EVERGREEN ECONOMICS

PY2011 Energy Savings Assistance Program Impact Evaluation

Final Report

STUDY ID: SDG0273.01

August 30, 2013

Table of Contents

Executive Summary ..... **i**

ESA Program Delivery Overview I
Evaluation Objectives II
Analysis Methods IV
Evaluation Results IV
Electric Impact Estimates ..... *iv*
Gas Impact Estimates ..... *vi*
Impact Results Discussion ..... *vii*
Conclusions and Recommendations X

1 Introduction **1**

1.1 Program Background 1
1.1.1 Program Measures *2*
1.1.2 Program Services *3*
1.2 Evaluation Background 4
1.3 Organization of Report 6

2 Research Methods **7**

2.1 Conformity with the California Evaluation Protocols 7
2.2 Regression Models 7
2.2.1 Basic Model Specification *7*
2.2.2 Measure Model Specification *8*
2.2.3 Whole House Model Specification *9*
2.2.4 Additional Regression Models *10*
2.2.5 Measure Groupings *11*
2.2.6 Weather Variables *13*
2.3 Demand Impacts 15
2.4 Data Screening 18
2.5 Phone Survey 19

3 Regression Model Results **21**

3.1 Participant Data Analysis 21
3.1.1 Participation Patterns Across Climate Zones *21*
3.1.2 Analysis of Participation Energy Consumption *29*
3.2 Electric Models (Basic Model Specification) 31
3.3 Gas Models (Basic Model Specification) 34
3.4 Measure and Whole House Models 36

4 Impact Estimates **37**

4.1 Electric Impact Estimates 37
4.2 Gas Impact Estimates 39
4.3 Impact Results Discussion 40
4.4 Demand Impact Estimates 43
4.5 Whole House Impact Estimates 44
4.5.1 Electric Impacts *44*
4.5.1 Gas Impacts *44*
4.6 Impact Estimate Comparison with DEER Values 45

5 Phone Survey Results 49
5.1 PHONE SURVEY SAMPLE DESIGN 49
5.2 PHONE SURVEY RESULTS 53
5.2.1 Home Characteristics 54
5.2.2 Home Cooling 58
5.2.3 Home Heating 62
5.2.4 Added or Replaced Items in Home 66
5.2.5 Change in Number of People in Home 68
5.3 PHONE SURVEY RESULTS SUMMARY 72
6 Conclusions and Recommendations 74
6.1 CONCLUSIONS 74
6.2 RECOMMENDATIONS 75
7 Appendices 79

Executive Summary

The Energy Savings Assistance (ESA) Program, formerly referred to as the Low Income Energy Efficiency (LIEE) Program, provides energy efficiency measures and services at no cost to qualifying low-income customers of California's four investor-owned utilities (IOUs), Pacific Gas and Electric Company (PG&E), Southern California Edison (SCE), Southern California Gas Company (SoCal Gas), and the San Diego Gas and Electric Company (SDG&E). The ESA Program is administered by the IOUs in their respective service areas. This report presents the findings of the impact evaluation of the ESA Program for Program Year 2011 (PY2011) conducted by Evergreen Economics, CIC Research, and Michaels Energy.

ESA Program Delivery Overview

Initially established by the California Public Utilities Commission (CPUC) in the early 1980s,[1] low-income energy efficiency programs provided a channel for low-income customers to receive services similar to those provided by the energy efficiency programs instituted in response to the energy crisis of the 1970s. Subsequent legislation through the early 2000s continued to allow for the provision of energy efficiency measures to low-income customers in California.[2] Following the 2001 California energy crisis and an unanticipated increase in energy prices in 2005, the CPUC took increasingly aggressive approaches to low-income efficiency programs, expanding services and marketing activities, funding and income eligibility levels.[3]

In D. 07-12-051, the CPUC committed to expanding low-income programs by making them available to more customers, improving their cost effectiveness and designing them in ways to make them a reliable energy resource. To achieve these objectives, it adopted a programmatic initiative to provide all eligible low-income customers the opportunity to participate in the ESA Program and to offer those who wish to participate all cost-effective energy efficiency measures in their residences by 2020. The IOUs' 2009-11 ESA programs were to be treated as resource programs by focusing on energy savings, while improving the customers' quality of life. Budgets were also increased substantially in order to treat 25 percent of the overall 2020 goals within the 2009-11 program period.

Both home owners and renters may participate in the ESA Program if they have an account with an IOU offering the ESA Program and meet low-income qualifications. Eligibility for the ESA Program is determined by income-level and household-size guidelines established by the CPUC, which are updated annually to account for inflation. As indicated above, in 2005's Decision 05-10-044, the CPUC expanded the criteria for low-income program eligibility to include customers at or below 200 percent (an increase from 175 percent) of the Federal Poverty Level guidelines, regardless of elderly or disability status.

