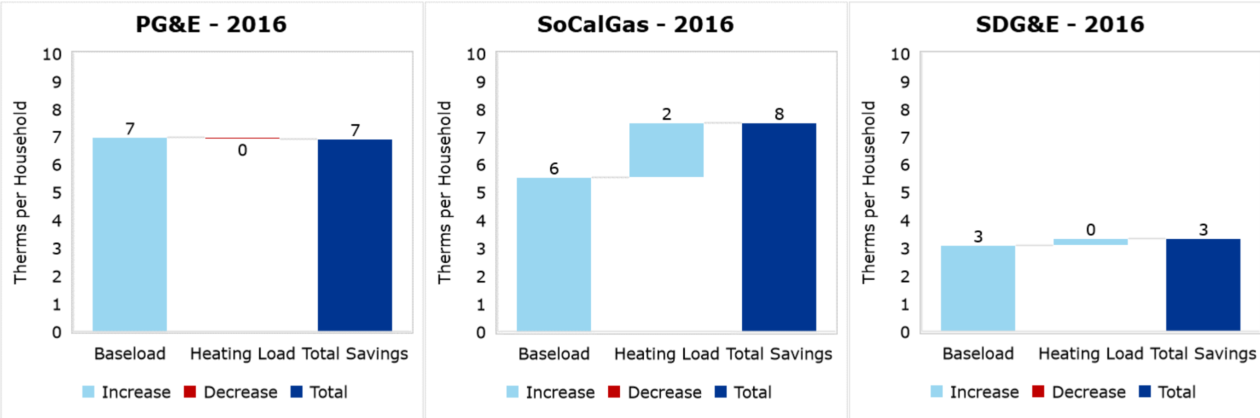
Figure 5-13 indicates that most common housing type receiving gas measures is single family houses followed by multifamily. Manufactured homes constitute the lowest proportions of housing type so the effect of low savings from this segment on overall savings is limited.

Figure 5-13. Percent of participants by housing type in the gas analysis



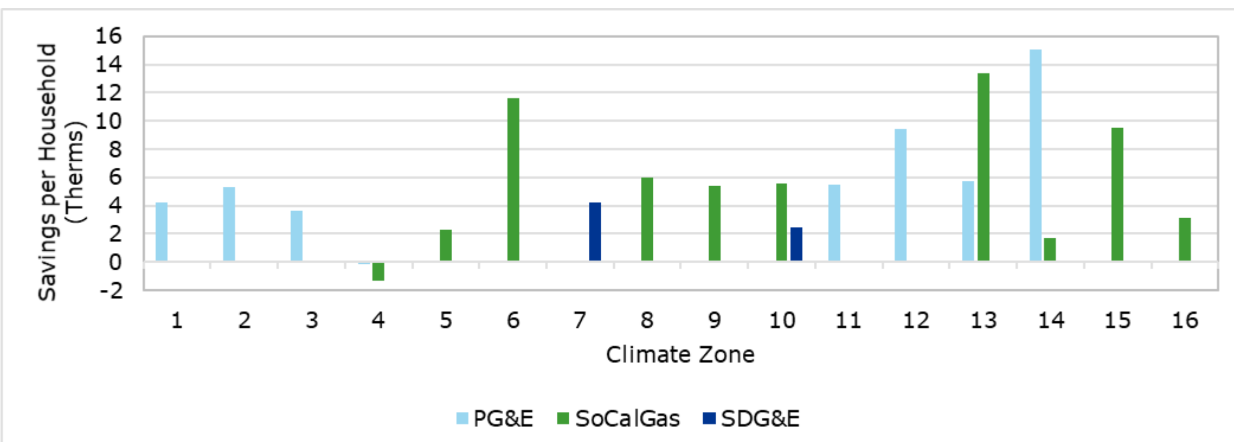
The components of gas savings per household indicate that most gas savings are due to baseload reductions for PG&E and SDG&E. Baseload savings are likely driven by water heating savings. Figure 5-14 indicates the 2016 estimates; the figures for all the years under consideration are in Appendix BAppendix BAppendix BAppendix BAppendix B. The figure indicates SCG's gas savings are also mostly due to baseload reduction, but there is also some heating load reduction for ESA homes in SCG's service territory. SCG installed more insulation, with its high savings, twice as frequently as PG&E, which probably accounts for the heat load savings noted.

Figure 5-14. Gas savings components, 2016



Estimates by climate zone indicate that most of PG&E's gas savings are among households located in homes in the central valley, particularly in climate zones 12 and 14 (Figure 5-15). Gas savings per household is lower among homes in the costal climate zones of 1, 2, and 3; the latter includes the San Francisco Bay Area. SCG does not have a regional trend in gas savings, but the highest gas savings per household is for houses located on the coast (climate zone 6), inland in the central valley climate zone 13 and among homes located in the desert climate zone of 15. SDG&E's savings are moderately higher for homes on the coast (climate zone 7) than inland (climate zone 10).

Figure 5-15. Gas savings by climate zone



5.2.2 Measure level

There are no estimated gas savings per household from HVAC related measures (Figure 5-16). Furnace repair and replace and furnace maintenance savings are negative. Shell measures, insulation, duct sealing and air sealing, on the other hand, provide gas savings per household. Insulation, particularly, provides the highest gas savings per household of between 20 and 50 therms. These point to shell measures as being the probable source of heating load reductions. The remaining measures are related to baseload reductions and offer mid-range gas savings per household ranging from 2 to 10 therms for the three hot water measures to 10 to 20 therms for high efficiency clothes washers.

Figure 5-16. Gas measure level savings per household

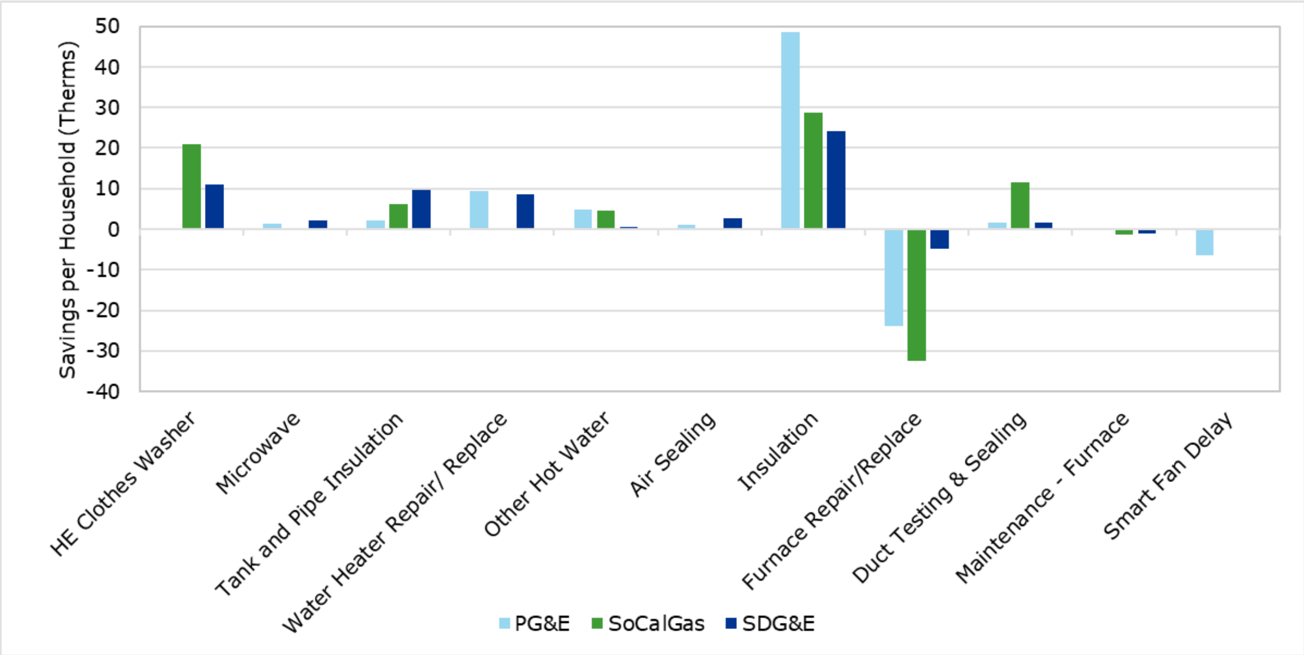
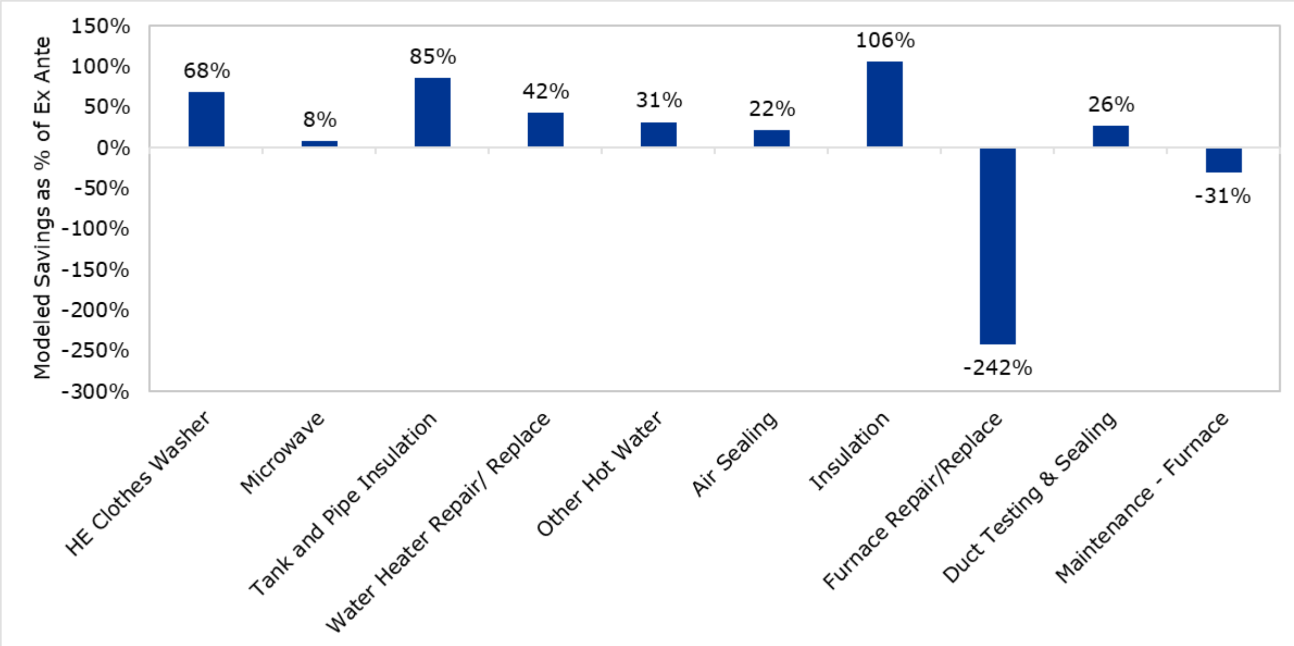


Figure 5-17 indicates that measures achieve 8% to 106% of their expected savings. Of the shell measures, only insulation achieves slightly more than its expected savings while the rest do not perform as well as expected. The HVAC measures produce negative savings.

Figure 5-17. Gas measure-level savings as percent of ex ante

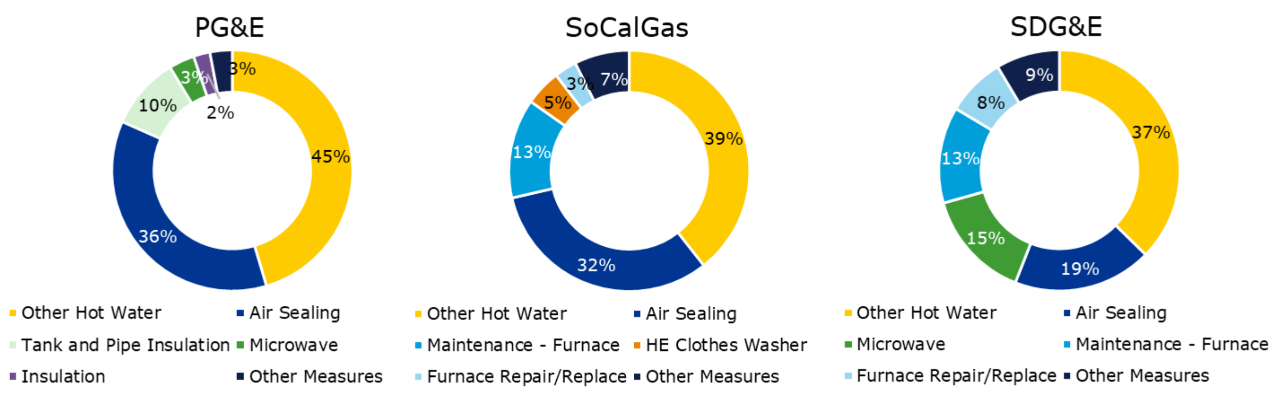


Note: We did not include smart fan delay in this figure as no ex ante savings were provided.

