Impact Evaluation of 2014 Marin Clean Energy Home Utility Report Program (Final Report)

California Public Utilities Commission

Date: 04/01/2016

CALMAC Study ID: CPU0126.01

LEGAL NOTICE

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

Table of contents

List of tables

List of figures

1 EXECUTIVE SUMMARY

This report summarizes the results of DNV GL's impact evaluation of the Marin Clean Energy (MCE) 2013-2014 Home Utility Reports (HUR) program.

1.1 Background

The HUR program provides comparative energy usage information similar to the Home Energy Reports (HER) programs run by Pacific Gas & Electric (PG&E) and other investor-owned utilities (IOUs). It also encourages customers to go to the MCE website for more customized information regarding contractors, financing, and rebates.

MCE structured the HUR program as a randomized controlled trial in which the eligible population was randomly assigned to the treatment and control groups. There were three waves of promotion during the period of time studied by this impact evaluation. Table 1 presents basic information about the three waves, including the number of households that received comparative energy usage reports (treatment customers), the frequency with which they received those reports, and the counts of control group customers.

Table 1: MCE HUR Program Waves, Frequency of Reports, and Program Start Dates

Wave	Frequency of Report / Target Group	Program Start Date	Control Customers	Treatment Customers
HUR-1	Monthly/Top usage quintile	Nov 2013	2,766	3,643
HUR-2	Monthly	Mar 2014	5,934	6,560
	Quarterly			6,586
HUR-3	Bi-monthly/Top two usage quintiles	Nov 2014	2,114	4,233

1.2 Research questions and objectives

The primary objective of this evaluation was to provide independent verification of electricity savings attributable to the HUR program. Specific research questions included the following:

- Is the experimental design employed by MCE acceptable?
- What are the energy savings for each HUR cohort (monthly, bi-monthly, and quarterly)?
- Are there downstream/upstream rebate program savings that could be jointly claimed by both the HUR program and PG&E rebate programs?

1.3 Study approach

To answer these research questions, DNV GL conducted an impact evaluation for the first 14 months of the 2013-2014 program cycle. This evaluation included two major tasks:

1. **Validate MCE's experimental design**. DNV GL reviewed MCE's experimental design of the HUR program to ensure the validity of this impact evaluation.
2. **Calculate program savings**. DNV Gl calculated the overall (unadjusted) savings, the joint upstream/downstream savings that could be claimed by both HUR and PG&E, and the final adjusted program savings (i.e., overall savings minus joint savings) to identify the savings attributable to the HUR program.

1.4 Key findings

Table 2 shows the estimated savings for the 2013-2014 HUR program, broken out by wave. DNV GL found that the MCE HUR program did not achieve any detectable electric savings in any of the three waves. In fact, DNV GL found slight increases in consumption across the span of each wave, though none of these estimates were statistically different than zero. The first HUR wave (HUR-1M) showed slight positive savings during 2014, but these savings were also not statistically significantly different from zero.[1]

Table 2: Program-Level Savings Estimates for 2013-2014

			Source			
Wave	**Evaluation Period**	**Unadjusted Savings**	**Tracked, Downstream Joint Savings**	**Untracked, Upstream Lighting Joint Savings**	**Adjusted Savings**	**Statistically Significant with 90% confidence?**
		Electric (MWh)				
HUR-1M	November 2013 - December 2014	-47.9	0.1	-	-47.9	No
	January 2014 – December 2014 (2014 only)	64.1	0.2	-	63.9	No
HUR-2M	January 2014 - December 2014	-15.2	5.3	-	-20.6	No
HUR-2Q	January 2014 - December 2014	-46.2	2.8	-	-49.0	No
HUR-3B	January 2014 - December 2014	-2.3	0.2	-	-2.5	No
Total	November 2013 - December 2014	-111.6	8.4	-	-120.0	No

Table 2 also shows the downstream joint savings, which were subtracted from the unadjusted savings total to produce the adjusted savings total; this adjustment was performed to address the potential for "double-counting" savings already claimed by PG&E programs. While there is evidence of joint upstream savings, as well, DNV GL did not calculate upstream savings or further adjust the results because: 1) the savings results are negative, 2) we are currently working with the IOUs to update the upstream savings algorithm, and 3) calculating the numbers would have no effect on the evaluation outcome.

Table 3 provides estimates of unadjusted and adjusted savings at the household level for the treatment group as compared to the control group. The per-customer savings make it clear that the magnitude of the negative savings is extremely small.

[1] Statistically significantly different than zero at 90% confidence indicates a relative precision of 90/99 or better. That is, the 90% confidence interval is less than the magnitude of the estimated savings. Generally HER results are expected to achieve precision on the order of 90/20 or better.

Table 3: Average Electric Savings per Household as a Percent of Consumption

Wave	Evaluation Period	Unadjusted kWh per Customer Savings	Adjusted kWh per Customer Savings	kWh per Customer Consumption	Unadjusted Savings as % of Consumption	Adjusted Savings as % of Consumption
HUR-1M	November 2013 - December 2014	-14.3	-14.3	11,228.8	-0.1%	-0.1%
	January 2014 - December 2014	16.8	16.7	9,464.2	0.2%	0.2%
HUR-2M	January 2014 - December 2014	-2.2	-3.0	4,945.6	0.0%	-0.1%
HUR-2Q	January 2014 - December 2014	-7.0	-7.5	4,945.6	-0.1%	-0.2%
HUR-3B	January 2014 - December 2014	-0.6	-0.6	7,796.9	0.0%	0.0%

While randomized control trials give highly precise and un-biased estimates of savings, they do not provide any insight into what worked or did not. In this case, a low-level overlap with an MCE school program (discussed in Chapter 2 of this report) and some shortcomings of the HUR program's experimental design (discussed in Chapter 4) could be contributing to the lack of savings. However, had the program generated the 1 to 3% savings that other behavioral programs have offered, those savings would have been detectable despite the program overlap and experimental design issues.

Ultimately, the success of a behavioral program is driven by the effectiveness of the reports and the willingness and ability of the targeted populations to decrease their energy consumption. Any of these factors, individually or in combination, could explain the lack of response to the HUR program.

2 INTRODUCTION

The California Public Utilities Commission (CPUC) engaged DNV GL to conduct an impact evaluation of the Marin Clean Energy (MCE) 2013-2014 Home Utility Reports (HUR) program. This impact evaluation uses HUR program tracking data provided by MCE and monthly consumption data provided to the CPUC by Pacific Gas & Electric (PG&E). The evaluation provides independent verification of electricity savings attributable to the HUR program.

2.1 HUR program description

2.1.1 Overview

MCE began implementing the HUR program in 2013. This direct engagement program delivers normative-comparative messages via direct mail in order to motivate customers to change their energy use behavior. The messaging provides information similar to that found in other comparative feedback reports (consumption information, comparison with similar neighbors, and customized tips for saving energy). The program also encourages customers to go to MCE's website for additional information regarding contractors, financing, and rebates.

2.1.2 Potential overlap with school program

In addition to the HUR program, MCE also implemented a school program that offered a specially crafted curriculum and provided students with a kit of energy-saving measures (5 CFLs, 1 showerhead, 1 aerator, and 1 filter whistle). Students were required to sign a pledge stating they would install the equipment. Early in the program, MCE dropped the kit measures because they were not cost-effective and required too much time to distribute.

This evaluation does not cover the MCE's school program; however, it is likely that some households with students participating in the school program also received the HUR direct mail, resulting in some low-level overlap between the programs. The school program was not tracked, so this overlap cannot be quantified. Even so, DNV GL believes it is unlikely that this overlap had substantial effect on the HUR program, for the following reasons:

- The school program had relatively limited impact.
- Because the treatment and control groups are randomly distributed across the area, there is no compelling reason to expect that the school program impacts would not be approximately randomly distributed across the treatment and control groups.
- Only where the school program efforts were redundant with HUR program efforts would we expect the overlap to moderate the HUR program savings estimates.

2.1.3 Experimental design

MCE implemented the HUR program using a randomized controlled trial (RCT) experimental design to facilitate estimating program savings. The RCT experimental design randomly assigns a population of interest to control and treatment groups. Only the treatment group receives program messaging/reports.

This approach effectively establishes a causal relationship between treatment and the effect, in this case a possible change in consumption. This approach produces an unbiased estimate of this change with a high level of statistical precision, and is widely considered "the gold standard" in program evaluation.

MCE engaged Planet Ecosystem (PEI) to develop the sample for the HUR program. PEI developed a universal group for the different waves using the criteria shown in Table 4.

Table 4: Criteria for HUR waves

HUR-1	HUR-2 and HUR-3
• MCE customers • Single-family homes in Marin County • Non-medical rate • Electric rate schedule is E1 or EL1 (CARE) • Latitude and longitude values are not outliers by more than 2 sigma • Have known square footage • Had 12 months of usage data at program start • Name field did not appear to be a small business	• MCE customers • Single-family homes in Marin and the city of Richmond • Non-medical rate • Electric rate schedule is E1, EL1, or E6 • Latitude and longitude values known • Have known square footage • Had 11 or 12 months of usage data at program start • Name field did not appear to be a small business

PEI applied additional restrictions to the universal group to develop the sample for the HUR waves. Households were only included in the randomization if they met the criteria shown in Table 5.

Table 5: Criteria for Inclusion in Sample

HUR-1	HUR-2	HUR-3
• Households in top usage quintile • Not in the treated or control group of the PG&E HER program • Home has at least 50 neighbors	• Not in the treated or control group of the PG&E HER campaign • Not in the treated or control groups for any other MCE HUR program • All usage quintiles • Home has at least 50 neighbors	• Not in the treated or control group of the PG&E HER campaign • Not in the treated or control groups for any other MCE HUR program • Usage for the previous 12 months placed the home in roughly the top two quintiles (top 40%) when compared to their neighbors • Home has at least 50 neighbors

Note: For HUR-1 and HUR-2, a neighbor is defined as any home in the universal group within 1 mile radius and with square footage within +/-10%. For HUR-3, a neighbor is defined as the nearest neighbor within the universal group, given a maximum radius of 2 miles and +/- 250 square feet.

Table 6 presents the three HUR waves with corresponding program start date and number of households in the treatment and control groups. The report counts of customers are based on the tracking data received from MCE.

Table 6: HUR Experimental Waves and Launch Dates

Wave	Frequency of Report/Target Group	Program Start Date	Control Customers	Treatment Customers
HUR-1	Monthly/Top usage quintile	Nov 2013	2,766	3,643
HUR-2	Monthly	Mar 2014	5,934	6,560
	Quarterly			6,586
HUR-3	Bi-monthly/Top two usage quintiles	Nov 2014	2,114	4,233

After the experimental design was set, MCE stopped sending reports to lower consumption quintiles in the HUR-1 wave. The best practice in these situations is to use the original design for the evaluation. Any savings that exist among those who did receive the reports should still be measured and fully accounted.

Because savings will be spread over the full number of treatment group households, the actual magnitude of average household savings may be smaller; this could have an effect on precision. Under the circumstances, however, it is better to accept the potential reduction in precision than potentially undermine the validity of the experiment altogether.

2.1.3.1 Random allocation process

MCE randomly assigned all three HUR waves to treatment and control groups with no additional stratification. After finalizing the HUR-1 selection, the treatment and control groups were found to be substantially unbalanced. As a result, for HUR-2 and HUR-3 waves, MCE repeated the random selection process several times until the treatment and control groups for both waves demonstrated balance among available parameters.

This situation reflects an ongoing experience in the area of behavioral programs, and represents a cautionary tale. While the savings estimation techniques will control for mean differences across the treatment and control samples (as with HUR-1), a balanced set of treatment and control groups is desirable. The solution to this problem, however, is not multiple random allocations to find a suitable balance.[2] The preferred approach is to use the available data to stratify the population and perform the random allocation within those strata. Taking this approach greatly increases the likelihood that the overall allocation will be balanced with respect to all or most characteristics, and makes it more likely that the samples will be amenable to analysis by subsets defined by those characteristics.

MCE supplied information on sampling procedures and results from statistical tests employed. Table 7 provides the results from MCE's randomization tests comparing treatment and control differences with respect to eight household characteristics (e.g., number of occupants, number of bedrooms, etc.). HUR-1 showed substantial imbalance in five out of eight household characteristics, while HUR-2 and HUR-3 showed no indication of statistical differences with respect to most of the parameters tested. In Section 4, these results are replicated for this evaluation.

Table 7: HUR Program Balance Test: t-test p-values

Household characteristics	HUR-1	HUR-2	HUR-3
Home area (sq. ft.)	0.00*	0.22	0.26
Number of occupants	0.30	0.27	0.46
Number of bedrooms	0.00*	0.43	0.50
Number of bathrooms	0.00*	0.03*	0.47
Zip code	0.02*	0.27	0.39
Home construction year	0.00*	0.37	0.20
Number of children	0.25	0.32	0.11
Number of adults	0.24	0.42	0.02*

2.2 Evaluation objectives and approach

The primary objective of this evaluation was to provide independent verification of electricity savings attributable to the HUR program. Specific research questions included the following:

- Is the experimental design employed by MCE acceptable?
- What are the energy savings for each HUR cohort (monthly, bi-monthly, and quarterly)?

[2] The SEEAction Report does put this method forward as an option, though in subsequent protocols the authors have responded to feedback and changed this recommendation. Citation in subsequent footnote.

- Are there downstream/upstream rebate program savings that could be jointly claimed by both the HUR program and PG&E rebate programs?

To answer these research questions, DNV GL conducted an impact evaluation for the first 14 months of the 2013-2014 program cycle. We began the study by reviewing the program's experimental design to verify the validity of this impact evaluation. Our assessment of the experimental design is discussed in Section 4.

Next, we estimated three categories of program savings:

1. **Overall (unadjusted) savings.** Using a fixed effects regression model, DNV GL compared the pre- to post-program difference for a treatment group to the pre- to post-program difference for a control group. The change that occurred in the treatment group was adjusted to reflect any change that occurred in the control group, in order to isolate changes attributable to the HUR program.
2. **Joint savings.** DNV GL estimated the savings achieved by the HUR program in concert with PG&E energy efficiency programs. This estimate is normally produced for two areas:
   - Downstream joint savings due to an increased participation by the treatment group versus the control group in PG&E's tracked energy efficiency programs due to the HUR program.
   - Upstream joint savings due to an increase in adoptions by the treatment group versus the control group of measures promoted in PG&E's Upstream Lighting Program (ULP). DNV GL did not produce an estimate for the upstream joint savings since there were no overall savings produced indicating the possibility of no savings occurring due to upstream programs.
3. **Adjusted savings**. DNV GL calculated the adjusted savings estimate by removing the joint savings (downstream only) from the overall savings to avoid double-counting savings potentially already claimed by PG&E.

The results of these savings calculations are presented in Section 5.

3 METHODOLOGY AND DATA SOURCES

3.1 Methodology

For this evaluation we used a fixed-effects regression model that is the standard for evaluating behavioral programs like HUR. The model produces a "difference of differences" calculation by comparing the pre- to post-program difference for the treatment group to the pre- to post-program difference for the control group. The change that occurs in the treatment group is adjusted to reflect any change that occurred in the control group, in order to isolate changes attributable to the program.