Customers may also be eligible to participate in the ESA Program if they have already been enrolled in one of the following low-income programs that require income verification:

---

[1] See CPUC D.92653, D.82-02-135, D.82-11-019 and D.82-11-086.
[2] See Pub. Util. Code § 2790, Pub. Util. Code § 382, SB 845, AB 1890, AB 1393 and SBX15.
[3] See CPUC D.01-05-033, D.01-08-065, D.05-10-044, D.06-12-036 and D.06-12-038.

- Bureau of Indian Affairs General Assistance
- CalFresh/Supplemental Assistance Program (SNAP)
- CalWORKS/Temporary Assistance for Needy Families (TANF)
- Head Start income Eligible (Tribal Only)
- Healthy Families A&B
- Low Income Home Energy Assistance Program (LIHEAP)
- Medicated/Medi-Cal
- National School Lunch Program (NSLP)
- Supplemental Security income (SSI)
- Tribal TANF
- Women, Infants, and Children Program (WIC)

Evaluation Objectives

The PY2011 ESA Impact Evaluation is one of four low-income program studies that the CPUC directed the IOUs to undertake in Decision 12-08-044. In this Decision, the CPUC directs the IOUs to conduct an impact evaluation of the ESA Program. To this end, the primary objective for the impact evaluation is to estimate first-year gas and electric energy savings, and coincident peak demand reduction attributable to the PY2011 ESA Program. The RFP issued for this study specifically directed that the energy impact estimates be provided in the following manner:

- In aggregate;
- By IOU service area;
- By average participant household;
- By measure and/or measure group; and
- Where possible and appropriate, by climate zone and housing type (multifamily, single family and mobile homes).

In addition to providing impact results, additional research goals were developed as part of the study's Final Research Plan to address issues that arose during the two prior impact evaluations of the ESA Program (covering PY2005 and PY2009), some of which were discussed as part of CPUC A.11-05-017 et al. Specific issues addressed in the current evaluation include the following:

1. **Data Screening.** For the PY2009 impact evaluation, data screens were used to remove those observations that represented either erroneous data entries or abnormally high usage points that would bias the billing model results. Although some data screening is necessary to estimate a billing model, concern was raised during the previous evaluation that the data screening process excluded too many high usage customers, and, had these observations been retained, ESA Program impacts would have increased substantially.[4] The issue was addressed in the PY2009 evaluation by re-estimating the billing regression model with less stringent

[4] See comments filed June 17, 2011 on the draft PY2009 impact evaluation by the CPUC Department of Ratepayer Advocates (DRA), The East Los Angeles Community Union (TELACU) and other community based organizations at http://docs.cpuc.ca.gov/efile/P/138446.pdf.

screens, and these results were reported to CPUC Energy Division and the IOUs in a memo and included as an appendix in the final PY2009 impact evaluation report.[5]

As discussed below, the PY2011 ESA Program impact evaluation uses a less stringent data screening process that only eliminates a small number of outlier observations.

2. **Savings estimates over time.** An important finding from the last several evaluations of the ESA Program is that savings tend to fluctuate over time. The fact that there are valid reasons why estimates might vary from year to year needs to be communicated better in the evaluation reports, and the reasons for these fluctuations better understood. The distribution of measures across customer usage groups and weather zones in a given year will change average savings levels, for example. This issue was explored in the PY2009 evaluation by examining changes in participation across usage categories and weather zones, and using information from phone surveys and on-sites. In the current evaluation, we provide a comparison of PY2011 impact estimates to the results from prior program year evaluations, as well as comparisons with both the *ex ante* and DEER savings values. Possible reasons for discrepancies across these sources are also discussed.

3. **Weather zones.** The PY2009 evaluation found that ESA participation had shifted to more moderate climate zones for some of the large weather-dependent measures. This led to lower impact estimates for these measures and resulted in a recommendation to focus ESA Program installation for these measures in the harsher climate zones. In the current evaluation, we continue to examine weather effects by analyzing how weather-normalized energy consumption changes between the pre-participation and post-installation periods for PY2011.

4. **Survey results.** The PY2009 impact evaluation included extensive phone survey and on-site data collection efforts, which provided some important insights into how customers use energy and the measures installed through the ESA Program. For instance, the surveys revealed that 34 percent of customers were not operating their evaporative coolers properly,[6] which helped explain why the impact estimates were lower than expected in the billing regression. For the current evaluation, a smaller and more targeted participant phone survey effort was conducted. The survey sample targeted customers who saw an increase in energy use after participating in the ESA Program, and questions explored possible reasons for the increase.

A Research Plan that addresses these issues was developed at the beginning of the PY2011 ESA Program Impact Evaluation. A Draft Research Plan was first posted on the CPUC website and a public workshop was held in San Francisco to present the plan and answer questions. Once the comment period ended, the plan was revised to address comments and a Final Research Plan was posted to the CPUC Energy Division website[7] on March 18, 2013.

