AMI Billing Regression Study

Final Report

February 23, 2016

Table of Contents

List of Tables

List of Figures

Executive Summary

The move toward advanced metering infrastructure (AMI) data has the potential to provide evaluators with a clearer understanding how households and businesses use energy. A single customer's metered data at one-hour intervals translate to over 700 data points per month. With this improved granularity, AMI data provide an important opportunity to develop impact estimates tailored to specific days or hours, rather than a daily average impact derived from monthly data.

One of the key areas where AMI data have the potential to improve accuracy is in billing regression models used to estimate program impacts. Billing regressions have traditionally relied on monthly consumption data, as these are typically all that have been available for estimating impacts at the program level. As most impact evaluations focus on developing annual savings estimates, monthly consumption data have been adequate for these models. With the advent of AMI data it is time to revisit the billing regression method and explore ways these traditional models (and other popular impact analysis tools) can be modified to take full advantage of the additional information available with hourly consumption data.

To explore the potential benefits of using AMI data, Southern California Edison (SCE), on behalf of SCE, PG&E, SDG&E and SoCal Gas, contracted with Evergreen Economics and SBW Consulting (the Evergreen team) to conduct an in-depth analysis using AMI data combined with additional customer data. The two primary goals of this study were to:

1. Use billing regression models and AMI data to estimate HVAC program impacts when both whole-house AMI data and HVAC end-use metered data are available, and;
2. Use billing regression models and AMI data to estimate HVAC program impacts when only whole-house AMI data are available (i.e., no HVAC metered data are available).

To conduct the study, the Evergreen team needed to identify additional data sources that could meet the original research objectives. Several alternative data sources were explored, and three sources were ultimately chosen that provided an opportunity to test billing regression models and meet (at least in part) the original goals of the study. The three data sources selected were the following:

- **Northwest Energy Efficiency Alliance Residential Building Stock Assessment Metering Study** – a publicly available dataset of 15-minute interval, whole house and end-use submetered consumption data for 103 homes in the Pacific Northwest.
- **SCE Residential Quality Installation (QI) Program Participant Data –** a dataset containing 1-hour interval whole house metered consumption on 2,039 homes that participated in the SCE QI Program between January 2012 and December 2014. The SCE QI Program is a California statewide program designed to achieve energy and demand savings through the installation of replacement split or packaged HVAC systems in accordance with industry standards.

- **PG&E Residential Quality Maintenance (QM) Program Participant Data -** a dataset containing 1-hour interval whole house metered consumption on 1,230 homes that participated in the PG&E QM Program between January 2012 and December 2014. The PG&E QM Program is part of a California statewide program designed to achieve energy and demand savings through assessment and optimization of existing residential HVAC units through ongoing maintenance.

While no one data source was ideal, together these sources provided the range of data needed to test the various billing analysis tools and investigate program impacts. A detailed summary of households in each of these datasets, including their energy consumption and weather, is provided in the report appendix.

Once the analysis datasets were compiled, the following billing analysis methods were tested using the AMI data:

- Random Coefficients Model
- Fixed Effects Model
- Princeton Scorekeeping Method (PRISM)
- Energy Charting and Metrics Tool (ECAM)

Of these approaches, the most attention was devoted to the random coefficients model that the Evergreen team adapted for use in the billing analysis. The random coefficients model involved a multi-stage process that first categorized ("binned") customers into groups based on energy use and weather. Once the binning assignment process was complete, separate models were estimated to predict energy consumption for each bin category. These estimates were then used to develop load shapes at the customer level, and predicted load shapes were compared with actual energy usage to estimate program savings.

As discussed throughout this report, we believe the random coefficients model represents a significant improvement over traditional billing regression models as it provides an efficient method for tailoring impacts to specific customer conditions (e.g., day types, seasons, customer types). The random coefficients model also proved to be extremely accurate when tested against a holdout sample of customers.

The results of the random coefficient model using each of the datasets is as follows:

- **SCE QI Program.** The random coefficients model was very accurate in predicting load shapes, with forecasted usage within 1 percent of actual usage for a holdout sample of customers. To estimate program savings, daily load shapes were estimated for the post-participation period and then compared with actual usage over the same time period. The difference between actual and forecasted usage was used as an estimate of energy savings. For the QI program, estimated annual savings was about 7 percent of total usage. Most of the savings occurred during peak hours (as would be expected), which provided additional support the model specification. Seasonal impacts were also calculated and showed larger savings in the summer months, which provides additional support for the model and illustrates how the model can be used to adjust savings estimates for different season, rather than using a single average annual savings value, which is the current standard practice.

- **PG&E QM Program.** The results of the QM Program analysis were similar to the QI program. The random coefficients model using QM data was very accurate in predicting load shapes, with forecasted usage also within 1 percent of actual usage for a holdout sample of customers. Program savings was estimated as the difference between forecasted and actual daily energy use in the post-participation period. Estimated QM Program savings was 3.6 percent annually. As with the QI model, most of the QM savings occurred during peak hours and during the summer months.
- **RBSA Data Analysis.** The random coefficients model was also used to estimate the HVAC load using the RBSA dataset, as this was the only dataset available that both whole house and HVAC metered data. There was a small sample of homes (n=61 for homes with central heating or cooling) within the RBSA that could be used to test how well the model could predict just the HVAC end use. Using this sample, the random coefficients model prediction was within about 1 percent of actual HVAC load on a daily basis.

The results from these three different tests of the random coefficients model were very encouraging. In each case, the model performed very well forecasting energy use for a holdout sample of customers, with estimates generally within approximately 1 percent of actual usage for the holdout group. Impact estimates were also generally in line with expectations for both the QI and QM programs.

In addition to its forecast accuracy, another important advantage of the random coefficients model is the ability to automatically generate load shapes for a wide range of conditions. The models tested in this report were able to automatically generate load shapes for daily, seasonal, and annual values. Given the structure of the binning process, additional load shapes for other subcategories can be easily generated. This is in contrast to the other traditional billing analysis methods where separate models typically need to be developed manually, which makes it difficult to develop load shapes and savings estimates for more than a few subcategories.

Recommended future research includes using the random coefficients model for a sample of commercial customers, expanding the analysis to include a comparison group in the model sample, and experimenting more with model parameters (e.g., binning process, set point temperatures) to determine what effect underlying assumptions for these factors are influencing the model results. Analyzing customer information in a separate model to correlate specific load shapes to customer characteristics is also recommended for future research.

1 Introduction

As the California investor-owned utilities (IOU's) transition to short-interval metering of households and businesses through the implementation of advanced metering infrastructure (AMI), a greater amount and richer source of data are becoming available to researchers. These data have the potential to address complex questions about customer behavior as well as the performance of HVAC, lighting, and other energy-consuming equipment. AMI data provide an important opportunity for evaluators to better understand the impact that energy efficiency programs have on energy consumption during specific hours of the day, rather than a daily average derived from monthly data. A single customer's metered data at one-hour intervals translate to over 700 data points per month.

The improved granularity of energy usage data has the potential to provide evaluators with a clearer understanding of how energy efficiency measures and other factors affect energy consumption. A common concern among economists and other analysts working with monthly (or daily) interval data is that the aggregation conceals more than it reveals. The availability of short-interval meter data allow for potentially more accurate and robust models.

One of the key areas where AMI data have the potential to improve accuracy is in billing regression models used to estimate program energy impacts. Billing regressions have traditionally relied on monthly consumption data, as these are typically all that have been available for estimating impacts at the program level. As most impact evaluations have traditionally been interested in developing annual savings estimates, monthly consumption data have been adequate for analyzing impacts. With the advent of AMI data, there is an important opportunity to revisit the billing regression method and explore ways these traditional models (and other popular impact analysis tools) can be modified to take full advantage of the additional information available with hourly consumption data.

To explore the potential benefits of using AMI data, Southern California Edison (SCE), on behalf of SCE, PG&E, SDG&E and SoCal Gas, contracted with Evergreen Economics and SBW Consulting (the Evergreen team) to conduct an in-depth analysis using AMI data combined with additional customer data. The two primary goals of this study were to:

3. Use billing regression models and AMI data to estimate HVAC program impacts when both whole-house AMI data and HVAC end-use metered data are available, and;
4. Use billing regression models and AMI data to estimate HVAC program impacts when only whole-house AMI data are available (i.e., no HVAC metered data are available).

To conduct the study, the Evergreen team needed to identify additional data sources that could meet the original research objectives. Several alternative data sources were explored, and three sources were ultimately chosen that provided an opportunity to test billing regression models and meet (at least in part) the original goals of the study. The three data sources we chose were the following:

- **Northwest Energy Efficiency Alliance Residential Building Stock Assessment Metering Study –** a publicly available dataset of 15-minute interval, whole house and end-use submetered consumption data for 103 homes in the Northwest. The study collected two full years of data for the sampled homes from April 2012 to September 2014. The resulting database includes both whole house and HVAC submetered data at 15-minute intervals. These data include submetered end use data but no program participation data, so there are no impacts to be measured. These data were modeling with the following approaches: random coefficients regression, Energy Charting And Metrics, and PRISM.

- **SCE Residential Quality Installation (RQI; QI) Program Participant Data –** a dataset containing 1-hour interval whole house metered consumption on 2,039 homes that participated in the SCE QI Program between January 2012 and December 2014. The SCE QI Program is a California statewide program designed to achieve energy and demand savings through the installation of replacement split or packaged HVAC systems in accordance with industry standards. Program data included household and program participation information including the home climate zone and date of participation in the program. The study team created the final analysis dataset by merging the program data and AMI data using a unique customer ID. These data are program data and include homes that were subjected to an energy efficiency intervention and thus, energy impacts can be measured. However, these data do not include submetered end use data. These data were modeled with the following approaches: random coefficients regression, fixed effects regression, and ECAM.

- **PG&E Residential Quality Maintenance (RQM; QM) Program Participant Data -** a dataset containing 1-hour interval whole house metered consumption on 1,230 homes that participated in the PG&E QM Program between January 2012 and December 2014. The PG&E QM Program is part of a California statewide program designed to achieve energy and demand savings through assessment and optimization of existing residential HVAC units as well as enrolling customers in an ongoing maintenance agreement with a qualifying contractor that performs two maintenance calls per year in the pre-cooling season and pre-heating season. Similar to the SCE QI program, the PG&E QM program data included household and program participation information including technology type, climate zone, and date of participation in the program. These program data include homes that were subjected to an energy efficiency intervention and thus, energy impacts can be measured. However, these data do not include submetered end use data. These data were used in modeling with the following approaches: random coefficients regression, fixed effects regression, and ECAM.

While no one data source was ideal, together these sources provided the range of data needed to test the various billing analysis tools and investigate program impacts.[1] A detailed summary of households in each of these datasets, including their energy consumption and weather, is provided in the report appendix.

The Evergreen team combined each data set with weather data obtained from the National Oceanic and Atmospheric Administration (NOAA) to develop datasets with both consumption and weather data. We selected weather station data based on proximity to each home's zip code, matching climate zone, and availability of complete hourly data.[2] The selection process resulted in hourly data for 95.5 percent of hourly observations. We performed additional analysis to identify unreasonably high or low temperature readings, based on the record high and low temperatures in each climate zone. Missing observations, and temperatures identified as unreasonable were imputed using the next closest weather station if available; otherwise, they were imputed with the average of the preceding and following temperature reads.

Once the analysis datasets were compiled, several billing analysis methods were explored using the AMI data. During the course of this research, however, it became apparent that an innovative new analysis method – the random coefficients model – has the potential to be a ground breaking impact evaluation approach that fully utilizes the benefits of the more granular AMI. As discussed throughout this report, we believe the random coefficients model represents a significant improvement over traditional billing regression models as it provides an efficient method for tailoring impacts to specific customer conditions (e.g., day types, seasons, customer types). The random coefficients model also proved to be very accurate when tested against a holdout sample of customers. Additionally, the automated method developed to group the customer data allows for multiple models to be run relatively easily for different subgroups of interest. Since the random coefficients model has not been used previously for impact evaluation, the decision was made collectively by Evergreen and the IOU/CPUC Study Team to devote the majority of project resources to testing the method and documenting the initial model exploration results.

In addition to the random coefficients model, this research also explored the following more traditional billing regression analysis methods:

- Fixed Effects Model
- Princeton Scorekeeping Method (PRISM)

[1] An example of an ideal dataset would have been a single program dataset with at least 12 months of pre- and post-intervention whole home energy consumption, sub-metered HVAC consumption, weather data, demographic and home characteristics data such as occupancy and home square footage, and detailed HVAC equipment information.
[2] The selection criteria chose the closest weather station to each home, but only considered weather stations that were located in the same climate zone as the home and had hourly data for 2012-2014. This resulted in 60 different weather stations being assigned to the 428 different zip codes. On average, the selected weather station was 17.8 miles from homes that participated in the SCE program and 12 miles from homes that participated in the PG&E program.

- Energy Charting and Metrics Tool (ECAM)

These other methods were given less attention in this research, however, to allow for more exploration of the random coefficients model. The results for each of these alternative methods are provided in the report appendix for context, as they allow for the random coefficient model to be compared with results obtained from more traditional billing analysis techniques.

The remainder of this report presents the random coefficients model, with comparisons to the other alternative analysis methods where appropriate. The report concludes with recommendations for future research.

2 Random Coefficients Model

The primary focus of our AMI analysis ultimately was the random coefficients model, which we believe has great potential for utilizing AMI data to estimate energy savings. Before describing the random coefficients model, it is important to first understand some of the limitations of the standard billing regression model (including the fixed effects specification) that has traditionally been used in impact evaluations.

The standard approach in regression analysis is to focus on the average response of the population of interest. With a billing regression, the model produces a regression line that represents the average energy use across all customers included in the model. The standard billing regression is often limited to one or two coefficient estimates to calculate savings for all customers included in the sample. Variations of the regression model can be developed that produce separate savings estimates for sub-groups of customers but these can be cumbersome to process, as each model needs to be developed and evaluated separately.

Instead of modeling only a single average energy use or energy savings for all customers, the random coefficients model looks at each customer's energy consumption over time and develops savings estimates tailored to specific customer types, and/or weather conditions. The term "random coefficients model" refers to a framework that provides a distribution of model parameters across customer types rather than a single average value. By focusing on the trajectory of each customer instead of the average across all customers, the random coefficients model is able to provide additional information about the changes in energy usage for individual customer types and/or weather conditions. The random coefficients modeling approach still produces a population-based model and savings estimates, equivalent to the standard billing regression approach, but the population model coefficients are aggregated from the customer-specific coefficients.