The percentage of most commonly installed measures explain why gas savings per households are mostly tied to baseload reduction. Homes in PG&E's service territory have baseload gas reductions because most measures installed (hot water savings and tank and pipe insulation) are baseload related (Figure 5-18). Even though insulation provides a remarkable amount of gas savings, it makes up only 2% of installed measures and is installed in only 5% of homes. We note, however, temperatures during the evaluation timeframe were mild compared to historical trends, reducing the potential for heating load reductions.

The most commonly installed measure by SDG&E (other hot water) similarly provides baseload gas savings. SDG&E also installs measures that either do not provide much gas savings (microwaves) or are estimated to increase gas use (HVAC related). The latter may improve comfort and result in an increase in gas use due to possible pent up demand (takeback). SCG also offers measures likely to affect baseload gas use. It also offers air sealing, which could be the source of the heating load reductions.

Figure 5-18. The five most frequent gas measures by IOU



5.2.3 Results of additional measure-level analysis task

The additional analysis task discussed in Section 5.1.3 primarily affected the gas measure-level results. Air-sealing and insulation, though both shell-related, are quite different in their installation pattern as well as their expected savings. For gas participants, over three quarters of SCG and PG&E households receive air sealing measures, whereas insulation is installed in roughly 5% of homes. Slightly under half of SDG&E gas participants receive air sealing measures but only a few hundred homes are insulated. Expected savings are at least double for insulation compared to air sealing. As with the electric results, separating these measures produces substantial and statistically significant savings for insulation, while air sealing results are mixed.

The other split measure was “domestic hot water other,” which includes pipe and tank insulation along with various water-flow related measures. Like the “enclosure” measure, the “pipe and tank insulation” measures have different expected savings and are installed less frequently. Almost all gas participants receive multiple water-flow related measures, grouped in the other hot water measure bundle, which has higher expected savings than the pipe and tank insulation measure. PG&E installs pipe and tank insulation in roughly 20% of homes and SCG installs this measure in about 5% of homes. Both of the split measures produce statistically significant savings for both PG&E and SCG. The water-flow savings are consistent across the PG&E and SCG but are lower and less consistent for SDG&E. SCG’s pipe and tank insulation savings are higher than PG&E’s.

5.3 Aggregate program impacts

This section shows the total evaluated electric (kWh) and gas (therms) program savings. As described in Section 3.5, the analysis is based on a subset of total program participants for whom there is sufficient data. Total program savings are based on modeled average household savings for each program year projected to the population of participating households. Table 5-3 shows the total program electric savings by year for each electric utility. Evaluated program savings are fractions of the program savings expected by the IOUs (presented in the tracking summary in Table 3-1). For the 2015–2017 period, PG&E expected savings of 116 GWh compared to estimated program reduction of 29 GWh. SCE expected program savings of 87 GWh and achieved an estimated reduction of 45 GWh. SDG&E’s smaller treatment population was expected to generate 11 GWh of savings over the three-year period and achieved a reduction of 3 GWh.

Table 5-3. Electric program savings by IOU and year

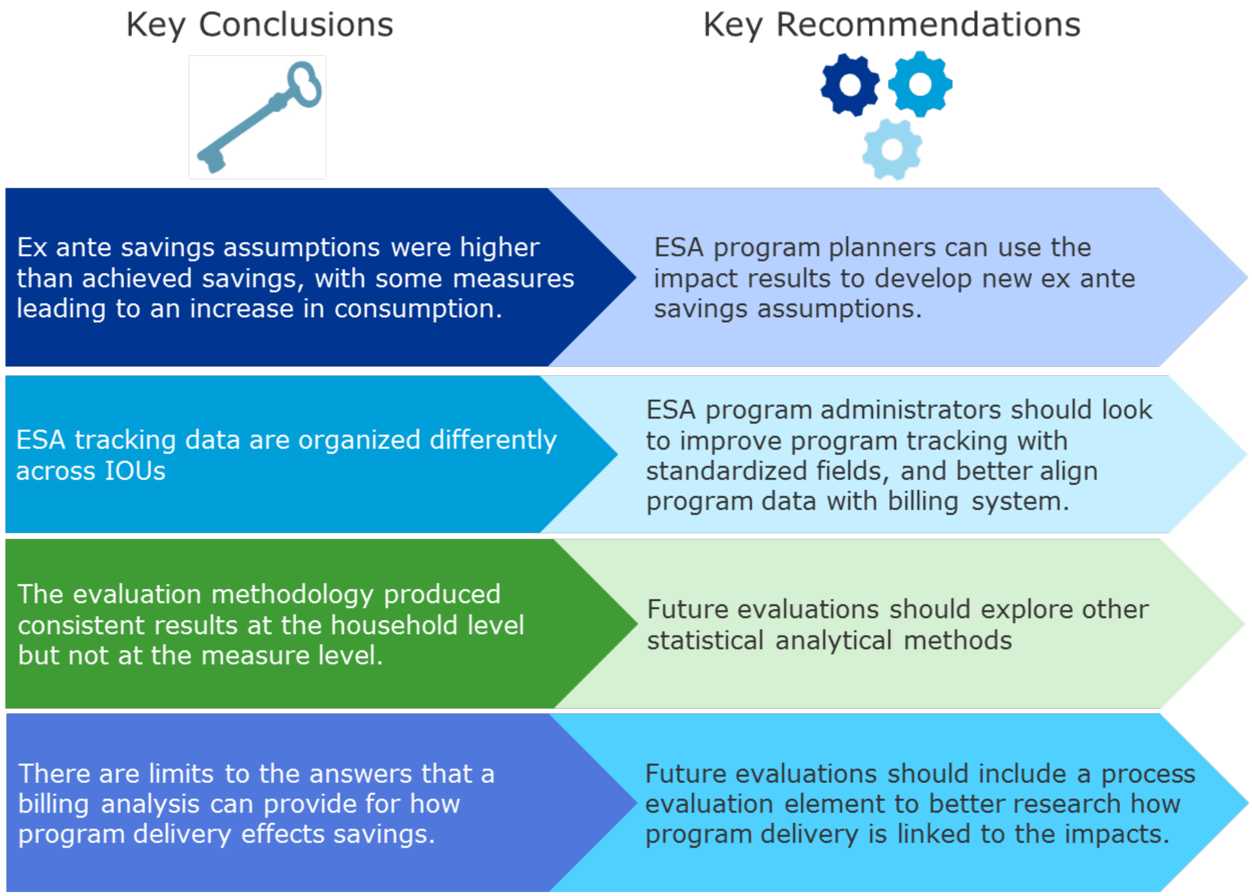
IOU	Year	Program Participant Households	Modeled Average Savings (kWh)	Total Program Savings (kWh)
PG&E	2015	96,878	125	12,071,906
	2016	71,705	90	6,461,277
	2017	83,272	131	10,917,437
SCE	2015	65,287	261	17,030,502
	2016	63,169	238	15,010,613
	2017	70,828	187	13,262,244
SDG&E	2015	21,413	48	1,018,996
	2016	20,325	67	1,353,092
	2017	21,620	30	640,831

Table 5-4 shows the total program gas savings by year for each gas utility. Total program gas savings are similarly lower compared expected savings. PG&E expected total gas savings of 5 million therms over three years (2015–2017) but achieved only 2 million therms. SCG expected gas program of 4 million therms but only achieved gas savings of 1.5 million therms over the same three-years. SDG&E expected total gas savings of about 600,000 therms from 2015 to 2017 and achieved gas savings of 187,000 therms.

Table 5-4. Gas program savings by IOU and year

IOU	Year	Program Participant Households	Modeled Average Savings (Therms)	Total Program Savings (Therms)
PG&E	2015	96,878	8	813,684
	2016	71,705	7	495,552
	2017	83,272	9	741,534
SCG	2015	71,202	6	395,017
	2016	65,600	8	492,448
	2017	82,271	7	567,476
SDG&E	2015	21,413	3	58,811
	2016	20,325	3	67,953
	2017	21,620	3	59,877

6 CONCLUSIONS AND RECOMMENDATIONS



6.1 ESA program results and implications for future years

The ex ante savings assumptions for ESA program years 2015–2017 were higher than the achieved savings found in this impact evaluation. At the household level, ESA program expected savings were overestimated by each IOU. Measure level expected savings are similarly high for most measures with the general exception of insulation and cooling related measures. One explanation for this outcome is inflated savings estimates in prior evaluations (Section 2.2) that were subsequently used as the planning assumptions for the program years 2015–2017. Other possible explanations for achieved savings below expected savings include possible issues related to program design and implementation practices. This evaluation did not include a process evaluation and, as a result, does not have data with which to assess the potential role of these issues.[43] Our evaluation indicates that some measures may lead to an increase in consumption. Program administrators understand some measures are justified on health and safety grounds and not

[43] Related recommendations are discussed in Section 6.3.

energy efficiency. To the extent that these measures still have positive ex ante savings assumptions, the consumption-increasing potential may not have been factored in sufficiently.

- **Recommendation 1**: ESA Program planners should use the results of this impact evaluation to develop new ex ante savings assumptions for measures that roll up to reasonable household level savings.
- **Recommendation 2**: ESA program planners should fully account for potential consumption-increase assumptions for measures that are installed for non-energy related benefits. This would include, for instance, flagging fixes to heating or cooling units where the unit was not working or not used prior to the visit. This will segregate off installations that increased consumption and improve overall program savings projections.

6.2 ESA program tracking data issues

The evaluation team encountered challenges cleaning the ESA program tracking data and ensuring that the tracking data aligns with CPUC reported values. Verifying the accuracy of the tracking data will improve future analyses. The verification should confirm that the tracking data matches what the IOUs report to the CPUC. Preparing the tracking data took considerable time and effort that could have been spent on the analysis. We provide a summary of challenges and suggested solutions for future evaluations below work:

- Tracking data should include ex ante unit and total measure savings; tracking data submissions for two IOUs did not include ex ante savings for reported measures. Attaching the supplemental information provided later was challenging
- Tracking data should provide counts for each measure that roll up to total measure quantities installed at each site, and ultimately to annually reported counts
- IOUs should ensure the number of households reported in the tracking data reflect the totals in the annual values
- Tracking data should provide measure categories and measure names that match the annual reporting categories
- If possible, IOUs should provide instructions on how tracking data counts and savings should be aggregated; one challenge in this evaluation involved determining how measures with units of “each” or “home” should be rolled up to obtain total counts
- In addition to participant housing type, tracking data should include climate zone and zip code

Matched comparison groups within housing type could not be created due to inconsistent definitions of MF between the ESA program and customer information tables.

- **Recommendation 3**: All relevant identifiers in the program tracking fields should be standardized such that information readily rolls up to program totals and matches the values reported to the CPUC.
- **Recommendation 4**: Program definitions and requirements should be aligned with billing information.

6.3 Future ESA evaluations

This evaluation developed a routinized and replicable evaluation methodology for the ESA program. The impact methodology applied in this evaluation is consistent with methods proposed for “California Pay for Performance” third party programs and can be adapted for advanced meter infrastructure (AMI) data. This is an ideal method for future evaluations to ensure consistent analysis methods across evaluation timeframes.

However, there are there are several research questions that could be addressed in future evaluations by using AMI data and conducting a process evaluation that could potentially yield useful information that informs the implementation of ESA programs in the future.

While measure level savings information is important for planning purposes, aggregated measure-level results must be consistent with household level savings. That is, a weighted sum of the measure-level savings must add up to the household level savings. Zero and negative savings cannot be ignored without adjusting positive savings estimates.

- **Recommendation 5**: Future evaluations should replicate the two-stage analysis approach followed here and expand the billing analysis from monthly consumption data to daily data using the available AMI data. AMI data will improve the quality of weather-normalization in the first stage modeling reducing variability and improving the quality of results in the second-stage model. Any new methods or approaches proposed for future evaluations should be required to replicate results using this approach and demonstrate the relation to and improvements relative to the current approach.
- **Recommendation 6**: Future impact evaluations should include a process evaluation of program delivery mechanisms to inform future impact evaluations.

7 APPENDICES

7.1 Appendix A: Matching results

7.1.1 Additional matching details

We provide further detail on the matching algorithm we use as well as matching results in this section. He Propensity score matching method used here follows the steps listed below:

- Select households' characteristics that are related to program participation
- Examine the distribution of these characteristics and exclude observations of the comparison group that do not overlap with those of participants' as a first round of identifying common support for matching
- Fit a logistic regression using these variables to estimate the probability of program participation
- Conduct a second round of trimming or common support identification based on propensity scores
- Select a matching method, the number of comparators in the many-to-one matching, and whether to match with or without replacement; match participant households' scores to comparison households based on these selections
- Conduct diagnostic checks to see selected matches are well-balanced

We match participant households in each IOUs service territory using monthly electric and gas use prior to program implementation. We also use climate zone information to stratify the data for matching. This involves implementing the matching procedure within three climate zone groups defined as inland, desert, and mild. Table 7-1 illustrates the climate zones used in the analysis and the distribution of participant households across the climate zone groups by IOU.