---

[5] See Appendix E of *Impact Evaluation of the 2009 Low-Income Energy Efficiency Program*. Prepared for SCE and the CPUC by ECONorthwest (June 16, 2011), available on www.calmac.org.
[6] *Impact Evaluation of the 2009 Low-Income Energy Efficiency Program*. Prepared for SCE and the CPUC by ECONorthwest (June 16, 2011), p. 34.
[7] http://www.energydataweb.com/cpuc/home.aspx

Analysis Methods

There are two primary analysis components of this impact evaluation:

1. A fixed effects billing regression model was used to develop energy savings estimates (both kWh and therms) at the measure level for each IOU. The billing regression model relied on detailed information regarding which measures were installed through the ESA Program, combined with weather data and monthly energy consumption for both gas and electricity. All of this information was supplied to the evaluation team by the IOUs.

2. A phone survey was conducted on a sample of 602 participants that exhibited an increase in usage in the period directly after program participation. The goal of this survey was to collect information on customer behavior that would help illuminate why energy use was increasing.

Details on both of these evaluation components (and related analysis tasks) are included in the main body of this report.

Evaluation Results

The results of the regression models are used to calculate impacts for each measure group by IOU, house type and (where possible) climate zone.

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 impact estimates using these assignments are discussed below by fuel type. In most cases, the impact estimate from the Basic Model was used whenever possible.

Electric Impact Estimates

Table ES-1, Table ES-2, and Table ES-3 show the electric impacts by measure group. For each measure, the *ex ante*, Basic Model and Measure Model estimates are provided, along with information on the impact estimates from the PY2009 ESA Program evaluation. Note that in cases where the regression models estimate zero or negative savings (e.g., an increase in usage rather than a decrease), the estimated impact has been set to zero in the table. Our engineering team reviewed those measures where the algorithm assigned the *ex ante* values to assess if the *ex ante* values appeared reasonable. In the case of the SCE values for AC Tune-up and Pool Pumps, an alternative value was calculated based on engineering estimates for these measures.

The final impact number assignment is shown in the highlighted column of each table. Using the final assigned values, the total average household savings is shown at the bottom of the table for each IOU. The far right column of the tables also shows the impact estimates from the PY2009 evaluation, both at the measure-group and household level. Note that impacts on a per unit level (rather than per household, where multiple units may be installed) are shown in the detailed impacts estimates provided in Appendix D.

Once the final savings values are assigned and the whole house savings calculated, the aggregated effect increases total household savings slightly from the PY2009 evaluation for SCE, while SDG&E and PG&E both experience decreases relative to the previous evaluation estimates.

Table ES-1: SDG&E Electric Impact Estimates (kWh)

Measure	Households Receiving Measure	Basic Model	Measure Model	Average *Ex Ante* Savings	Final Assignment	Final Source	PY2009 Savings Estimate
Room AC	305	27.40	99.88	42.11	27.40	Basic Model	50
Central AC	30	N/A	N/A	38.66	38.66	*Ex ante*	50
AC Tune-up	59	N/A	N/A	229.13	229.13	*Ex ante*	326
CFLs	16,434	N/A	N/A	112.11	112.11	*Ex ante*	93
Ducts	937	55.72	1.36	0.00	55.72	Basic Model	-
Clothes Washer	1,667	123.05	86.94	528.57	123.05	Basic Model	788
Hardwired lighting	6,623	34.61	0.00	115.05	115.05	*Ex ante*	100
Insulation	800	85.53	359.74	94.90	85.53	Basic Model	104
Lighting	20,825	36.99	30.35	60.48	36.99	Basic Model	346
Microwave	1,852	0.00	66.52	175.91	66.52	Measure Model	-
Refrigerator	1,808	640.42	399.40	722.11	640.42	Basic Model	697
HW Conservation	1,334	85.19	60.30	172.03	172.03	*Ex ante*	24
WH Repair/Replace	5	0.00	0.00	0.00	0.00	*Ex ante*	-
Weatherization	16,703	0.00	0.00	49.59	49.59	*Ex ante*	63
Average household savings		**119.71**	**92.92**	**346.35**	**278.57**		**303**

Table ES-2: PG&E Electric Impacts (kWh)