The fixed-effects equation is:

$$E_{it} = \mu_i + \lambda_t + \beta P_{it} + \varepsilon_{it}$$

Where:

E_{it} = Average daily energy consumption for account i during month t
P_{it} = Binary variable: one for households in the treatment group in the post period month t, zero otherwise
λ_t = Binary variable: one for a specific month/year, zero otherwise
μ_i = Account level fixed effect
ε_{it} = Regression residual

This model produces estimates of average monthly savings using the following equation:

$$\bar{S}_t = \hat{\beta}_t$$

Where:

$\bar{S}_t$ = Average treatment related consumption reduction during month t
$\hat{\beta}_t$ = Estimated parameter measuring the treatment group difference in the post period month t

The model also includes site-specific and month/year fixed effects. The site-specific effects control for mean differences between the treatment and control groups that do not change over time. The month/year fixed effects control for change over time that is common to both treatment and control groups. The monthly post-program dummy variables pick up the average monthly effects of the treatment. Households that move are dropped from the model. The total savings are a sum of the monthly average savings combined with the count of households still eligible for the program in that month. Households that actively opt out of the program remain in the model as long as they remain in their house. In this respect, the treatment can be considered "intent to treat." This model is consistent with best practices as delineated in State and Local Energy Efficiency Action Network's Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations.[3]

3.1.1 Downstream rebate joint savings

One possible effect of the HUR program is to increase rebate activity in other PG&E energy efficiency programs. The RCT experimental design facilitates the measurement of this effect. We compared the

[3] State and Local Energy Efficiency Action Network. 2012. *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*. Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkeley National Laboratory. http://behavioranalytics.lbl.gov.

average savings from rebate measures installed by the treatment group with the savings from measures installed by the control group. An increase in treatment group rebate program savings represents savings caused by the HUR program jointly with the rebate programs. While these joint savings are an added benefit of the HUR program, it is essential that these joint savings are only reported once. The most common and simple approach is to remove all joint savings from the HUR program savings rather than remove program-specific joint savings from all of the associated rebate programs. This has been the approach used historically to adjust the savings from the IOU HER programs.

The savings estimates from the fixed effects regressions include all differences between the treatment and control group in the post-report period. Joint savings are picked up by the regressions and included in the overall savings estimate. These joint savings are also included in PG&E rebate program tracking databases and are claimed as part of those programs' savings unless further actions were taken to remove them. Savings from the HUR program are adjusted using joint savings to avoid double counting of savings.

DNV GL used the following approach for rolling up individual rebate's savings and calculating joint savings overall:

- Use accepted deemed savings values (those being used to claim the savings for the rebate program).
- Start accumulating savings from the installation date moving forward in time.
- Assign daily savings on a load-shape-weighted basis (more savings when we expect the measure to be used more).
- Maintain the load-shape-weighted savings over the life of the measure.

This approach takes the deemed annual savings values and transforms them into realistic day-to-day savings values given the installation of that measure. We determined the daily share of annual savings using hourly 2011 DEER load shapes[4] for PG&E. [5] These load shapes indicate when a measure is used during the year and, by proxy, when efficiency savings would occur.[6]

Savings for each installed measure start to accrue at the time of installation (or removal for refrigerator recycling). We calculated average monthly household rebate program savings for the treatment and control groups including zeroes for the majority of households that do not take part in any rebate program. An increase in average per-household tracked program savings among the treatment group versus the control group indicates joint savings.

3.1.2 Upstream joint savings

Upstream joint savings are similar to downstream joint savings, except that upstream savings are not tracked at the customer level. PG&E upstream savings still represent a source of savings that MCE HUR could potentially double count. Unlike tracked programs, it is not possible to directly compare all treatment and control group member activity. This makes it more challenging to determine if the HUR program does increase savings in upstream programs.

The alternative to the downstream census-level approach is to do a comparison of treatment and control group uptake of the upstream program measures on a sample basis. This approach also takes advantage of the RCT experimental design, which provides the structure to produce an un-biased estimate of upstream

[4] DEER load shapes are in an 8760 hourly format. DNV GL aggregated the hourly shares to daily shares in order to estimate daily savings.

[5] http://deeresources.com/DEER2011/download/DEER2011-UpdatedImpactProfiles-v2.zip

[6] This is more accurate and equitable than subtracting out the first year savings values that are used in DEER, because most measures are not in place from the first day to the last day of the year.

savings. PG&E conducted in-home surveys in 2013 to assess uptake of upstream measures (specifically, CFLs and flat-screen TVs). The surveys included samples of treatment and control customers from their HER program. However, given that the HUR program has produced very little (if any) savings, there is no practical evidence that joint savings due to upstream programs are occurring. As such there was no need to apply a double-counting adjustment for upstream savings to the final savings.

3.2 Data Sources and Disposition

This section describes the data used in DNV GL's impact evaluation of the HUR program.

3.2.1 Data sources

Program Participants

MCE provided HUR participant account numbers and the corresponding customer account numbers in PG&E's customer database. Additional information such as zip codes, house square footage, number of bedrooms/bathrooms, treatment assignment, and other household characteristics were also provided. These data served as the roster of program participants for the HUR evaluation.

Monthly Billing Data

DNV GL used the PG&E monthly billing data for HUR customer consumption from November 2012 to December 2014. The billing data included account numbers, premise numbers, billing cycle start and end dates, consumption reads, net metering flags, and the type of reading (actual reading/estimated reading).

Downstream Program Tracking Data

DNV GL used PG&E program tracking data to collect information on MCE HUR customers who participated in PG&E downstream rebate programs after the inception of the HUR program. PG&E tracking data included participant information, account numbers, program name, measures installed, installation dates, and claimed savings. This dataset facilitated calculating downstream joint savings for the HUR program.

3.2.2 Data disposition

The impact evaluation relies on consumption data from the PG&E monthly billing data system. Consumption data are closely tied to the billing function and are generally considered accurate. On the other hand, missed reads, estimated reads, and corrections do occur, and may undermine the validity of some readings. In non-RCT billing analysis evaluations, it is common to apply a range of consumption data checks in an attempt to limit invalid data. This can lead to the removal of customers from the analysis because of limitations in their billing data. In an RCT analysis, we would expect anomalies to appear in the same proportion in the treatment and control groups, and thus there is no need to remove such records. For this evaluation, the two primary groups removed from the analysis were net metering customers and customers with insufficient data.

Table 8 provides an overview of the data issues identified in the billing data. The incidence of issues is small across treatment and control groups and both fuel types. For large reads (>10,000 kWh per month for electric), large monthly consumption was observed in less than 0.5% of the households overall. One site with consumption over 10,000 kWh per month was excluded from the analysis. This site was a special case of a mobile home trailer park serving more than 40 mobile home units.

Around 1 to 3% of the households were also identified as net metered sites. Customers who installed solar panels and switched to net metering posed a dilemma for this evaluation. This is due to the way that net

metering is addressed in the billing data, which creates challenges for either including them in the analysis or fully understanding the extent of the issue. For example, if the solar households were included in the analysis it would be necessary to incorporate household-level energy production data.[7] Otherwise, potential differences in solar energy production could be conflated with program-related savings, biasing the results up or down. For this evaluation, all net-metered customers were left out of the analysis.

For most cases, potential data issues are small and proportionally balanced between the treatment and control groups. These findings indicate that data issues are infrequent and that the treatment/control difference inherent in the RCT structure controlled for the majority of the issues that existed.

Table 8: Summary of Billing Data

Summary	Electric	
	Control	Treatment
HUR-1 Sites	**2,766**	**3,643**
Negative Reads	2%	1%
Extreme Reads	0%	0%
Net metered sites	3%	3%
No consumption in pre or post	0%	1%
No Issues	97%	96%
HUR-2M Sites	**5,934**	**6,560**
Negative Reads	1%	1%
Extreme Reads	0%	0%
Net metered sites	2%	2%
No consumption in pre or post	0%	0%
No Issues	98%	98%
HUR-2Q Sites	**5,934**	**6,586**
Negative Reads	1%	1%
Extreme Reads	0%	0%
Net metered sites	2%	2%
No consumption in pre or post	0%	0%
No Issues	98%	98%
HUR-3 Sites	**2,114**	**4,233**
Negative Reads	0%	1%
Extreme Reads	0%	0%
Net metered sites	2%	2%
No consumption in pre or post	1%	1%
No Issues	96%	96%

Table 9 through Table 11 summarizes the count of households with respect to natural attrition due to change in occupancy for each HUR wave. Each table provides the count of active households for the treatment group that was used to calculate total program savings. The estimates of monthly savings produced by this impact evaluation reflect the consumption data of the active households remaining in the

[7] It is instructive to compare solar-installing households to HER opt-outs with respect to their effect on the analysis results. The removal of opt-outs from the treatment group would likely remove households with lower savings effects thus artificially increasing the savings estimate for those households remaining in the treatment group. This potential upward bias in the savings result is a clear reason for including these households despite their opting out. The solar-installing households have a less clearly defined HER program savings effect so it is more difficult to assess the effect of their removal on the HER savings of remaining households. More importantly, energy generated by solar systems would dwarf the amount of HER program savings at most households . The decision to remove these households is based on a lack of clear evidence of a biasing effect in the savings estimate and the concern that their inclusion would be practically speaking infeasible and would have the potential to introduce bias.

program (treatment or control group). At the end of program year 2014, overall attrition rate ranged from 1% to 4% for treatment and control groups across the three HUR waves.

DNV GL used the end-date electric account read periods to establish the number of active households. The below tables also provide the number of move-outs per month and the cumulative number of accounts used for both the treatment and control groups to determine active households.

Table 9: Household Attrition by HUR-1 Wave

Month	Control Group			Treatment Group		
	Open Accounts	Closed Accounts		Open Accounts	Closed Accounts	
		Cumulative	Monthly		Cumulative	Monthly
Nov-13	2,745	0.8%	0.0%	3,600	1.2%	0.0%
Dec-13	2,745	0.8%	0.0%	3,600	1.2%	0.0%
Jan-14	2,745	0.8%	0.0%	3,598	1.3%	0.1%
Feb-14	2,745	0.8%	0.0%	3,598	1.3%	0.0%
Mar-14	2,745	0.8%	0.0%	3,598	1.3%	0.0%
Apr-14	2,726	1.5%	0.7%	3,574	1.9%	0.7%
May-14	2,715	1.9%	0.4%	3,551	2.6%	0.6%
Jun-14	2,700	2.4%	0.6%	3,522	3.4%	0.8%
Jul-14	2,672	3.5%	1.0%	3,483	4.6%	1.1%
Aug-14	2,658	4.1%	0.5%	3,445	5.7%	1.1%
Sep-14	2,646	4.5%	0.5%	3,427	6.3%	0.5%
Oct-14	2,625	5.4%	0.8%	3,400	7.1%	0.8%
Nov-14	2,615	5.8%	0.4%	3,383	7.7%	0.5%
Dec-14	2,609	6.0%	0.2%	3,368	8.2%	0.4%

Note: The monthly counts provided exclude sites with net metering

Table 10: Household Attrition by HUR-2 Wave

Month	Control Group			Treatment Group (Monthly Recipients)			Treatment Group (Quarterly Recipients)		
	Open Accts	Closed Accounts		Open Accts	Closed Accounts		Open Accts	Closed Accounts	
		Cumulative	Monthly		Cumulative	Monthly		Cumulative	Monthly
Mar-14	5,862	1.2%	0.0%	6,479	1.3%	0.0%	6,512	1.1%	0.0%
Apr-14	5,819	2.0%	0.7%	6,479	1.3%	0.0%	6,512	1.1%	0.0%
May-14	5,766	2.9%	0.9%	6,478	1.3%	0.0%	6,509	1.2%	0.0%
Jun-14	5,715	3.8%	0.9%	6,478	1.3%	0.0%	6,506	1.2%	0.0%
Jul-14	5,652	5.0%	1.1%	6,473	1.3%	0.1%	6,450	2.1%	0.9%
Aug-14	5,600	6.0%	0.9%	6,440	1.9%	0.5%	6,393	3.0%	0.9%
Sep-14	5,555	6.8%	0.8%	6,381	2.8%	0.9%	6,344	3.8%	0.8%
Oct-14	5,518	7.5%	0.7%	6,344	3.4%	0.6%	6,299	4.6%	0.7%
Nov-14	5,489	8.1%	0.5%	6,292	4.3%	0.8%	6,250	5.4%	0.8%
Dec-14	5,456	8.8%	0.6%	6,257	4.8%	0.6%	6,216	6.0%	0.5%

Note: The monthly counts provided exclude sites with net metering

Table 11: Household Attrition by HUR-3 Wave

Month	Control Group			Treatment Group		
	Open Accounts	Closed Accounts		Open Accounts	Closed Accounts	
		Cumulative	Monthly		Cumulative	Monthly
Nov-14	2,054	2.9%	0.0%	4,109	3.0%	0.0%
Dec-14	2,048	3.2%	0.3%	4,096	2.7%	0.3%

Note: The monthly counts provided exclude sites with net metering

4 EXPERIMENTAL DESIGN VALIDATION

As part of this evaluation, DNV GL reviewed the experimental design of the HUR program to ensure validity of this impact evaluation. Statistical t-tests were applied by testing pre-existing differences in energy consumption and household characteristics between the treatment and control groups. Results from the t-tests are presented for each wave.

4.1 HUR-1 wave

Figure 1 shows the monthly difference in electric consumption between the treatment and control groups, along with the upper and lower limits at a 90% confidence interval. Differences greater than zero indicate higher consumption by the treatment group. Results show that electric consumption of the treatment group is significantly higher relative to the control group. These results confirm that the treatment and control groups are unbalanced. The fact that the two samples are substantially more different during the winter months is important. The savings estimation approach used for this evaluation corrects for mean differences across the whole pre-report period, not individual monthly differences. This will be reflected in the monthly savings estimates.

Figure 1: Electric Consumption Differences between Treatment and Control, HUR-1 Wave

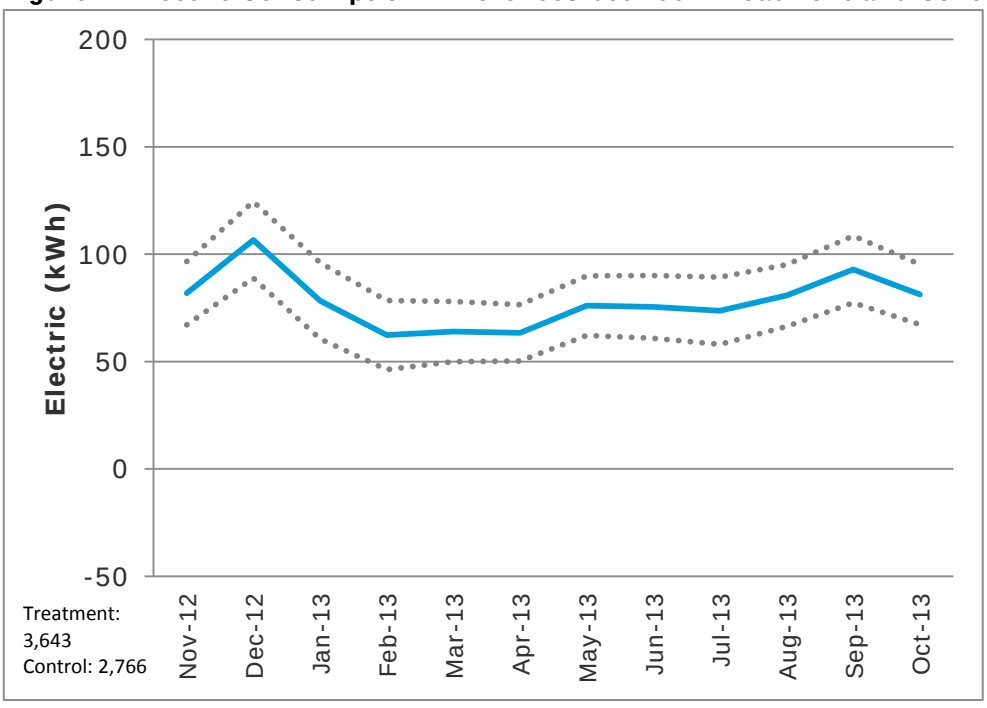


Table 12 provides a comparison of different household characteristics between the treatment and control groups. The test of differences also showed statistically significant differences in several household characteristics between the treatment and control groups, such as number of bedrooms/bathrooms, number of adults, construction year, and house size.