Table 7-1. Climate zones used in stratified matching

Climate Zone Group	Title 24 Climate Zones	Percent Program Participants in Each Climate Group			
		PG&E	SCE	SCG	SDG&E
Mild/Coastal	1,2,3,4,5,6,7,16	36%	8%	9%	55%
Inland	8,9,10,11,12,12,14	64%	88%	85%	45%
Desert	15	0%	4%	6%	0%

7.1.2 Matching results

We use two metrics, standardized difference of the mean and the ratio of the variance of matched-comparison and participant households, to check that the selected matches are well-balanced and appropriate for analysis. The mean and the variance fully characterize the distribution of consumption among the two groups, and the two metrics provide a good indication of the condition of balance. A standardized difference value that exceeds 0.2 or 20% indicates a significant imbalance as does a variance ratio that is 2 or more, or less than 0.5. Values of standardized means differences that are close to 0 and ratios that are close to 1 indicate well-matched samples.

Table 7-2 to Table 7-7 provide the value of the metrics for normalized monthly energy consumption before and after matching participant households to matched-comparison households.[44] We note an imbalance prior to matching with the value of standardized difference means equaling up to 0.2. All matched datasets, on the other hand, have a value of zero. The variance of the ratios of the two groups, which are also close to 1, indicate good balance post matching.

Table 7-2. SCE electric matched-comparison balance test

Month	Unmatched		Matched	
	Standardized Difference (D)	Variance Ratio (R)	Standardized Difference (D)	Variance Ratio (R)
1	0.1	0.9	0.0	1.0
2	0.1	0.9	0.0	1.0
3	0.1	0.9	0.0	1.0
4	0.1	0.9	0.0	1.0
5	0.1	0.9	0.0	1.0
6	0.1	0.9	0.0	1.0
7	0.1	0.9	0.0	1.0
8	0.1	0.9	0.0	1.0
9	0.1	0.9	0.0	1.0
10	0.1	0.9	0.0	1.0
11	0.1	0.9	0.0	1.0
12	0.1	0.9	0.0	1.0

Table 7-3. SCG gas matched-comparison balance test

Month	Unmatched		Matched	
	Standardized Difference (D)	Variance Ratio (R)	Standardized Difference (D)	Variance Ratio (R)
1	0.1	0.9	0.0	1.0
2	0.1	0.9	0.0	1.0
3	0.1	0.9	0.0	1.0
4	0.1	0.9	0.0	1.0
5	0.0	1.0	0.0	1.0
6	0.0	1.1	0.0	1.0
7	0.1	1.1	0.0	1.0
8	0.0	1.1	0.0	1.0
9	0.0	1.1	0.0	1.0
10	0.0	1.0	0.0	1.0

[44] Matching results for data from Phase 2 of the study (based on program years 2016 and 2017) is presented in this section. Phase 1 of the study is based on data from 2014 until mid-2016 and matched data for this phase indicates the same good balance shown here.

Month	Unmatched		Matched	
	Standardized Difference (D)	Variance Ratio (R)	Standardized Difference (D)	Variance Ratio (R)
11	0.1	0.9	0.0	1.0
12	0.1	0.9	0.0	1.0

Table 7-4. SDG&E electric matched-comparison balance test

Month	Unmatched		Matched	
	Standardized Difference (D)	Variance Ratio (R)	Standardized Difference (D)	Variance Ratio (R)
1	0.0	1.0	0.0	0.9
2	0.0	1.1	0.0	0.9
3	0.0	1.0	0.0	0.9
4	0.0	1.1	0.0	0.9
5	0.1	1.1	0.0	0.9
6	0.1	1.1	0.0	0.9
7	0.0	1.0	0.0	0.9
8	0.0	1.0	0.0	0.9
9	0.0	1.0	0.0	0.9
10	0.1	1.0	0.0	0.9
11	0.0	1.0	0.0	0.9
12	0.0	1.0	0.0	0.9

Table 7-5. SDG&E gas matched-comparison balance test

Month	Unmatched		Matched	
	Standardized Difference (D)	Variance Ratio (R)	Standardized Difference (D)	Variance Ratio (R)
1	0.0	0.9	0.0	1.0
2	0.0	1.0	0.0	1.0
3	0.0	1.0	0.0	1.0
4	0.0	1.0	0.0	1.0
5	0.0	1.0	0.0	1.0
6	0.1	1.1	0.0	1.0
7	0.1	1.1	0.0	1.0
8	0.1	1.1	0.0	1.0
9	0.1	1.1	0.0	1.0
10	0.1	1.1	0.0	1.0
11	0.0	0.9	0.0	0.9
12	0.0	0.9	0.0	0.9

Table 7-6. PG&E electric matched-comparison balance test

Month	Unmatched		Matched	
	Standardized Difference (D)	Variance Ratio (R)	Standardized Difference (D)	Variance Ratio (R)
1	0.1	1.1	0.0	0.9
2	0.1	1.0	0.0	1.0
3	0.1	1.0	0.0	1.0
4	0.1	1.0	0.0	1.0
5	0.1	1.1	0.0	1.0
6	0.1	1.2	0.0	1.0
7	0.2	1.2	0.0	1.0
8	0.2	1.2	0.0	1.0
9	0.1	1.2	0.0	1.0
10	0.1	1.1	0.0	1.0
11	0.1	1.0	0.0	1.0
12	0.1	1.1	0.0	0.9

Table 7-7. PG&E gas matched-comparison balance test

Month	Unmatched		Matched	
	Standardized Difference (D)	Variance Ratio (R)	Standardized Difference (D)	Variance Ratio (R)
1	0.1	1.0	0.0	1.0
2	0.0	1.0	0.0	1.0
3	0.0	1.0	0.0	1.0
4	0.0	0.9	0.0	1.0
5	0.1	0.9	0.0	1.0
6	0.1	0.9	0.0	1.0
7	0.0	0.9	0.0	1.0
8	0.1	0.9	0.0	1.0
9	0.1	0.9	0.0	1.0
10	0.0	0.9	0.0	1.0
11	0.0	1.0	0.0	1.0
12	0.1	1.0	0.0	1.0

We also provide a visual demonstration of the condition of matches using plots of the mean monthly normalized energy consumption of participant and matched comparison households in Figure 7-1. We provide a plot of the distribution of six matched datasets in Figure 7-1 demonstrating that the participant and comparison group households are well matched.

Figure 7-1. Distribution of matched-comparison households

7.2 Appendix B: Additional impact results

Figure 7-2 and Figure 7-3 provide electric and gas savings components for SCE by household and year.

Figure 7-2 indicates that, except for 2017, PG&E's per household electric results are primarily baseload savings with a cooling consumption reduction of only 10% to 15% and no heating consumption reduction.

Figure 7-2. PG&E electric savings components over time

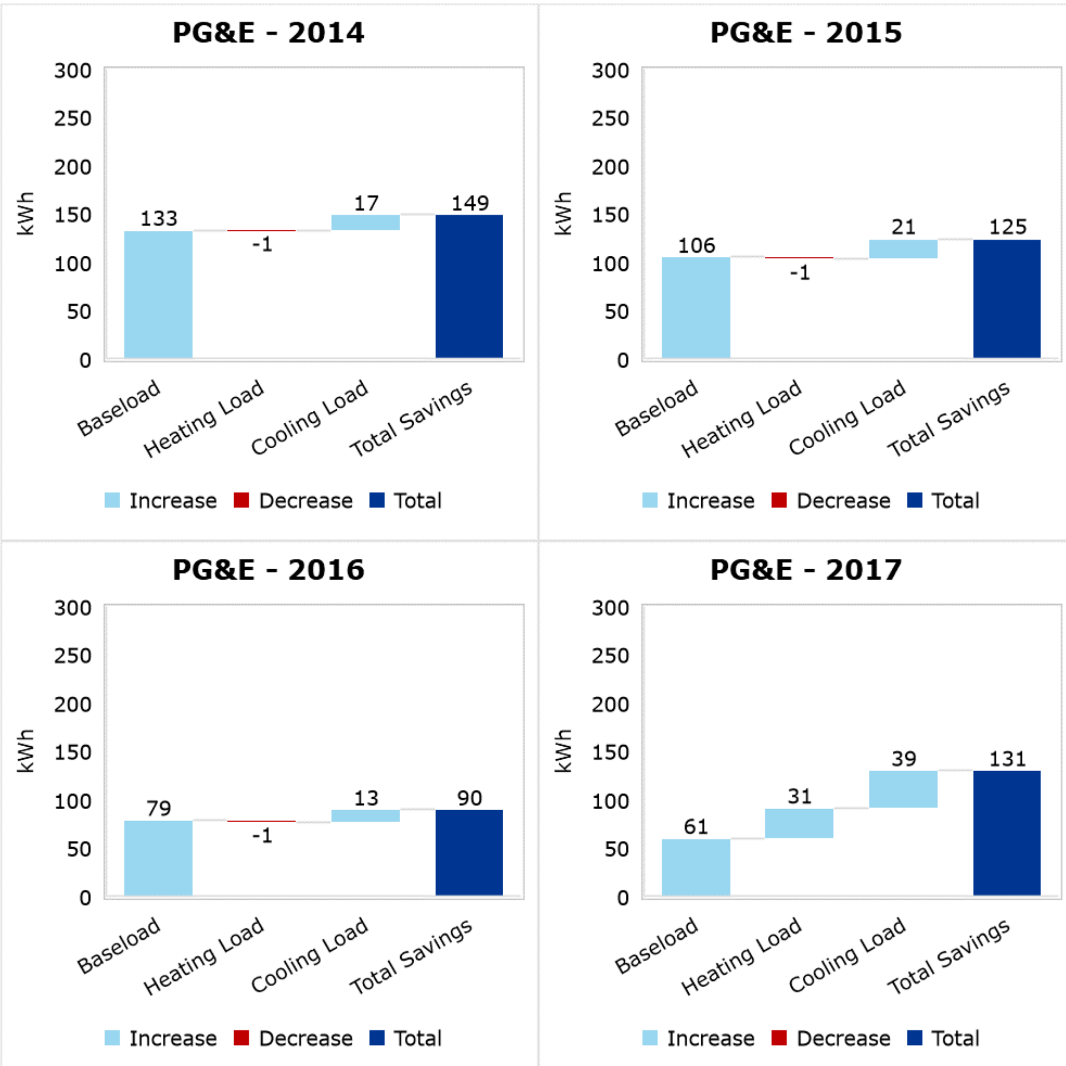


Figure 7-3 illustrates that SCE electric savings are balanced, with about two-thirds of electric savings coming from baseload reduction and one-third from cooling.

Figure 7-3. SCE electric savings components over time

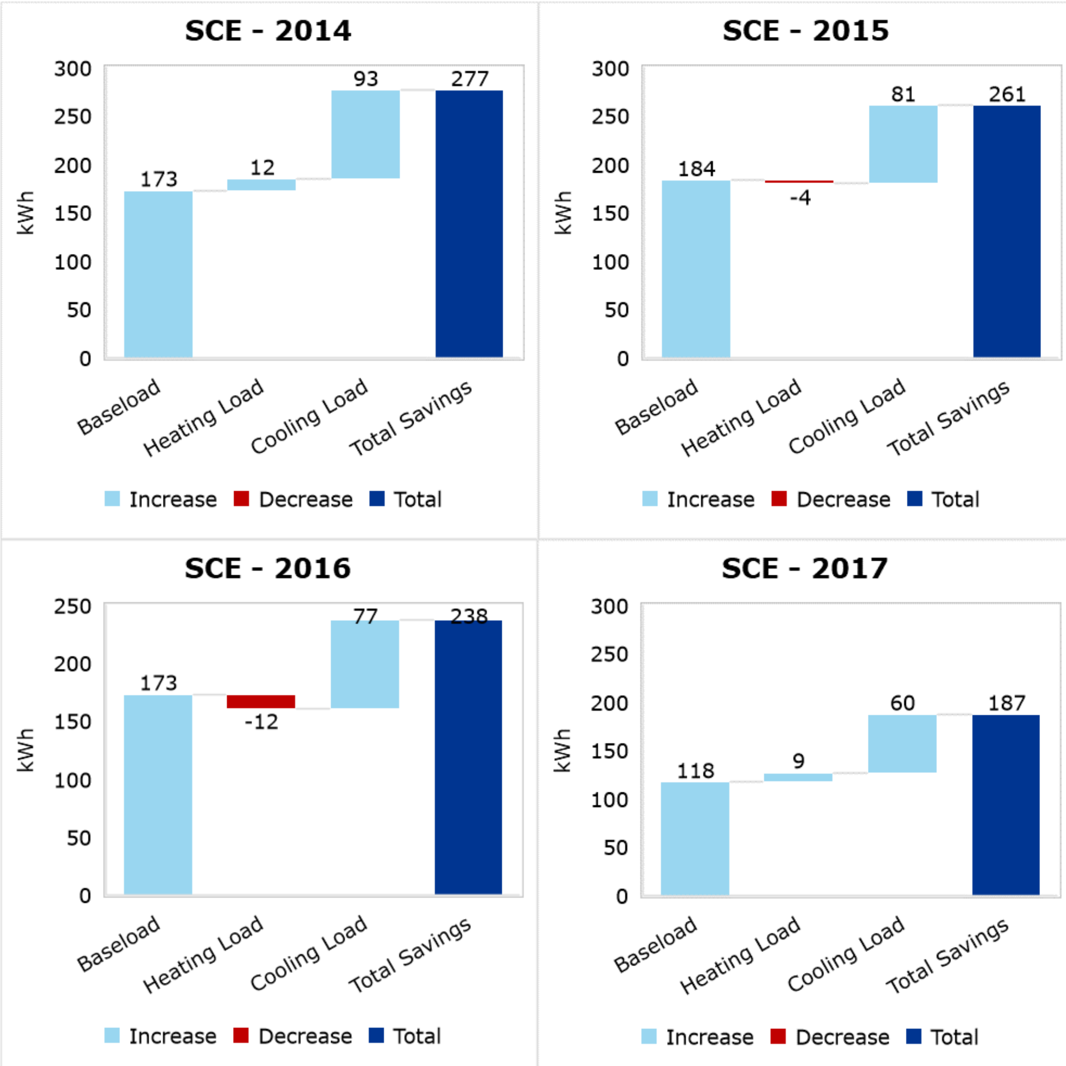


Figure 7-4 indicates that SDG&E's electric savings per household over the four years is split between baseload and heating load reductions, except in 2017 where savings are primarily achieved in the baseload component.