Measure	Households Receiving Measure	Basic Model	Measure Model	Average *Ex Ante* Savings	Final Assignment	Final Source	PY2009 Savings Estimate
Central AC	79	141.04	116.53	317.35	141.04	Basic Model	50
AC Tune-up	12,143	0.00	0.00	230.04	230.04	*Ex ante*	326
CFLs	99,402	0.00	0.00	75.29	75.29	*Ex ante*	--
Ducts	3,007	112.26	10.59	94.33	112.26	Basic Model	--
Evaporative Cooler	5,841	0.00	0.00	262.15	262.15	*Ex ante*	502
Hardwired lighting	87,276	1.85	0.00	145.74	145.74	*Ex ante*	100
Insulation	6,290	145.41	0.00	46.69	145.41	Basic Model	104
Lighting	26,414	0.75	0.00	140.47	140.47	*Ex ante*	346
Refrigerator	16,773	655.36	427.92	766.89	655.36	Basic Model	697
HW Conservation	11	0.00	0.00	273.30	273.30	*Ex ante*	24
Weatherization	64,837	3.51	0.00	9.99	3.51	Basic Model	63
Room AC	3,175	0.00	0.00	111.56	111.56	*Ex Ante*	50
Average household savings		**113.11**	**64.47**	**381.46**	**366.90**		**402**

Table ES-3: SCE Impact Estimates (kWh)

Measure	Households Receiving Measure	Basic Model	Measure Model	Average *Ex Ante* Savings	Final Assignment	Final Source	PY2009 Savings Estimate
Room AC	927	0.00	57.51	69.47	57.51	Measure Model	50
Central AC	4,869	309.18	160.69	150.41	160.69	Measure Model	-
AC Tune-up	32	0.00	0.00	1265.00	257.00	Engineering Est.	326
CFL	67,872	71.25	82.25	25.44	71.25	Basic Model	93
Central Heat Pumps (CHP)	137	N/A	N/A	695.24	695.24	*Ex ante*	-
Ducts	4,490	0.00	20.65	0.00	20.65	Measure Model	-
Evaporative Cooler	15,928	239.16	448.48	481.87	448.48	Measure Model	502
Evaporative Cooler Tune-up	9	N/A	8236.20	37.13	37.13	*Ex ante*	-
Lighting	3,390	38.73	145.09	161.33	145.09	Measure Model	346
Pool Pump	1,908	0.00	0.00	1686.00	1088.00	Engineering Est.	-
Refrigerator	16,714	773.99	768.14	704.03	773.99	Basic Model	697
HW Conservation	505	720.97	1255.32	83.00	83.00	*Ex ante*	24
Weatherization	722	0.00	0.00	51.14	51.14	Ex ante	63
Average household savings		**230.31**	**270.46**	**253.38**	**279.26**		**247**

Gas Impact Estimates

The gas impact estimates are shown in Table ES-4, Table ES-5 and Table ES-6, and use the same savings assignment algorithm discussed above for the electric measures. Note that in cases where the Basic or Measure Model resulted in negative savings (an increase in usage), a savings value of zero is assigned to that measure for that model. At the household level, average household savings increased substantially for all three utilities relative to the PY2009 evaluation.

Table ES-4: SDG&E Gas Savings (therms)

Measure	Households Receiving Measure	Basic Model	Measure Model	Average *Ex Ante* Savings	Final Assignment	Final Source	PY2009 Savings Estimate
Ducts	930	14.54	13.48	0.00	14.54	Basic Model	-
Furnace Repair/Replace	3,666	0.00	0.00	0.00	0.00	*Ex Ante*	-
Furnace Clean & Tune	6,551	9.81	4.02	0.00	9.81	Basic Model	
Clothes Washer	1,585	15.88	14.42	35.88	15.88	Basic Model	-
Insulation	732	26.66	5.35	9.17	26.66	Basic Model	10
Pilot Light Change Out	985	15.10	18.50	11.85	15.10	Basic Model	-
HW Conservation	11,125	0.00	0.00	15.49	15.49	*Ex ante*	7
WH Repair/Replace	1,236	6.80	0.00	0.00	6.80	Basic Model	-
Weatherization	9,113	3.24	0.85	5.01	3.24	Basic Model	4
Average household savings		**13.14**	**6.87**	**21.99**	**26.06**		**8**

Table ES-5: PG&E Gas Savings (therms)

Measure	Households Receiving Measure	Basic Model	Measure Model	Average *Ex Ante* Savings	Final Assignment	Final Source	PY2009 Savings Estimate
Ducts	3,578	17.17	12.10	32.75	17.17	Basic Model	0
Furnace Repair	2,197	0.00	0.00	3.21	3.21	*Ex ante*	0
Furnace Replace	1,218	0.00	0.00	3.31	3.31	*Ex ante*	0
Insulation	7,165	44.50	22.13	61.05	44.50	Basic Model	10
HW Conservation	80,871	0.00	0.00	13.92	13.92	*Ex ante*	7
WH Repair/Replace	1,326	5.58	0.00	11.68	5.58	Basic Model	0
Weatherization	69,656	0.00	0.00	9.46	9.46	*Ex ante*	4
Average household savings		**3.82**	**1.99**	**23.29**	**21.50**		**9**

Table ES-6: SoCal Gas Savings (therms)