Table 12: Differences in Household Characteristics between Treatment and Control, HUR-1 Wave

Characteristics	Treatment			Control			Treatment - Control		
	Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
No. of adults	3,633	2.4	0.02	2,768	2.5	0.02	-0.1*	0.03	0.000
No. of bathrooms	3,638	2.6	0.01	2,770	2.5	0.01	0.1*	0.01	0.000
No. of bedrooms	3,638	3.0	0.02	2,770	3.2	0.02	-0.3*	0.03	0.000
No. of children	3,633	0.6	0.01	2,768	0.6	0.02	0.0	0.02	0.101
House construction year	3,638	1965	0.34	2,770	1963	0.39	2.9*	0.52	0.000
No. of occupants	3,638	1.7	0.01	2,770	1.7	0.01	0.0	0.02	0.123
House square footage	3,638	2,317	12.61	2,770	2,070	12.39	247*	18.04	0.000

**Statistically significant at 90% confidence level*

Results from the randomization tests for the HUR-1 wave suggest that, on the average, households in the treatment group use 9% more electricity and 13% more gas relative to the control group. Also, households in the treatment group have relatively larger and newer homes. On the other hand, the treatment group also has fewer adults and few bedrooms. While it is unfortunate that the sample is not balanced in many aspects, using the pooled fixed effects model with a difference-in-differences structure to estimate savings should control for pre-existing differences between the treatment and control groups with respect to consumption and any unobserved heterogeneity across households that are fixed over time.

4.2 HUR-2 wave

Figure 2 and Figure 3 show the results from the randomization test on consumption for the HUR-2S and the HUR-2R waves. Consumption in all months is not statistically significantly different than zero. HUR-2S and HUR-2R pre-period energy consumptions are balanced between the treatment and control groups.

Figure 2: Electric Consumption Differences between Treatment and Control, HUR-2S Wave

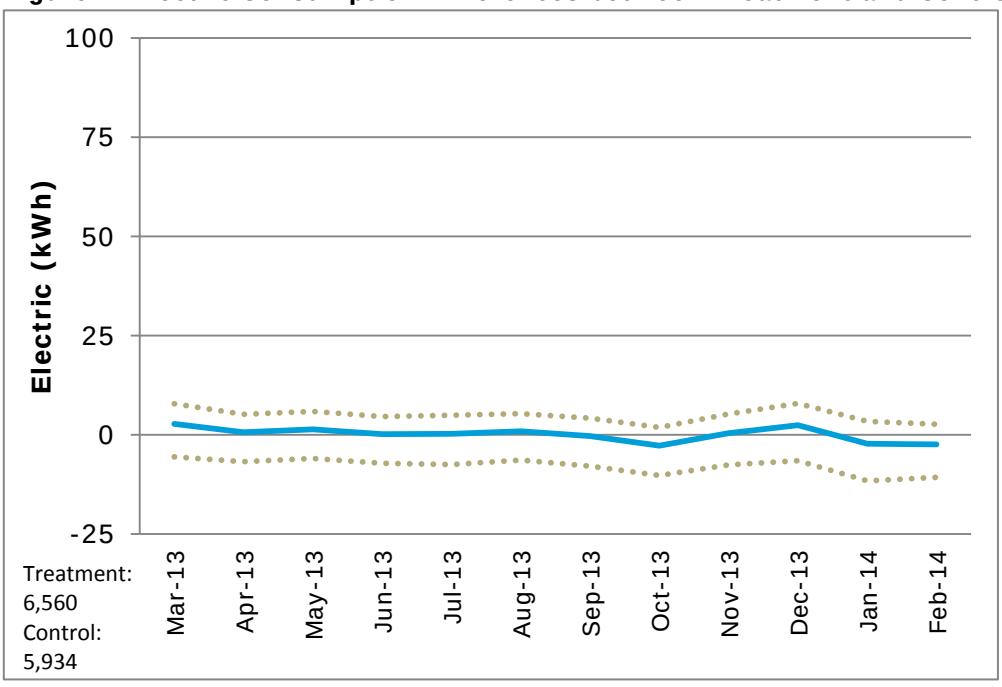


Figure 3: Electric Consumption Differences between Treatment and Control, HUR-2R Wave

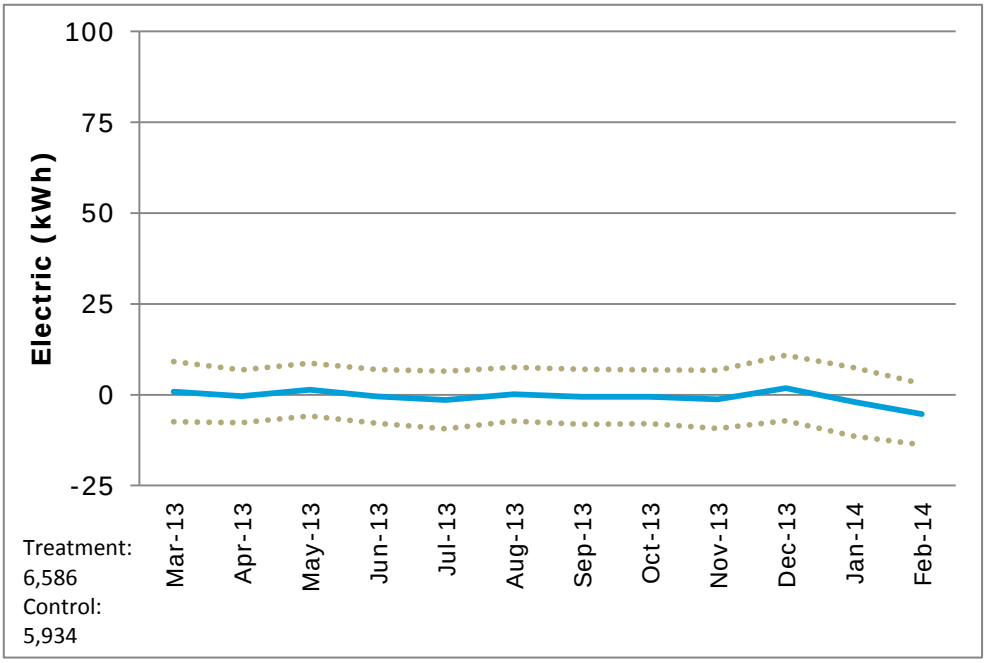


Table 13 and Table 14 provide the comparisons of household characteristics for HUR-2 standard (HUR-2S) and HUR-2 reduced (HUR-2R) frequencies. Despite the fact that samples were chosen using multiple "random" assignments, the results show small but statistically significant differences in some household characteristics between the treatment and control groups for HUR-2S and HUR-2R. The observed imbalance in household characteristics for the HUR-2 wave is not expected to bias results produced in this evaluation for the same reasons stated above.

Table 13: Differences in Household Characteristics between Treatment and Control, HUR-2 Standard

Characteristics	Treatment			Control			Treatment - Control		
	Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
No. of adults	6,563	2	0.0	5,928	2	0.0	0.0	0.0	0.731
No. of bathrooms	6,576	2	0.0	5,944	2	0.0	0.0*	0.0	0.046
No. of bedrooms	6,576	2	0.0	5,944	2	0.0	0.0	0.0	0.716
No. of children	6,563	0	0.0	5,928	0	0.0	0.0	0.0	0.292
House construction year	6,576	1,958	0.7	5,944	1,958	0.6	0.0	0.9	0.997
No. of occupants	6,576	2	0.0	5,944	2	0.0	0.0*	0.0	0.064
House square footage	6,576	1,699	8.9	5,944	1,677	9.2	21.8*	12.8	0.088

**Statistically significant at 90% confidence level*

Table 14: Differences in Household Characteristics between Treatment and Control, HUR-2 Reduced

Characteristics	Treatment			Control			Treatment - Control		
	Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
No. of adults	6,579	2	0.0	5,928	2	0.0	0.0	0.0	0.962
No. of bathrooms	6,599	2	0.0	5,944	2	0.0	0.0*	0.0	0.009
No. of bedrooms	6,599	2	0.0	5,944	2	0.0	-0.1*	0.0	0.010
No. of children	6,579	0	0.0	5,928	0	0.0	0.0*	0.0	0.019
House construction year	6,599	1,960	0.3	5,944	1,958	0.6	-1.4*	-0.3	0.026
No. of occupants	6,599	2	0.0	5,944	2	0.0	0.0	0.0	0.957
House square footage	6,599	1,717	9.0	5,944	1,677	9.2	-39.7*	-12.9	0.002

**Statistically significant at 90% confidence level*

4.3 HUR-3 wave

Figure 4 shows the results from the randomization test on energy consumption for the HUR-3 wave, and Table 15 provides a comparison of household characteristics between the treatment and control groups. Results show that electric consumption for each month in the pre-period are similar, and only one out of the seven household characteristics had significant differences between treatment and control groups.

Figure 4: Electric Consumption Differences between Treatment and Control, HUR-3 Wave

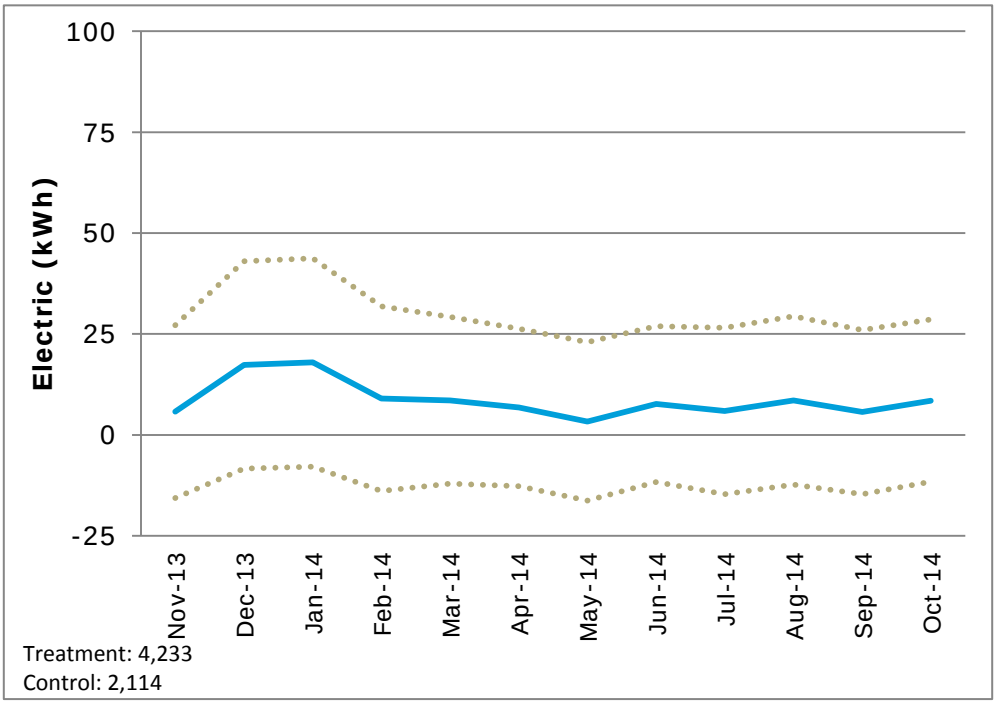


Table 15: Differences in Household Characteristics between Treatment and Control, HUR-3 Wave

Characteristics	Treatment			Control			Treatment - Control		
	Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
No. of adults	4,105	2.1	0.02	2,061	2.1	0.02	0.0	0.03	0.143
No. of bathrooms	4,225	2.2	0.01	2,110	2.2	0.01	0.0	0.02	0.198
No. of bedrooms	4,225	2.4	0.02	2,110	2.4	0.03	0.0	0.03	0.332
No. of children	4,105	0.4	0.01	2,061	0.4	0.02	0.0	0.02	0.158
House construction year	4,225	1957	0.75	2,110	1958	0.52	-1.1	1.13	0.322
No. of occupants	4,225	1.7	0.01	2,110	1.7	0.01	0.0	0.02	0.633
House square footage	4,225	1,686	11.48	2,110	1,640	15.78	46.3*	19.71	0.019

**Statistically significant at 90% confidence level*

DNV GL's validation does not exactly replicate the results from the balance test that MCE provided (Table 7). The general findings, however, are consistent. HUR-1 shows substantial imbalance, while the later waves that were explicitly chosen for balance show less imbalance. The discrepancies are likely due to differences in the exact sample used in the analysis for each wave. In addition, unlike the original balance check, for this evaluation we split the HUR-2 groups by report frequency.

5 RESULTS: SAVINGS ESTIMATES

This chapter presents the final reported savings estimates for the 2013-2014 MCE HUR program.

- Section 5.1 reports the *overall average savings,* which represent the unadjusted effect of the HUR program on treatment group consumption.
- Section 5.2 reports the *joint savings* estimates, which identify the downstream joint savings included in the overall savings estimate that are reported by other PG&E programs.
- Section 5.3 combines these estimates, removing the joint savings from the overall savings, and producing a 2013-2014 HUR program savings estimate that does not double-count energy savings from other energy efficiency programs.

5.1 HER program overall savings estimates

Figure 5 through Figure 8 provide graphic illustrations of the monthly electric savings for 2013-2014 for each HUR wave. The average monthly savings across all waves are between -20 kWh (effectively no savings) and 11 kWh per household. Only a few individual months are statistically different than zero, and three of those months have negative savings. These plots indicate that there is no evidence of savings resulting from the efforts of the MCE HUR program.

Given the average per-customer consumption for the different waves, a savings of 1% in each month would be approximately 9, 4, and 6 kWh for waves 1, 2 and 3, respectively. Three months in HUR-1 pass this minimum benchmark, and no months pass the benchmark for the other two waves. Furthermore, given the relatively small counts in each wave, even savings of 1% would be at best borderline with respect to statistical significance, and far below expected precision levels.

Figure 5 is quite different from the other three figures. While the second and third waves are smooth and flat, the first wave has substantial variability over the months of the year. This appears to be an outcome of the poorly balanced treatment and control groups. As discussed in Section 4, the treatment group has substantially higher usage than the control group in general, but particularly so in the winter. The fixed-effects savings estimate approach corrects for the difference on an average annual basis over the whole pre-report period. That means, on an annual basis, the savings estimates produced are un-biased. However, on a monthly basis, it is not a surprise that during the winter months the results show negative savings. The annual average correction does not fully correct for the higher usage treatment group in those months. On the other hand, the mean correction over-corrects in the summer month. The graph does have the expected shape given the shape of the pre-period difference. Ultimately, the annual estimate of savings from the post period is appropriately adjusted for the limitation of the RCT. The best 12-month period in the post period for HUR-1 estimates 0.1% savings, and is not statistically significantly different than zero.