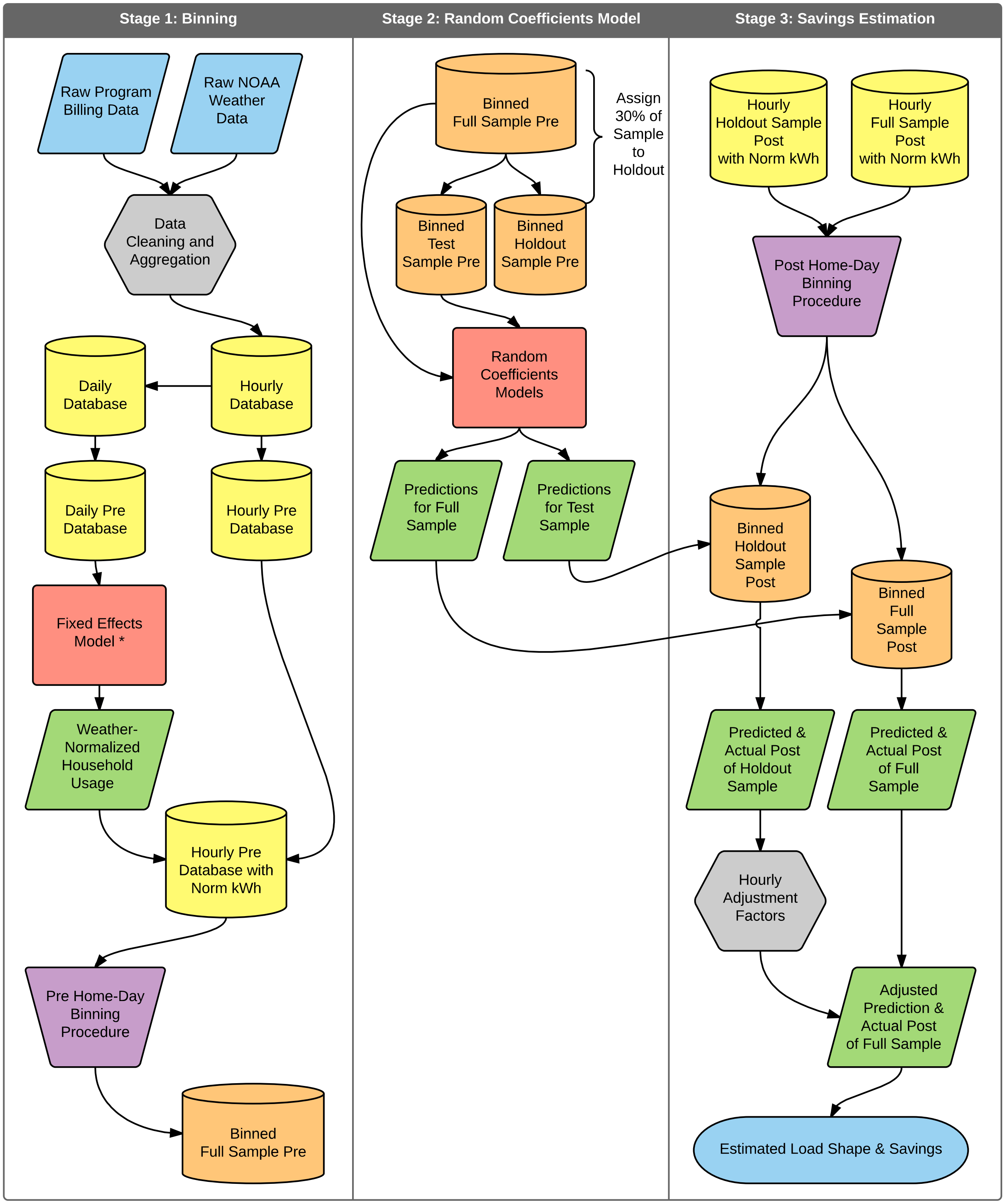
The random coefficients model works by explicitly accounting for two separate sources of variability commonly found in interval energy-use data. The first, within-subject variability, represents the variation in energy usage *throughout the day* by an individual customer. The second, among-subject variability, represents the variation in energy use *across customers and varying weather conditions* experienced by each customer.

In summary, the traditional billing regression model will produce a single impact estimate that is then applied universally to all the participants in the program. The random coefficients model, in contrast, produces savings estimates that are tailored to individual customers based on energy use and weather conditions. Both approaches produce average program-level savings estimates that are suitable for most program evaluations. The random coefficients model achieves this by aggregating up the individual customer-level savings values. As discussed more below, these disaggregated values can be used to provide a richer picture of energy impacts by providing separate savings estimates based on customer type, day type (weekend vs. weekday) and season.

2.1 Random Coefficients Model Development Process

Figure 1 outlines the steps followed to develop the random coefficients model, which combines first stage data categorization process (referred to as "binning" below) and a second stage random coefficients model to estimate the hourly energy use within each bin in both the pre-participation and post-participation period. The process used to describe each of these stages is described in more detail following the figure.

Figure 1: Random Coefficients Model Approach



* The fixed effects model in the first stage is a simple billing model that estimates the weather-normalized household usage for each individual home. This estimate is used to identify low, medium, and high use households for the binning procedure.

2.1.1 First Stage: Binning Process

In the first stage of our modeling approach, we use a fixed effects regression model to estimates of daily baseload electricity use for each home, controlling for outside air temperature.[3] The fixed effects model specification is as follows:

$$DailykWh_{i,t} = a_i + b_1(CDD_{i,t}) + b_2(HDD_{i,t}) + e_t$$

$Where:$

$DailykWh_{i,t}$ = Daily kWh consumption for customer i on day t.

$CDD_{i,t}$ = Cooling degree days (CDD) for customer i on day t.

$HDD_{i,t}$ = Heating degree days (HDD) for customer i on day t.

a_i = Customer specific constant (i.e., baseload weather normalized consumption)

b_1, b_2 = Coefficients estimated in the regression model

$e_{i,t}$ = Random error assumed normally distributed

A characteristic of fixed effects models is the estimation of a specific constant, or intercepts parameter, α_i, for every customer site. This constant varies by customer site and accounts for time-invariant effects on consumption. In the model specification above, the constant can be interpreted as site-specific baseload consumption after controlling for variation in outside air temperature (CDD and HDD, using a base temperature of 65 degrees Fahrenheit). Using statistical software we estimate this constant and obtain and estimate of baseload energy use for each customer site. We then ranked the homes in ascending order of baseload energy use and assigned each home to one of 20 "home groups" based on each home's weather normalized home usage, prior to program participation. In this way we group homes with similar energy consumption together. Each home group represents about five percent of total daily electricity (baseload) usage for the homes in our sample. Because of this, the number of homes in each bin varies, but the amount of daily kWh each bin represents is approximately the same. [4]

[3] Before running the fixed effects model, we removed days with fewer than 24 observations (one per hour) from the SCE QI Program and PG&E QM Program datasets to ensure that inclusion of incomplete days does not bias our estimates for daily consumption.

[4] Homes that are vacant in the pre-period due to long vacations, tenant turnover in rental properties, or other reasons will naturally fall into the lowest home group. If the home is not vacant during the post-period, the home's total usage will increase greatly and may mask program savings. The opposite is expected to be true as well. This is not a limitation of the binning procedure, but is a limitation on any analysis conducted with these data. If we had access to more information about these buildings (e.g., occupancy, owned vs. rental property, vacation vs. permanent residence), we could incorporate it into the binning procedure to limit any bias it may have on the resulting program savings estimates. In order to limit this potential for bias, we removed homes with average daily consumption of less than 5 kWh per day during the study period.

Next, we characterized every day that each home experienced in terms of the weather and day type. To create weather groups, we computed the cooling degree hours (CDH) for each hourly observation using a base temperature of 65 degrees[5] Fahrenheit, and then took the average of these hourly values to create a single cooling degree day (CDD) value for each home on each day (i.e., each "home-day") in the study period. Next we rounded the CDD up to the next integer and assigned it to a CDD group. For example, an hour with an outdoor temperature of 66.2°F would have a CDH of 1.2 (66.2°F – 65.0°F = 1.2). If the average of all 24 CDH was 1.4, it was rounded up to 2 and would be assigned to CDD group 2. For annual models, we repeated this process to assign days to heating degree day (HDD) groups, again using a base temperature of 65 degrees Fahrenheit. Categorizing days using outdoor temperature in this manner explicitly incorporates temperature into our modeling approach. To reflect possible differences in energy usage between weekends and weekdays, we also binned home-days based on the day type. Weekends were assigned to "day type" group 1 and weekdays were assigned to day type group 0.

Lastly, we combined all groups to create home-day bins containing only one type of home on one type of day. These bins describe the home-days in our sample based on the home group (baseline weather normalized energy usage), weather group (CDD and/or HDD), and day type group (weekday versus weekend).[6] Each home remained assigned to just one home group, but because temperature and day type changes day-to-day, each home had home-days that were assigned to many different bins.

This binning process has the following benefits:

- Each bin has only one type of home on one type of day. This means that variation in CDD is controlled for in the bins so it does not need to be included as a variable in the model specification. The same is true for all other binning factors like HDD, day type, and each home's baseline energy usage.[7]
- We are modeling home-days rather than households so we are able to exclude individual days with missing observations from the data. For example, we can remove specific days with less than a complete 24 hours of hourly data (e.g., remove three days from home 113's data because they have 22 hourly observations), rather than limiting the analysis to homes with flawless data throughout the study period (e.g., remove home 113 entirely).

[5] Future research will explore identifying the balance point on a per-building basis, rather than just assuming a constant 65 degrees.

[6] For example, consider a home-day on the weekend from a home whose baseline usage is in the 22nd percentile, on a day with a CDD of 1.4 and a HDD of 7.8. This home-day is part of home group 5, CDD 2, HDD 8, and day type 1; it is therefore assigned to home-day bin 05-02-08-1.

[7] If additional information is known about these households, such as HVAC size or conditioned floor area, then this information could be used to further refine the binning process. This would help avoid the possibility of grouping together houses that have similar total consumption but are very different in other factors (e.g., equipment holdings, envelope, occupancy) that affect energy use.

- When binning annual observations by household group, weather, and day type, only one model is required. The output is generated at the bin-level so the model allows creation of load shapes and savings estimates for each specific bin (i.e., a specific combination of home group, weather, and day type), group (e.g., home group 20 on all possible days, the highest users), or at the program-level (i.e., incorporating all bins), without the need to run additional models.
- Participant households with no post-period observations are still useful when constructing models of the pre-period because they are simply a series of home-days. Their pre-period observations can be grouped with similar home-days of households that do have post-period data. Later in the analysis, households with no post-period observations are automatically excluded from the impact estimates.

2.1.2 Second Stage: Random Coefficients Model

For the next stage, we randomly selected 70 percent of homes to be used in a random coefficients regression model to develop predicted hourly load shapes for each home-day bin using pre-period data. The remaining 30 percent of homes were set aside as a holdout sample to test the performance of the predicted load shape.[8] In this way we ensure that the predictive power of the model is tested against data that were not used to develop the model.

We computed the average hourly kWh value for the homes in each home-day bin selected for modeling. These average hourly values of kWh represent the average load shape for each home-day bin in the final regression model. For large datasets like the annual SCE QI model, which has thousands of observations in a single bin, this approach cuts down on processing time without introducing bias for the resulting coefficients. If processing time is not a concern, all observations can be included in the model.

We specified a random coefficients model because this approach allows us to simultaneously model the daily load shape (i.e., hourly kWh usage) of each home-day bin while accounting for covariance with other home-day bin load shapes. Unlike a typical fixed effects regression, which produces a single set of coefficients and household-specific constants, the random coefficients model produces a vector of regression coefficients for each home-day bin. Our final random coefficients model specification is as follows:

[8] Comparing the model to a random sample of multiple hold out groups will be explored in more detail in the next phase of this research.

$$kW\_Hr_{i,t} = \sum_{j=1}^{5} \beta_{j,i}(ChangeH_{i,t}) + \sum_{k=1}^{5} \beta_{k,i}(ChangeH_{i,t} * H_{i,t}) + \varepsilon_i$$

Where:

$kW\_Hr_{i,t}$ = Mean kW consumption for homes in bin i during hour t.

$ChangeH_{i,t}$ = An array of dummy variables (0,1) representing hourly changepoints, taking a value of 1 if an hourly observation falls between two changepoints. In our final model, we use the changepoints 5am, 8am, 3pm, 6pm, 8pm, and midnight.

$ChangeH_{i,t} * H_{i,t}$ = An array of variables that interact the dummy changepoint variables with the hour of the day.

$\beta_{j,i}, \beta_{k,i}$ = Coefficients estimated in the model for homes in bin i.

ε_i = Random error, assumed normally distributed.

Using the above model specification we generate coefficients for:

- The pre-period for the 70 percent modeling sample. We use these coefficients to test the model's ability to predict pre-period consumption in the 30 percent hold out sample.
- The pre-period using 100 percent of homes. Once we are satisfied with the model predictions compared to the holdout sample, the full sample is then used to estimate the model to take full advantage of all available data. We use these coefficients to develop predicted post-period consumption in the absence of the energy efficiency program intervention.

These coefficients are used to test the model and develop savings estimations as explained in the following section.

2.1.3 Third Stage: Savings Estimation

The first step in our savings estimation approach is to test the predictive ability of our model. We compare the hold-out sample predicted pre-period hourly kWh values, developed using the coefficients from the 70 percent modeling sample, with the actual pre-period hourly kWh values of our holdout sample. If our model is performing well, the predicted pre-period hourly kWh and actual pre-period hourly kWh should be similar, with any difference representing the error that exists in our modeling approach. We create an hourly adjustment factor from this comparison to account for any error, which we use later in the process to improve our modeling predictions.

We then subject the post-period data to the same binning process as we did to the pre-period data (in the first stage). Each individual home remains in the same weather-normalized usage group that they were assigned to in the pre-period, which helps isolate the effect of the program intervention occurring in the post-period by holding the expected general usage constant throughout the analysis period.[9] Next, each day is assigned to a weather group (by CDD and/or HDD) and day type group (i.e., weekdays versus weekends).

After assigning each home-day in the post-period to a home-day bin, in an identical fashion to the pre-period data, we import the predicted hourly pre-period kWh values for each home-day bin in the random coefficients regression model. We then multiply each prediction by our adjustment factor to correct for any error we found in our modeling approach (from the holdout sample). This process gives us a predicted estimate of each household's consumption during each hour of the post-period if they had not participated in the program.

We compare the predicted post-installation hourly kWh values (based on the pre-period consumption model and post-period weather data) with the actual post-period hourly kWh values. This is essentially comparing predicted household consumption, had the program participation not occurred, to actual post-period consumption on days with the same weather conditions and day types. When actual post-period consumption falls below the predicted hourly kWh, this indicates energy savings during that hour attributable to the program. In essence, the estimated program savings is the difference between the predicted post-period hourly kWh and the actual post-period hourly kWh adjusted for any error found in the first step of the savings estimation.

[9] Households with no pre-period observations are automatically removed because they do not have a baseline coefficient group (from the initial fixed effects regression model), making it impossible to assign them to a kWh-bin. Households with no post-period observations are retained and used to improve the pre-period models, but they are also automatically removed when calculating impacts.

Example of Daily Savings Estimates

Unlike a fixed effects regression, the relationship between program savings and each control variable (e.g., CDD) is not explained with a single coefficient or series of interaction terms. Instead, the random coefficients model produces a full set of coefficients for each bin, where each bin has a different combination of values for the control variables. The resulting output reflects the dynamic relationships between program savings and each control variable (e.g., separate CDD coefficients for each home-day bin). The following example demonstrates how the random coefficient model's daily savings estimates for each bin are used to estimate the program-level average daily kWh savings.