Measure	Households Receiving Measure	Basic Model	Measure Model	Average *Ex Ante* Savings	Final Assignment	Final Source	PY2009 Savings Estimate
Ducts	2,629	15.37	0.00	0.00	15.37	Basic Model	-
Furnace Repair/Replace	15,644	0.00	0.00	0.00	0.00	*Ex ante*	-
Furnace Clean & Tune	20,016	5.65	15.55	2.70	5.65	Basic Model	-
Clothes Washer	4,648	30.88	30.96	27.30	30.88	Basic Model	-
Insulation	8,225	26.51	17.49	7.76	26.51	Basic Model	10
Pilot Light Conversion	109	N/A	N/A	44.31	44.31	Ex ante	
HW Conservation	113,312	3.31	5.43	7.00	5.43	Measure Model	7
WH Repair/Replace	1,812	3.52	1.30	0.00	3.52	Basic Model	-
Weatherization	108,402	3.98	2.74	4.00	3.98	Basic Model	4
Average household savings		**11.31**	**12.90**	**12.58**	**13.40**		**11**

Impact Results Discussion

Despite the variation in impact estimates across program years and utilities, the current evaluation impact estimates are relatively close to the original *ex ante* values. Table ES-7 shows the realization rates at the household level, which is simply the estimated household savings using the current evaluation estimates divided by the estimated *ex ante* household savings. With the exception of the SDG&E electric measures, in general the evaluation estimates are reasonably consistent with the *ex ante* values. The realization rate metric is somewhat misleading in this application, however, as some of the evaluation assigned values were in fact the *ex ante* values, which move the realization rate closer to 1.0. Therefore, the realization rate as calculated here should not be interpreted as a confirmation of the *ex ante* values, as several of the *ex ante* values are used in the calculation. Nevertheless, the realization rate metric does show that the savings values recommended by the evaluation team are fairly close to the original savings estimates provided by the IOUs.

Table ES-7: ESA Impact Evaluation Realization Rates

	Evaluation Savings	*Ex Ante* Savings	Realization Rate
Electricity (kWh)			
SDG&E	278.57	346.35	0.80
PG&E	366.90	381.46	0.96
SCE	279.26	253.38	1.10
Gas (therms)			
SDG&E	26.06	21.99	1.19
PG&E	21.50	23.29	0.92
SoCal Gas	13.40	12.58	1.07

While there is some consistency with current evaluation savings estimates and the *ex ante* values at the household level, there are some obvious differences in savings estimates for individual measures. The electric impact models provide a range of savings estimates – some of which have internal consistency while other measures show significant variation across utilities, previous evaluation results, and individual *ex ante* values. While we attempted to explore reasons for these differences, it was not possible with the current budget and timeline to explore in-depth all the possible reasons for variations across models, utilities, and the results from the previous evaluation.

It is also important to note that – as discussed in the previous impact evaluation – there are legitimate reasons for savings numbers to vary both across time and utilities. In particular, with regard to comparing evaluation estimates across time, one must not conclude from these differences that one set of estimates is 'correct' or 'more accurate' than the other; the estimates may be equally accurate but reflect different baseline, program, or market conditions inherent in the different evaluation periods.

Table ES-8 shows the current PY2011 impact estimates compared with the whole house savings estimates from prior evaluation years. Since 2000, there has been a wide range of savings estimates for both gas and electricity at the household level. For electricity, the current impact estimates are lower than those from PY2009 and PY2005, but in line with estimates from PY2000 thru PY2002. For gas, the current impact estimates are significantly higher than those from PY2009 and generally consistent with impacts from earlier evaluations.

Table ES-8: Impact Estimate Comparison with Prior Evaluations

	PY2011 Evaluation	PY2009 Evaluation	PY2005 Evaluation	PY2002 Evaluation	PY2001 Evaluation	PY2000 Evaluation
Electric Savings (kWh)						
PG&E	367	402	433	399	236	240
SCE	279	247	435	286	203	153
SDG&E	279	303	342	370	215	89
Gas Savings (therms)						
PG&E	21	9	19	9	18	28
SDG&E	26	8	14	4	13	13
SoCal Gas	13	11	17	17	20	26

There are a multitude of factors that can result in different levels of savings across program years and utilities, and some of the more prevalent influences are discussed below.

Energy consumption. Households that use more energy may have the potential for greater energy savings, depending on what end uses are driving energy consumption. Differences in household energy use across both utilities and evaluation periods may account for some of the differences observed in the estimated energy savings. Additionally, it is not just the levels of energy use that are important, but also the degree to which energy consumption changes between pre-participation and post-participation periods. Changes in energy use between these two periods (and the degree to which this inter-period change differs from changes in other utilities and time periods) will also result in different impact estimates.