Figure 5: Average Monthly kWh Savings per Household in HUR-1

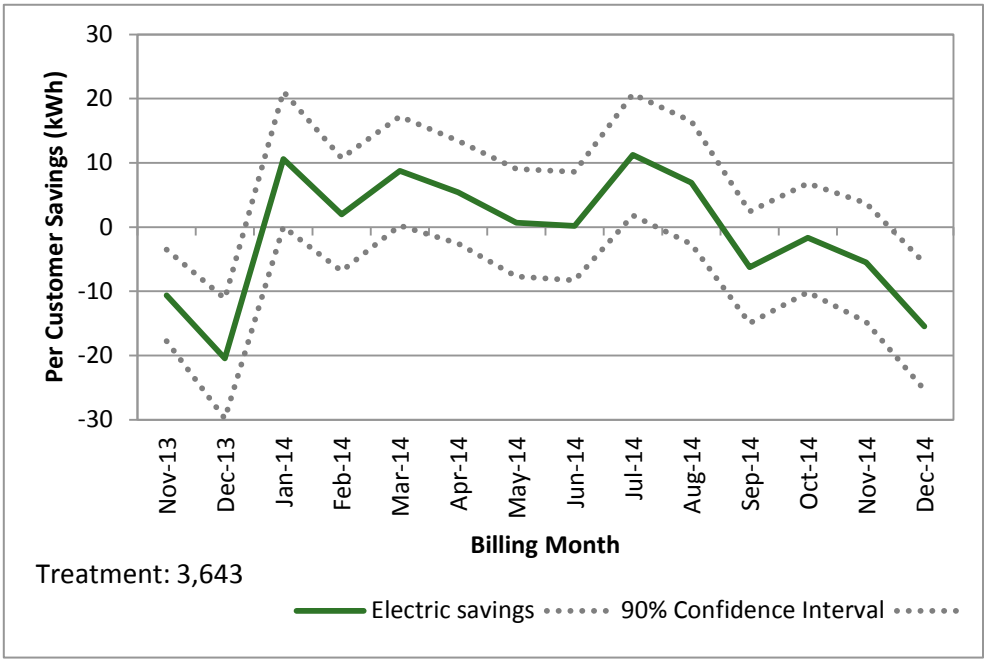


Figure 6: Average Monthly kWh Savings per Household in HUR-2 Standard

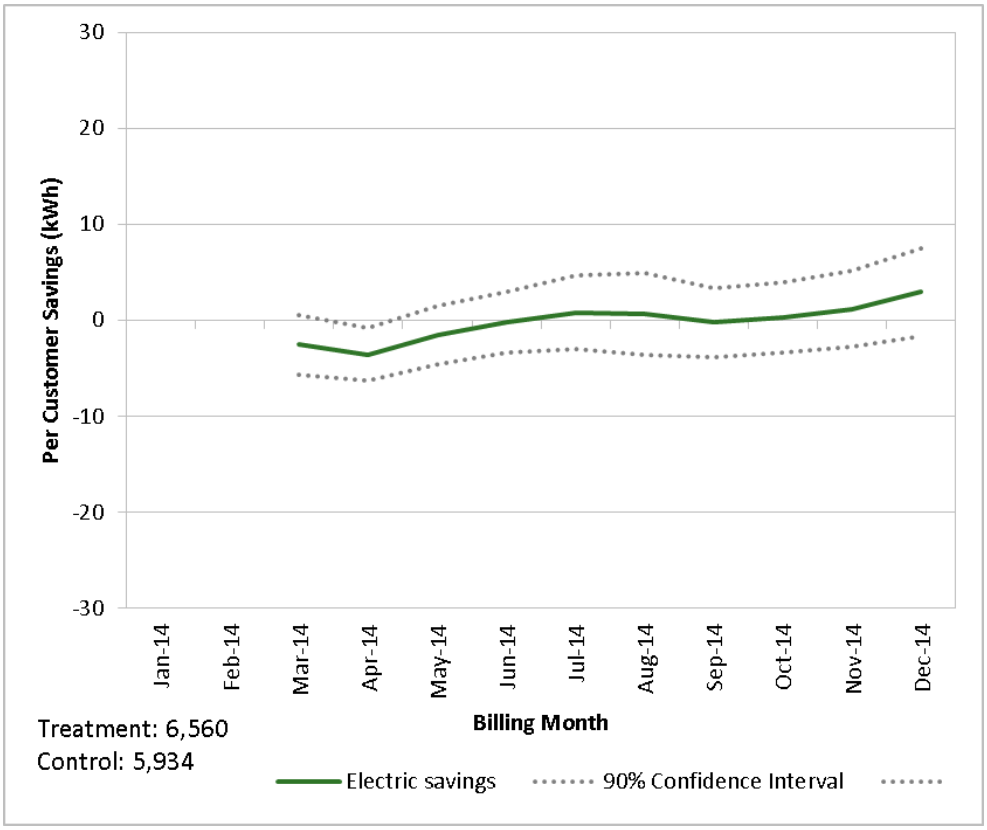


Figure 7: Average Monthly kWh Savings per Household in HUR-2 Reduced

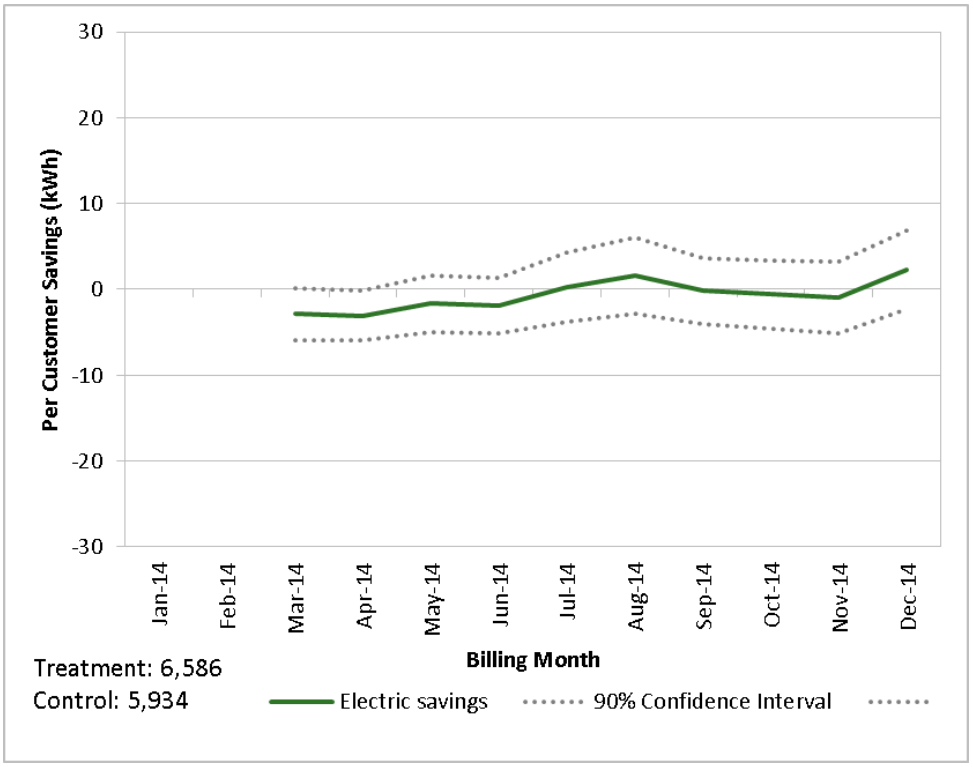


Figure 8: Average Monthly kWh Savings per Household in HUR-3

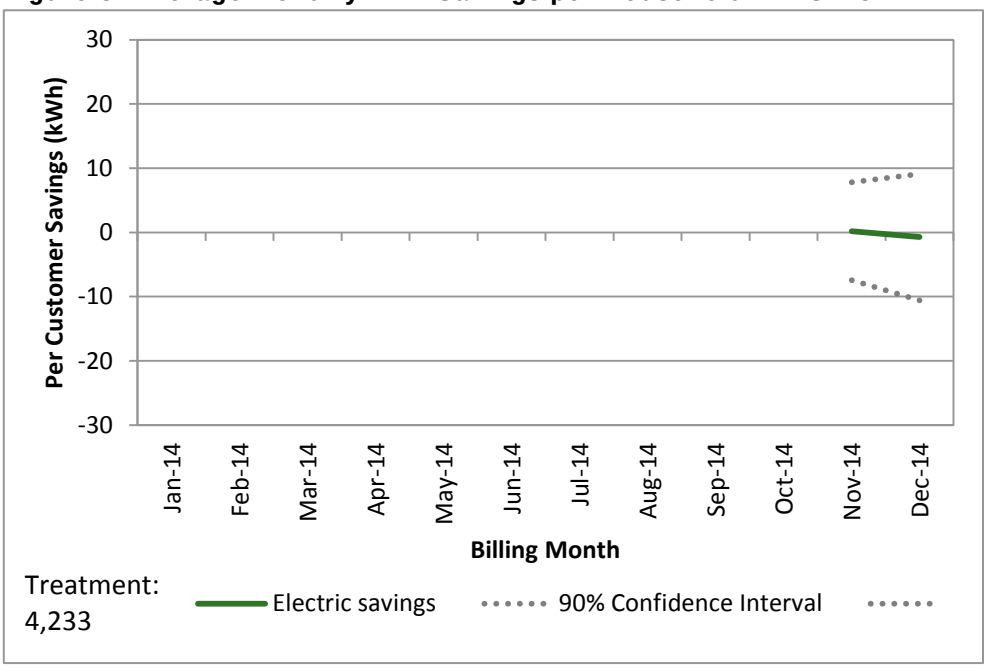


Table 16 and Table 17 provide the monthly electric savings in tabular form, along with the count of treatment group households for each month. In combination, these numbers generate the total monthly estimated electric savings for the HUR program. The total rows at the bottom of the tables provide the total and annual savings along with confidence intervals for the aggregate numbers.

Table 16: Household Counts and Average Monthly Unadjusted per Household Electric Savings

Month	Count of treatment households				Savings per household			
	HUR-1	HUR-2		HUR-3	HUR-1	HUR-2		HUR-3
		M	Q			M	Q	
Nov-13	3,600				(10.6)			
Dec-13	3,600				(20.5)			
Jan-14	3,598				10.6			
Feb-14	3,598				2.0			
Mar-14	3,598	6,479	6,512		8.7	(2.5)	(2.9)	
Apr-14	3,574	6,479	6,512		5.4	(3.6)	(3.0)	
May-14	3,551	6,478	6,509		0.7	(1.5)	(1.6)	
Jun-14	3,522	6,478	6,506		0.2	(0.2)	(1.9)	
Jul-14	3,483	6,473	6,450		11.3	0.8	0.2	
Aug-14	3,445	6,440	6,393		6.9	0.7	1.6	
Sep-14	3,427	6,381	6,344		(6.3)	(0.2)	(0.2)	
Oct-14	3,400	6,344	6,299		(1.7)	0.3	(0.6)	
Nov-14	3,383	6,292	6,250	4,109	(5.5)	1.2	(0.9)	0.2
Dec-14	3,368	6,257	6,216	4,096	(15.5)	2.9	2.3	(0.7)
Total					(14.3)	(2.2)	(7.0)	(0.6)

Table 17: Total Unadjusted Electric Savings

Month	Unadjusted Program Savings (kWh)			
	HUR-1	HUR-2		HUR-3
		M	Q	
Nov-13	(38,338)			
Dec-13	(73,646)			
Jan-14	38,051			
Feb-14	7,040			
Mar-14	31,443	(16,485)	(18,853)	
Apr-14	19,423	(23,093)	(19,752)	
May-14	2,441	(10,010)	(10,708)	
Jun-14	577	(1,270)	(12,168)	
Jul-14	39,236	5,252	1,421	
Aug-14	23,830	4,447	10,347	
Sep-14	(21,485)	(1,568)	(1,252)	
Oct-14	(5,640)	1,802	(3,724)	
Nov-14	(18,687)	7,382	(5,686)	713
Dec-14	(52,104)	18,342	14,178	(3,035)
Total	(47,859)[ns]	(15,202)[ns]	(46,198)[ns]	(2,322)[ns]

[ns] *Not statistically significant at 90% confidence level*

5.2 HER program joint savings: downstream rebates

Table 18 shows some of the broad categories in which HUR may have influenced uptake in PG&E rebate programs. HUR-3 did not have much rebate activity, probably because the program started late in the year. Otherwise, the most common type of program rebates were related to lighting, while refrigerator and clothes-washer-related activities were very similar across the different waves.

Table 18: Types of Rebates

Wave	Group	PG&E Rebates			
		Refrigerator	Lighting	Clothes Washer	Other
HUR-1 monthly	Treatment	31	107	33	19
	Control	30	74	33	13
HUR-2 monthly	Treatment	46	55	49	20
	Control	41	44	48	16
HUR-2 quarterly	Treatment	49	97	39	24
HUR-3 monthly	Treatment	5	0	2	3
	Control	3	0	0	0

Figure 9 through Figure 12 show the monthly downstream savings per HUR group. Most of the months show savings that are not statistically different from zero. HUR-3 shows the least amount of monthly variation in savings due to having only two data points. HUR-1 shows the most amount of monthly variation in savings.