Figure 7-4. SDG&E electric savings components over time

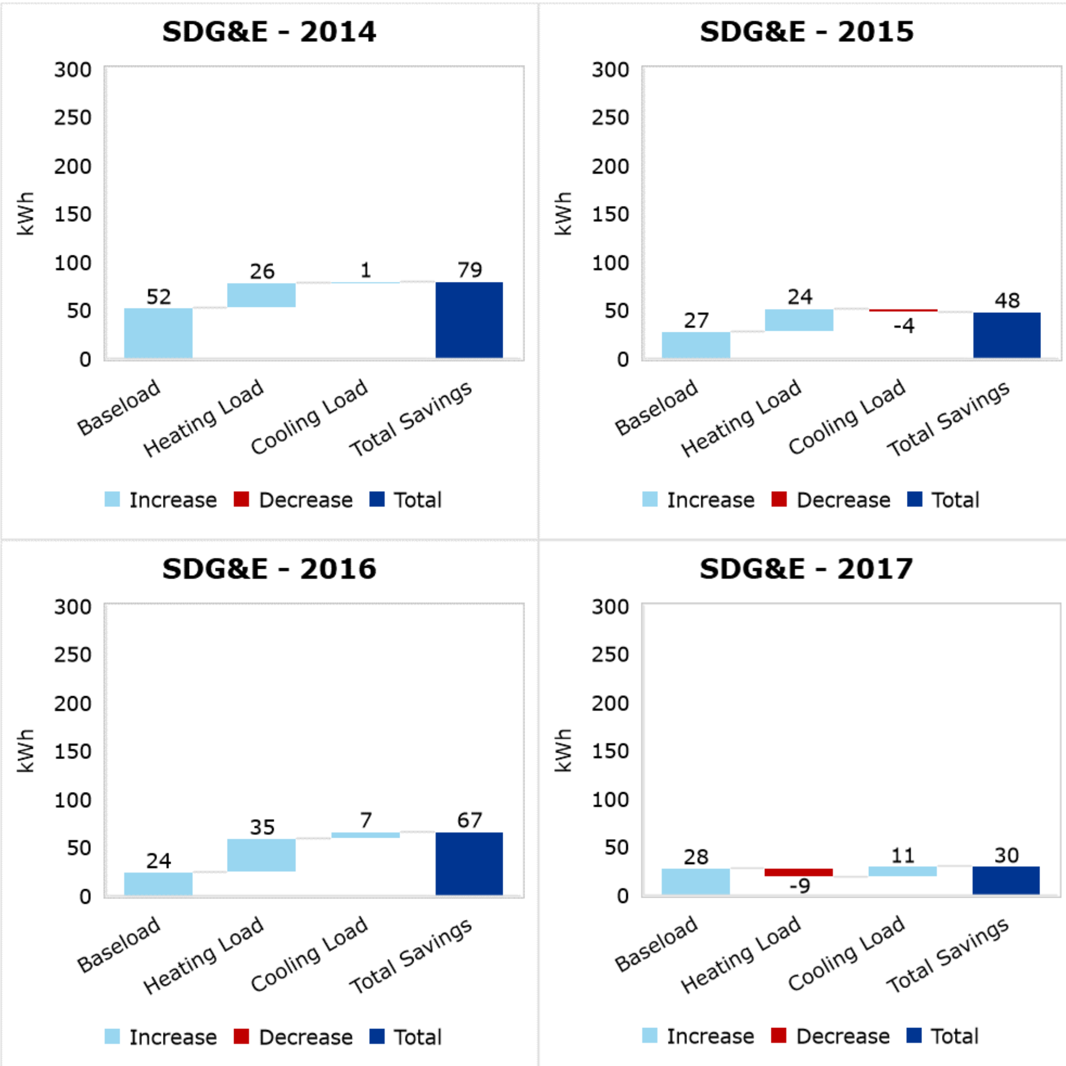


Figure 7-5 presents PG&E's gas savings components for 2014 through 2017. Except for 2015, PG&E's savings are almost entirely due to baseload reductions but also included heating savings.

Figure 7-5. PG&E gas savings components over time

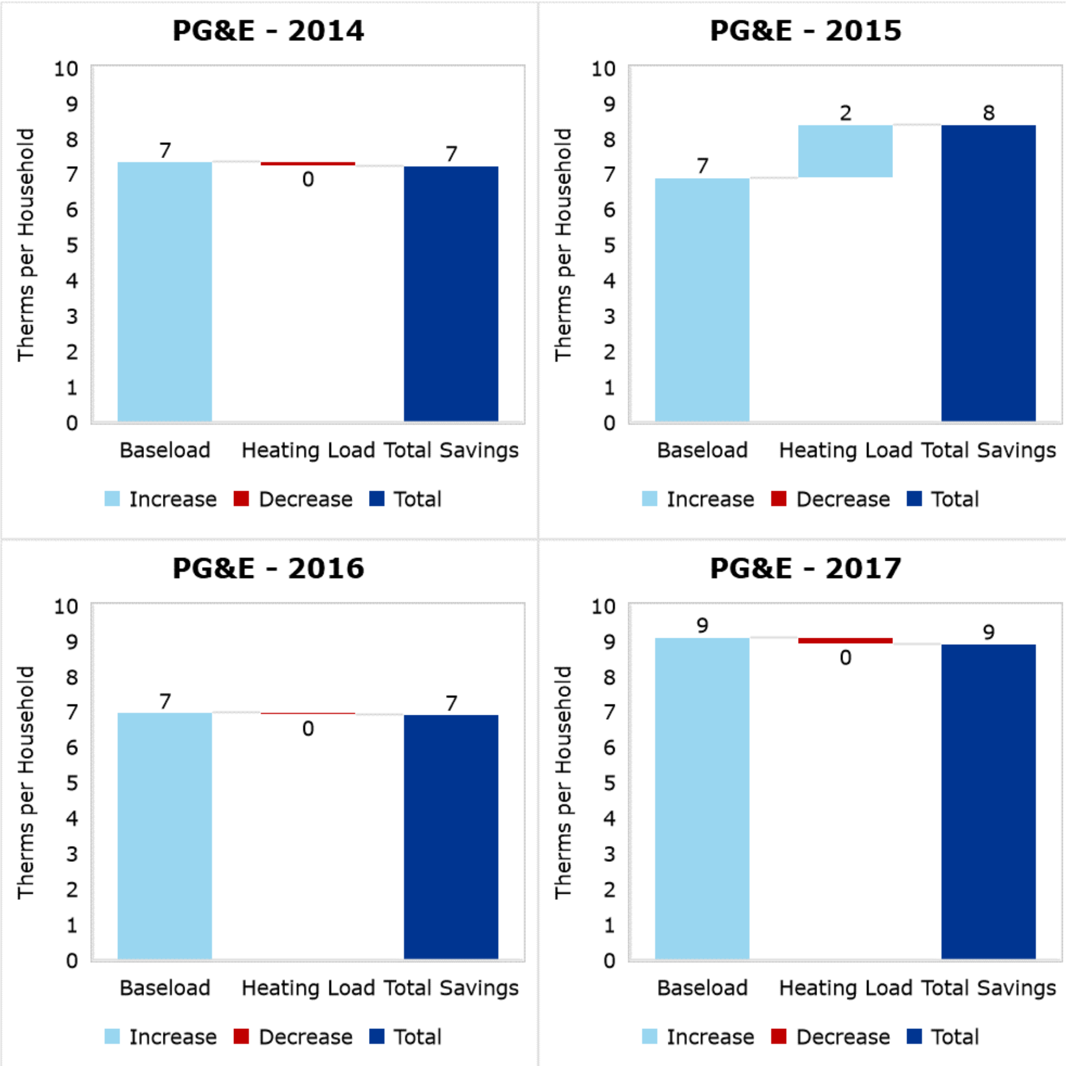


Figure 7-6 shows that the savings components for SCG from 2014 through 2017 are relatively consistent.