Household composition and home characteristics. One of the most important factors determining energy use is the number of occupants within a home. Those households with more people typically use more energy (all else equal). Similarly, differences in the household structures themselves will lead to differences in energy impacts. Homes with larger or older structures will likely have a greater potential for energy savings, as will homes in disrepair (requiring more energy to heat and cool) or older appliances (requiring more energy to run).

Weather. Weather has an important influence on energy savings, particularly for those measures where energy use and savings will vary with changes in temperature. In the current evaluation, weather is incorporated directly into the savings calculations for those measures where we can reasonably expect savings to vary with changes in temperature. The discussion later in this report illustrates how weather has changed between the current and prior evaluations, both in terms in the amount of heating degree and cooling degree days, as well as the distribution of participants across climate zones. Also note that – while the climate zones have been defined to have similar weather within each zone – there is still often significant variation in temperatures within a climate zone, particularly for those zones that include the hottest and coldest areas.

Measure mix. The amount of total household savings will vary by the types and quantity of measures installed. This is important to remember when considering that many of the savings estimates from the regression models are for groups of measures, such as weatherization and hot water conservation. While these are by necessity modeled as a single group in the regression (to mitigate the estimation problems associated with collinearity), customers may have different amounts of individual measure components installed within each measure group. These differences in measure group composition will lead to differences in savings estimates across utilities and across evaluations.

Different estimation methods. For the current evaluation, we have used the same model specification and data screening process for each utility, so different analysis methods will not explain differences in the current estimates across utilities. The current models, however, are different than what were used in the previous two impact evaluations (PY2009 and PY2005), which in turn were different than the models used in the earlier evaluations (PY2000, PY2001, PY2002). We attempted to develop impact estimates in the current evaluation using the same model specification from the 2009 evaluation, but this was abandoned due to high collinearity issues and because many of the measure-level impact estimates were showing either no energy savings or increased energy use. While we believe that the current models are an improvement over earlier evaluations, the different specifications will result in different energy savings estimates.

Savings small relative to overall energy consumption. For many of the measures installed in the ESA program, the amount of savings expected is small relative to overall household consumption. This is particularly true for some of the most common measures such as CFLs, lighting, weatherization, and hot water conservation. Given the small amount of savings, it is challenging to develop rigorous estimates that are consistent across utilities and evaluations from prior years – even when the exact same model specifications are used. The small amount of savings involved, combined with a lack of information on other influencing factors (discussed above) can result in the ESA savings being overwhelmed in the regression model by these other forces.

Conclusions and Recommendations

General conclusions that can be drawn from the impact analysis results include the following.

Savings from the ESA Program measures is a small fraction of overall household energy consumption. Savings from the ESA program on average ranges from three to nine percent of overall energy consumption. This low level of savings makes developing savings estimates (particularly at the measure level) particularly challenging. These challenges are compounded by the wide array of external factors that can influence energy use. As discussed throughout the report, the small amount of program savings is sometimes overwhelmed by these other non-program factors in the billing regression and result in estimates of no savings or increased energy use for some measures.

The final impact estimates are generally consistent with the *ex ante* savings values. The final recommended impact values for both electric and gas measures resulted in total household savings that were fairly close to the original *ex ante* savings values. For electricity, household realization rates ranged from 80 to 110 percent of *ex ante* savings. For gas, realization rates ranged from 92 to 119 percent. Note that this consistency with the *ex ante* values is due in part to how the final impact numbers were assigned from either the regression models or *ex ante* values. Since the *ex ante* values were used as the final impact estimates in cases where the regression models did not produce a reliable estimate, the potential for differences with the *ex ante* values was naturally reduced.

The impact estimates deviate from the previous evaluation and from DEER values. For electric measures, estimated savings in the current evaluation are lower than estimates from PY2009, while gas estimates in the current evaluation are significantly higher. In the case of the gas savings, this may be due to significantly more heating degree days in the current evaluation relative to the last. The current impact estimates are within the range of those observed in previous evaluations going back to 2001, however, as there is substantial variation in household savings estimates over the years. The current evaluation estimates were also different from DEER values for the same measures, although no trend of being consistently higher or lower than DEER at the measure level was observed.

Impact estimates will naturally vary across years due to a variety of factors. Differences across customer groups in terms of energy use, geographic location, measure mix, demographics, economic situation, and condition of the home will all lead to differences in impact estimates for the ESA Program. We should not expect these estimates to be the same across time or across service territories due to the large number of potential influencing factors. In the current evaluation, differences from the prior evaluation may also be due to the utilization of a different regression model and data screening process. While identifying these influencing factors is straightforward, determining the relative importance of each of these factors on the change in savings values between years is not possible without significantly more evaluation resources being devoted to making a

detailed comparison of participation patterns between years. Given that the primary objective of this impact evaluation is to develop impact estimates for the current program year, a more detailed analysis was not attempted beyond the comparisons presented earlier in this report.