Figure 9: Monthly kWh Joint Savings per Household in HUR-1

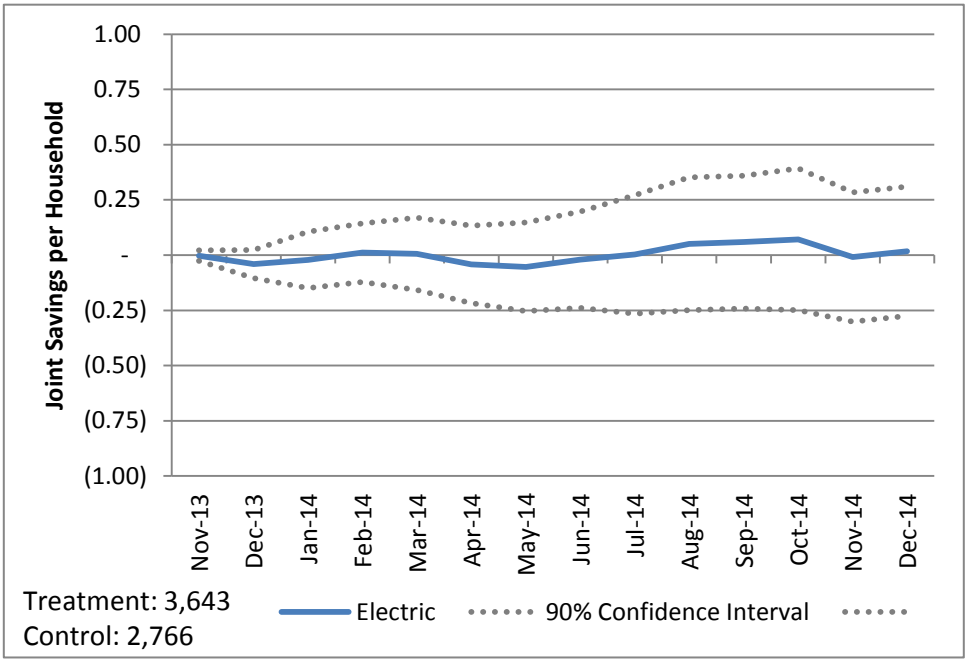


Figure 10: Monthly kWh Joint Savings per Household in HUR-2 Standard

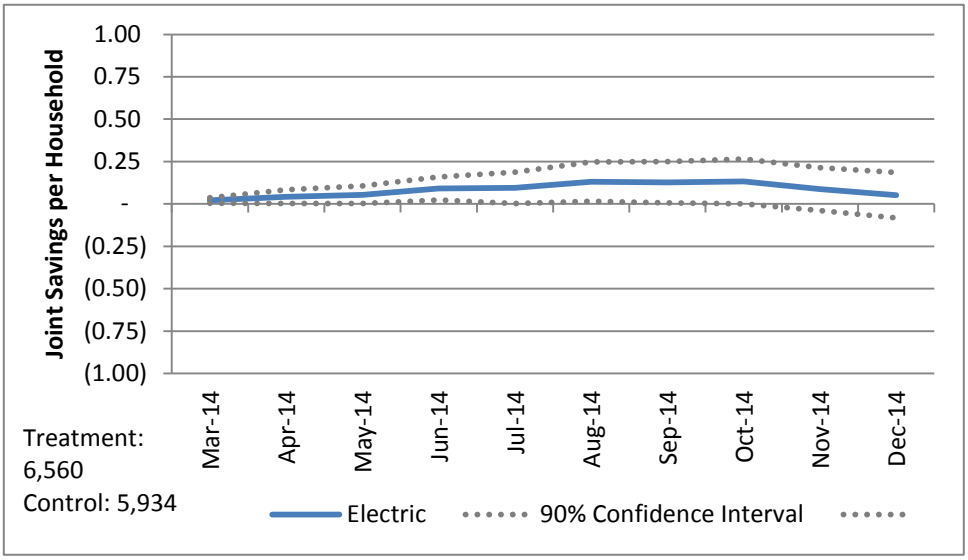


Figure 11: Monthly kWh Joint Savings per Household in HUR-2 Quarterly

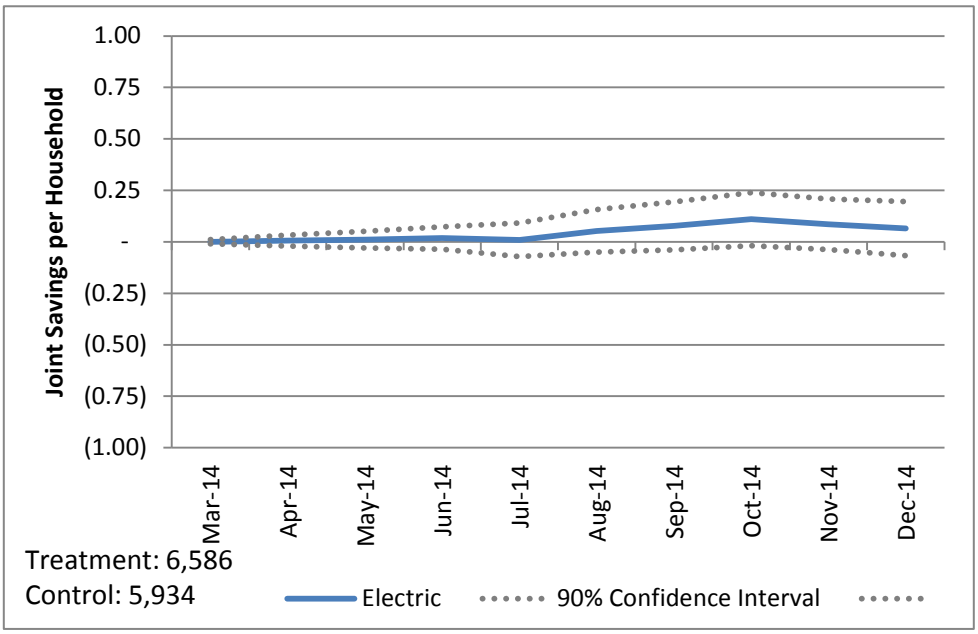
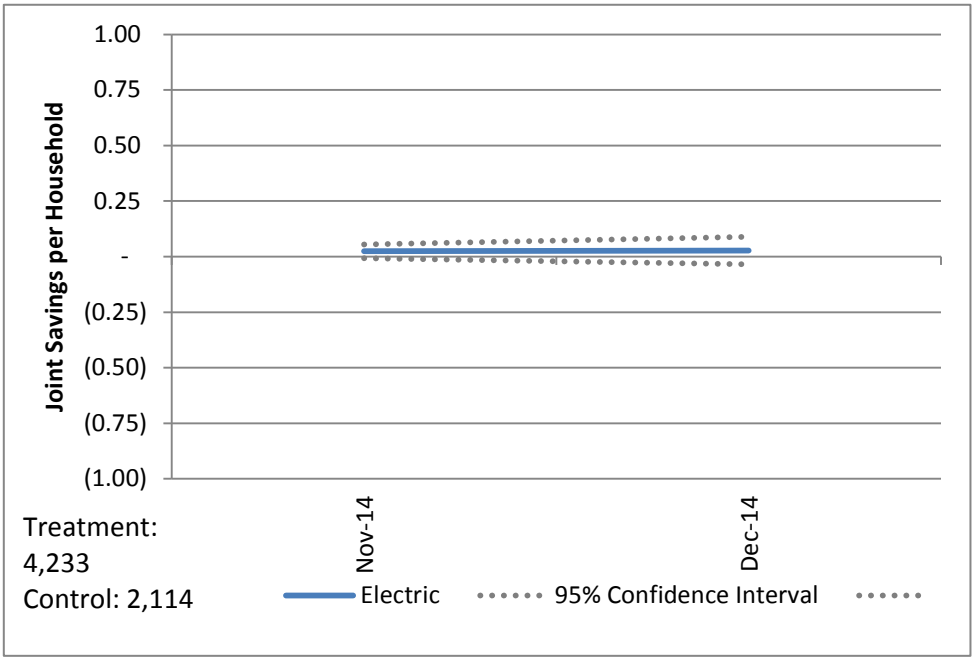


Figure 12: Monthly kWh Joint Savings per Household in HUR-3



5.3 Per-household savings and total program savings

Table 19 provides the final per-household kWh savings for the MCE HUR program. The results are identical to those provided in section 5.1. None of unadjusted results are statistically significantly different than zero. This means there are no claimable savings from this program.

Table 19: Per Household kWh Savings and Percent Savings

Wave	Evaluation Period	Unadjusted kWh per Customer Savings	Adjusted kWh per Customer Savings	Per kWh Customer Consumption	Unadjusted Savings as % of Consumption	Adjusted Savings as % of Consumption
HUR-1	November 2013 - December 2014	-14.3	-14.3	11,228.8	-0.1%	-0.1%
	January 2014 - December 2014	16.8	16.7	9,464.2	0.2%	0.2%
HUR-2M	January 2014 - December 2014	-2.2	-3.0	4,945.6	0.0%	-0.1%
HUR-2Q	January 2014 - December 2014	-7.0	-7.5	4,945.6	-0.1%	-0.2%
HUR-3	January 2014 - December 2014	-0.6	-0.6	7,796.9	0.0%	0.0%

Table 20 provides the total savings estimates for the program. Once again, none of these results are statistically significantly different than zero. This means the savings are effectively zero.

Table 20: Total Unadjusted and Adjusted MWh Program Savings

		Source				
Wave	**Evaluation Period**	**Unadjusted Savings**	**Tracked, Downstream Joint Savings**	**Untracked, Upstream Lighting Joint Savings**	**Adjusted Savings**	**Statistically Significant with 90% confidence?**
		Electric (MWh)				
HUR-1	November 2013 - December 2014	-47.9	0.1	-	-47.9	No
	January 2014 - December 2014	64.1	0.2	-	63.9	No
HUR-2M	January 2014 - December 2014	-15.2	5.3	-	-20.6	No
HUR-2Q	January 2014 - December 2014	-46.2	2.8	-	-49.0	No
HUR-3	January 2014 - December 2014	-2.3	0.2	-	-2.5	No
Total	November 2013 - December 2014	-111.6	8.4	-	-120.0	No

6 CONCLUSIONS

This report finds that the MCE HUR program produced no savings. This finding is definitive given the experimental design within which the program was organized, and the standards set by the CPUC for the evaluation of these programs. Below we discuss secondary findings that put the overall results in context.

Interaction with the school program – The school-based program run by MCE may have had some minor interaction with the HUR program. It is reasonable to expect that those school program effects would have been small and equally distributed across treatment and control groups. However, if savings efforts from the school program were in any way redundant with HUR efforts, then the presence of the school program effect in both treatment and control groups could have a dampening effect on HUR savings. There is a possibility, then, that the interaction with the school program could explain a small part of the HUR program shortfall on expected savings. If this was the case, and it is only a hypothesis, it would have no effect on the interpretation of the final results reported here. Behavior programs are required to be implemented in an RCT experimental design precisely because of the difficulty of measuring variable and small-magnitude behavior-related savings. The school program did not undermine the RCT, but may have contributed in a small way to the lack of savings measured therein.

RCT issues – The MCE HUR program produced poorly balanced treatment and control groups for the first program wave. This is not ideal, but the fixed-effects estimation approach addresses this problem by correcting for mean consumption differences across the two groups. The mean difference correction in this approach would not correct for different distributions of monthly consumption across the year related to, for instance, different levels of electric heat during the winter. This is a likely explanation for the negative savings in the first couple months of the program. Despite this, savings would be evident in annual averages if they existed. As it stands, the highest annual average during the 14 months the program's existence, calendar year 2014, showed an estimated savings of 0.2% that was not statistically significant. If the program had produced savings, they would have been evident despite the poorly balanced samples.

Relatively small wave population sizes – Given the size and structure of the different waves, if the program had produced actual savings of 1%, it is unlikely that those savings would have been statistically significant. In general, the goal for these kinds of programs is a relative precision of approximately 90/20, which is a substantially higher hurdle. This program would have benefited from a careful power analysis to guide the size of the sample given the expected levels of savings. On the other hand, a power analysis will only be useful if the estimate of expected savings is realistic and delivered.

Population-specific characteristics – Marin Clean Energy is an energy provider that markets itself as providing "better, cleaner, healthier energy." It is possible that the population mix that MCE serves is not ideally suited to behavior programs. If MCE customers are already aware of many of the potentially wasteful energy consumption habits, they may have already adjusted their behavior to address them. If this were the case, the reports would not produce additional saving in these areas.

This evaluation finds no evidence that the MCE HUR program produced savings. While there are some minor details to the program implementation and the experimental design that may contribute to this result, it is extremely unlikely that these issues individually or in combination would mask any consistent evidence of a program effect. Instead, these results clearly point to an overall lack of program effect.

APPENDIX A. RANDOMIZATION TESTS

The tables in this appendix highlight the randomization t-tests that DNV GL performed to evaluate the consumption balance between the various control and treatment groups. HUR-1 shows the most imbalanced results, while the others are all fairly well-balanced.

Table 21: Randomization Test for HUR-1

Fuel	Month	Treatment			Control			Treatment - Control		
		Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
Electric (kWh)	Nov-12	3,565	944	6.33	2,613	862	5.89	81.9	8.9	0.000
	Dec-12	3,303	1,117	7.86	2,530	1,010	6.92	106.7	10.8	0.000
	Jan-13	3,601	1,132	7.79	2,752	1,053	6.95	78.3	10.8	0.000
	Feb-13	2,935	958	6.91	2,253	896	6.65	62.3	9.8	0.000
	Mar-13	3,598	940	6.04	2,752	876	5.76	64.0	8.6	0.000
	Apr-13	3,607	863	5.66	2,752	799	5.33	63.3	8.0	0.000
	May-13	3,608	876	5.90	2,752	800	5.71	76.0	8.4	0.000
	Jun-13	3,239	860	6.14	2,474	785	6.24	75.5	8.9	0.000
	Jul-13	3,613	918	6.47	2,753	844	6.77	73.7	9.5	0.000
	Aug-13	3,615	885	6.03	2,753	804	6.20	80.8	8.8	0.000
	Sep-13	3,325	892	6.70	2,535	799	6.34	92.9	9.5	0.000
	Oct-13	3,621	886	5.99	2,754	805	5.80	81.2	8.5	0.000

Table 22: Randomization Test for HUR-2S

Fuel	Month	Treatment			Control			Treatment - Control		
		Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
Electric (kWh)	Mar-13	6543	464.32	3.36	5882	467.07	3.83	2.75	5.07	0.59
	Apr-13	6545	412.01	2.96	5887	412.65	3.41	0.64	4.49	0.89
	May-13	6546	401.84	2.94	5887	403.25	3.41	1.41	4.48	0.75
	Jun-13	6331	385.67	2.92	5688	385.82	3.37	0.15	4.44	0.97
	Jul-13	6547	413.22	3.14	5887	413.45	3.53	0.23	4.71	0.96
	Aug-13	6547	397.92	2.88	5886	398.83	3.39	0.91	4.43	0.84
	Sep-13	6320	401.17	3.00	5678	400.83	3.49	-0.35	4.58	0.94
	Oct-13	6548	414.44	3.12	5886	411.71	3.36	-2.73	4.58	0.55
	Nov-13	6145	432.58	3.16	5527	432.96	3.69	0.38	4.83	0.94
	Dec-13	6545	522.61	3.63	5902	525.07	4.10	2.46	5.46	0.65
	Jan-14	6547	537.48	3.84	5907	535.22	4.18	-2.26	5.67	0.69
	Feb-14	6102	459.77	3.42	5500	457.39	3.77	-2.38	5.08	0.64

Table 23: Randomization Test for HUR-2R

Fuel	Month	Treatment			Control			Treatment - Control		
		Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
Electric (kWh)	Mar-13	6564	466.22	3.27	5882	467.07	3.83	0.85	5.01	0.87
	Apr-13	6566	413.07	2.88	5887	412.65	3.41	-0.42	4.44	0.92
	May-13	6566	401.86	2.85	5887	403.25	3.41	1.39	4.41	0.75
	Jun-13	6360	386.29	2.98	5688	385.82	3.37	-0.47	4.48	0.92
	Jul-13	6567	414.91	3.26	5887	413.45	3.53	-1.45	4.80	0.76
	Aug-13	6567	398.69	2.95	5886	398.83	3.39	0.14	4.48	0.98
	Sep-13	6323	401.41	3.05	5678	400.83	3.49	-0.58	4.61	0.90
	Oct-13	6567	412.25	3.00	5886	411.71	3.36	-0.54	4.49	0.90
	Nov-13	6115	434.24	3.22	5527	432.96	3.69	-1.28	4.88	0.79
	Dec-13	6566	523.23	3.69	5902	525.07	4.10	1.84	5.50	0.74
	Jan-14	6565	537.18	3.91	5907	535.22	4.18	-1.97	5.72	0.73
	Feb-14	6115	462.72	3.46	5500	457.39	3.77	-5.33	5.11	0.30

Table 24: Randomization Test for HUR-3

Fuel	Month	Treatment			Control			Treatment - Control		
		Count	Mean	Std Error	Count	Mean	Std Error	Difference	Std Error	Pr > |t|
Electric (kWh)	Nov-13	3,813	692	7.60	1,908	686	10.38	5.7	13.0	0.659
	Dec-13	4,208	861	9.44	2,099	844	11.44	17.3	15.6	0.267
	Jan-14	4,210	877	9.51	2,099	859	11.41	17.9	15.7	0.253
	Feb-14	3,992	748	8.14	1,977	739	10.99	9.0	13.9	0.518
	Mar-14	4,209	725	7.49	2,100	716	9.48	8.6	12.5	0.494
	Apr-14	4,211	662	7.16	2,100	655	8.70	6.8	11.8	0.566
	May-14	4,209	641	7.10	2,099	638	9.10	3.3	11.9	0.779
	Jun-14	4,207	608	7.09	2,100	601	8.59	7.6	11.7	0.515
	Jul-14	4,208	626	7.61	2,102	620	9.09	5.9	12.5	0.636
	Aug-14	4,181	626	7.67	2,090	618	9.28	8.5	12.7	0.500
	Sep-14	4,201	615	7.41	2,099	609	9.21	5.7	12.3	0.647
	Oct-14	4,208	640	7.43	2,103	631	8.84	8.5	12.2	0.487

APPENDIX B. COMBINED RESULTS

The tables in this appendix provide the monthly unadjusted, downstream, and adjusted savings for each wave of the 2013-2014 HUR program.