Table 1 shows the average daily savings estimate for each group and bin from an example model of an energy efficiency program during weekdays in summer months.[10] The columns show households grouped by their weather normalized energy usage in the pre-period for each home (highest users on the right) and the rows show days grouped by the temperature via cooling degree-days[11] (hottest days on the bottom). Each cell shows the estimated program savings (kWh per day) for a specific home-day bin. We automatically color-coded the cells with the highest kWh savings in dark blue and the lowest kWh savings in dark red; colorless cells fall in the middle of this spectrum. Within each household group, there are home-days from a wide range of temperatures, each with their own savings estimate. Similarly, each group of days with similar temperatures (i.e., CDD) includes home-days from a range of households (i.e., high, mid, and low users), which experience a wide range of daily kWh savings. Negative savings (dark red cells) could indicate that the cooling equipment was either broken or unused in the pre-period.

[10] The tables in this section are derived form a hypothetical example to illustrate the capabilities of the random coefficients model, the tables do not reflect specific analysis presented in the results.
[11] The row labeled "5" for instance is made up of home-days where CDD=5.

Table 1: Example Program Savings (kWh per day) by Bin

		Household Groups																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Cooling Degree-Day Groups	1	4.5	4.9	-5.5	-2.8	1.6	4.9	30.9	-3.5	-4.0	2.6	5.2	3.9	-7.7	2.5	25.5	-8.8	-15.9	0.0	-0.8	0.0	-0.8
	2	1.5	-0.2	-3.8	-8.9	-7.2	-5.3	-6.8	-10.0	-8.0	-0.4	-7.2	-6.9	0.6	-2.7	-6.9	-7.5	-16.0	-7.7	-15.1	-13.5	-6.7
	3	-3.0	-2.7	-2.8	-8.0	-2.4	-6.4	-5.0	-10.0	-9.8	-3.2	-4.3	-4.6	-0.8	-5.9	-6.1	-3.3	-0.7	-3.6	-2.0	14.8	-4.7
	4	-2.8	-4.8	-8.1	-3.4	-2.0	-3.3	-3.9	-4.5	-2.0	-0.9	-1.1	-2.4	5.1	-6.1	7.9	4.0	5.1	3.8	2.5	18.8	-2.0
	5	-3.6	-2.0	-4.1	-1.4	0.5	0.4	-1.1	-3.7	-0.8	-1.0	-2.2	2.0	5.4	2.5	4.2	0.6	13.0	6.9	10.3	22.0	1.5
	6	-2.6	-3.5	-4.2	-0.4	-1.2	0.8	0.6	-2.7	4.3	2.2	0.0	7.0	0.6	7.6	8.7	3.0	10.4	3.3	21.7	23.7	2.7
	7	-4.0	-1.9	-4.2	0.8	-2.1	-0.3	1.2	-3.9	1.3	1.8	3.3	2.4	4.4	1.3	11.4	7.1	10.5	11.4	17.4	30.7	3.0
	8	-4.5	-2.1	-4.7	-1.0	-0.6	0.6	-1.0	-2.3	4.2	1.4	2.3	6.3	1.8	-1.2	5.8	6.6	11.5	9.2	15.3	25.4	2.0
	9	-4.8	-3.5	-3.5	-0.9	-4.2	2.2	3.5	-1.1	5.6	2.6	6.8	4.5	1.5	0.8	8.4	9.1	12.4	6.2	19.0	37.8	2.8
	10	-5.7	-2.9	-2.8	-0.5	1.7	2.9	1.3	0.0	5.3	1.8	5.4	2.9	3.2	0.8	6.6	9.5	15.7	11.1	21.1	30.4	3.0
	11	-6.2	-3.7	-5.0	-2.2	-1.6	1.7	2.9	0.8	7.7	1.6	4.7	3.7	5.9	-0.1	10.6	9.3	13.1	6.0	17.0	24.9	2.2
	12	-6.6	-3.6	-2.5	-0.7	-0.1	1.8	5.6	3.0	9.1	4.5	11.6	6.7	6.6	-0.2	15.0	15.3	14.2	17.0	22.4	41.3	5.2
	13	-7.5	-3.6	-3.9	1.2	0.8	3.8	5.6	4.5	8.2	6.9	11.4	8.3	8.0	8.6	13.4	18.6	13.9	15.9	26.4	33.2	5.4
	14	-7.6	-0.2	-4.2	4.1	5.2	5.4	5.1	8.2	8.7	7.3	13.1	12.4	8.1	9.4	16.2	21.6	19.6	21.1	29.3	46.1	8.3
	15	-7.7	-2.0	-2.8	1.1	5.9	5.7	3.0	7.1	7.9	8.8	11.3	10.6	9.7	9.8	20.2	17.2	20.5	15.7	22.9	28.0	7.2
	16	-8.4	-1.3	1.1	6.2	6.7	5.1	6.4	7.1	11.6	9.8	14.2	14.6	10.5	13.6	20.0	16.8	18.0	16.3	16.9	20.9	7.9
	17	-8.3	-2.8	-5.3	1.3	7.4	9.6	11.0	11.4	15.5	8.9	11.1	14.6	17.3	14.8	14.3	22.6	20.9	22.1	41.3	35.4	8.7
	18	-8.8	-1.8	3.8	0.0	6.2	6.5	18.5	10.9	13.8	7.9	10.3	14.2	17.3	15.1	26.5	18.5	20.7	13.9	33.1	29.9	8.9
	19	-10.7	-0.6	2.5	1.9	-0.2	8.4	17.9	11.2	14.2	16.7	15.7	15.0	19.1	22.9	31.7	23.8	27.4	19.9	23.6	47.9	11.3
	20	-11.6	-4.1	-3.7	0.4	11.8	9.8	12.3	17.9	8.6	10.7	11.6	15.3	18.0	18.7	25.9	22.4	18.1	24.9	31.9	31.0	9.2
	21	-8.3	-8.0	6.2	-1.4	14.1	14.8	20.9	12.2	10.1	6.4	10.3	14.4	20.4	20.7	23.1	22.6	20.4	16.7	30.4	24.5	9.7
	22	-4.9	-8.3	12.2	-3.2	21.8	9.3	23.5	15.5	13.7	16.1	11.7	26.6	18.4	16.3	27.0	28.5	26.2	22.1	58.7	37.1	13.9
	23	-7.8	-7.3	10.6	-7.9	-2.4	18.0	23.7	13.7	14.0	9.8	10.3	5.3	5.1	14.0	16.3	19.8	12.2	23.9	26.3	30.6	7.7
	24	-7.6	-5.9	9.6	4.7	3.3	10.5	23.7	19.4	19.7	12.8	22.1	16.9	25.6	28.0	15.8	27.8	20.7	23.8	27.5	33.0	11.6
	25	-5.5	-2.2	16.2	11.2	20.8	8.4	12.0	12.8	8.4	21.5	46.9	13.8	12.8	18.8	8.9	34.0	21.6	12.3	30.7	37.7	10.3
Total		-6.3	-1.7	-0.4	0.4	3.3	4.3	5.7	4.7	8.5	6.0	8.9	8.3	8.4	7.5	14.5	14.4	16.8	14.4	22.3	32.1	6.0

Table 2 shows the count of home-days in the post-period that was assigned to each bin based on observed energy use (also from an example model). As with the previous table, we automatically color-coded the cells with the highest count in dark blue and the lowest count in dark red, white cells fall somewhere in the middle of this spectrum. Hence, this table shows the actual distribution of participant households and the weather they experienced in the summer of 2014 (i.e., the post-period for this model). In this example, there are more mid-temperature days with CDDs ranging from 7 to 17 than especially high or low temperature days within each of the household groups, as well as overall.[12] Each type of home-day is not equally represented, hence, the average program-level savings cannot simply be the average of the bin-level savings shown above. Instead, the program-level savings is a weighted average of the bin-level savings, weighted by the number of home-days in each bin.

[12] Two of the home-day bins do not have any observations in the post-period, these are HH18:CDD1 and HH20:CDD1. In this example, there are a large number of home-days assigned to household group 1 and CDD group 25 because there were many days with CDD of 25 or greater. Some options for addressing this issue are to add more CDD groups (adds to processing time) or redefine the groups to assign a wider range of CDD to each bin (increases variation in the other bins). We explored these options for the annual models (RBSA, QI, and QM).

Table 2: Example Number of Home-Days in Each Bin

Cooling Degree-Day Groups	Household Groups: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
1	3	9	13	8	10	5	7	9	3	7	3	4	4	5	2	4	4	0	7	0	107
2	12	45	61	40	42	25	27	40	13	34	13	15	19	22	11	19	20	5	32	2	497
3	12	41	49	36	25	23	29	36	18	35	15	21	15	21	15	20	21	8	27	6	473
4	50	100	118	87	67	58	69	80	34	78	41	46	43	51	39	47	53	25	66	12	1,164
5	56	109	106	94	65	68	122	108	96	102	72	122	70	88	86	79	90	68	67	53	1,721
6	83	119	126	109	103	78	114	92	79	83	76	94	67	104	81	74	69	61	70	44	1,726
7	147	240	231	207	174	165	224	213	177	191	144	240	131	184	172	150	152	127	134	95	3,498
8	255	310	295	293	261	219	250	193	174	211	196	211	175	208	190	155	149	145	143	84	4,117
9	235	322	314	310	262	232	250	207	182	237	168	253	170	165	180	153	142	105	132	79	4,098
10	348	444	348	358	334	291	291	216	245	309	269	281	232	233	233	193	194	172	152	107	5,250
11	488	517	411	423	327	371	360	258	326	350	300	298	283	247	231	185	231	190	158	82	6,036
12	254	299	279	254	220	205	230	197	179	217	175	218	171	169	164	147	143	133	120	79	3,853
13	386	370	317	299	250	256	261	204	209	257	222	220	205	218	177	159	170	155	140	83	4,558
14	439	457	347	380	324	309	324	271	313	315	310	301	290	281	279	202	249	239	166	150	5,946
15	344	315	251	267	234	217	242	178	204	229	223	207	205	198	192	178	165	190	116	113	4,268
16	359	316	243	233	223	218	233	157	185	214	208	170	189	190	165	147	154	170	117	89	3,980
17	414	340	281	277	215	235	243	181	215	196	205	172	202	181	140	117	147	127	101	45	4,034
18	210	169	151	146	96	120	135	97	123	118	113	90	113	86	76	66	79	100	50	36	2,174
19	292	183	176	155	109	147	151	118	154	145	120	127	135	107	104	88	116	107	56	57	2,647
20	86	58	59	45	36	41	46	28	40	49	37	42	37	35	31	33	27	33	14	15	792
21	107	51	43	29	27	28	49	30	43	33	25	34	39	21	29	29	33	42	10	28	730
22	261	117	103	83	87	78	87	84	82	82	83	78	74	79	73	61	72	75	53	69	1,781
23	87	34	42	19	22	27	35	30	24	29	13	32	17	14	19	17	12	29	13	26	541
24	161	75	69	56	49	57	52	52	44	40	39	35	44	33	30	39	33	32	30	33	1,003
25	2,229	348	541	225	254	251	167	498	242	160	200	59	184	71	138	130	234	170	341	490	6,932
Total	7,318	5,388	4,974	4,433	3,816	3,724	3,998	3,577	3,404	3,721	3,270	3,370	3,114	3,011	2,857	2,492	2,759	2,508	2,315	1,877	71,926

Computing Standard Errors

As discussed previously, in the second stage random coefficients modeling we estimate individual regression models for each of the approximately 1,000 home-bins, based on the average kWh usage for each hour within each bin. While the output of each regression model includes the statistical error between the actual and predicted kWh for each home-bin, these errors represent the different between the *mean* hourly kWh and the predicted *mean* hourly kWh for each home-bin. What is needed is the error between the actual hourly kWh and predicted mean hourly kWh usage for each hour of each actual home-day in the pre-period. To obtain the correct values, we computed the standard errors for each hour-of-the-day in each home day bin by:

1. Squaring each error between the actual and predicted hourly kWh usage
2. Summing the squared errors for each hour of each home-day bin
3. Computing the variance for each hour of each home-day bin
4. Computing the standard deviation of each hour of each home-day bin as the square root of the variance
5. Computing the standard error of each hour of each home-day bin by dividing the standard deviation by the square root of the number of home-days within each bin

With 24 hours per home-day and approximately 1,000 home-day bins, we computed approximately 24,000 standard errors. Finally, because this exercise was conducted using the pre-data, we computed the relative standard error for each hour of the day for each home bin as the ratio of the standard error to predicted hourly kWh usage. These relative standard errors are then applied to the post-data to compute an estimate of the standard error for each hour of each home-day in the post-period.

2.2 Random Coefficients Model Results

This section presents the results of the random coefficients models using the NEEA RBSA program data, SCE's QI program data and PG&E's QM program data. We used the NEEA RBSA data initially to develop and test the validity of the random coefficients model. Next, we turned to the SCE QI and PG&E QM program data to estimate program impacts. For each of these sources, we provide detailed results of two models, a seasonal model and an annual model.

The seasonal models for the SCE QI and PG&E QM programs use data from weekdays only during summer months. This model is intended to demonstrate that the random coefficients model is able to predict energy load and consumption for homes with cooling equipment on warm days, a quality that is particularly important when assessing programs where all the savings come from changes in the energy usage of air conditioning equipment.

Similarly, the seasonal models for RBSA include a summer cooling model that uses data from summer months in homes with cooling equipment and a winter heating model that uses data from winter months in homes with heating equipment. Unlike the SCE QI and PG&E QM seasonal models, these use data from both weekdays and weekends.

The more complex annual models predict energy load and consumption for participating homes over the entire year on both weekdays and weekends, considering both heating and cooling needs. The annual model generates an annualized savings estimate by detecting all savings during warm summer days when the air conditioning is used, but unlike the summer model, it also looks for evidence of savings on colder days (e.g., heating thermostat setback, replacement/repair of heat pumps used for both heating and cooling, fan savings), or savings on unseasonably warm days in non-summer seasons.

All of these models follow the same basic procedure described previously:

1) Bin the hourly observations into groups of home-days;
2) Run the random coefficients model to predict hourly consumption; and
3) Estimate program savings by comparing the model's predictions of energy usage without the program to what households actually used after participating.

We have provided a series of tables and charts for each model including:

- A table describing the groups used in the binning procedure;
- A graph comparing our pre-period model to the actual consumption of a holdout group in the pre-period;
- A graph comparing our post-period consumption predictions in the absence of program participation (based on the pre-period model and observed post-period weather) with actual consumption in the post-period; and
- A graph depicting hourly estimates for program savings with error bars.

Since the RBSA data do not involve any efficiency programs that save energy (recall that this is a general population building stock assessment), the RBSA discussion does not include estimates of post-period consumption or hourly savings estimates. The RBSA models are useful, however, as they provide an initial test of the random coefficients method and demonstrate how well the model performs with very few households, a variety of HVAC equipment, and in regions with much lower average temperatures than the QI and QM programs.

Evergreen experimented with a large variety of alternative model specifications and filters as part of this research. This included modeling homes by climate zone, including lagged temperatures in the binning procedure, selecting different change points, using a 30 percent holdout sample from each bin, estimating savings for a full year of typical meteorological year (TMY3) weather data[13], and many other specifications and filters. Additional information

[13] Typical meteorological year (TMY) data are collations of data for a particular location over several years (15–30 years) that form a representative typical year of weather, rather than a specific year with extreme weather events. The TMY3 data are derived from the 1991-2005 National Solar Radiation Data Base (NSRDB) archives.

about our findings from the research into alternative specifications is provided in the report appendix..