A significant number of ESA participant households are using more energy after participation. Despite the new measures and energy education received through the program, a significant number of households were found to be consuming more energy after participation. For electricity, more than half all of all participants exhibited weather-normalized increases in energy use during either heating or cooling periods. Similarly, approximately 60 percent of gas participants increased their gas consumption in the post-participation period. Because this increase appears to be independent of weather, it is especially challenging to address in the billing regression and may lead to biased impact estimates. The phone survey did not provide any additional information as to what might be causing this increase in energy use. Since the vast majority of participants were already on the CARE rate prior to ESA enrollment, it is unlikely that the lower CARE rate is a factor in increased energy use for the time period examined.

Whole house impacts estimated from the household-level regression models produced lower estimates. The results from the Whole House fixed effects models that estimate total savings (rather than savings for individual measures) produced generally lower house-level savings values than simply aggregating up the measure-level savings from the Basic and Measure Models. This is due in part to the ability with the Basic/Measure models to remove impact estimates showing an increase in energy use and replacing them with the *ex ante* values, which by definition will increase the overall savings estimate. Since measure-level detail is not available in the Whole House model, it is not possible to make these types of post-model adjustments.

While it was hoped that having a whole house variable for savings would help address the possibility of collinearity among the measure variables, this advantage appears to have been outweighed by a lower ability to disentangle the program effects from other factors influencing energy consumption. This is particularly challenging given the number of homes observed to have an increase in energy use in the post-participation period (particularly with PG&E). Given this context, it is not surprising that the Whole House model (which utilizes less program information) produces lower savings estimates than the Basic Model that utilizes more information on what was installed through the program.

Customers may be unaware that they are using more energy. The phone survey targeting households with increased energy use did not provide any clear answers on what might be driving the higher consumption. Respondents generally reported that they were using their heating and cooling systems about the same as they did prior to participation. For those that said they used the systems more, the most common reason for using heating and cooling systems more had to do with changes in weather (e.g., hotter or cooler weather). As shown in the analysis of weather-normalized energy use, changes in weather are not sufficient to explain all of the increase in usage. Other factors, such as having more people home during the day, did not appear to be a significant factor in explaining increased use. While participants have been adding new appliances to their homes, these appear mostly to be replacing older units and therefore should be using less energy. These findings raise the possibility that – despite the new measures and energy education – consumers are using more energy and (perhaps more importantly) they are unaware that they are consuming more energy. The issue of whether they were truly unaware was not explored directly in the phone survey, however.

From the evaluation conclusions, we offer the following recommendations for the ESA Program.

Continue using billing regression to estimate program impacts. Despite some of the challenges discussed in this report, we recommend that the fixed effects billing regression model continue to be used to estimate impacts for the ESA Program using data from the participant population. The fixed effects model provides a means for producing statistically reliable and unbiased estimates of savings that account for both differences across households and time periods.

For future impact evaluations utilizing a billing regression, developing multiple model specifications provides more flexibility. If billing regression is to be used in future ESA Program evaluations, we recommend an approach that combines results from the Basic and Measure Model specifications presented here. While this does rely on evaluator judgment to make some impact assignments, the approach is ultimately more flexible than relying on the results of a single model. In the current evaluation, having multiple models resulted in impact estimates for some measures that could not have been provided using the Basic Model alone.

If variations in impact estimates over time are not acceptable, consider using DEER deemed values to estimate savings. The wide swings in savings estimates – both across utilities and evaluation time periods – has raised concern among some reviewers. Possible reasons for these discrepancies have been discussed in the last two impact evaluations, and variations will continue in the future. It is also stressed again here that the exact cause of these differences will likely remain unknowable without an enormous data collection effort that collects statistically representative data on home and customer demographics within each utility service territory by housing type, climate zone, and possibly additional household characteristics such as family size and home vintage. Short of a massive data collection effort, the root causes of energy savings variation across utilities and program years will likely remain unknowable.

As argued in this report, we do not believe that the variation in savings estimates is necessarily a bad thing. Nevertheless, if more consistency in the impact estimates is desired, then using deemed savings values from DEER in place of a billing regression should be considered. This deemed approach will reduce uncertainty with respect to savings estimates across utilities within a program year, as well as produce more stable savings estimates across program years. Using DEER, however, does not allow for the possibility that the low-income population is significantly different in terms of energy savings relative to the general population. While testing this theory is beyond the scope of this project, it may be worth reducing the uncertainty in savings estimates by using DEER even if that database is not an entirely accurate representation of the savings achieved in the low-income sector.

Weather variables should be calculated using hourly (rather than daily) temperature data. The calculations of CDD and HDD using hourly temperature data allow for a more accurate representation of days that heating or cooling equipment might be used. In this evaluation, the hourly method resulted in significantly more cooling degree days and only slightly more heating degree days then the traditional daily method. Given that the hourly method is more accurate and easy to calculate, we recommend that it be used for future impact evaluations of this program.