Table 25: Combined Results for HUR-1 Electric Savings

Month	kWh per Household			Count of Treatment Group Participants	Adjusted Program Savings (MWh)
	Unadjusted Savings	Joint Savings - Tracked Downstream Programs	Adjusted Savings		
13-Nov	-10.6	0	-10.6	3,600	-38
13-Dec	-20.5	0	-20.4	3,600	-74
14-Jan	10.6	0	10.6	3,598	38
14-Feb	2	0	1.9	3,598	7
14-Mar	8.7	0	8.7	3,598	31
14-Apr	5.4	0	5.5	3,574	20
14-May	0.7	-0.1	0.7	3,551	3
14-Jun	0.2	0	0.2	3,522	1
14-Jul	11.3	0	11.3	3,483	39
14-Aug	6.9	0.1	6.9	3,445	24
14-Sep	-6.3	0.1	-6.3	3,427	-22
14-Oct	-1.7	0.1	-1.7	3,400	-6
14-Nov	-5.5	0	-5.5	3,383	-19
14-Dec	-15.5	0	-15.5	3,368	-52

Table 26: Combined Results for HUR-2 Monthly Electric Savings

Month	kWh per Household			Count of Treatment Group Participants	Adjusted Program Savings (MWh)
	Unadjusted Savings	Joint Savings - Tracked Downstream Programs	Adjusted Savings		
14-Mar	-2.5	0	-2.6	6,479	-17
14-Apr	-3.6	0	-3.6	6,479	-23
14-May	-1.5	0.1	-1.6	6,478	-10
14-Jun	-0.2	0.1	-0.3	6,478	-2
14-Jul	0.8	0.1	0.7	6,473	5
14-Aug	0.7	0.1	0.6	6,440	4
14-Sep	-0.2	0.1	-0.4	6,381	-2
14-Oct	0.3	0.1	0.2	6,344	1
14-Nov	1.2	0.1	1.1	6,292	7
14-Dec	2.9	0.1	2.9	6,257	18
Total	-2.2	0.8	-3		-20.6

Table 27: Combined Results for HUR-2 Quarterly Electric Savings

Month	kWh per Household			Count of Treatment Group Participants	Adjusted Program Savings (MWh)
	Unadjusted Savings	Joint Savings - Tracked Downstream Programs	Adjusted Savings		
14-Mar	-2.9	0	-2.9	6,512	-19
14-Apr	-3	0	-3	6,512	-20
14-May	-1.6	0	-1.7	6,509	-11
14-Jun	-1.9	0	-1.9	6,506	-12
14-Jul	0.2	0	0.2	6,450	1
14-Aug	1.6	0.1	1.6	6,393	10
14-Sep	-0.2	0.1	-0.3	6,344	-2
14-Oct	-0.6	0.1	-0.7	6,299	-4
14-Nov	-0.9	0.1	-1	6,250	-6
14-Dec	2.3	0.1	2.2	6,216	14
Total	-7	0.4	-7.5		-49

Table 28: Combined Results for HUR-3 Electric Savings

Month	kWh per Household			Count of Treatment Group Participants	Adjusted Program Savings (MWh)
	Unadjusted Savings	Joint Savings - Tracked Downstream Programs	Adjusted Savings		
14-Nov	0.2	0	0.1	4,109	1
14-Dec	-0.7	0	-0.8	4,096	-3
Total	-0.6	0.1	-0.6		-2.5

APPENDIX C. CARE VS. NON-CARE

The Energy Division asked DNV GL to compare savings between CARE and non-CARE customers. Because customers were marked as CARE or non-CARE at a monthly level, we created three different thresholds to assign customers to the CARE or non-CARE categories. The following tables show how the CARE customers are distributed for each HUR wave using the three different thresholds.

Table 29: No. of customers with CARE rates for HUR-1

HER sample	# of household		% of households	
	Control	Treatment	Control	Treatment
CARE (definition 1*)	87	114	3%	3%
Non-CARE	2679	3527	97%	97%
Total	2766	3641	100%	100%
CARE (definition 2**)	59	81	2%	2%
Non-CARE	2707	3560	98%	98%
Total	2766	3641	100%	100%
CARE (definition 3***)	46	55	2%	2%
Non-CARE	2720	3586	98%	98%
Total	2766	3641	100%	100%

**CARE definition 1: Customers with CARE rate for at least 1 billing month in 2014*
***CARE definition 2: Customers with CARE rate for at least 6 billing months in 2014*
****CARE definition 3: Customers with CARE rate for at least 10 billing months in 2014*

Table 30: No. of customers with CARE rates for HUR-2

HER sample	Monthly / Standard				Quarterly / Reduced			
	Control	Treatment	Control	Treatment	Control	Treatment	Control	Treatment
CARE (definition 1*)	1040	905	18%	14%	1040	879	18%	13%
Non-CARE	4893	5654	82%	86%	4893	5707	82%	87%
Total	5933	6559	5933	6559	5933	6586	5933	6586
CARE (definition 2**)	868	748	15%	11%	868	733	15%	11%
Non-CARE	5065	5811	85%	89%	5065	5853	85%	89%
Total	5933	6559	5933	6559	5933	6586	5933	6586
CARE (definition 3***)	745	629	13%	10%	745	629	13%	10%
Non-CARE	5188	5930	87%	90%	5188	5957	87%	90%
Total	5933	6559	5933	6559	5933	6586	5933	6586

**CARE definition 1: Customers with CARE rate for at least 1 billing month in 2014*
***CARE definition 2: Customers with CARE rate for at least 6 billing months in 2014*
****CARE definition 3: Customers with CARE rate for at least 10 billing months in 2014*

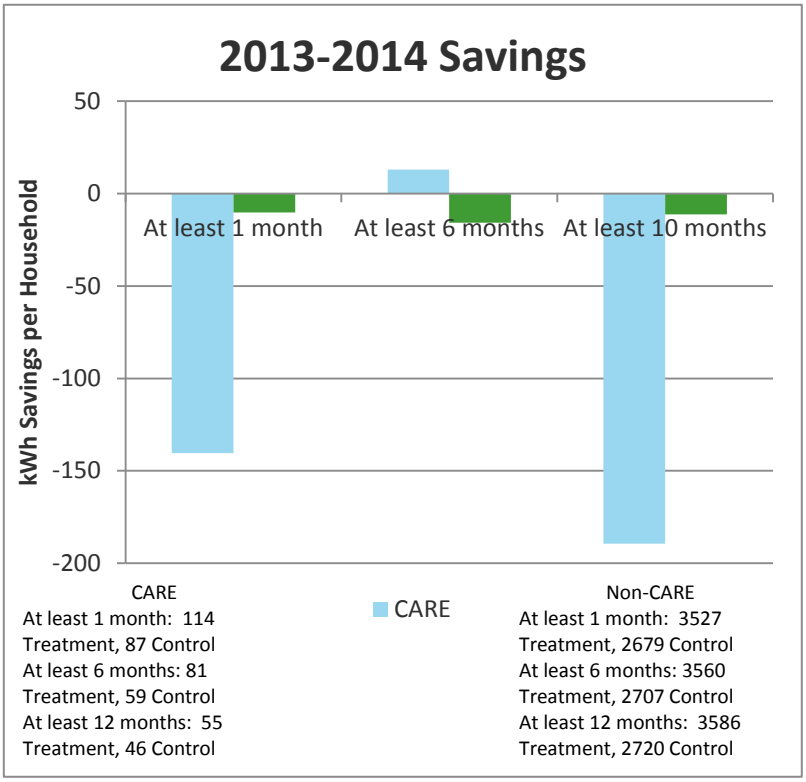
Table 31: No. of customers with CARE rates for HUR-3

HER sample	# of household		% of households	
	Control	Treatment	Control	Treatment
CARE (definition 1*)	377	776	18%	18%
Non-CARE	1729	3440	82%	81%
Total	2106	4216	100%	100%
CARE (definition 2**)	325	670	15%	16%
Non-CARE	1781	3546	84%	84%
Total	2106	4216	100%	100%
CARE (definition 3***)	287	611	14%	14%
Non-CARE	1819	3605	86%	85%
Total	2106	4216	100%	100%

**CARE definition 1: Customers with CARE rate for at least 1 billing month in 2014*
***CARE definition 2: Customers with CARE rate for at least 6 billing months in 2014*
****CARE definition 3: Customers with CARE rate for at least 10 billing months in 2014*

Figure 13 shows the Care and non-CARE savings for HUR-1. The first graph shows the total savings over the course of the program, while the second graph shows the total savings over the 2014 calendar year. Both graphs show that using the six-month threshold yielded the best results.

Figure 13: CARE and Non-CARE Electric Savings for HUR-1



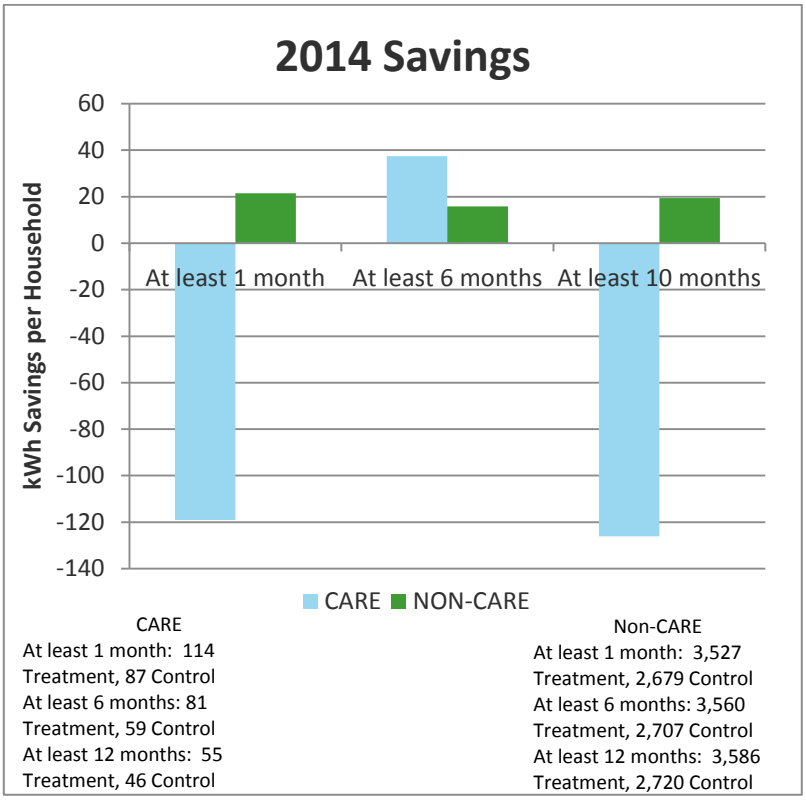


Figure 14 and Figure 15 show the total savings for the two HUR-2 subgroups. Among CARE customers, the quarterly subgroup generated more savings than the monthly subgroup. Among non-CARE customers, the monthly subgroup generated more savings than the quarterly subgroup.