2.2.1 Residential Building Stock Assessment Model Results

Data from NEEA's RBSA were used first to develop and test the random coefficients model approach. The RBSA data include detailed premise, temperature, submetering, and whole-home metering for a sample of homes in the Pacific Northwest. These homes did not participate in any energy efficiency programs, so there are no impacts to be measured. However, these data allowed the Evergreen team to validate the model with data from homes in a cooler climate, as well as compare our model's predictions of the HVAC load to the actual submetered HVAC data.

The sample for the RBSA model included 99 homes in the Pacific Northwest, with both whole-home and submetered HVAC consumption data at 15-minute intervals from April 2012 to September 2014.[14] For this model, we aggregated the consumption data into hourly intervals. While the random coefficients model can handle 15-minute intervals, we believe hourly intervals are sufficient for this research, and make the model more comparable with the subsequent models developed for the SCE QI program and PG&E QM program.

Since this model includes all seasons and day types, we binned the home-days to four-dimensional bins, and the binning results are summarized in Table 3. For the annual RBSA random coefficients model, we included 5 household groups, 9 CDD groups with a cap of 26 CDD, 14 HDD groups with a cap of 70 HDD, and two day types.[15] Unlike the homes in the SCE QI or PG&E QM data, the RBSA homes experienced many very cold days and few hot days. For this reason, we used a larger number of HDD bins that cover a much wider range of temperatures than the CDD bins. We assigned each rounded HDD to groups of 5 HDDs and capped the HDD groups at 70, including all days with HDDs greater than 70 to HDD group 70. This was done to limit the total number of bins and thereby reduce processing time, but for program evaluations we suggest binning up to the true maximum HDD and CDD in the data.

[14] This analysis excludes four of the 103 RBSA households due to a high number of missing observations in the metering data.
[15] When modeling only homes with cooling equipment in the summer, the true maximum HDD was much lower, so we binned by HDD and CDD with 9 bins each with a cap of 26 CDD. When modeling only homes with heating equipment in the winter, we did not include any CDD bins because the true maximum CDD was 0.

Table 3: Summary of RBSA Annual Binning

Group	Description	Number of Groups
Homes	Usage – weather normalized annual energy usage grouped by percentile, with 1/5th of the total assigned to each group in order from smallest to largest	5
Days	CDD – average of CDH rounded up to a whole number, assigned three CDDs per group from 0-26 with all days higher than CDD 26 put into the last group	9
	HDD – average of HDH rounded up to a whole number, assigned 5 HDDs per group from 10-70 with all days higher than HDD 70 put into the last group and all days lower than HDD 10 put into the first group	14
	Day Type – flag for weekends that separates them from weekdays	2
Total[16]	Home-Day Bins	1,260

Figure 2 shows the actual average hourly total kWh consumption (purple) and predicted hourly total kWh consumption (yellow) of the holdout sample over the entire study period (April 2012 through September 2014). The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. Most of these homes are located in regions with moderate heating and very limited cooling needs; thus, it is not surprising that the annual whole-home load shape resembles the winter and fall load shapes of homes in the SCE QI and PG&E QM annual models. This graph shows that the random coefficients model prediction does differ from the actual load in most hours, but the overall difference is only 1 percent over 24 hours.

[16] Some of the combinations of temperatures were not present in the data; this is expected for the HDD and CDD bins in particular. For example, the data did not include any days that have hourly temperatures ranging from 10˚F -90˚F, so there are no home-days assigned to both CDD 25 and HDD 55. Our final model includes 221 bins.

Figure 2: RBSA Average Daily Predictions vs. Actual Use of Holdout Sample, All Seasons

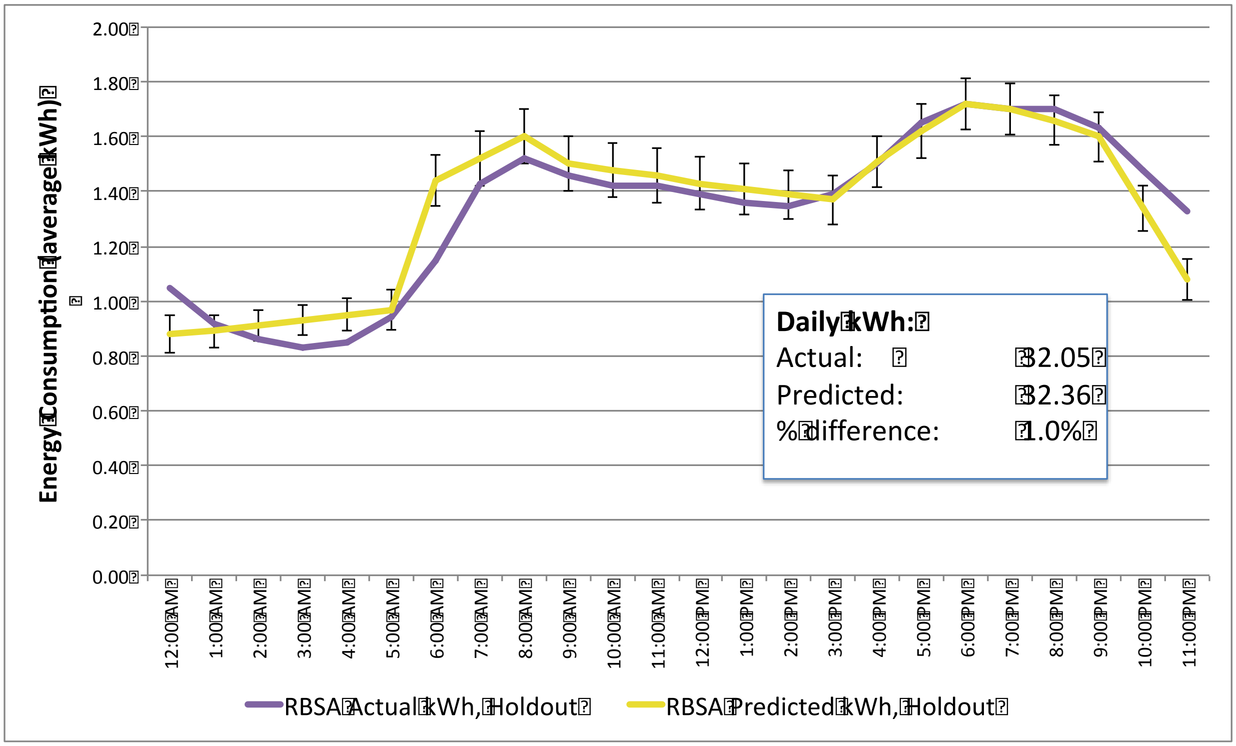
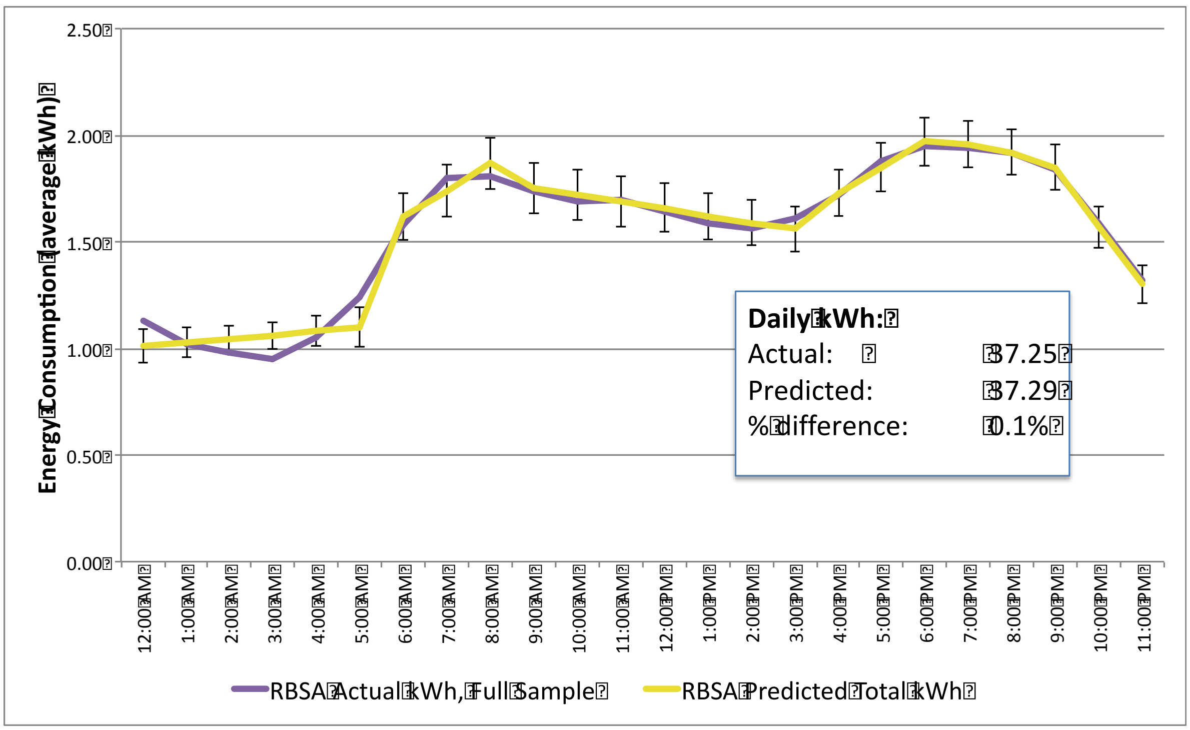


Figure 3 compares the predicted load shape with the actual load shape for all bins combined, including all 99 households (i.e., no holdout group) over the entire study period. The modeled load shape (yellow) aligns even better with actual load shape (purple), with a difference of about 0.1 percent over 24 hours. The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. This result is to be expected, as the data we are comparing are the same data used to estimate the model, as opposed to the holdout comparison in which we compared the model results with a sample of customers that was not used in the model estimation.

Figure 3: RBSA Average Daily Predictions vs. Actual Energy Use of Full Sample, All Seasons



The next two figures show the results of the random coefficients model when we restricted the sample to only include summer months for the 60 homes in the RBSA data with cooling equipment (Figure 4) or only winter months for the 61 homes in the RBSA data with heating equipment (Figure 5).[17] The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. The hourly predictions of the holdout sample are not as precise as when we were using data from all 99 homes due to the small sample sizes, but with the full sample the model does a good job of predicting

[17] This summer model binned homes by household group, CDD, groups of HDD, and day type while the winter model binned homes by household group, HDD, and day type.

consumption in both of these cases. This demonstrates that the random coefficients model is capable of modeling homes with diverse weather conditions and HVAC equipment.

Figure 4: RBSA Predictions versus Actual of Homes with Cooling in Summer Months

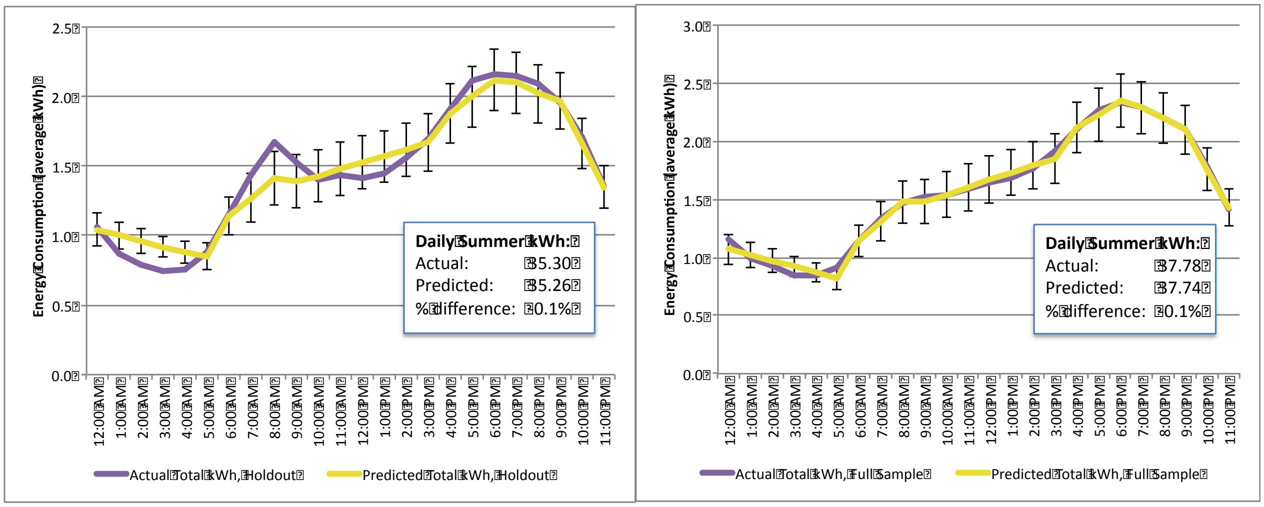
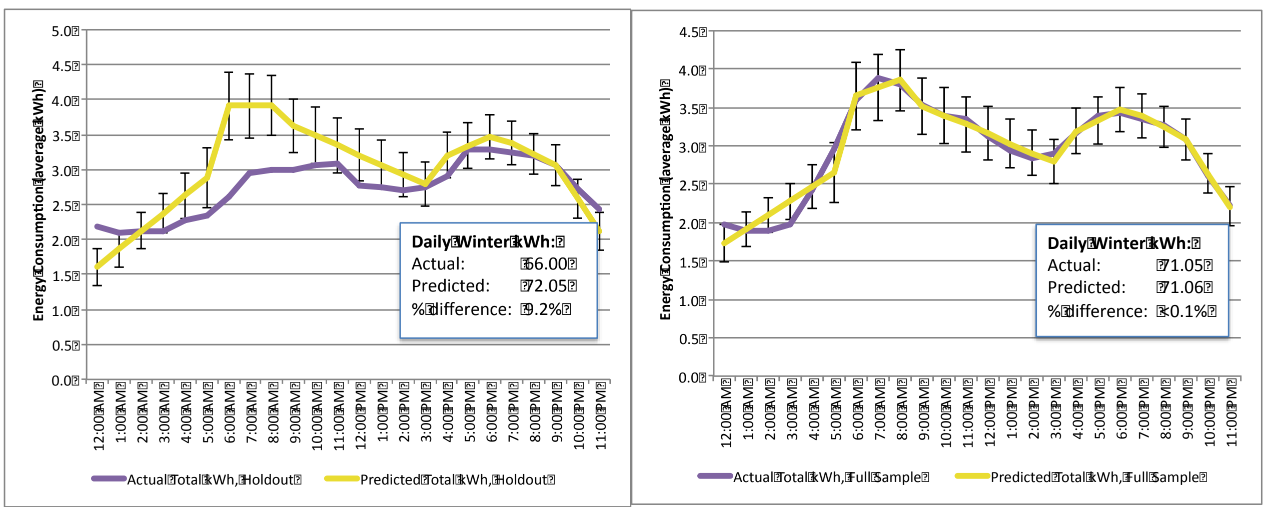


Figure 5: RBSA Predictions vs. Actual Use of Homes with Heating in Winter Months



HVAC Load Prediction Results

The RBSA's submetered HVAC data gave us a unique opportunity to test the random coefficients model's ability to predict HVAC consumption (both heating and cooling loads) using whole house consumption data only. The method used to develop HVAC load predictions is detailed below.