Allow more time for the impact evaluation. The time allocated for this evaluation was very short (five months), with a research plan finalized on March 1 and a final report produced by August 31. For comparison, the previous impact evaluation took 20 months. While the current impact evaluation

was completed in the time allotted, this was accomplished by having a very focused approach that did not allow for exploring additional research questions when they arose. For example, more time might have allowed for additional analysis of the survey data, or even a short follow up survey to explore other aspects of energy use that might have shed more light on increased energy consumption. Similarly, there was not enough time to conduct a more in-depth comparison of the impact estimates between the 2009 and 2011 evaluations to determine how changes in participation patterns, measure mix, and weather might have contributed to differences in impact estimates between the two years. Adding three to six months to the impact evaluation timeline would allow for a more in-depth and flexible approach that provides more insights into the ESA Program savings estimates.

Conduct a more rigorous analysis of participation patterns across evaluation years. As mentioned above, the current evaluation did not have enough time to conduct a rigorous comparison of participation patterns between PY2009 and PY2011. While this evaluation did provide some information on weather conditions and participation across climate zones between the two evaluation years, the primary focus was in developing defensible savings estimates for the current evaluation year. Additional analysis on changes in participation patterns in terms of measure mix, housing type, energy use, weather conditions, and geographic distribution would likely provide additional insights as to the factors driving the variation in savings estimates across program years. We recommend additional time and budget be allocated for this analysis in the next ESA Program impact evaluation.

Continue with current evaluation cycle timing. The last several impact evaluations have focused on a single program year and have occurred every 2-3 years, and we recommend that this cycle continue. Given that the savings levels will change regularly due to weather, measure mix, and participant characteristics, the evaluation should also be conducted at regular intervals in order to reflect this variation. This is especially important when the impact evaluation results are used to set the *ex ante* savings values for future program years. If impact evaluations are done less often, or are done for multiple evaluation years combined, then some of the inherent variability will be lost due to the timing and structure of the evaluation cycle. This may result in less accurate impact estimates moving forward, particularly if the market is shifting and the programs are locked in to using fixed impact estimates for a longer period of time until a new impact evaluation can be completed. Having the evaluations done more often (instead of every five years, as has been suggested) will provide flexibility to adjust the energy savings estimates as needed to reflect changing demographics and market conditions.

Remember lessons from previous evaluations. Finally, reviewers raised a couple of issues that relate to analysis methods that were explored in the previous impact evaluation. These are methods that were recommended by reviewers of this current report as possible methods to consider in the future:

- **Billing regression using additional survey data.** A common approach for obtaining additional customer information for use in a billing model is to conduct a phone survey of program participants that asks detailed questions about their home and factors that may have changed since participating in the program. This approach was used in the PY2009 ESA impact evaluation but did not yield useful results for the impact analysis. While in theory it might be valuable to have survey data that provide additional explanatory variables in the billing regression, in practice this did not result in an improved billing model in the PY2009 evaluation. Consequently, we do not recommend this approach for the billing regression in

future evaluations and instead recommend that the billing models rely on the ESA participant population.

- **Billing regression using on-site data.** Customer on-sites can be used to collect additional information on home characteristics that can be used as additional variables in a billing regression model. This method was also used in the PY2009 impact evaluation and did not provide credible impact estimates. The on-sites are also expensive to conduct, especially if a large enough sample is needed to be representative for a billing regression. We also do not recommend conducting on-sites in future ESA Program evaluation if their primary purpose is to collect data to support a billing regression. The on-sites may be useful for other purposes, however, such as providing additional information on baseline conditions, customer attitudes toward efficiency and energy use, whether or not installed equipment is being used properly, and other factors that affect energy consumption.
- **Billing regression using a control group of non-participants.** The PY2009 evaluation also developed a billing regression that utilized a control group of low-income non-participants, where the PY2010 participants were used as a non-participant control group for PY2009. The theory underlying this method is that the control group customers will have similar patterns of energy use as participants and therefore will control for external events such as economic conditions within the model.[8] Selecting a well-matched control group is challenging at best, however, and particularly difficult in the low-income population given the variability across program years. Using the control group did not produce useful billing regression results in the previous evaluation, and we are not optimistic that these challenges can be overcome in future evaluations without significantly more resources being devoted to identifying an appropriate control group. Despite these concerns, future evaluations may want to explore the potential benefits of using a control group if there is a way to ensure that the control group matches the participant population on key demographic variables (e.g., home type, energy use, geographic location, vintage, etc.). Exploring the use of several alternative control groups in the billing regression may also prove useful, as this was not attempted in the previous impact evaluation.

[8] The control group also helps account for free ridership in the model, which is less of a concern with the low-income population where free ridership rates are likely very low.