Figure 14: CARE and Non-CARE Electric Savings for HUR-2 Monthly

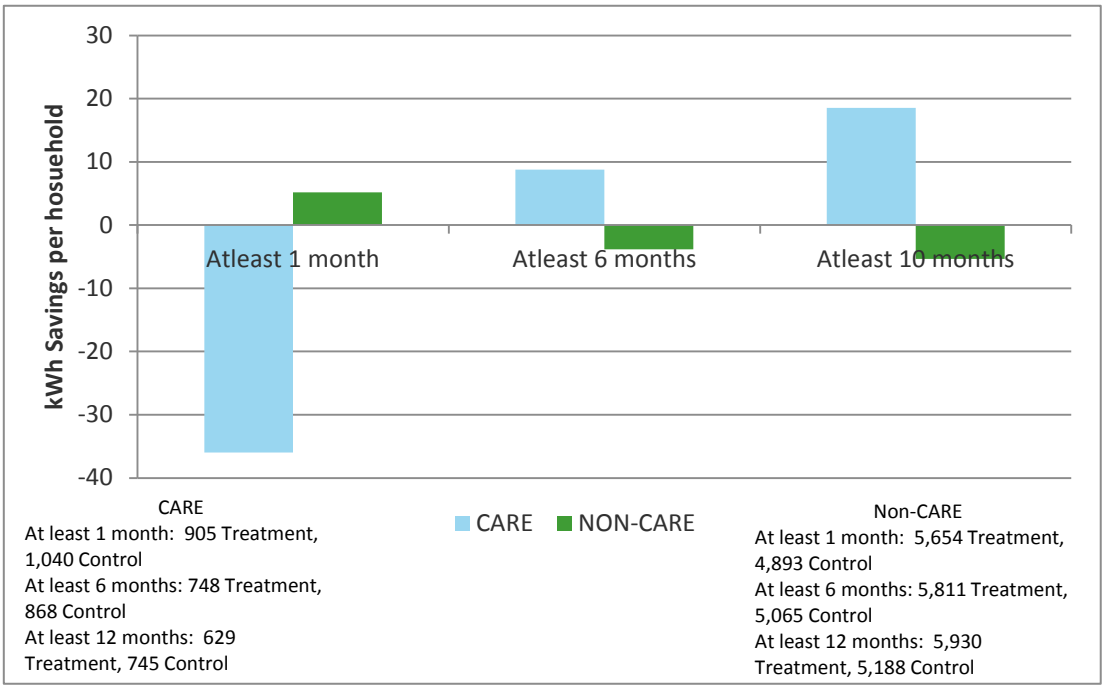


Figure 15: CARE and Non-CARE Electric Savings for HUR-2 Quarterly

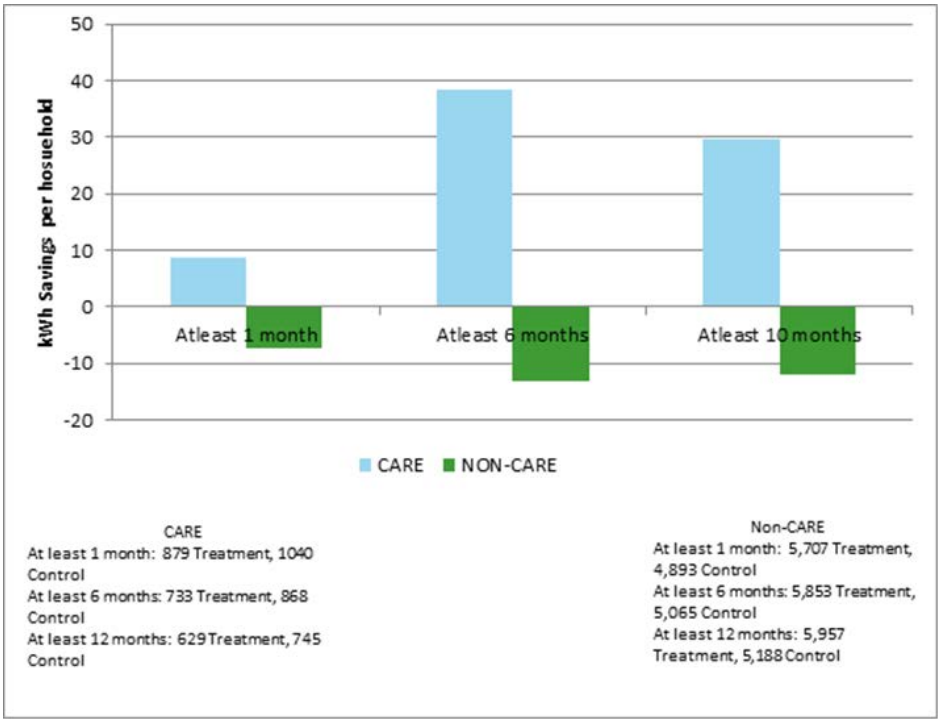
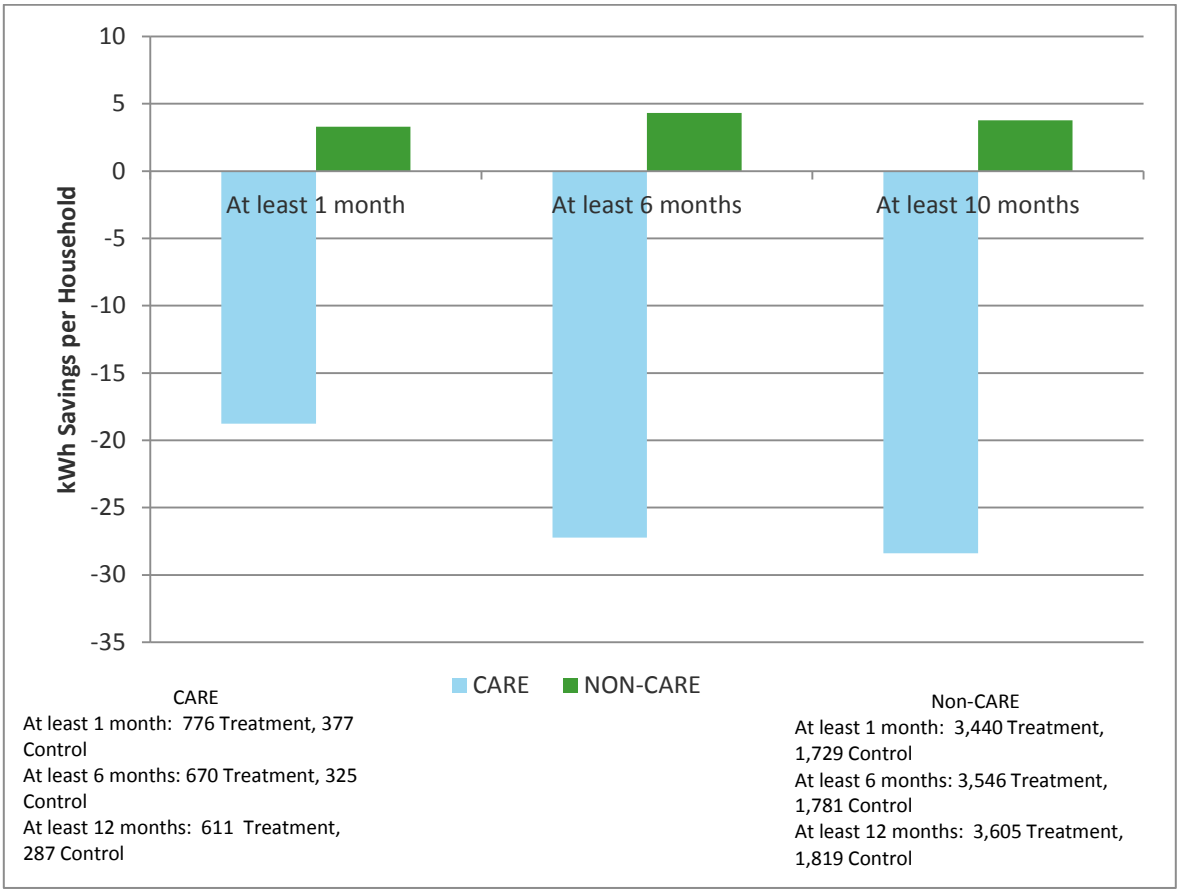


Figure 16 shows HUR-3 CARE and non-CARE customers' total savings. CARE customers did not generate any savings, while non-CARE customers did; however, HUR-3 only had two months of data.

Figure 16: CARE and Non-CARE Electric Savings for HUR-3



Appendix AA. Standardized High Level Savings

The tables in Appendix AA summarizing natural gas savings make use of the unit MTherms – 1,000 Therms – rather than MMTherms – 1,000,000 Therms – for formatting purposes.

Gross Lifecycle Savings (MWh)

Report Name	PA	Standard Report Group	Ex-Ante Gross	Ex-Post Gross	GRR	% Ex-Ante Gross Pass Through	Eval GRR
RES 3.1	PG&E	Home Energy Reports					
RES 3.1	**PG&E**	**Total**					
RES 3.1		***Statewide***					
RES 3.2	SCE	Home Energy Reports					
RES 3.2	**SCE**	**Total**					
RES 3.2		***Statewide***					
RES 3.3	SDG&E	Home Energy Reports					
RES 3.3	**SDG&E**	**Total**					
RES 3.3		***Statewide***					
RES 3.4	MCE	Home Utility Reports					
RES 3.4	**MCE**	**Total**					
RES 3.4		***Statewide***					

Net Lifecycle Savings (MWh)

Report Name	PA	Standard Report Group	Ex-Ante Net	Ex-Post Net	NRR	% Ex-Ante Net Pass Through	Ex-Ante NTG	Ex-Post NTG	Eval Ex-Ante NTG	Eval Ex-Post NTG
RES 3.1	PG&E	Home Energy Reports		107,704						
RES 3.1	**PG&E**	**Total**		**107,704**						
RES 3.1		***Statewide***		***107,704***						
RES 3.2	SCE	Home Energy Reports		3,496						
RES 3.2	**SCE**	**Total**		**3,496**						
RES 3.2		***Statewide***		***3,496***						
RES 3.3	SDG&E	Home Energy Reports		3,575						
RES 3.3	**SDG&E**	**Total**		**3,575**						
RES 3.3		***Statewide***		***3,575***						
RES 3.4	MCE	Home Utility Reports		0						
RES 3.4	**MCE**	**Total**		**0**						
RES 3.4		***Statewide***		***0***						

Gross Lifecycle Savings (MW)

Report Name	PA	Standard Report Group	Ex-Ante Gross	Ex-Post Gross	GRR	% Ex-Ante Gross Pass Through	Eval GRR
RES 3.1	PG&E	Home Energy Reports					
RES 3.1	**PG&E**	**Total**					
RES 3.1		***Statewide***					
RES 3.2	SCE	Home Energy Reports					
RES 3.2	**SCE**	**Total**					
RES 3.2		***Statewide***					
RES 3.3	SDG&E	Home Energy Reports					
RES 3.3	**SDG&E**	**Total**					
RES 3.3		***Statewide***					
RES 3.4	MCE	Home Utility Reports					
RES 3.4	**MCE**	**Total**					
RES 3.4		***Statewide***					

Net Lifecycle Savings (MW)

Report Name	PA	Standard Report Group	Ex-Ante Net	Ex-Post Net	NRR	% Ex-Ante Net Pass Through	Ex-Ante NTG	Ex-Post NTG	Eval Ex-Ante NTG	Eval Ex-Post NTG
RES 3.1	PG&E	Home Energy Reports		19.5						
RES 3.1	**PG&E**	**Total**		**19.5**						
RES 3.1		***Statewide***		***19.5***						
RES 3.2	SCE	Home Energy Reports		0.8						
RES 3.2	**SCE**	**Total**		**0.8**						
RES 3.2		***Statewide***		***0.8***						
RES 3.3	SDG&E	Home Energy Reports								
RES 3.3	**SDG&E**	**Total**								
RES 3.3		***Statewide***								
RES 3.4	MCE	Home Utility Reports								
RES 3.4	**MCE**	**Total**								
RES 3.4		***Statewide***								

Gross Lifecycle Savings (MTherms)

Report Name	PA	Standard Report Group	Ex-Ante Gross	Ex-Post Gross	GRR	% Ex-Ante Gross Pass Through	Eval GRR
RES 3.1	PG&E	Home Energy Reports					
RES 3.1	**PG&E**	**Total**					
RES 3.1		***Statewide***					
RES 3.2	SCE	Home Energy Reports					
RES 3.2	**SCE**	**Total**					
RES 3.2		***Statewide***					
RES 3.3	SDG&E	Home Energy Reports					
RES 3.3	**SDG&E**	**Total**					
RES 3.3		***Statewide***					
RES 3.4	MCE	Home Utility Reports					
RES 3.4	**MCE**	**Total**					
RES 3.4		***Statewide***					

Net Lifecycle Savings (MTherms)

Report Name	PA	Standard Report Group	Ex-Ante Net	Ex-Post Net	NRR	% Ex-Ante Net Pass Through	Ex-Ante NTG	Ex-Post NTG	Eval Ex-Ante NTG	Eval Ex-Post NTG
RES 3.1	PG&E	Home Energy Reports		3,017						
RES 3.1	**PG&E**	**Total**		**3,017**						
RES 3.1		***Statewide***		***3,017***						
RES 3.2	SCE	Home Energy Reports								
RES 3.2	**SCE**	**Total**								
RES 3.2		***Statewide***								
RES 3.3	SDG&E	Home Energy Reports		124						
RES 3.3	**SDG&E**	**Total**		**124**						
RES 3.3		***Statewide***		***124***						
RES 3.4	MCE	Home Utility Reports								
RES 3.4	**MCE**	**Total**								
RES 3.4		***Statewide***								

Gross First Year Savings (MWh)

Report Name	PA	Standard Report Group	Ex-Ante Gross	Ex-Post Gross	GRR	% Ex-Ante Gross Pass Through	Eval GRR
RES 3.1	PG&E	Home Energy Reports					
RES 3.1	**PG&E**	**Total**					
RES 3.1		***Statewide***					
RES 3.2	SCE	Home Energy Reports					
RES 3.2	**SCE**	**Total**					
RES 3.2		***Statewide***					
RES 3.3	SDG&E	Home Energy Reports					
RES 3.3	**SDG&E**	**Total**					
RES 3.3		***Statewide***					
RES 3.4	MCE	Home Utility Reports					
RES 3.4	**MCE**	**Total**					
RES 3.4		***Statewide***					

Net First Year Savings (MWh)

Report Name	PA	Standard Report Group	Ex-Ante Net	Ex-Post Net	NRR	% Ex-Ante Net Pass Through	Ex-Ante NTG	Ex-Post NTG	Eval Ex-Ante NTG	Eval Ex-Post NTG
RES 3.1	PG&E	Home Energy Reports		107,704						
RES 3.1	**PG&E**	**Total**		**107,704**						
RES 3.1		***Statewide***		***107,704***						
RES 3.2	SCE	Home Energy Reports		3,496						
RES 3.2	**SCE**	**Total**		**3,496**						
RES 3.2		***Statewide***		***3,496***						
RES 3.3	SDG&E	Home Energy Reports		3,575						
RES 3.3	**SDG&E**	**Total**		**3,575**						
RES 3.3		***Statewide***		***3,575***						
RES 3.4	MCE	Home Utility Reports		0						
RES 3.4	**MCE**	**Total**		**0**						
RES 3.4		***Statewide***		***0***						

Gross First Year Savings (MW)

Report Name	PA	Standard Report Group	Ex-Ante Gross	Ex-Post Gross	GRR	% Ex-Ante Gross Pass Through	Eval GRR
RES 3.1	PG&E	Home Energy Reports					
RES 3.1	**PG&E**	**Total**					
RES 3.1		***Statewide***					
RES 3.2	SCE	Home Energy Reports					
RES 3.2	**SCE**	**Total**					
RES 3.2		***Statewide***					
RES 3.3	SDG&E	Home Energy Reports					
RES 3.3	**SDG&E**	**Total**					
RES 3.3		***Statewide***					
RES 3.4	MCE	Home Utility Reports					
RES 3.4	**MCE**	**Total**					
RES 3.4		***Statewide***					

Net First Year Savings (MW)

Report Name	PA	Standard Report Group	Ex-Ante Net	Ex-Post Net	NRR	% Ex-Ante Net Pass Through	Ex-Ante NTG	Ex-Post NTG	Eval Ex-Ante NTG	Eval Ex-Post NTG
RES 3.1	PG&E	Home Energy Reports		19.5						
RES 3.1	**PG&E**	**Total**		**19.5**						
RES 3.1		***Statewide***		***19.5***						
RES 3.2	SCE	Home Energy Reports		0.8						
RES 3.2	**SCE**	**Total**		**0.8**						
RES 3.2		***Statewide***		***0.8***						
RES 3.3	SDG&E	Home Energy Reports								
RES 3.3	**SDG&E**	**Total**								
RES 3.3		***Statewide***								
RES 3.4	MCE	Home Utility Reports								
RES 3.4	**MCE**	**Total**								
RES 3.4		***Statewide***								

Gross First Year Savings (MTherms)

Report Name	PA	Standard Report Group	Ex-Ante Gross	Ex-Post Gross	GRR	% Ex-Ante Gross Pass Through	Eval GRR
RES 3.1	PG&E	Home Energy Reports					
RES 3.1	**PG&E**	**Total**					
RES 3.1		***Statewide***					
RES 3.2	SCE	Home Energy Reports					
RES 3.2	**SCE**	**Total**					
RES 3.2		***Statewide***					
RES 3.3	SDG&E	Home Energy Reports					
RES 3.3	**SDG&E**	**Total**					
RES 3.3		***Statewide***					
RES 3.4	MCE	Home Utility Reports					
RES 3.4	**MCE**	**Total**					
RES 3.4		***Statewide***					

Net First Year Savings (MTherms)

Report Name	PA	Standard Report Group	Ex-Ante Net	Ex-Post Net	NRR	% Ex-Ante Net Pass Through	Ex-Ante NTG	Ex-Post NTG	Eval Ex-Ante NTG	Eval Ex-Post NTG
RES 3.1	PG&E	Home Energy Reports		3,017						
RES 3.1	**PG&E**	**Total**		**3,017**						
RES 3.1		***Statewide***		***3,017***						
RES 3.2	SCE	Home Energy Reports								
RES 3.2	**SCE**	**Total**								
RES 3.2		***Statewide***								
RES 3.3	SDG&E	Home Energy Reports		124						
RES 3.3	**SDG&E**	**Total**		**124**						
RES 3.3		***Statewide***		***124***						
RES 3.4	MCE	Home Utility Reports								
RES 3.4	**MCE**	**Total**								
RES 3.4		***Statewide***								

Appendix AB. Standardized Per Unit Savings

Per Unit (Quantity) Gross Energy Savings (kWh)