The random coefficients model produces hourly predictions of whole-home energy consumption for each of the 1,260 bins, each containing a single type of home on a single type of day. We split these predictions into two groups for each type of home:

1. Baseline weather days with little to no need for heating or cooling, in the lowest possible CDD and HDD groups.[18] In this model, baseline days are days with temperatures between 55 degrees and 67 degrees Fahrenheit.
2. Variable weather days with some need for heating and/or cooling, in any of the other CDD and HDD groups. In this model, variable weather days are days with temperatures below 55 degrees Fahrenheit and/or above 67 degrees Fahrenheit.

If we compare the predictions for these two different types of days for the same group of homes (i.e. homes with similar weather normalized energy usage), the difference in these predictions is the model's predicted weather-dependent consumption. We assumed all of the weather-dependent consumption is attributable to HVAC equipment.[19]

Table 4 provides an example of how this calculation was performed using homes in group 3 (mid-users) on weekdays. The first row shows the consumption for these homes is predicted to be 31.77 kWh on baseline weather days (days with CDD of 0-2 and HDD of 0-10). The second row shows a different day with assumed heating load; this day has the same CDD of 0-2 but higher HDD of 50-55. The model has predicted the homes represented in this second row will consume an average of 94.36 kWh on this type of day. Comparing this value to the average consumption on the baseline weather days of 31.77 kWh, we estimate that the HVAC (weather-dependent) consumption on this type of cold day is 62.59 kWh. The next two rows show predictions for additional groups of variable weather days, including a day with cooling load and a day with mild heating load.

Table 4: Example Calculation of Predicted HVAC Consumption

Home Group	Day Type	CDD	HDD	Average Predicted Daily Whole-Home kWh	Estimated HVAC kWh
3	Weekday	0-2	0-10	31.77	-
3	Weekday	0-2	50-55	94.36	62.59
3	Weekday	14-17	0-10	72.62	40.85
3	Weekday	2-5	0-10	33.43	1.66

[18] It would be ideal to isolate days with CDD of zero and HDD of zero to use as a baseline. However, this would require that the data include many days with an average hourly temperature of precisely 65 (or whatever baseline temperature is chosen), which was not realistic for the homes in the RBSA data. Instead, we used days with CDD of 0-2 and HDD of 0-10 because these are the most moderate temperature days observed, with very limited cooling or heating required.
[19] It is possible that some of the incremental increase in energy usage should not be attributed to HVAC. For example, we would expect some increase in energy usage for lighting as the days get shorter in the winter.

Figure 6 shows the average predicted hourly HVAC load (yellow) and actual hourly HVAC load (purple), as measured by the RBSA submeters. This analysis is using the full sample of 99 homes in all months (as shown in Figure 3), not just those with electric heating in winter months.

Since homes naturally heat up in the afternoons when the sun is out, we would expect the majority of heating load on cool days to occur in the mornings when homes are naturally cold and the occupants are likely to be home and awake. Similarly, we would expect the majority of cooling load on warm days to occur in the late afternoon and evenings when the home has been building up heat all day and occupants are likely to be home. Overall, the average daily HVAC load over the study period (all seasons) will include energy usage from some combination of heating and cooling. Due to the moderate temperatures in the Northwest, we expected these homes would have more prominent heating than cooling, corresponding to a larger HVAC peak in the morning than the afternoon or evening.

The main peak in actual HVAC load for these homes occurs in the morning from 5:00 a.m. to 8:00 a.m., around the time of the first peak in whole-home consumption (Figure 3). The whole-home model predictions underestimated the HVAC load during these hours by 14.1 percent. When looking at whole-home usage, the model predicts a second and slightly larger peak in the late afternoon or evening from 3:00 p.m. to 8:00 p.m. The actual HVAC load also has a second peak during these hours, but it is smaller than the morning peak. The model overestimated HVAC load during these hours by 13.3 percent. When we look at the actual consumption data, the HVAC load makes up 41 percent of whole-home energy consumption from 5:00 a.m. to 8:00 a.m. and 27 percent from 3:00 p.m. to 8:00 p.m., with the remainder of energy usage coming from other end uses. On a daily basis, the modeled HVAC kWh load is still a good predictor of actual HVAC kWh load (purple), overestimating the HVAC load by only 1.1 percent over 24 hours. In general the model performed well, given that it only had data from 99 homes with a variety of heating and cooling equipment, and it is predicting HVAC load using only whole-home energy consumption data.

Figure 7 presents the model predicted and actual load from the full sample of 99 RBSA homes for whole home load, HVAC load and non-HVAC load.

Figure 6: RBSA Average Daily Predicted Hourly HVAC Load, All Seasons

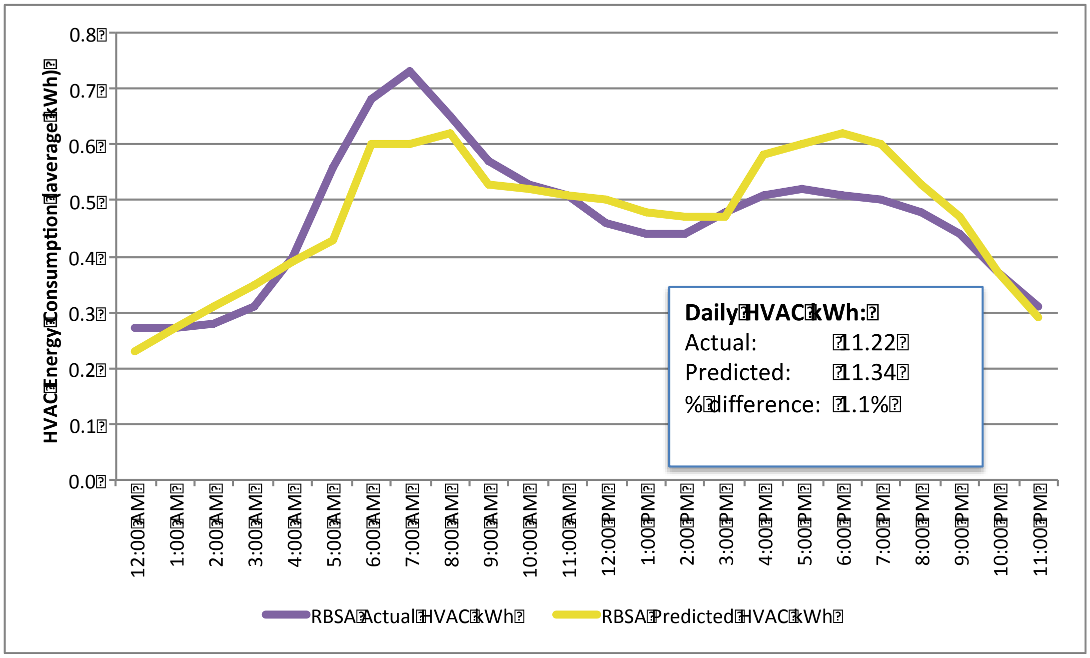


Figure 7: RBSA Predictions vs. Actual Use – Whole Home Load; HVAC Load; Non-HVAC Load

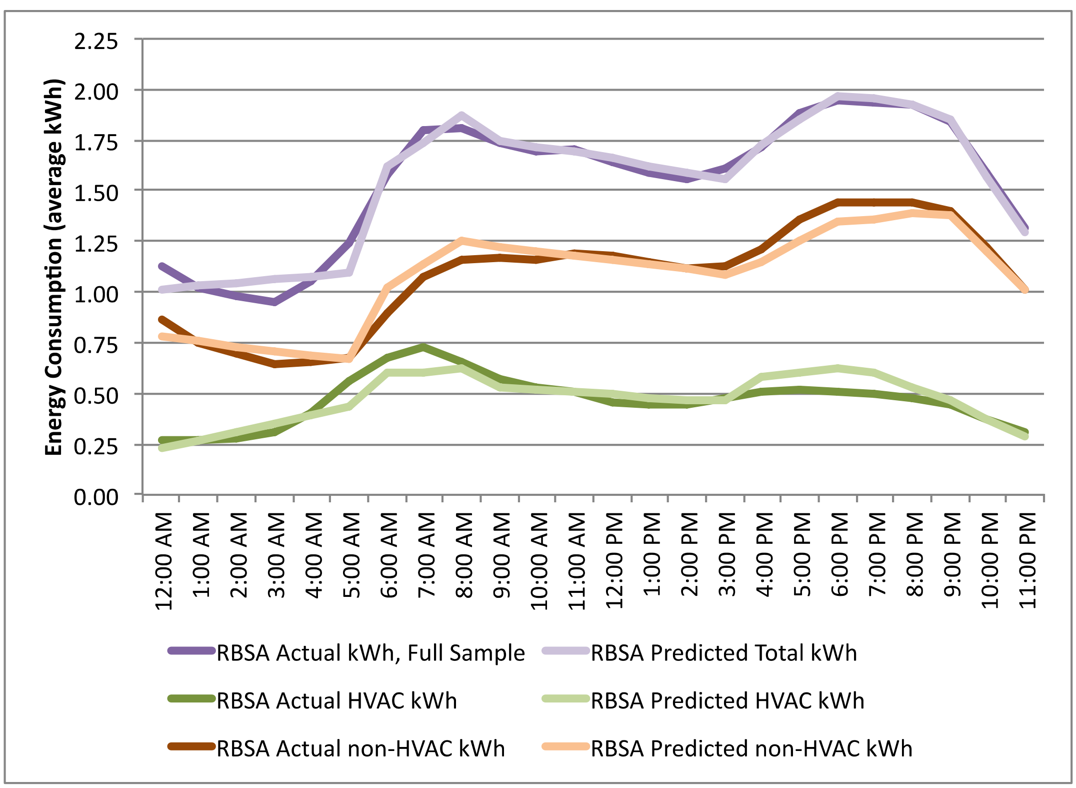
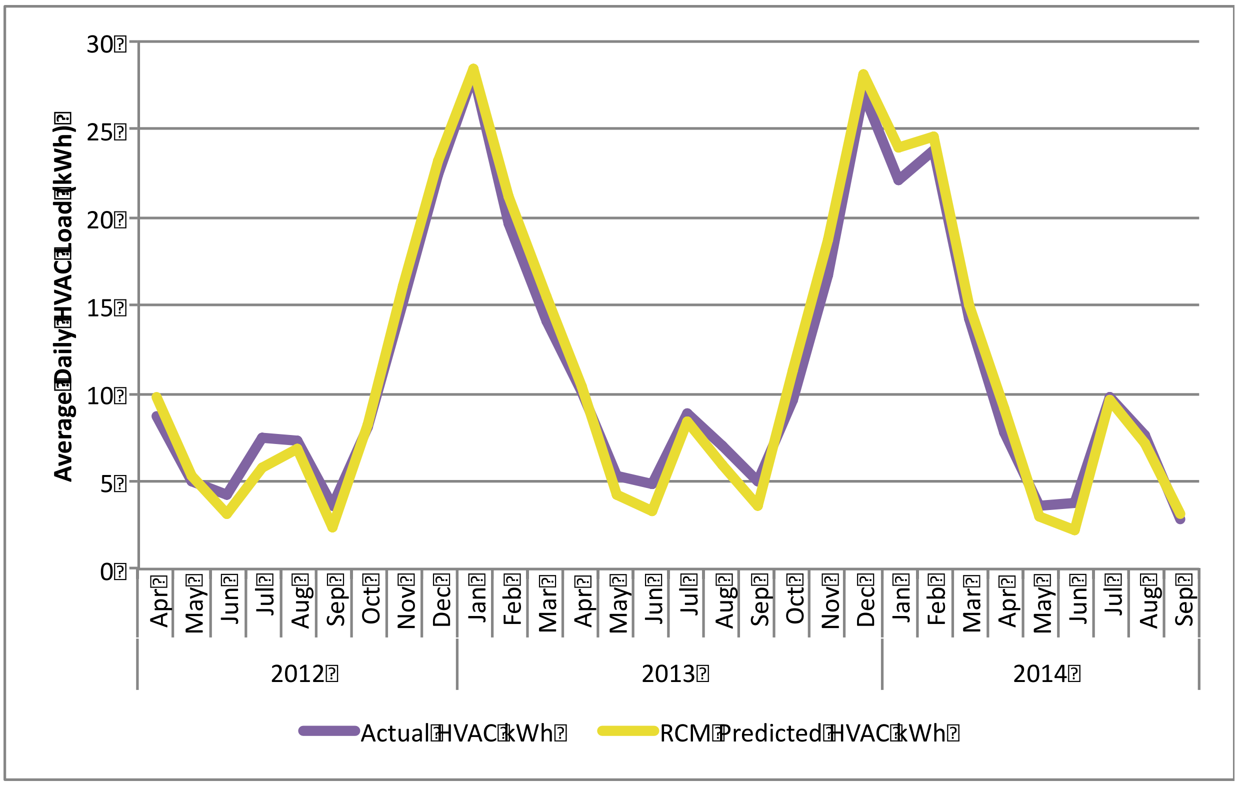


Figure 8 shows the actual average daily HVAC kWh consumption (purple) and the model's prediction of the average daily HVAC kWh (yellow) for each month during the test period of April 2012 to September 2014. Again, these predictions are based on whole-home energy consumption, not the submetered HVAC energy consumption. The root mean squared error (RMSE) for these predictions is 1.27 kWh, which is about 5 percent of average daily electricity usage by the HVAC equipment[20]. The model overestimated HVAC consumption during winter months and underestimated it during summer months. In general the model performed very well, given that it only had data from 99 homes (with a variety of heating and cooling equipment), and the model is predicting HVAC load using only whole-home energy consumption data.

Figure 8: RBSA Predicted Monthly HVAC Consumption



The promising results from the RBSA random coefficients modeling gave us confidence in the approach. To determine how well the modeling approach could predict load shapes and develop savings estimates from actual energy efficiency program data, we applied the random coefficients modeling approach to the SCE QI and PG&E QM program data, and present the results in the following sections.

[20] The RMSE and the mean squared error (MSE) are reported for selected results in this report, and are fairly consistent across model applications.

2.2.2 SCE Quality Installation Program Results

Once we tested the random coefficients model with the RBSA data and determined how well it could perform, the next step was to estimate the model using data from an efficiency program where impacts could be measured. The first program examined is the SCE's Quality Installation Program. Each home in the SCE QI Program replaced an existing HVAC system[21] using an installation contractor who received additional training in quality installation practices through the program. The contractor is responsible for ensuring that the air conditioning unit is sized property for the home and installs the new unit based on strict ENERGY STAR requirements, and connects it to the ductwork/distribution system.[22] The quality installation theoretically improves cooling delivery (from reduced runtime and/or power draw) by preventing common problems that occur during installation, that cause the new unit to operate below its energy efficiency specification.