Figure 7-6. SCG gas savings components over time

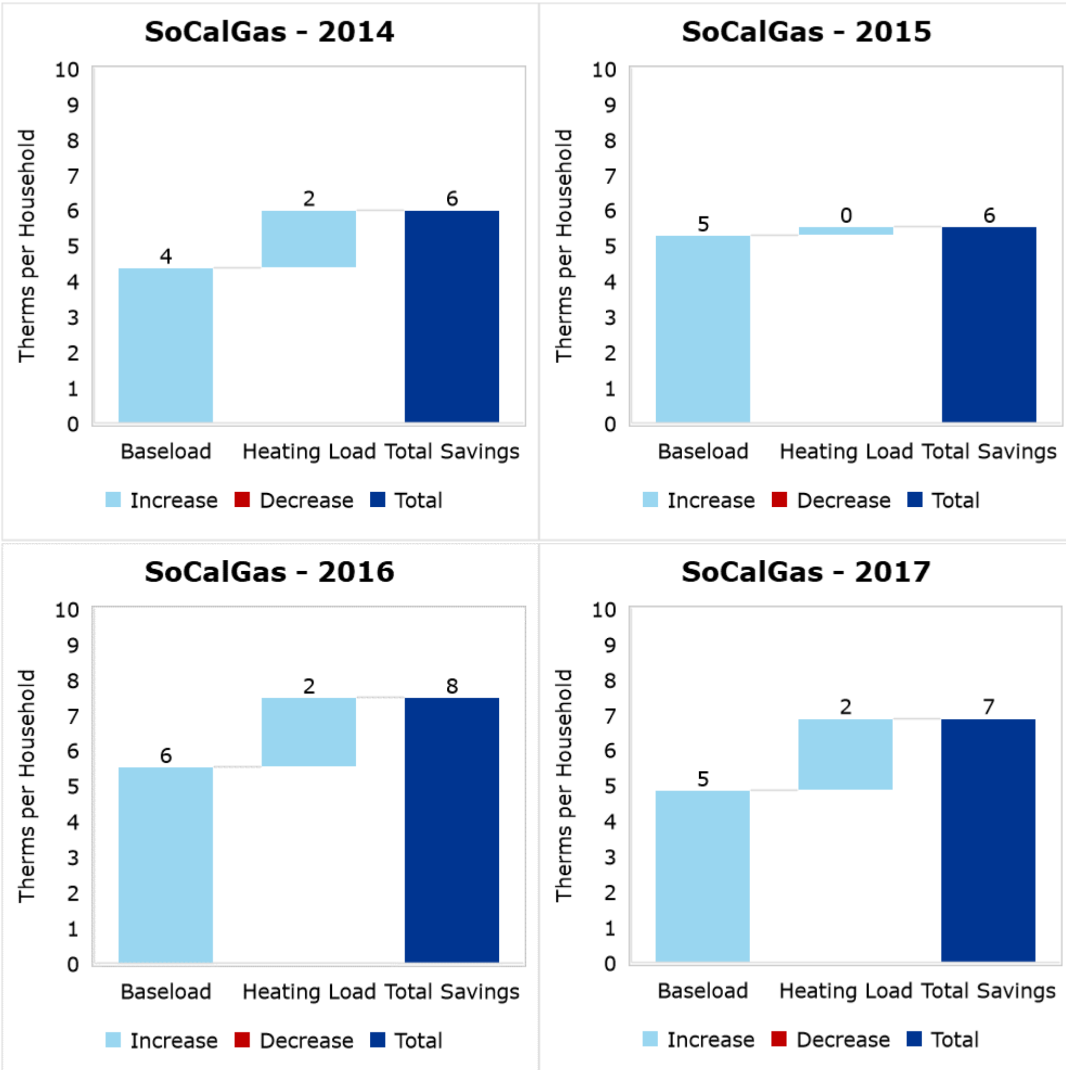


Figure 7-7. SDG&E gas savings components over time

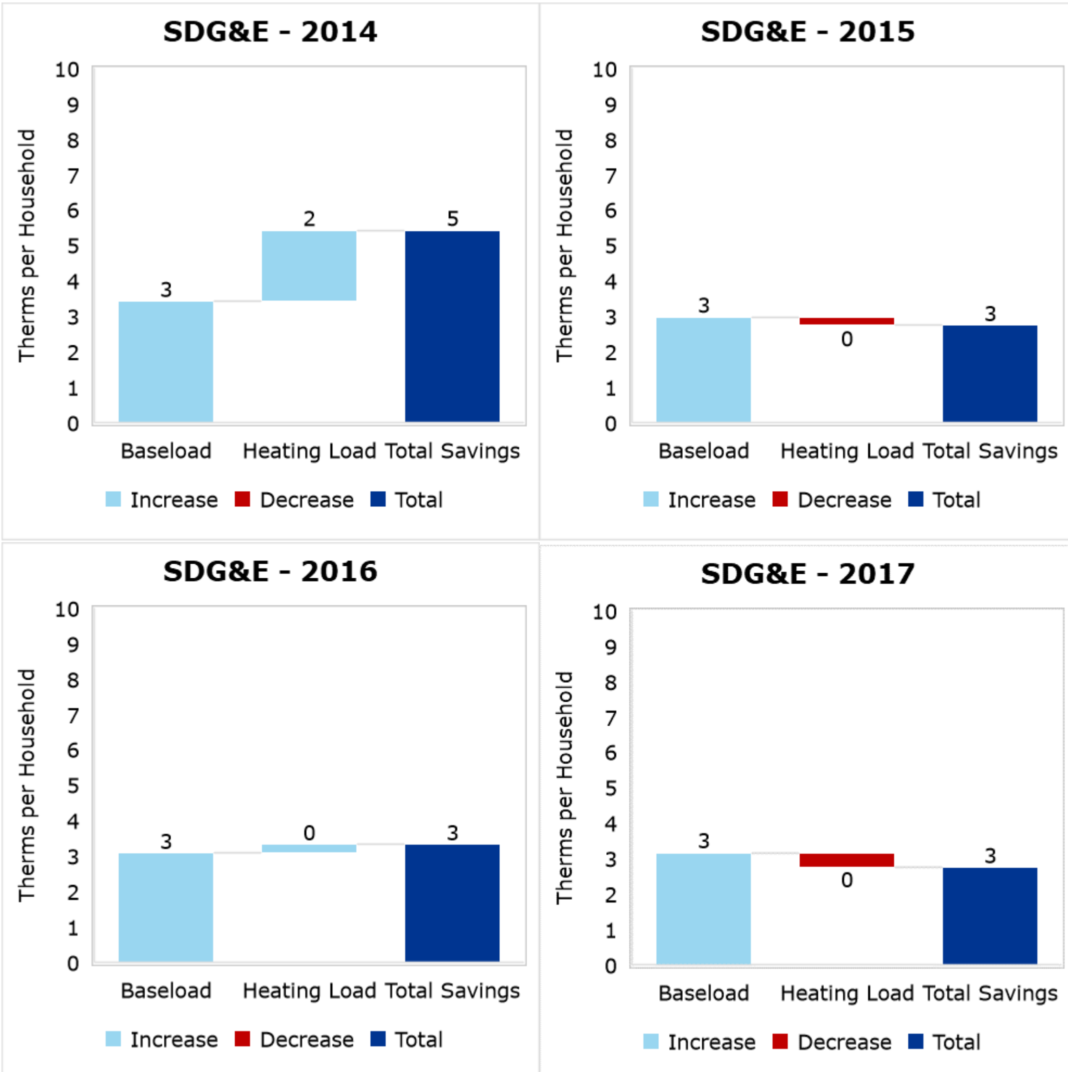


Figure 7-8. Electric per household savings components by housing type

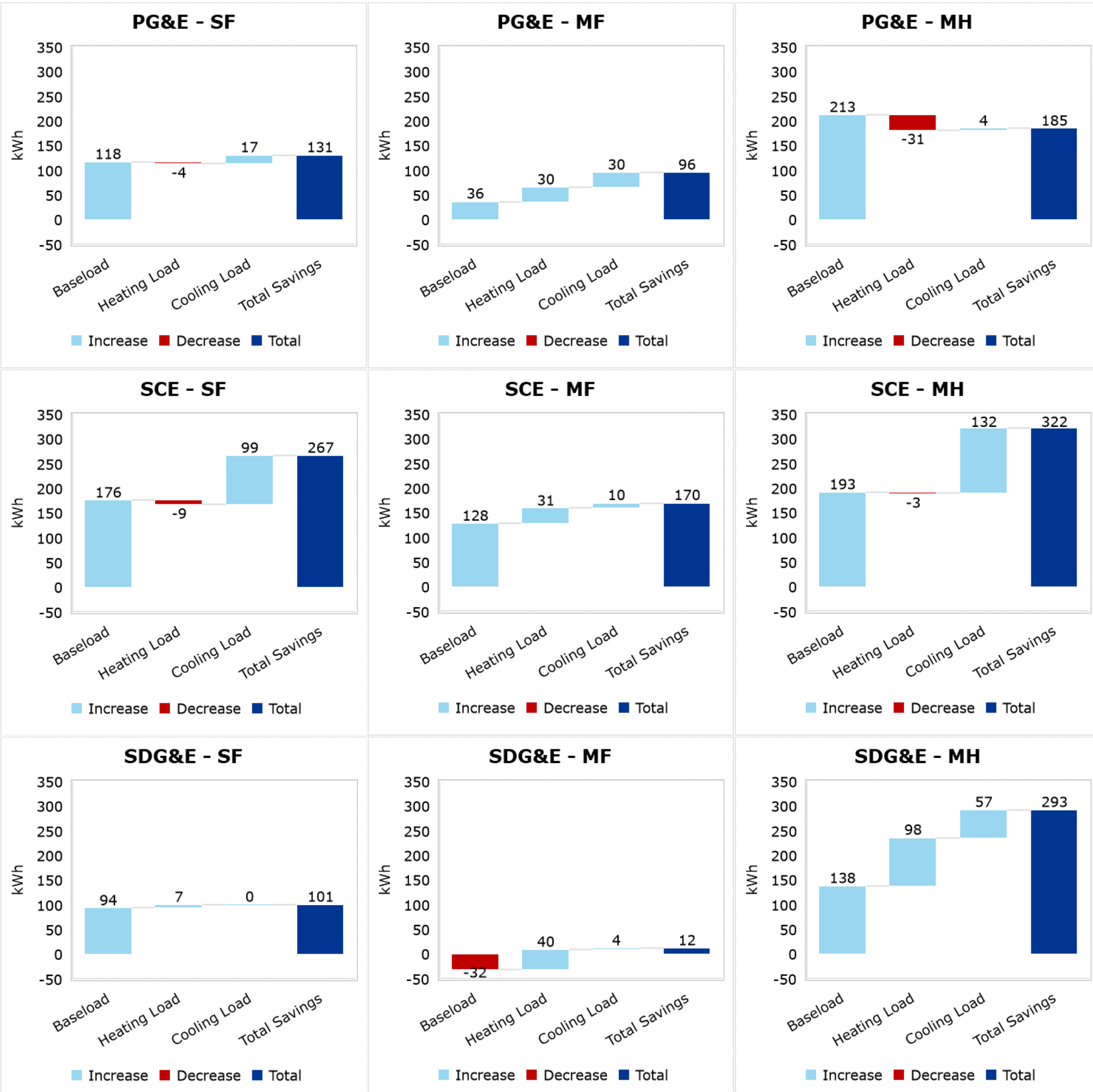
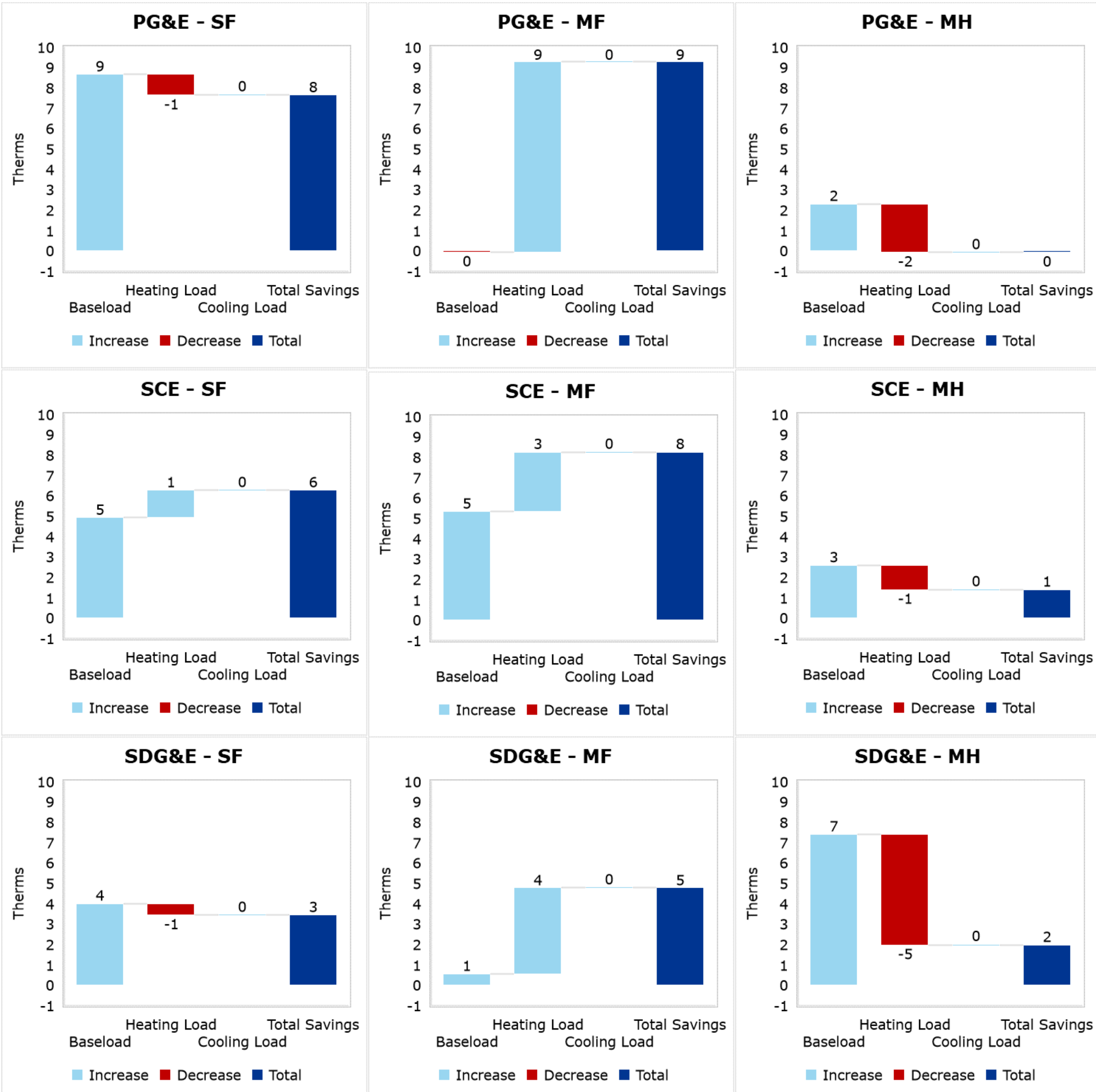


Figure 7-9. Gas per household savings components by housing type



7.3 Appendix C: kWh to kW conversion Factors

Table 7-3 lists the current kWh to kW ESA kWh to kW conversion factors used by PG&E, SCG, and SDG&E. We have highlighted the PG&E's LED measures in red because we believe they are too high and may be incorrect.