Report Name	PA	Standard Report Group	Pass Through	% ER Ex-Ante	% ER Ex-Post	Average EUL (yr)	Ex-Post Lifecycle	Ex-Post First Year	Ex-Post Annualized
RES 3.1	PG&E	Home Energy Reports	0		0.0%	1.0			
RES 3.2	SCE	Home Energy Reports	0		0.0%	1.0			
RES 3.3	SDG&E	Home Energy Reports	0		0.0%	1.0			
RES 3.4	MCE	Home Utility Reports	0		0.0%	1.0			

Per Unit (Quantity) Gross Energy Savings (Therms)

Report Name	PA	Standard Report Group	Pass Through	% ER Ex-Ante	% ER Ex-Post	Average EUL (yr)	Ex-Post Lifecycle	Ex-Post First Year	Ex-Post Annualized
RES 3.1	PG&E	Home Energy Reports	0		0.0%	1.0			
RES 3.2	SCE	Home Energy Reports	0		0.0%	1.0			
RES 3.3	SDG&E	Home Energy Reports	0		0.0%	1.0			
RES 3.4	MCE	Home Utility Reports	0		0.0%	1.0			

Per Unit (Quantity) Net Energy Savings (kWh)

Report Name	PA	Standard Report Group	Pass Through	% ER Ex-Ante	% ER Ex-Post	Average EUL (yr)	Ex-Post Lifecycle	Ex-Post First Year	Ex-Post Annualized
RES 3.1	PG&E	Home Energy Reports	0		0.0%	1.0	77.1	77.1	77.1
RES 3.2	SCE	Home Energy Reports	0		0.0%	1.0	48.0	48.0	48.0
RES 3.3	SDG&E	Home Energy Reports	0		0.0%	1.0	239.6	239.6	239.6
RES 3.4	MCE	Home Utility Reports	0		0.0%	1.0	0.0	0.0	0.0

Per Unit (Quantity) Net Energy Savings (Therms)

Report Name	PA	Standard Report Group	Pass Through	% ER Ex-Ante	% ER Ex-Post	Average EUL (yr)	Ex-Post Lifecycle	Ex-Post First Year	Ex-Post Annualized
RES 3.1	PG&E	Home Energy Reports	0		0.0%	1.0	2.2	2.2	2.2
RES 3.2	SCE	Home Energy Reports	0		0.0%	1.0			
RES 3.3	SDG&E	Home Energy Reports	0		0.0%	1.0	8.2	8.2	8.2
RES 3.4	MCE	Home Utility Reports	0		0.0%	1.0			

Appendix AC. Recommendations

Study ID	Study Type	Study Title	Study Manager
Res 3	Impact Evaluation	Validation and Impact Evaluation of IOU's 2014 Home Energy Reports Program	CPUC

ommendat	Program or Database	Summary of Findings	Additional Supporting Information	Best Practice / Recommendations	Recommendation Recipient	Affected Workpaper or DEER
1	HER	DNV GL and the IOUs are using different assumptions on the distribution of savings from measures installed under IOU rebate programs.	N/A	DNV GL is working with the IOUs and their consultants to standardize the approach used in joint savings analysis.	DNV GL, PG&E, SCE and SDG&E	N/A
2	HER	DNV GL and the IOUs are using different approaches in calculating joint savings at the peak.	N/A	DNV GL proposes leveraging CA statewide lighting report to estimate peak savings from efficient bulbs. DNV GL is working with the IOUs and their consultants to standardize the approach.	DNV GL, PG&E, SCE and SDG&E	N/A
3	HER	DNV GL's inability to replicate the climate zone heat waves identified in PG&E HER early impact study while seeming to leverage data from the same underlying sources and approaches, presents evidence that peak periods using the DEER definition is sensitive to small changes.	N/A	DNV GL proposes to employ a separate definition of peak period for comparison with the current peak definition. DNV GL is working with the IOUs and their consultants to standardize this process.	DNV GL, PG&E, SCE and SDG&E	N/A
4	HER	The IOUs are using slightly different approaches in peak demand savings that can produce substantially different results.	N/A	Estimate or continue to estimate demand savings at the wave-level instead of calculating demand savings at the climate zone-level. DNV GL is working with the IOUs and their consultants to standardize the approach used in calculating peak demand savings.	DNV GL, PG&E, SCE and SDG&E	N/A

5	HER	Discrepancies between DNV GL program saving estimates and saving estimates reported in the IOU's early impact evaluation reports are mostly due to differences in billing month assignments.	N/A	Standardize the billing month assignment. Use or continue to use the mid-point when assigning billing months to standardize the approach and minimize the sources of discrepancies in the results.	DNV GL, PG&E, SCE and SDG&E	N/A
6	HER	Rebate savings from program participation of inactive customers were counted in joint savings calculation for PG&E HER early impact study.	N/A	DNV GL recommends calculating joint savings based on rebate participation of customers that are still active in 2014.	PG&E	N/A
7	HER	Combining households from all Gamma waves (or Wave One) can produce results that are substantially different.	N/A	DNV GL recommends splitting out Gamma and Wave One sub-waves in the PG&E HER rebate analysis so that the treatment group is compared to the corresponding control group and for consistency with the approach used in energy savings calculation	PG&E	N/A
8	HER	Early impact evaluation of PG&E HER reported standard errors for the aggregated savings that were based on a regression model at the wave-level where an overall post-treatment indicator was specified	N/A	The standard errors of the annual savings should be calculated using the combined monthly parameter standard errors weighted by the monthly counts.	PG&E	N/A

Appendix BA. Public Comments on 2014 MCE HER Evaluation

No.	From	Section	Comments	Response
1	PEI	MCE HUR-1	As advised previously, program treatment actually commenced in December, 2013; not November, 2013 – this needs to be corrected and results restated, the impact is material and statistically significant. i. The first batch of reports (only 659 participants) were mailed on November 27, 2013, and arrived with customers after Thanksgiving weekend, after December 1, 2013. ii. The remaining participants' reports were mailed on December 15, 2013; indeed, these reports likely did not arrive until after the holidays. iii. Program treatment cannot be considered as underway until at least December, 2013; the evaluation period must reflect this and results corrected accordingly.	Our approach started estimating monthly savings during the month when the first mailing started, in this case November 2013. The pre and post period assignment we applied to the MCE HUR program is identical to the approach we used when evaluating the HER program for the other three IOUs. Also in Table 20, DNV GL showed program savings for HUR-1 for two evaluation periods (Nov2013-Dec 2014 and Jan2014-Dec2014) and none of the savings were statistically significant.
2	PEI	MCE HUR-1	As acknowledged in the draft report, there is a known and measurable imbalance in the HUR 1 program's treatment and control population characteristics, despite the utilization of proven and approved (using Evaluators' Protocols) randomization techniques, and including the use of third party data. i. Although valid techniques are used in the draft report eliminate the impact of these imbalances, the causes of this imbalance warrant determination and are likely related to the data problems the program was subject to.	DNV GL provides oversight role in sampling design for behavioral programs to ensure that the experimental design adhere to CPUC standards and also to avoid problems with the setup of the experiment. MCE/PEI conducted the randomization and DNV GL's involvement started after the HUR program was already underway. Any imbalance in the characteristics of the treatment and control groups under the assumption that the randomization was correctly done is purely by chance. Addressing data problems used by the program that may have caused the imbalance in the sample is outside the scope of this evaluation.
3	PEI	MCE HUR 2 Program	a. The HUR 2 program was designed and operated as 4 cohorts; the evaluation	Ideally, the sampling design of the program must be approved first by ED

			should reflect this structure and not consolidate the results into 2 cohorts. Without using this 4 cohort design structure, a misleading picture is presented, and valuable insights into the nature of the Marin population's responsiveness are needlessly lost. i. The program was specifically structured using stratified random sampling (using Evaluators' Protocols) to facilitate measurement of all 4 cohorts; each have their matching treated and control populations: Higher and Lower usage consumers were further designated as Monthly and Quarterly treatment recipients. 1. Indeed, in the program filing the design intent to focus on high users, and to be distinguished from low users, was made clear as a key aspect to program design. ii. Notably, an evaluation of all 4 cohorts provides the insight that the 2 continued cohorts had a statistically significant savings rate of 2.4%, while the discontinued cohorts had the reverse. 1. Treatment for the HUR 2 ML and HUR 2 QL cohorts were discontinued, as instructed by our client, and should be separately analyzed from the HUR 2 MH and HUR 2 QH cohorts. 2. Indeed, programs are encouraged to undertake such discontinuance decisions to maximize their TRC results (in both reducing costs as well as limiting negative impacts), consistent with D. 10-04-029 and ex-post evaluations. iii. Evaluating the 4 cohorts (rather than consolidating them into 2 cohorts) will not "compromise the integrity of the RCT evaluation;" such an assertion also ignores the fact that the experiment was specifically and deliberately structured using stratified random sampling to gain insight into Marin residents' responsiveness characteristics	(or reviewed by DNV GL) before program implementation. This is the process that PG&E, SCE and SDG&E follow when implementing additional experimental wave or implementing a new behavioral program. For the HUR program, DNV GL's involvement started after the HUR program was already underway. The sampling documentation provided by MCE should have specifically described the four different cohorts and the stratified sampling procedure employed. Based on MCE's sampling documentation, we quote "HUR2 only: Initially targeting all usage quintiles, after three months, customers in the bottom two quintiles were dropped from treated and control groups in order to optimize savings." This statement suggests that the original intent of the HUR-2 experiment is to target all usage quintiles and that the experiment was changed during the course of the program. Behavioral program design of the other IOUs either targets all usage quintiles or specific usage segments. The target groups are determined at the outset as well as the frequency of mailings and planned discontinuation of the treatment. DNV GL evaluates the program based on the sample design ED/DNV GL reviewed and includes all participants that were included in the randomization in its impact evaluation to maintain the integrity of the experiment. After the evaluation period, MCE/PEI submitted own impact evaluation for a subset of the HUR-2 wave where they showed savings rate of 2.4%. DNV GL only reviewed the report but did not have access to program codes and datasets used in MCE/PEI evaluation.

			(again, a key experimental objective, and in compliance with Evaluators' Protocols). 1. Detailed data/information on participants' segregation in each of the 4 cohorts has been provided for the evaluation (in 2015). 2. Should evaluators now decide that discontinuance of HUR 2 ML and HUR 2 QL cohorts' treatment are in some way a problem for their measurement (which warrants explanation; treatment discontinuance is readily separable from performance measurement), then this position should have no impact to the evaluation of the separate HUR 2 MH and HUR 2 QH cohorts. a. It is unclear how ex-post evaluators who were not part of the program's design can take the position that the program was designed as, and can only be evaluated as, 2 cohorts. iv. The draft evaluation points to the likely different nature of Marin residents' responsiveness; a deft acknowledgement of why the program was structured in 4 cohorts. It seems counter-productive to then consolidate the evaluation into 2 cohorts and lose the opportunity to report the program's key insight. 1. In order to inform and properly design future behavioral programs, such as to achieve TRC goals, it is imperative to understand such market characteristics; structuring evaluations such that they ignore market characteristics is	One major concern with the analysis is that the estimated savings of 2.4% were only based on the chosen sample groups for which the treatment continued. PEI/MCE's evaluation also showed that when all households in HUR-2 were included in the analysis, savings were not statistically significant. These two results suggest that the sample groups removed in the analysis are households that had positive impact on consumption on average after receiving the treatment. In addition, the analysis period covered in PEI/MCE's analysis included January to August of 2015 when the current evaluation is for 2014 program cycle.
4	PEI	MCE HUR 3	a. As advised previously, program treatment actually commenced in December, 2014; not November, 2014 – this needs to be corrected. b. It is not clear why one month's performance of this cohort warrants inclusion in this evaluation.	See response # 1. DNV GL included this wave the same way we include new waves for other IOUs that started late in 2014.

5	PEI	Data Issues	a. There is considerable evidence that the usage data provided by PG&E over the course of the program differs materially from that utilized in the evaluation, and possesses quality issues. i. When applied to the usage data supplied throughout the program, different configurations of the difference-in-differences RCT method produce different results; this strongly indicates the presence of measurement error in usage data. ii. The evaluation draft reports roughly 3% lower usage than the data provided throughout the program – potentially swamping the expected savings rates of approximately 2%. iii. The evaluation draft has substantially different customer counts and eliminations from those actually in the program; substantially unbalanced/disproportionally different between control and treated population counts (e.g. HUR 2 difference in participant counts: control +4.2%; treated 0.0%). iv. The report's conclusions on imbalances in pre-treatment differ from the data utilized in the program, suggesting differences in usage datasets. v. The usage data utilized throughout the program shows substantive correlation in errors; this implies that standard errors are incorrect.	DNV GL used program participation data from MCE and customer/consumption datasets provided by PG&E to CPUC for the entire program period to conduct the ex-post evaluation. The data preparation process we applied to MCE HUR program is identical to the approach we used in program evaluation for the other IOUs. While it is unfortunate that there were issues in the consumption data provided by PG&E to MCE for the program, these issues are outside the scope of our evaluation.
6	PEI		4. data issues: Poor data quality and delays likely retarded results; this needs to be characterized and acknowledged in the evaluators report, and data processes and standards adjusted to eliminate such problems (that remain after 3 years). i. Should the data provided to the program and messaged to customers have been	DNV GL is doing an ex-post evaluation where we look at the program as a whole. Any issues with program implementation are reflected in our results. In this case, this issue with implementation and timing of reports help explain poor program impacts of MCE HUR.

			inaccurate (see 4.a.), customer confidence, and thus responsiveness, will have been materially impacted as messaging hasn't match customer actions or bills. These inaccuracies warrant measurement and reference in the report. ii. Usage data utilized throughout the program was provided on a 'cumulative, rollingbatch' basis, and is quite stale by the time it is delivered in HURs; this has meant some customers receive HURs that report on behaviors exhibited over 2 months prior to the report being read (that is, one season late). iii. These data issues themselves should be noted in the evaluators report, as should the need to improve the data flow process to CCAs and RENs, as they increasingly undertake such programs.	
7	MCE		Marin Clean Energy (MCE) respectfully requests that the evaluation be updated to reflect the lack of guidance provided by Energy Division regarding the proper procedure for conducting a residential behavioral evaluation program. While MCE acknowledges that there were faults and challenges in program design and implementation which may have led to the finding of no savings, MCE also feels that early guidance from third party evaluators could have assisted in identifying and trouble shooting these challenges early on. When MCE prepared and submitted its 2013 – 2014 Application, MCE furnished a work paper, which pointed to the Home Utility Report (HUR) program as the primary means of claiming savings for the program. This work paper was submitted to the *ex ante* team and the analyst in Energy Division assigned to work with MCE's portfolio. At this time, Energy Division did not advise MCE that there would need to be involvement from a third party evaluator	This comment is not an evaluation issue and is outside the scope of DNV GL's analysis.

			contracted with the CPUC in order to claim savings for the HUR program. In 2014, Planet Ecosystems Inc., the consultant with whom MCE contracted to implement the HUR program, reported findings of energy reductions of 2% being observed in the control group. MCE sought guidance from Energy Division on the appropriate method for claiming savings in quarterly reports. While MCE originally sought guidance in May of 2014, it wasn't until November of 2014 that Energy Division made MCE aware of the requirement to involve an external consultant in order to be able to claim savings from the program. At this time, MCE quickly worked with DNV-GL to incorporate them into the process of program design and evaluation. MCE requests the final evaluation include acknowledgment that program results may have differed if Energy Division had indicated the need for a third party evaluation team involvement in program design in a timely manner.	

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.