Note that the savings discussed here for the QI Program is measured as the difference in predicted and actual usage in the post-installation period. In the analysis presented here, the entire difference is attributed to the QI Program. In a formal impact evaluation for California, we would need to adjust this estimate to account for the difference between the existing equipment and the standard efficiency baseline designated for the QI Program.[23] We would also need to account for other external factors that may be affecting energy use between the periods, which may be done by using an appropriate comparison group of non-participating customers. The use of a comparison group will be addressed more in the next phase of analysis. Due to the data limitations discussed above, our initial test of the random coefficients model was limited to a sample of participants only.

Basic SCE QI Program Model – Weekdays, Summer Only

The sample of homes for this model includes all homes in the SCE QI Program with AMI data for the 2013 and 2014 cooling seasons (i.e., the summer months of July through September). Since the HVAC units installed through the SCE QI Program are all air conditioners, we expected the majority of program savings to occur during the cooling season. This initial model uses only weekday data to avoid possible differences in energy usage between weekends and weekdays. The resulting dataset includes 1,379 homes dispersed across 9 different climate zones.

Since this model only includes weekdays during summer months, the home-days are assigned to two-dimensional home-day bins that do not include bins for HDD or day type. For this

[21] Eligible homes installed a new package or split system air conditioner or heat pump that is 15 SEER or greater.
[22] ANSI/ACCA 5 QI-2010: HVAC Quality Installation Specification
[23] It should also be noted that California (through AB 802) is moving toward a new evaluation protocol where 'meter-based' savings would be calculated, and consequently the existing equipment conditions would be used to measure savings. The method demonstrated here for the QI program is consistent with AB 802 approach.

model, we used 20 home groups and 25 CDD groups, resulting in 500 home-day bins. We assigned each rounded CDD to its own CDD group but capped the CDD at 25, assigning all days with CDD greater than 25 to CDD group 25. This was done to limit the total number of bins and thereby reduce processing time, but for program evaluations, we suggest binning up to the true maximum CDD in the data. In order to isolate days with expected cooling, we removed all days with a CDD of zero. Expanding the weather bins will be explored more in the next phase of this analysis.

Table 5: Summary of QI Summer Weekday Binning

Group	Description	Number of Groups
Homes	Usage – weather normalized annual energy usage grouped by percentile, with 1/20th of the total assigned to each group in order of smallest to largest	20
Days	CDD – average of CDH rounded up to a whole number, assigned one CDD per group from 1-25 with all days higher than CDD 25 put into the last group (CDD 25)	25
Total	Home-Days	500

Table 6 shows the count of home-days in the SCE QI summer model post-period that was assigned to each bin. We automatically color-coded the cells with the highest count in dark blue and the lowest count in dark red, white cells fall somewhere in the middle of this spectrum. This table shows the actual distribution of participant households and the weather they experienced in the summer of 2014 (i.e., the post-period for this model). We see more mid-temperature days with CDDs ranging from 7 to 17 than especially high or low temperature days within each of the household groups, as well as overall. We also see more home days in the home groups at the lower end of the usage spectrum.[24] This is because each home group represents about 5 percent of total baseline electricity usage for the homes in our sample. Because of this, the number of homes in each group varies but the amount of daily kWh each home group represents is approximately the same.

[24] In particular, we see a large number of home-days in home group 1 and CDD 25. One possible explanation for this is that people commonly go on vacation during summer months, leaving the home unoccupied with minimal energy usage on hot days. Those who spend very little time at home in the summer are more likely than others to end up being identified as a low-users (columns on the left).

Table 6: Number of Home-Days in Each Bin

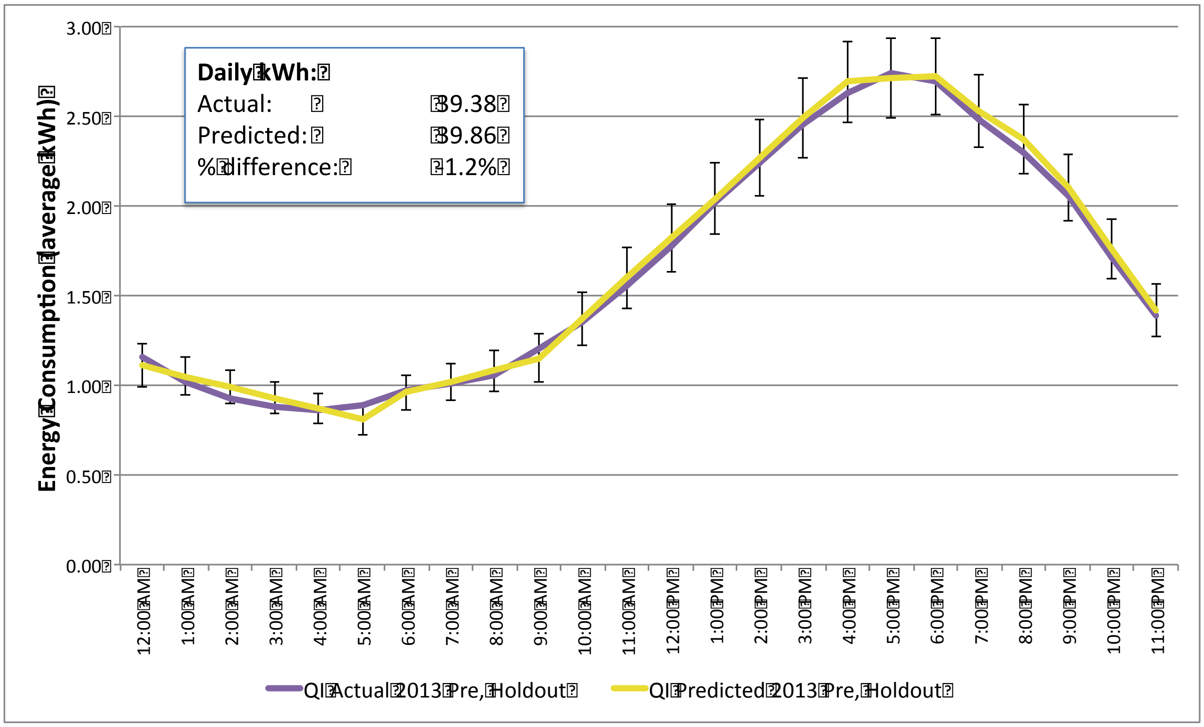
Cooling Degree-Day Groups	Household Groups: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
1	3	9	13	8	10	5	7	9	3	7	3	4	4	5	2	4	4	0	7	0	107
2	12	45	61	40	42	25	27	40	13	34	13	15	19	22	11	19	20	5	32	2	497
3	12	41	49	36	25	23	29	36	18	35	15	21	15	21	15	20	21	8	27	6	473
4	50	100	118	87	67	58	69	80	34	78	41	46	43	51	39	47	53	25	66	12	1,164
5	56	109	106	94	65	68	122	108	96	102	72	122	70	88	86	79	90	68	67	53	1,721
6	83	119	126	109	103	78	114	92	79	83	76	94	67	104	81	74	69	61	70	44	1,726
7	147	240	231	207	174	165	224	213	177	191	144	240	131	184	172	150	152	127	134	95	3,498
8	255	310	295	293	261	219	250	193	174	211	196	211	175	208	190	155	149	145	143	84	4,117
9	235	322	314	310	262	232	250	207	182	237	168	253	170	165	180	153	142	105	132	79	4,098
10	348	444	348	358	334	291	291	216	245	309	269	281	232	233	233	193	194	172	152	107	5,250
11	488	517	411	423	327	371	360	258	326	350	300	298	283	247	231	185	231	190	158	82	6,036
12	254	299	279	254	220	205	230	197	179	217	175	218	171	169	164	147	143	133	120	79	3,853
13	386	370	317	299	250	256	261	204	209	257	222	220	205	218	177	159	170	155	140	83	4,558
14	439	457	347	380	324	309	324	271	313	315	310	301	290	281	279	202	249	239	166	150	5,946
15	344	315	251	267	234	217	242	178	204	229	223	207	205	198	192	178	165	190	116	113	4,268
16	359	316	243	233	223	218	233	157	185	214	208	170	189	190	165	147	154	170	117	89	3,980
17	414	340	281	277	215	235	243	181	215	196	205	172	202	181	140	117	147	127	101	45	4,034
18	210	169	151	146	96	120	135	97	123	118	113	90	113	86	76	66	79	100	50	36	2,174
19	292	183	176	155	109	147	151	118	154	145	120	127	135	107	104	88	116	107	56	57	2,647
20	86	58	59	45	36	41	46	28	40	49	37	42	37	35	31	33	27	33	14	15	792
21	107	51	43	29	27	28	49	30	43	33	25	34	39	21	29	29	33	42	10	28	730
22	261	117	103	83	87	78	87	84	82	82	83	78	74	79	73	61	72	75	53	69	1,781
23	87	34	42	19	22	27	35	30	24	29	13	32	17	14	19	17	12	29	13	26	541
24	161	75	69	56	49	57	52	52	44	40	39	35	44	33	30	39	33	32	30	33	1,003
25	2,229	348	541	225	254	251	167	498	242	160	200	59	184	71	138	130	234	170	341	490	6,932
Total	7,318	5,388	4,974	4,433	3,816	3,724	3,998	3,577	3,404	3,721	3,270	3,370	3,114	3,011	2,857	2,492	2,759	2,508	2,315	1,877	71,926

In order to test the reliability of our model, we randomly selected 30 percent of homes as a holdout sample and then modeled the remaining 70 percent of homes. In addition to the 30 percent holdout group used to test the accuracy of the model's predictions of pre-period consumption, we set aside all observations from households in the summer of 2014 before their new unit was installed.

One concern with comparing pre- and post-period data without a control group is the idea that there are unseen changes that occur from one period to the next that are not related to program participation. In this dataset, program participation occurred throughout the summer, so the pre-period days in the summer of 2014 reflect a wide range of home types and day types. In theory, if a model constructed with pre-period days in 2013 is able to accurately predict the pre-period days in 2014, then we can conclude that the pre-period model generates reasonable estimates for the full summer of 2014, both pre- and post-participation.

Figure 9 shows the comparison of the predicted average load shape across all homes from the model (yellow) with the actual average load shape for the 30 percent holdout group (purple). The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. As demonstrated in this graph, the model does a very good job of predicting energy use in the holdout group, with a difference between estimated and actual usage of about 1 percent over 24 hours.

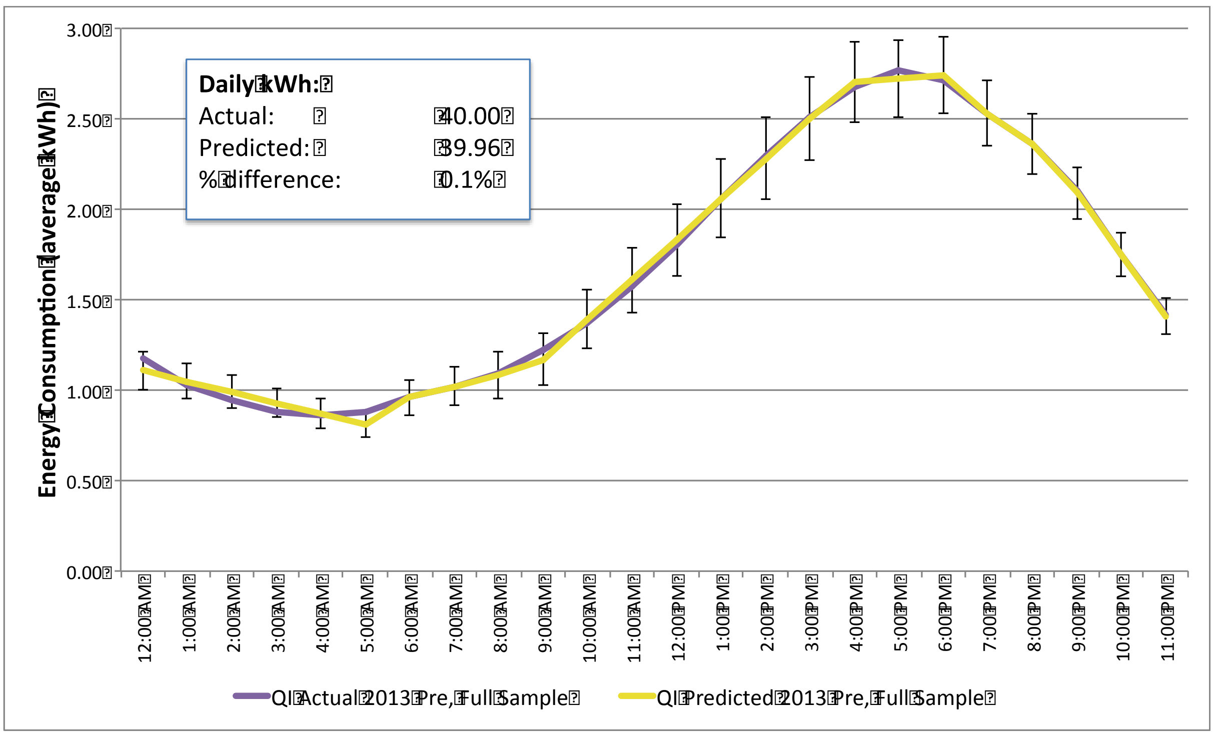
Figure 9: SCE QI Summer Predictions versus Actual of Holdout Homes, 2013 Pre-Period



RMSE = 0.044, MSE = 0.002

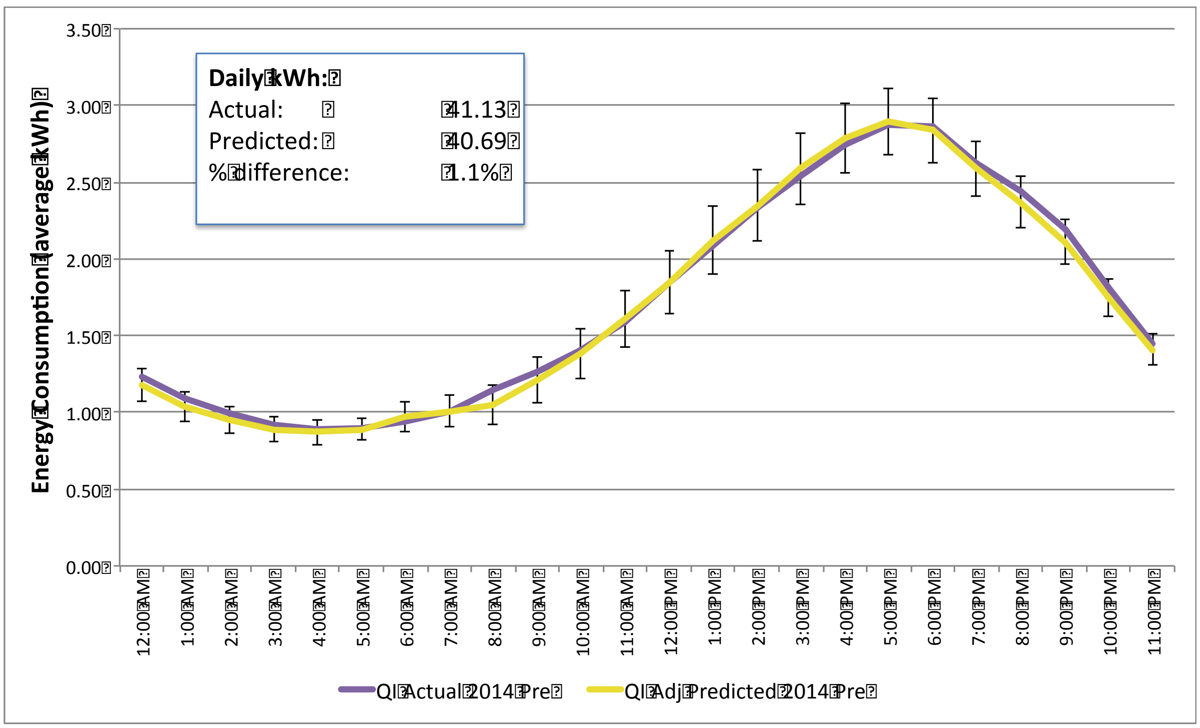
Figure 10 presents the comparison of the 2013 pre-installation predicted load shape from the model with the actual 2013 pre-installation load shape for all kWh-CDD-bins combined for all 1,379 households in 9 different climate zones (i.e., no holdout group). The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. The modeled 2013 pre-installation period load shape (yellow) aligns even closer with the actual 2013 pre-installation load shape (purple), with a difference of about 0.1 percent over 24 hours. This result is to be expected, as the data we are comparing are the same data used to estimate the model. This is in contrast to the holdout analysis where estimates are compared to actual usage for customers that were not included in the model sample.