Table 7-8. PG&E kWh to kW conversion factors

Measure Category	Conversion Factor
Air Sealing/Envelope	0.00019
Central A/C Tune-up	0.00018
Compact Fluorescent Lights (CFL)	0.00013
Duct Testing and Sealing	0.00016
Evaporative Coolers (Replacement)	0.00032
Exterior Hard-wired CFL fixtures	0.00013
Exterior Hard-wired LED fixtures	0.0008
Faucet Aerator	0.0002
HE Clothes washer	0.00018
Interior Hard wired CFL Fixtures	0.00013
Interior Hard-wired LED fixtures	0.0018
Low Flow Shower Head	0.0002
New - LED A-Lamps	0.0018
New - Smart Power Strips - Tier 2	0.00014
Refrigerators	0.00014
Room A/C Replacement	0.00018
Smart Power Strips - Tier 1	0.00014
Thermostatic Shower Valve	0.00022
Torchiere (CFL)	0.00013
Torchiere LED	0.0018
Vacancy Sensor	0.00013
Water Heater Blanket	0.00022
Water Heater Pipe	0.00017

Table 7-9. SCE kWh to kW conversion factors

Measure Category	Conversion Factor
Air Sealing/Envelope	0.00012
Central A/C Replacement	0.00015
Compact Fluorescent Lights (CFL)	0.00013
Duct Testing and Sealing	0.00016
Evaporative Cooler (Installation)	0.00015
Exterior Hard wired CFL fixture	0.00013
Exterior Hard-wired LED fixture	0.00011
Faucet Aerators	0.00012
Heat Pump Replacement	0.00045
LED Reflector Lamp	0.00011
Low Flow Shower Head	0.00012
New - LED A-Lamps	0.00011
New - Smart Power Strips - Tier 2	0.00014
Pool Pumps	0.00031
Refrigerators	0.00012
Room A/C Replacement	0.00015
Smart Power Strip - Tier 1	0.00014
Torchiere	0.00013
Torchiere - LED	0.00011
Water Heater Blanket	0.00012
Water Heater Pipe Insulation	0.00012

Table 7-10. SDG&E kWh to kW conversion factors

Measure Category	Conversion Factor
Air Sealing	0.00021
Attic Insulation	0.00019
Central A/C Tune-up	0.00019
Compact Fluorescent Lights (CFLs)	0.00012
Duct Testing and Sealing	0.00021
Exterior Hard wired CFL fixtures	0.00007
Faucet Aerator	0.00014
High Efficiency Clothes Washer	0.00013
Interior Hard wired CFL fixtures	0.00007
LED Night Lights	0.00009
Low Flow Showerhead	0.00014
Microwaves	0.0002
New - LED Diffuse Bulb (60W Replacement)	0.00012
New - LED Reflector Bulb	0.00012
New - Smart Power Strips - Tier 2	0.00014
Refrigerators	0.00012
Room A/C Replacement	0.00019
Smart Strip	0.00014
Thermostatic Shower Valve	0.00012
Torchiere	0.00012
Water Heater Blanket	0.00013

7.4 Appendix D: Research plan for ESA impact evaluation

The final ESA Low-Income Impact Evaluation Research Plan, publicly posted on September 28, 2017, is a pdf attachment with the following file name.

ESA Impact Eval
Research Plan.pdf

7.5 Appendix E: IESR table

Study ID	Study Type	Study Title	Study Manager
	Impact Evaluation	Energy Savings Assistance (ESA) Program Impact Evaluation Program years 2015–2017	SoCalGas

Recommendation	Program or Database	Summary of Findings	Additional Supporting Information	We Practice / Recommendations	Recommendation Recipient
1	ESA	The ex ante savings assumptions for ESA program years 2015–2017 were higher than the achieved savings found in this impact evaluation.		ESA Program planners should use the results of this impact evaluation to develop new ex ante savings assumptions for measures that roll up to reasonable household level savings.	All IOUs
2	ESA			ESA program planners should fully account for potential consumption-increase assumptions for measures that are installed for non-energy related benefits. This would include, for instance, flagging fixes to heating or cooling units where the unit was not working or not used prior to the visit. This will segregate off installations that increased consumption and improve overall program savings projections.	All IOUs

Recommendation	Program or Database	Summary of Findings	Additional Supporting Information	We Practice / Recommendations	Recommendation Recipient
3	ESA	The evaluation team encountered challenges cleaning the ESA program tracking data and ensuring that the tracking data aligns with CPUC reported values. Verifying the accuracy of the tracking data will improve future analyses. The verification should confirm that the tracking data matches what the IOUs report to the CPUC.		All relevant identifiers in the program tracking fields should be standardized such that information readily rolls up to program totals and matches the values reported to the CPUC.	All IOUs
4	ESA			Program definitions and requirements should be aligned with billing information.	All IOUs

Recommendation	Program or Database	Summary of Findings	Additional Supporting Information	We Practice / Recommendations	Recommendation Recipient
5	ESA	The method used is an ideal method for future evaluations to ensure consistent analysis methods across evaluation timeframes. There are several research questions that could be addressed in future evaluations by using AMI data and conducting a process evaluation that could potentially yield useful information that informs the implementation of ESA programs in the future		Future evaluations should replicate the two-stage analysis approach followed here and expand the billing analysis from monthly consumption data to daily data using the available AMI data. AMI data will improve the quality of weather-normalization in the first stage modeling reducing variability and improving the quality of results in the second-stage model. Any new methods or approaches proposed for future evaluations should be required to replicate results using this approach and demonstrate the relation to and improvements relative to the current approach	All IOUs
6	ESA			Future impact evaluations should include a process evaluation of program delivery mechanisms to inform future impact evaluations.	All IOUs

7.6 Appendix F: DNV GL's responses to comments

The following table provides DNV GL's responses to specific comments.

	Comment	Response
Comment from Carol Yin	Would it be possible for the evaluation team to include an appendix with recommendations presented using the table from the CPUC Energy Division Impact Evaluation Standard Reporting Guidelines? Thank you! https://pda.energydataweb.com/api/downloads/1399/IESR_Guidelines_Memo_FINAL_11_30_2015.pdf	Added as an Appendix
Comment from Jason Symonds	Would it be possible to present Figure 5-1, 5-3, 5-9 and 5-10 in the report as savings per household as % of total household consumption? For example, on Figure 5-1, PG&E in 2015 saved 125 kWh per household. But what % of total household consumption is that 125 kWh? These percentages will allow a comparison of household energy savings (kWh and therms) across years and IOUs. Thank you for your time and effort on this report and being responsive to this request.	Thank you for the comment, we have added Figure 5-2 and 5-11 to the report.

	Comment	Response
Comment from GreenFan, Inc	Please see the attached PDF document for all of my comments. The primary conclusion of the ESA Evaluation states that “Ex-ante savings assumptions were higher than achieved savings, with some measures leading to an increase in consumption.” This conclusion is incorrect due to the reasons provided in my comments. The AC Tune-up (ACT) and Smart Fan Control (SFC) measures cannot “lead to an increase in consumption” because these measures increase cooling and/or heating capacity and thermal comfort, extend off-cycle time, and reduce HVAC energy use. Therefore, the ESA Evaluation study should be revised to include at least two explanations regarding the increase in energy consumption for HVAC measures: 1) customer “take back” due to repair of HVAC systems that would cause cooling and heating energy use to increase, and 2) issues with billing regression methodology being unable to accurately measure HVAC energy savings due to issues with data cleaning, selecting appropriate comparison groups, model specification, model validation, and savings being too small relative to whole building electricity and gas consumption (i.e., signal-to-noise ratio issues).	Response to GreenFan, Inc. and Mowris: The following comments respond to these parallel comments. We thank you for your comments and hear your concerns regarding “AC Tune-up” and “Smart Fan Delay.” We have not included the requested changes to the report conclusions. The report is clear with respect to the purpose of the measure-level estimates. The report is clear on the methods with which the measure-level parameters are estimated. A billing analysis was used to produce site-level estimates for the ESA program. At 2-4% of overall consumption, this is low by IPMVP standards, but IPMVP does not expect populations in the multiple tens of thousands. We provide measure-level results at the request of the IOUs, with consistent acknowledgement that these estimates are not definitive, but provide a rough idea how measures contribute to an overall change in consumption. Additionally, the AC Tune-up results reported in Figure 5-6 are positive, not negative. We acknowledge in the report that certain measures may include the effects of take-back in the estimates. The report reports the average change in consumption associated with specific measures and acknowledges that other non-energy benefits may justify those measures. We have recommended that non-working measures be flagged so that this take-back effect can be disentangled. If we had any evidence that the concerns discussed in Section 6.2 had an effect on measure-level results, we would have acknowledged this. Concerns regarding the effect of CARE subsidies do not seem relevant under the circumstances. Both the ESA participants and the comparison group were drawn from the CARE population producing no conflating effect in a difference in difference structure. To the contrary, the increased consumption levels, in general, that we would expect across the CARE population given the discounted cost of energy would tend to increase the magnitude of savings to the extent that they occurred. The general comments regarding the value of a billing analysis approach, and the requested textual changes, do not take into consideration a full reading of the report. The discussion of measure-level results contains clear language relating to the potential shortcomings of the measure-level results. The requested changes could not be made without making fundamental changes throughout the report that we do not believe are justified by the arguments in these comments.

	Comment	Response
Comment from Robert Mowris	The primary conclusion of the ESA Evaluation states that “Ex ante savings assumptions were higher than achieved savings, with some measures leading to an increase in consumption.” This conclusion is incorrect due to reasons provided in the comments. The Smart Fan Delay (“SFD”) measure cannot “lead to an increase in consumption,” because the SFD does not affect the thermostat setpoint or the thermostat call for cooling duration; and for 93% of gas furnaces, the SFD will decrease the thermostat call for heating duration by satisfying the thermostat sooner. The SFD measure increases cooling capacity, heating capacity, thermal comfort, extends off-cycle time, and reduces HVAC energy use. Therefore, the ESA Evaluation study should be revised to include an explanation regarding negative savings for the SFD measure to indicate that the billing regression methodology is unable to accurately measure HVAC energy savings because of issues with the baseline (due to repairs and customer “take-back”) and savings are too small relative to whole building electricity and gas consumption (i.e., small signal-to-noise ratio). Please remove the bar labeled “Smart Fan Delay” from Figure 5-6 (p. 46), Figure 5-7 (p. 47), and Figure 5-14 (p. 53). Please see the attached PDF file.	See above comment response
Comments from NRDC	We appreciate the consideration that the evaluation has paid to highlighting the need to ensure that ex-ante savings assumptions for the upcoming ESA applications are appropriate. Upon examination, we agree that the evaluation methodology is accurate and is designed to primarily get an estimate of measure of savings at the household level. We are concerned, however, that the conversation to this point about the low achieved savings relative to ex ante savings assumptions has seemed to preemptively conclude that this was a result of program planning assumptions being too high—without a discussion about what other factors might be responsible (including program design and implementation). Achieved savings are a function of ex-ante assumptions, program design, and implementation practices, and a full examination and discussion of all of these factors is a critical prerequisite before concluding that program planning assumptions were too high. Accordingly, “Recommendation 1” should be examined for embedded premature conclusions about how new ex ante savings assumptions for measures should “roll up to reasonable household level savings,” and any future proposal to lower the new ex ante savings assumptions without careful and comprehensive consideration should be scrutinized. The evaluation correctly identifies that more research is required to understand the program implementation process (i.e., a process evaluation). Without the recommended process evaluation, it is impossible to understand why ex-ante savings weren’t achieved and the level of accuracy of ex-ante estimates. Future evaluations should better investigate why ex-ante savings weren’t realized; this can be done through a thorough process evaluation and targeted site visits to understand how the implemented measures are performing. Any update to ex-ante savings assumptions for the upcoming ESA applications should carefully consider these future evaluations as well as the limits of the impact evaluation findings.	Thank you for your comment. Your points are noted. The conclusions and recommendations were updated to reflect these comments.

	Comment	Response
OpenEE	The use of billing regression analysis for the impact evaluations of the ESA programs is an appropriate approach for evaluating success. However, the savings claims for ESA could also be based on a method that is transparent, repeatable, and consistent with the approaches being used for the impact evaluation. Meter-based efficiency savings claims could reduce the variability in the ex ante and ex post results; and provide actionable intelligence on program performance to optimize interventions. Meter-based quantification could also be the basis for paying contractors for successful engagements with customers to reduce energy consumption, without having to design elaborate schemes for incentivizing specific measures (but can still track which combinations of measures are being installed). ESA program planners can use the more reliable whole building impact results to target low income customers with a propensity to save rather than spending time on new ex ante savings assumptions for measure-level results that were not as reliable. We recommend that the CPUC and IOUs adopt embedded measurement and verification as a principle for program design. Programs should be designed and implemented with the end goal of achieving whole building impacts for the portfolio of interventions. Regression (billing or hourly) analysis embedded in the programs should be consistent and transparent. CalTRACK is a standardized approach currently being used in residential applications in California (www.caltrack.org). AMI data, and the CalTRACK hourly methods, can be part of the embedded M&V expectations to quantify and value the load impacts of these programs and enable them to contribute to grid needs as a resource and support statewide goals. Respectfully submitted, Carmen Best, OpenEE	Thank you for the comment

7.7 Appendix G: PDF comments

GreenFan® Inc.

6125 Bear Claw Lane, Bozeman, MT 59715 ♦ john@greenfan.co ♦ 406-570-9494

Date: April 12, 2019

From: John Walsh, EE, President

Re: Comments on the ESA Program Impact Evaluation Program Years 2015–2017 managed by Southern California Gas Company

To: Loan Nguyen, Southern California Gas Company, LNguyen@semprautilities.com

---

Please consider the following comments regarding the Energy Savings Assistance (ESA) Program Impact Evaluation Program years 2015–2017 ("ESA Evaluation"). The ESA Evaluation mistakenly provides negative savings for the PG&E "Smart Fan Delay" (SFD) measure installed. See Figures 5-6 and 5-7 (pp.46-47) and Figure 5-14 (p.53).[1] The SFD measure improves HVAC efficiency and saves energy by detecting HVAC system type and mode of operation based on signals present on the thermostat or equipment terminals and provides an energy efficient variable fan-off time delay based on HVAC system type, mode of operation, and duration of the cooling or heating cycle. The EFC® delivers more sensible cooling or heating capacity to the space by exceeding standard thermostat setpoint differential temperatures, lengthen off-cycle durations, improve thermal comfort, and heating on-cycle durations. For some gas furnace systems, the EFC® provides high speed fan operation to satisfy the thermostat heating differential sooner. The EFC® can be installed on systems with a fixed fan-off time delay for cooling or heating or pre-existing cooling-only enhanced time delay and is cost effective for most prototypes and climate zones. The Smart EFC™ provides common wire functionality and reliable power for Smart Communicating Thermostats plus energy savings technology embodied in the EFC®.