Figure 10: SCE QI Summer Predictions versus Actual of Full Sample, 2013 Pre-Period



As another test of the model's reliability, we used the model that was constructed with only 2013 pre-period data to generate predictions of the 2014 pre-period home-days (yellow), then compared these to the actual 2014 pre-period (purple) shown in Figure 11. The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. Again, the model does a very good job of predicting energy use in the test period, with a difference between estimated and actual usage of about 1 percent over 24 hours. Given the very accurate prediction for 2013, the model appears to be capable of accurately predicting energy usage in the summer of 2014 in the absence of program participation.

Figure 11: SCE QI Summer Predictions versus Actual of 2014 Pre-Period, Days Not Included in the Modeling Sample



RMSE = 0.043, MSE = 0.002

Table 7 shows how we derived the post-period predictions depicted in the remaining figures (Figure 12 and Figure 13). The values in the first two columns (A and B) come from the holdout sample model of the pre-period, depicted in Figure 9. These two values were used to calculate the adjustment factor shown in column C, and this adjustment corrects for the small discrepancy observed between the predicted and actual values in the initial pre-period model. The values in columns E and F come from the full sample model of the post-period, shown in Figure 10. Column G shows the adjusted post-period prediction, calculated by multiplying the post-period model prediction in each hour (F) by the adjustment factor (C). The values in column H indicate the estimated kWh savings; this is the difference between our adjusted post-period model prediction (G) and the actual post-period observations (E).

Table 7: SCE QI Basic Summer Model Results and Calculations (Hourly)

	Holdout Pre-Period Model			Post-Period Model Output		Results of Calculations	
Hour	Actual (kWh)	Predicted (kWh)	Adj Factor (%)	Actual (kWh)	Predicted (kWh)	Adj Prediction (kWh)	kWh Savings (kWh)
-	A	B	C=A/B	E	F	G=F*C	H=G-E
12:00 AM	1.16	1.11	1.05	1.04	1.14	1.19	0.15
1:00 AM	1.02	1.05	0.97	0.92	1.08	1.05	0.13
2:00 AM	0.93	0.99	0.94	0.83	1.02	0.96	0.13
3:00 AM	0.88	0.93	0.95	0.79	0.95	0.90	0.11
4:00 AM	0.86	0.87	0.99	0.78	0.89	0.88	0.10
5:00 AM	0.89	0.81	1.10	0.82	0.83	0.91	0.09
6:00 AM	0.97	0.96	1.01	0.89	0.98	0.99	0.10
7:00 AM	1.01	1.02	0.99	0.92	1.04	1.03	0.11
8:00 AM	1.06	1.08	0.98	0.97	1.10	1.08	0.11
9:00 AM	1.20	1.15	1.04	1.06	1.20	1.25	0.19
10:00 AM	1.35	1.37	0.99	1.18	1.43	1.41	0.23
11:00 AM	1.56	1.60	0.98	1.35	1.67	1.63	0.28
12:00 PM	1.78	1.82	0.98	1.56	1.91	1.87	0.31
1:00 PM	2.02	2.04	0.99	1.79	2.15	2.13	0.34
2:00 PM	2.24	2.27	0.99	2.00	2.38	2.35	0.35
3:00 PM	2.45	2.49	0.98	2.19	2.62	2.58	0.39
4:00 PM	2.63	2.69	0.98	2.35	2.83	2.77	0.42
5:00 PM	2.74	2.71	1.01	2.43	2.84	2.87	0.44
6:00 PM	2.69	2.72	0.99	2.35	2.85	2.82	0.47
7:00 PM	2.48	2.53	0.98	2.17	2.62	2.57	0.40
8:00 PM	2.30	2.37	0.97	2.02	2.44	2.37	0.35
9:00 PM	2.06	2.10	0.98	1.80	2.16	2.12	0.32
10:00 PM	1.71	1.76	0.97	1.50	1.81	1.76	0.26
11:00 PM	1.39	1.42	0.98	1.23	1.45	1.42	0.19
Total	**39.38**	**39.86**		**34.94**	**41.39**	**40.89**	**5.95**

Figure 12 shows the adjusted model prediction of 2014 post-period consumption for all households that participated in the SCE QI Program. This prediction is based on the pre-period consumption model and post-period weather data; it represents the expected load shape for these households in absence of SCE QI Program participation. The error of each hourly consumption prediction is depicted as a 95 percent confidence interval in bars around each estimate. The error of the hourly predictions is smallest during the early hours in the morning and greatest during the evening peak. This is likely due to greater variation in overall energy usage across households in each bin in the afternoon and evening hours when homes have more activity, than the early morning hours when there is less activity across all homes.

Figure 12: SCE QI Summer Predictions of 2014 Post-Period with Error Bars

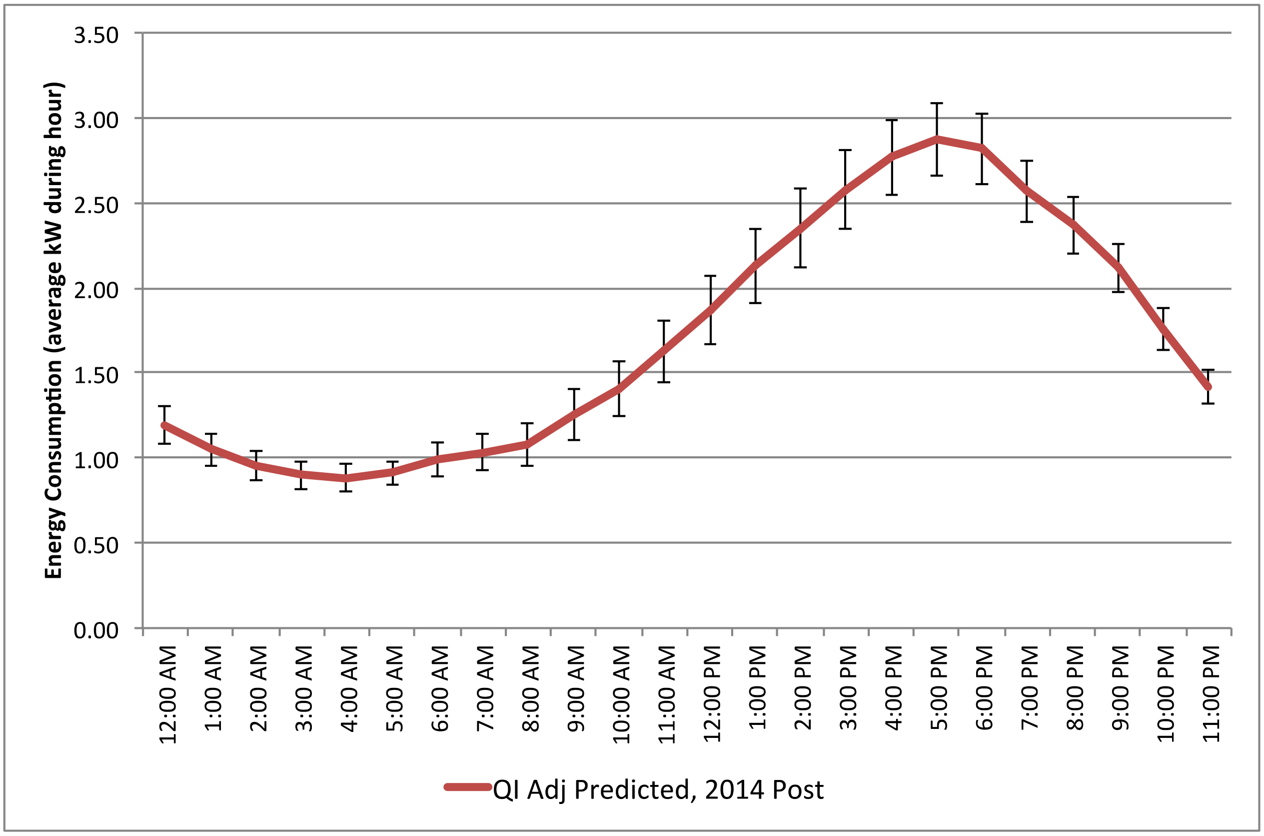
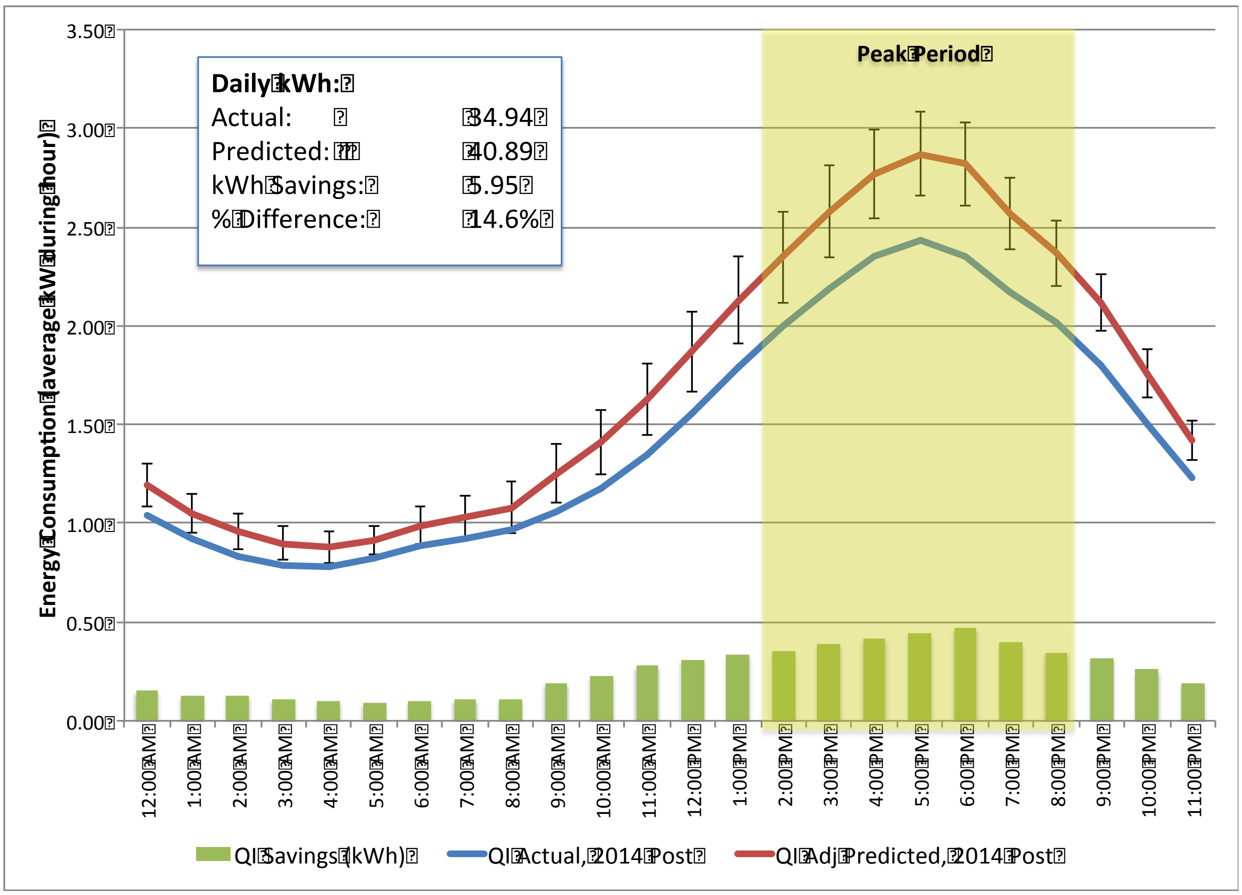


Figure 13 compares the post-period predicted load shape (red) with the actual post-period load shape (blue) across all households. Whenever the actual post-period load shape falls below the predicted post-period load shape, this indicates that savings were realized during that hour (green bars). After adjusting for the error in the model, based on the sample of homes used, the modeling approach finds approximately 15 percent savings during summer months attributable to the HVAC units installed through the SCE QI Program.[25] Note that this approach estimates that the majority of savings is realized during the later part of the day including during the peak hour period between 2:00 pm and 8:00 pm,[26] highlighted in yellow. The timing of the savings that coincides with expected peak HVAC use is also encouraging evidence of the model's ability to predict hourly consumption.

Figure 13: SCE QI Summer Predictions versus Actual, 2014 Post-Period



[25] It is not possible to separate the estimated savings impacts for SCE's QI program activities (i.e., quality installation) from the impacts of the new HVAC system. To determine separate impacts, one would need to compare these results to a control group consisting of customers who replaced their HVAC system but did not use a QI participating contractor for the installation.

[26] We use the residential peak period of 2:00 pm to 8:00 pm, as defined for SCE's residential Time-Of-Use rate plan; https://www.sce.com/wps/portal/home/residential/rates/residential-plan

Figure 14 shows the summer weekday hourly kWh savings estimates from the previous figure with bars depicting 95 percent confidence intervals around each estimate. The savings are statistically significant in every hour except 8:00 a.m.