The SFD does not affect the amount of energy required to cool a space and actually reduces the energy required to heat a space for about 93% of gas furnace heating systems.[2] When operating with a gas furnace, the SFD operates the fan at higher speed during the heating cycle (after HX reaches operating temperature) to satisfy the thermostat sooner and reduce gas furnace operation. The SFD adds fan-only time after the cooling or heating cycle based on the duration of the cycle which uses a percentage of the energy consumed while the heating or cooling cycle is performed. The SFD also adjusts the variable fan-off delay based on Fault Detection Diagnostics (FDD) of low cooling or heating capacity (low charge, dirty air filters, etc.) and severe weather. The SFD causes the thermostat to undershoot the setpoint for cooling and overshoot the thermostat set point for heating. This overshoot/undershoot causes the room temperature to take longer to drift back to the point where the thermostat calls for another heating or cooling cycle. For cooling, the SFD uses about 6.6% more electricity for the variable fan-off delay (based on 10 minute average AC operation); and delivers more cooling capacity, over satisfies the thermostat setpoint,

---

[1] The SFD is a patented product manufactured by GreenFan® Inc. sold under the trademarked name Efficient Fan Controller® (EFC®) with U.S. Patents 8,763,920, 9,328,933, 9,500,386, 9,671,125, 9,797,405, and 9,995,493.

[2] Market research based on 5,582 gas furnace units installed in California indicates that 6.7% of heating fans are not enabled to high speed fan operation by the SFD (pp. 26-29, R. Mowris, P. Jacobs. 2016. EFC workpaper Work Paper EFC173PHVC138).

and saves energy by lengthening the AC compressor off-cycle duration. For heating, the SFD uses about 31% more electricity for high-speed fan operation and variable fan-off delays (weighted average).[3] For all heating systems, the SFD variable fan-off delay delivers more heating capacity at the end of the cycle to over satisfy the thermostat setpoint and saves energy by increasing the gas furnace off-cycle time. The ex ante SFD weighted average extra fan energy is 17.6 kWh/year for gas furnace heating. The ex ante SFD weighted average cooling savings are 115.8 kWh/year. Therefore, the net ex ante SFD weighted average electricity savings are 98.2 kWh/year (115.8 – 17.6 = 98.2 kWh/y). Net ex ante savings are based on California housing stock data, Intertek test data, and calibrated DOE2 simulations of SFM, MFM and DMO prototypes using DEER 2017 eQUEST version 3.65.[4] <u>The SFD does not affect the thermostat setpoint or the thermostat call for cooling duration; and for 93% of units, the SFD will decrease the thermostat call for heating duration by satisfying the thermostat sooner. Therefore, the SFD cannot lead to an increase in cooling electricity or natural gas consumption.</u>

The SFD automatically detects the type of system to which it is attached and not only saves gas in furnaces, but also electricity in heat pumps, hydronic, and electric heating systems. These additional savings have not been evaluated in this report. Southern California Gas Company and DNVGL risk liabilities using billing analysis to evaluate HVAC measures in the ESA program due to repairs that are made to AC units which creates issues when determining baseline consumption. Furthermore, the IPMVP and other studies point out issues with billing analysis methodologies to evaluate energy efficiency measures when savings are less than 10% of whole building consumption.[5] This is particularly troubling when the ESA Evaluation billing analysis methodology does not yield results consistent with independently verified tests performed by Intertek, an ASHRI-certified testing laboratory (http://www.intertek.com/testing/). Intertek tests

---

[3] Ibid. 31% extra fan energy based on Intertek and field tests of fan power for split-system and packaged gas furnaces and weighted average of 92.3% enabled to high speed, 6.7% enabled to low speed, 69.3% base 90-second or less fixed fan-off delay, and 30.7% base 120 second or greater fixed fan-off delay.

[4] Intertek tests indicate savings of 3.8 to 32% depending on cooling cycle duration and heating savings of 5 to 30.1% depending on heating cycle duration. See Mowris, R. Jacobs, P. 2016. Efficient Fan Controller® (EFC®) for Residential HVAC Systems. Work Paper EFC173PHVC138. Prepared by Verified® Inc. and Building Metrics Inc.. Intertek. 2015. Performance Evaluation Based on Intertek Test Data of the GreenFan EFC Installed on Split and Packaged Air Conditioners with gas Furnaces. 101756555DAL-001B. Intertek. 2018. Performance Evaluation Based on Intertek Test Data of the GreenFan EFC Installed on Heat Pump and Hydronic Split Systems. 102791047DAL-001A. California housing stock: 65% single family, 31% multifamily, and 4% mobile home. See page 15, Figure 1.10: California Housing Stock by Type, 2010-2014 Average: Multifamily, Single-Family, and Mobile/Manufactured Homes/Other. California Department of Housing and Community Development. 2017. California's Housing Future: Challenges and Opportunities Public Draft. http://www.hcd.ca.gov/policy-research/plans-reports/docs/California%27s-Housing-Future-Main-Document-Draft.pdf (Available on request)

[5] See IPMVP. 2002. International Performance Measurement & Verification Protocol Concepts and Options for Determining Energy and Water Savings, Volume I, Revised March 2002, DOE/GO-102002-1554 (p. 27 of 31505.pdf) at https://www.nrel.gov/docs/fy02osti/31505.pdf. IPMVP Option C "is intended for projects where savings are expected to be large enough to be discernible from the random or unexplained energy variations that are normally found at the level of the whole facility meter. The larger the saving, or the smaller the unexplained variations in the baseyear, the easier it will be to identify savings. Also the longer the period of savings analysis after ECM installation, the less significant is the impact of short term unexplained variations. Typically savings should be more than 10% of the baseyear energy use if they are to be separated from the noise in baseyear data."
Also see E. Ziemba et al. 2017. Cleaning Up the Mess of Energy Billing Data: An investigation of Differences in Billing Analysis Results Caused by Data Cleaning Methodologies. Opinion Dynamics, Boston, MA. 2017 International Energy Program Evaluation Conference, Baltimore, MD. https://www.opiniondynamics.com/wp-content/uploads/2017/08/2017_IEPEC_Paper_Cleaning-up-the-mess-of-billing-data_Ziemba.pdf

of the SFD technology indicate cooling savings of 3.8 to 32% depending on cooling cycle duration and heating savings of 5 to 30.1% depending on heating cycle duration.[6] Publishing incorrect evaluation results of the SFD technology might cause a setback of several years or decades and deprive ratepayers of a cost effective HVAC energy efficiency measure that saves both electricity and natural gas.

The SFD technology cannot cause negative energy savings because the SFD increases cooling and/or heating capacity and thermal comfort, extends the off-cycle time, and reduces HVAC energy use. Therefore, the ESA Evaluation study should be revised to include an explanation regarding negative savings for the SFD measure to indicate that the billing regression methodology is unable to accurately measure HVAC energy savings because of issues with the baseline (due to repairs and customer "take-back") and savings are too small relative to whole building electricity and gas consumption (i.e., small signal-to-noise ratio).

The following revisions are recommended prior to finalizing the ESA Evaluation.

Section 1.6 (p. 7) Conclusions and Recommendations should be revised as follows.

Please Replace: "1. Ex ante savings assumptions were higher than achieved savings, with some measures leading to an increase in consumption. ESA program planners should use the impact results to develop new ex ante savings assumptions."

With: "1. Ex ante savings assumptions were higher than achieved savings, with some measures ~~leading to~~ appearing to cause an increase in consumption due to HVAC repairs which cause issues for the baseline using billing analysis methodologies. Future evaluations should use a different M&V strategy such as calibrated simulation modeling to provide more accurate results for HVAC measures. ESA program planners should ~~use the impact results to~~ develop new ex ante savings assumptions using workpapers that provide all key assumptions including Measure Analysis Software Control (MASControl2) to generate calibrated building energy simulation prototypes and post processing procedures to provide a more appropriate and accurate evaluation methodology."

Please Replace: "3. The evaluation methodology produced consistent year-over-year results at the household level." Future Evaluations should use daily AMI consumption data for more robust results."

With: "3. The evaluation methodology produced consistent year-over-year results at the household level but not at the measure level specifically for HVAC measures." Future Evaluations should use daily AMI consumption data for more robust results and a different M&V strategy such as calibrated simulation modeling to provide more accurate results for HVAC measures."

---

[6] Intertek. 2015. Performance Evaluation Based on Intertek Test Data of the GreenFan EFC Installed on Split and Packaged Air Conditioners with gas Furnaces. 101756555DAL-001B. Intertek. 2018. Performance Evaluation Based on Intertek Test Data of the GreenFan EFC Installed on Heat Pump and Hydronic Split Systems. 102791047DAL-001A.

Please Replace: "4. There are limits to the answers that a billing analysis can provide for how program delivery affects (sic) savings. Future Evaluations should include a process evaluation to better research how program delivery is linked to impacts."

With: "4. There are limits to the answers that a billing analysis can provide for how program delivery ~~effects~~ affects savings. Future Evaluations should include a process evaluation to better research how program delivery is linked to impacts. Billing analysis cannot be used to evaluate HVAC measures due to issues with data cleaning, selecting appropriate comparison groups, model specification, model validation, and savings being too small relative to whole building electricity and gas consumption (i.e., signal-to-noise ratio issues), making billing analysis especially problematic. A different M&V strategy such as calibrated simulation modeling will provide more accurate results for HVAC measures."

Please remove the bar labeled "Smart Fan Delay" from Figure 5-6 (p. 46), Figure 5-7 (p. 47), and Figure 5-14 (p. 53).

In summary, I believe the SFD is an innovative cooling and heating energy efficiency technology that provides cost effective savings for HVAC, specifically cooling and gas heating savings. Therefore, I believe the SFD should be given a fair evaluation using correct methodologies to avoid inadvertent errors that might cause lost opportunities for cost effective electric cooling and gas heating savings.

Thank you for considering these comments and recommended changes to the ESA Evaluation.

TO:	Loan Nguyen, Southern California Gas Company
FROM:	Mohit Chhabra, Miles Muller; Natural Resources Defense Council
DATE:	April 12, 2019
RE:	**Comments on 2015–2017 ESA Impact Evaluation Study**

We appreciate the consideration that the evaluation has paid to highlighting the need to ensure that ex-ante savings assumptions for the upcoming ESA applications are appropriate. Upon examination, we agree that the evaluation methodology is accurate and is designed to primarily get an estimate of measure of savings at the household level.

We are concerned, however, that the conversation to this point about the low achieved savings relative to ex ante savings assumptions has seemed to preemptively conclude that this was a result of program planning assumptions being too high—without a discussion about what other factors might be responsible (including program design and implementation). Achieved savings are a function of ex-ante assumptions, program design, and implementation practices, and a full examination and discussion of all of these factors is a critical prerequisite before concluding that program planning assumptions were too high. Accordingly, "Recommendation 1" should be examined for embedded premature conclusions about how new ex ante savings assumptions for measures should "roll up to reasonable household level savings," and any future proposal to lower the new ex ante savings assumptions without careful and comprehensive consideration should be scrutinized.

The evaluation correctly identifies that more research is required to understand the program implementation process (i.e., a process evaluation). Without the recommended process evaluation, it is impossible to understand why ex-ante savings weren't achieved and the level of accuracy of ex-ante estimates. Future evaluations should better investigate why ex-ante savings weren't realized; this can be done through a thorough process evaluation and targeted site visits to understand how the implemented measures are performing. Any update to ex-ante savings assumptions for the upcoming ESA applications should carefully consider these future evaluations as well as the limits of the impact evaluation findings.

Sincerely,

Mohit Chhabra
Senior Scientist
Natural Resources Defense Council
111 Sutter St., 21st Floor
San Francisco, CA 94104
(415) 875-6100
mchhabra@nrdc.org

Miles Muller
Clean Energy Legal Fellow
Natural Resources Defense Council
111 Sutter St., 21st Floor
San Francisco, CA 94104
(415) 875-8254
mmuller@nrdc.org

VERIFIED® INC.
P.O. Box 2159, Olympic Valley, CA 96146 ♦ robert@verified.co ♦ 530-448-6249

Date: April 11, 2019 (revised 04-13-19)

From: Robert Mowris, P.E., VERIFIED® Inc.

Re: Comments on the ESA Program Impact Evaluation Program Years 2015–2017

To: Loan Nguyen, Southern California Gas Company, LNguyen@semprautilities.com

Please consider the following comments regarding the Energy Savings Assistance (ESA) Program Impact Evaluation Program years 2015–2017 ("ESA Evaluation"). The ESA Evaluation mistakenly provides negative savings for PG&E measures "AC Tune-up" (ACT) and "Smart Fan Delay" (SFD) installed by PG&E. See Figures 5-6 and 5-7 (pp.46-47) and Figure 5-15 (p.53).The ESA Evaluation uses a billing analysis methodology, but does not fully explain how the ACT or SFD measures were evaluated (i.e., regression coefficients, selecting appropriate comparison groups, model specification and validation, etc.). The following citation from the International Performance Measurement & Verification Protocol (IPMVP), indicates that billing analysis methodologies referred to as "Option C" should not be used to evaluate individual measures unless energy savings are greater than 10% of total kWh or total therm billing data for a given site.[1]

> Option C "is intended for projects where savings are expected to be large enough to be discernible from the random or unexplained energy variations that are normally found at the level of the whole facility meter. The larger the saving, or the smaller the unexplained variations in the baseyear, the easier it will be to identify savings. Also the longer the period of savings analysis after ECM installation, the less significant is the impact of short term unexplained variations. Typically savings should be more than 10% of the baseyear energy use if they are to be separated from the noise in baseyear data."

A report published by Ziemba in the 2017 International Energy Program Evaluation Conference (IEPEC) discusses the following important "core elements to billing analysis approaches – 1) data cleaning and preparation, 2) selecting a comparison group, and 3) model specification and validation," and the "influence of data cleaning and preparation related to billing analysis results, specifically the influence related to aligning billing periods."[2] Section 6.2 of the ESA Evaluation (p. 58) indicates the "evaluation team encountered challenges cleaning the ESA program tracking data and ensuring that the tracking data aligns with CPUC reported values" and issues "verifying the accuracy of the tracking data."

ESA low income participants receive a 30 to 35% discount on their bill through the CARE program (http://www.cpuc.ca.gov/general.aspx?id=976) which is another reason why it is difficult to

[1] See IPMVP. 2002. International Performance Measurement & Verification Protocol Concepts and Options for Determining Energy and Water Savings, Volume I, Revised March 2002, DOE/GO-102002-1554 (p. 27 of 31505.pdf) at https://www.nrel.gov/docs/fy02osti/31505.pdf.

[2] E. Ziemba et al. 2017. Cleaning Up the Mess of Energy Billing Data: An investigation of Differences in Billing Analysis Results Caused by Data Cleaning Methodologies. 2017 IEPEC. https://www.opiniondynamics.com/wp-content/uploads/2017/08/2017_IEPEC_Paper_Cleaning-up-the-mess-of-billing-data_Ziemba.pdf

establish a correct counterfactual baseline. So there are three reasons why the HVAC measures appear to cause an increase in consumption: 1) small savings less than 10% causing signal-to-noise issues; 2) HVAC repairs causing issues with establishing correct counterfactual baselines; and 3) 30-35% bill subsidies causing "take back" which tends to increase ex post consumption. These issues cause significant uncertainty with respect to ESA Evaluation results for HVAC measures including the ACT and SFD measures.

The ACT measure ex ante savings represent about 9.4% of total Air Conditioning (AC) electricity (kWh) based on the 2015-17 ESA program average refrigerant charge adjustment of 7.3% (based on data for 24,215 AC units) and independent tests performed by Intertek, an AHRI-certified laboratory.[3] AC represents only 7% of total electric usage per Table ES-1 (p. 3) of the 2010 KEMA RASS Study.[4] Therefore, the ACT measure savings are 0.66% of the total electric bill (i.e., 0.66% = 7% * 9.4%). According to IPMVP, ACT cooling savings of 0.66% are too small "to be separated from the noise in baseyear data" using billing analysis.

The SFD measure ex ante savings are 9.5 to 13.3% of AC electricity (kWh) based on Intertek test data and calibrated DOE2 simulations of SFM, MFM and DMO prototypes using DEER 2017 eQUEST version 3.65 http://www.deeresources.com/.[5] However, AC is only 7% of the total electric usage per Table ES-1 (p. 3) of the 2010 KEMA RASS Study electric. Therefore, the SFD measure cooling savings are 0.7 to 0.9% of the total electric bill. The SFD measure also saves 13.3 to 15% of gas furnace heating (therms) based on Intertek test data and calibrated DOE2 simulations of SFM, MFM and DMO prototypes. Gas heating is only 37% of total gas usage per Table ES-6 (p. 9) of the 2010 KEMA RASS Study electric. Therefore, the SFD measure heating savings are only 4.9 to 5.6% of the total gas bill. According to IPMVP, SFD cooling savings of 0.7 to 0.9% and heating savings of 4.9 to 5.6% are too small "to be separated from the noise in baseyear data" using billing analysis.

Page 53 of the ESA Evaluation indicates that "HVAC measures produce negative savings." However, page 40 indicates that "Sometimes there is enough of a heating or cooling signal to include a heating and/or cooling slope in the model, but the overall model does not fit the data well" (underline added). The ESA Evaluation was unable to accurately measure HVAC energy savings

[3] Ex ante ACT savings of 9.42% are based on Intertek tests of average EER* impact due to -5 to -10% refrigerant charge faults for non-TXV and TXV systems. R. Mowris et al. 2014. (pp. 7-9) R. Mowris et al. 2015. Laboratory Measurements and Diagnostics of Residential HVAC Installation and Maintenance Faults. EEDAL '15 Conference. https://e3p.jrc.ec.europa.eu/publications/proceedings-8th-international-conference-energy-efficiency-domestic-appliances-and-0. R. Mowris, E. Jones, R. Eshom, K. Carlson, J. Hill, P. Jacobs, J. Stoops. 2015. Laboratory Test Results of Commercial Packaged HVAC Maintenance Faults. Prepared for the CPUC. Prepared by Robert Mowris & Associates, Inc. (RMA). http://www.calmac.org/publications/RMA_Laboratory_Test_Report_2012-15_v3ES.pdf.

[4] Figure ES-1: Statewide Electricity Consumption per Household 6208 kWh where AC is 7% of total. Source: 2009 California Residential Appliance Saturation Study, Executive Summary, Prepared for California Energy Commission, Prepared by KEMA, Inc., October 2010, CEC-200- 2010-004-ES. https://www.energy.ca.gov/2010publications/CEC-200-2010-004/CEC-200-2010-004-ES.PDF

[5] Intertek tests indicate savings of 3.8 to 32% depending on cooling cycle duration and heating savings of 5 to 30.1% depending on heating cycle duration. See Mowris, R. Jacobs, P. 2016. Efficient Fan Controller® (EFC®) for Residential HVAC Systems. Work Paper EFC173PHVC138. Prepared by Verified® Inc. and Building Metrics Inc. (Available on request). Intertek. 2015. Performance Evaluation Based on Intertek Test Data of the GreenFan EFC Installed on Split and Packaged Air Conditioners with gas Furnaces. 101756555DAL-001B. Intertek. 2018. Performance Evaluation Based on Intertek Test Data of the GreenFan EFC Installed on Heat Pump and Hydronic Split Systems. 102791047DAL-001A.

due to issues with data cleaning, selecting appropriate comparison groups, model specification and validation, and savings being too small relative to whole building electricity and gas consumption (i.e., signal-to-noise ratio issues).

The SFD measure is always installed with the ACT measure, and the ACT measure includes repairs performed by technicians that make HVAC systems operate properly where they did not operate properly before repairs are made. Repairs include: tightening or replacing leaking refrigerant system Schrader valves and replacing relays, capacitors, contactors, transformers, and/or thermostat batteries. Repairs performed before installing ACT and SFD measures cause the following issues regarding billing analysis (cited by Ziemba 2017): “1) data cleaning and preparation, 2) selecting a comparison group, and 3) model specification and validation,” and 4) the “influence of data cleaning and preparation related to billing analysis results.”

The ACT measure includes condenser coil cleaning, air filter replacements, and refrigerant charge adjustments which improve cooling capacity by reducing AC compressor operation and AC compressor power compared to dirty coils/filters and improper refrigerant charge. The SFD delivers additional cooling and heating capacity to the conditioned space by providing a variable fan-off delay based on the cooling or heating cycle duration and dynamically adjusts the variable fan-off delay based on fault detection diagnostics of low cooling or heating capacity and severe weather. For cooling the SFD variable fan-off delay uses about 6.6% more electricity than the baseline (based on 10 minute average AC operation) to deliver additional cooling capacity, over satisfy the thermostat setpoint differential, and save energy by lengthening the duration of the AC compressor off cycle. For heating, the SFD uses about 31% more electricity for high-speed fan operation and variable fan-off delays.[6] For about 93% of gas furnace heating systems, the SFD operates the fan at a higher speed during the heating cycle to satisfy the thermostat sooner and reduce gas furnace operation. For all heating systems, the SFD variable fan-off delay delivers additional heating capacity at the end of the cycle to over satisfy the thermostat setpoint differential and save energy by lengthening the duration of the gas furnace or heating system off cycle. The SFD does not affect the thermostat setpoint or the duration of the thermostat call for cooling.

The repairs noted above cause customers to use HVAC systems more after the ESA ACT and SFD measures are installed, and this "take back" causes issues defining an appropriate baseline billing period (Ziemba 2017). Using non-participant billing data as a baseline would also cause issues since non-participants would not have had their systems repaired. Furthermore, the HVAC savings are small relative to whole building consumption, making billing analysis especially problematic for AC Tune-up and Smart Fan Delay measures. A different M&V strategy such as calibrated simulation modeling will provide more accurate results. The EFC workpaper (Mowris 2016) used Measure Analysis Software Control (MASControl2) to generate calibrated building energy simulation prototypes and post processing procedures which may provide a more appropriate and accurate evaluation methodology. Building prototypes can be calibrated to the ESA post billing data to appropriately evaluate PG&E ESA HVAC measures.

The primary conclusion of the ESA Evaluation states that “Ex ante savings assumptions were higher than achieved savings, with some measures leading to an increase in consumption.” This conclusion is incorrect due to reasons provided in these comments. The ACT and SFD measures

[6] Market research indicates that 6.7% of heating fans are not enabled to high speed fan operation by the SFD.

cannot “lead to an increase in consumption” because these measures increase cooling and/or heating capacity and thermal comfort, extend off-cycle time, and reduce HVAC energy use. Therefore, the ESA Evaluation study should be revised to include at least two explanations regarding the increase in energy consumption for HVAC measures: 1) customer “take back” due to repair of HVAC systems that would cause cooling and heating energy use to increase, and 2) issues with billing regression methodology being unable to accurately measure HVAC energy savings due to issues with data cleaning, selecting appropriate comparison groups, model specification, model validation, and savings being too small relative to whole building electricity and gas consumption (i.e., signal-to-noise ratio issues).

The following revisions are recommended prior to finalizing the ESA Evaluation to improve the study and avoid unintended negative consequences.

Section 1.6 (p. 7) Conclusions and Recommendations should also be revised as follows with the suggested underlined text to eliminate or avoid errors and omissions.

Please Replace: “1. Ex ante savings assumptions were higher than achieved savings, with some measures leading to an increase in consumption. ESA program planners should use the impact results to develop new ex ante savings assumptions.”

With: “1. Ex ante savings assumptions were higher than achieved savings, with some measures ~~leading to~~ appearing to cause an increase in consumption due to: 1) savings less than 10% causing signal-to-noise issues; 2) HVAC repairs causing issues with establishing correct counterfactual baselines; and 3) 30-35% low income customer bill subsidies through the CARE program (http://www.cpuc.ca.gov/general.aspx?id=976) causing “take back” which tends to increase ex post consumption. These issues cause significant uncertainty with respect to establishing a correct counterfactual baseline using billing analysis methodologies. Future evaluations should use a different M&V strategy such as calibrated simulation modeling to provide more accurate results for HVAC measures. ESA program planners should ~~use the impact results to~~ develop new ex ante savings assumptions using workpapers that provide all key assumptions including Measure Analysis Software Control (MASControl2) to generate calibrated building energy simulation prototypes and post processing procedures to provide a more appropriate and accurate evaluation methodology.”

Please Replace: “3. The evaluation methodology produced consistent year-over-year results at the household level.” Future Evaluations should use daily AMI consumption data for more robust results.”

With: “3. The evaluation methodology produced consistent year-over-year results at the household level but not at the measure level specifically for HVAC measures.” Future Evaluations should use daily AMI consumption data for more robust results and a different M&V strategy such as calibrated simulation modeling to provide more accurate results for HVAC measures.”

Please Replace: “4. There are limits to the answers that a billing analysis can provide for how program delivery affects (sic) savings. Future Evaluations should include a process evaluation to better research how program delivery is linked to impacts.”

With: "4. There are limits to the answers that a billing analysis can provide for how program delivery ~~effects~~ affects savings. Future Evaluations should include a process evaluation to better research how program delivery is linked to impacts. Billing analysis cannot be used to evaluate HVAC measures due to issues with data cleaning, selecting appropriate comparison groups, model specification, model validation, and savings being too small relative to whole building electricity and gas consumption (i.e., signal-to-noise ratio issues), making billing analysis especially problematic. A different M&V strategy such as calibrated simulation modeling will provide more accurate results for HVAC measures."

ABOUT DNV GL

Driven by our purpose of safeguarding life, property, and the environment, DNV GL enables organizations to advance the safety and sustainability of their business. We provide classification and technical assurance along with software and independent expert advisory services to the maritime, oil and gas, and energy industries. We also provide certification services to customers across a wide range of industries. Operating in more than 100 countries, our 16,000 professionals are dedicated to helping our customers make the world safer, smarter, and greener.