Figure 14: SCE QI Summer Hourly Savings (Weekday) Estimates with Error Bars

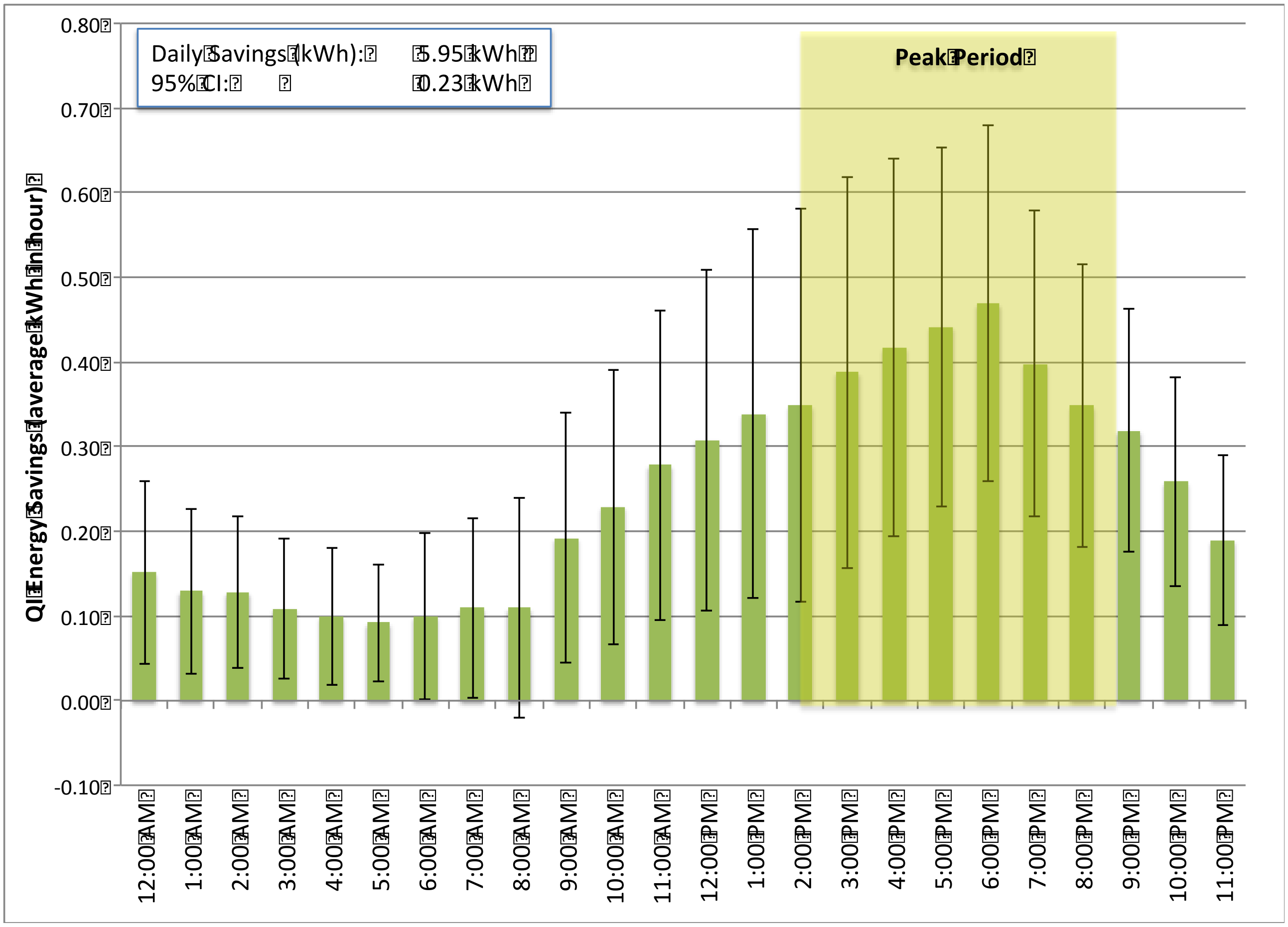


Table 1 shows the average daily savings estimate for each group and bin in the SCE QI Summer weekday model. The columns show households grouped by their weather normalized energy usage in the pre-period for each home (highest users on the right) and the rows show days grouped by the temperature via cooling degree-days (hottest days on the bottom). Each cell shows the estimated program savings (kWh per day) for a specific home-day bin. We automatically color-coded the cells with the highest kWh savings in dark blue and the lowest kWh savings in dark red; colorless cells fall in the middle of this spectrum. Within each household group, there are home-days from a wide range of temperatures, each with their own savings estimate. Similarly, each group of days with similar temperatures (i.e., CDD) includes home-days from a range of households (i.e., high, mid, and low users), which experience a wide range of daily kWh savings. For the SCE QI summer model, savings trends upward as temperature and household energy increase. For the low usage groups with negative savings (i.e., energy use increases after participation), this could be indicating that existing equipment was broken or not being used prior to participation in the program.

Table 8: Program Savings (kWh per day) by Bin

		Household Groups																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Cooling Degree-Day Groups	1	4.5	4.9	-5.5	-2.8	1.6	4.9	30.9	-3.5	-4.0	2.6	5.2	3.9	-7.7	2.5	25.5	-8.8	-15.9	0.0	-0.8	0.0	-0.8
	2	1.5	-0.2	-3.8	-8.9	-7.2	-5.3	-6.8	-10.0	-8.0	-0.4	-7.2	-6.9	0.6	-2.7	-6.9	-7.5	-16.0	-7.7	-15.1	-13.5	-6.7
	3	-3.0	-2.7	-2.8	-8.0	-2.4	-6.4	-5.0	-10.0	-9.8	-3.2	-4.3	-4.6	-0.8	-5.9	-6.1	-3.3	-0.7	-3.6	-2.0	14.8	-4.7
	4	-2.8	-4.8	-8.1	-3.4	-2.0	-3.3	-3.9	-4.5	-2.0	-0.9	-1.1	-2.4	5.1	-6.1	7.9	4.0	5.1	3.8	2.5	18.8	-2.0
	5	-3.6	-2.0	-4.1	-1.4	0.5	0.4	-1.1	-3.7	-0.8	-1.0	-2.2	2.0	5.4	2.5	4.2	0.6	13.0	6.9	10.3	22.0	1.5
	6	-2.6	-3.5	-4.2	-0.4	-1.2	0.8	0.6	-2.7	4.3	2.2	0.0	7.0	0.6	7.6	8.7	3.0	10.4	3.3	21.7	23.7	2.7
	7	-4.0	-1.9	-4.2	0.8	-2.1	-0.3	1.2	-3.9	1.3	1.8	3.3	2.4	4.4	1.3	11.4	7.1	10.5	11.4	17.4	30.7	3.0
	8	-4.5	-2.1	-4.7	-1.0	-0.6	0.6	-1.0	-2.3	4.2	1.4	2.3	6.3	1.8	-1.2	5.8	6.6	11.5	9.2	15.3	25.4	2.0
	9	-4.8	-3.5	-3.5	-0.9	-4.2	2.2	3.5	-1.1	5.6	2.6	6.8	4.5	1.5	0.8	8.4	9.1	12.4	6.2	19.0	37.8	2.8
	10	-5.7	-2.9	-2.8	-0.5	1.7	2.9	1.3	0.0	5.3	1.8	5.4	2.9	3.2	0.8	6.6	9.5	15.7	11.1	21.1	30.4	3.0
	11	-6.2	-3.7	-5.0	-2.2	-1.6	1.7	2.9	0.8	7.7	1.6	4.7	3.7	5.9	-0.1	10.6	9.3	13.1	6.0	17.0	24.9	2.2
	12	-6.6	-3.6	-2.5	-0.7	-0.1	1.8	5.6	3.0	9.1	4.5	11.6	6.7	6.6	-0.2	15.0	15.3	14.2	17.0	22.4	41.3	5.2
	13	-7.5	-3.6	-3.9	1.2	0.8	3.8	5.6	4.5	8.2	6.9	11.4	8.3	8.0	8.6	13.4	18.6	13.9	15.9	26.4	33.2	5.4
	14	-7.6	-0.2	-4.2	4.1	5.2	5.4	5.1	8.2	8.7	7.3	13.1	12.4	8.1	9.4	16.2	21.6	19.6	21.1	29.3	46.1	8.3
	15	-7.7	-2.0	-2.8	1.1	5.9	5.7	3.0	7.1	7.9	8.8	11.3	10.6	9.7	9.8	20.2	17.2	20.5	15.7	22.9	28.0	7.2
	16	-8.4	-1.3	1.1	6.2	6.7	5.1	6.4	7.1	11.6	9.8	14.2	14.6	10.5	13.6	20.0	16.8	18.0	16.3	16.9	20.9	7.9
	17	-8.3	-2.8	-5.3	1.3	7.4	9.6	11.0	11.4	15.5	8.9	11.1	14.6	17.3	14.8	14.3	22.6	20.9	22.1	41.3	35.4	8.7
	18	-8.8	-1.8	3.8	0.0	6.2	6.5	18.5	10.9	13.8	7.9	10.3	14.2	17.3	15.1	26.5	18.5	20.7	13.9	33.1	29.9	8.9
	19	-10.7	-0.6	2.5	1.9	-0.2	8.4	17.9	11.2	14.2	16.7	15.7	15.0	19.1	22.9	31.7	23.8	27.4	19.9	23.6	47.9	11.3
	20	-11.6	-4.1	-3.7	0.4	11.8	9.8	12.3	17.9	8.6	10.7	11.6	15.3	18.0	18.7	25.9	22.4	18.1	24.9	31.9	31.0	9.2
	21	-8.3	-8.0	6.2	-1.4	14.1	14.8	20.9	12.2	10.1	6.4	10.3	14.4	20.4	20.7	23.1	22.6	20.4	16.7	30.4	24.5	9.7
	22	-4.9	-8.3	12.2	-3.2	21.8	9.3	23.5	15.5	13.7	16.1	11.7	26.6	18.4	16.3	27.0	28.5	26.2	22.1	58.7	37.1	13.9
	23	-7.8	-7.3	10.6	-7.9	-2.4	18.0	23.7	13.7	14.0	9.8	10.3	5.3	5.1	14.0	16.3	19.8	12.2	23.9	26.3	30.6	7.7
	24	-7.6	-5.9	9.6	4.7	3.3	10.5	23.7	19.4	19.7	12.8	22.1	16.9	25.6	28.0	15.8	27.8	20.7	23.8	27.5	33.0	11.6
	25	-5.5	-2.2	16.2	11.2	20.8	8.4	12.0	12.8	8.4	21.5	46.9	13.8	12.8	18.8	8.9	34.0	21.6	12.3	30.7	37.7	10.3
Total		-6.3	-1.7	-0.4	0.4	3.3	4.3	5.7	4.7	8.5	6.0	8.9	8.3	8.4	7.5	14.5	14.4	16.8	14.4	22.3	32.1	6.0

Annual SCE QI Program Model

The preceding discussion only included summer weekdays in the sample, as these days were most likely to show up as savings in our initial model test. Given the success of the summer weekday model, the next step was to expand the model to include all days and develop annual impact estimates.

To develop annual savings estimates, the sample of homes includes all homes in the QI Program with AMI data for periods before and after they participated in the program. Unlike the summer weekday model, the annual model uses all months and day types (i.e., weekdays and weekends) requiring a more complex binning procedure. The resulting dataset includes 2,038 homes dispersed among 9 climate zones.

Since this model includes all seasons and day types, we binned the home-days with four-dimensional bins. Specifically, we used 20 home groups, 9 CDD groups, 9 HDD groups, and 2 day type groups, resulting in 3,240 possible home-day bins. We assigned each day to a CDD group and an HDD group that included a range of three degree-days each, up to a maximum of 26. Using multiple degree-days per group and setting a maximum value limits the total number of bins and thereby reduces processing time. For program evaluations, we suggest binning up to the true maximum CDD and HDD in the data and using narrower ranges of degree-days in each group to limit the variation in each bin, and thus improve the precision of the model's estimates.

Table 9: Summary of QI Annual Binning

Group	Description	Number of Groups
Homes	Usage – weather normalized annual energy usage grouped by percentile, with 1/20th of the total assigned to each group in order of smallest to largest	20
Days	CDD – average of CDH rounded up to a whole number, assigned three CDD per group from 0-26 with all days higher than CDD 26 put into the last group	9
	HDD – average of HDH rounded up to a whole number, assigned three HDD per group from 0-26 with all days higher than HDD 26 put into the last group	9
	Day Type – flag for weekends that separates them from weekdays	2
Total[27]	Home-Day Bins	3,240

Table 10 and Table 11 present the count of home-days in the post-period for the SCE QI annual model on weekdays and weekends respectively. As with previous tables, these tables show the actual distribution of participant households and the weather they experienced in the post period. In the annual model day-types are binned by combinations of both CDD and HDD, and the table is labeled with the upper limit of each day-type (e.g. the day type bin CDD 2 includes all days with CDD between 0 and 2). Similar to the SCE QI summer model, we see more moderate days with CDD or HDD ranging from 6 to 17 than especially high or low temperature days within each of the household groups. Again, there are more home days in the home groups at the lower end of the usage spectrum, because each home group represents about 5 percent of total baseline electricity usage for all the homes in our sample.

[27] Some bins have zero home-days. This is expected as certain combinations of groups are not present in the data, In particular combinations of HDD and CDD groups because there were no days with extreme temperature ranges. For example, the data did not include any days with a temperature range from 40˚F-90˚F, so there are no home-days assigned to both CDD 25 and HDD 25. Our final pre-period model includes 1,098 bins and the post-period includes 1,083 bins.

Table 10: Number of Home-Days in Each Bin (Annual - Weekdays)

Day Type / Household Group		Weekdays																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
CDD	HDD																					
2	2	154	130	175	170	130	143	86	160	99	89	142	68	48	77	57	105	25	83	84	54	2,079
	5	1,663	1,552	1,461	1,283	1,200	1,241	1,118	1,197	1,077	954	1,074	919	790	722	738	758	497	721	561	412	19,938
	8	3,117	3,003	2,624	2,585	2,357	2,373	2,299	2,287	2,120	1,808	1,884	1,901	1,759	1,596	1,618	1,487	1,170	1,271	1,097	784	39,140
	11	3,548	3,167	2,783	2,694	2,551	2,434	2,226	2,246	2,119	1,915	1,835	1,893	1,912	1,692	1,680	1,600	1,402	1,330	1,124	868	41,019
	14	2,048	2,058	1,812	1,784	1,708	1,601	1,570	1,486	1,376	1,242	1,143	1,316	1,322	1,199	1,145	1,048	929	823	704	545	26,859
	17	1,483	1,374	1,259	1,213	1,224	1,119	995	958	921	844	724	864	949	851	778	737	696	558	476	411	18,434
	20	778	655	614	534	610	541	506	398	429	424	384	418	471	400	341	348	352	275	203	197	8,878
	23	330	336	265	244	264	204	256	179	197	165	141	224	210	168	150	149	140	108	90	81	3,901
	26	554	611	472	505	453	299	458	413	339	290	246	394	337	309	283	254	230	177	160	98	6,882
5	2	1,344	1,302	1,179	1,116	959	966	974	1,024	925	767	818	795	693	641	676	647	427	541	496	333	16,623
	5	3,566	3,167	2,504	2,605	2,350	2,188	2,389	2,235	2,201	1,736	1,828	1,932	1,729	1,523	1,583	1,463	1,191	1,197	1,169	740	39,296
	8	2,105	1,836	1,724	1,591	1,535	1,542	1,365	1,379	1,252	1,222	1,252	1,153	1,071	1,012	1,033	916	813	865	698	542	24,906
	11	943	693	892	920	787	702	520	739	463	634	557	476	461	587	503	464	462	365	288	230	11,686
	14	227	147	203	208	178	157	133	170	97	143	125	105	97	137	114	102	105	72	79	51	2,650
8	2	3,615	3,083	2,641	2,620	2,332	2,229	2,300	2,299	2,108	1,848	1,977	1,868	1,570	1,470	1,509	1,440	1,109	1,291	1,137	760	39,206
	5	2,175	1,672	1,597	1,544	1,485	1,424	1,281	1,288	1,144	1,164	1,144	1,077	1,017	1,022	970	912	818	758	676	500	23,668
	8	473	364	441	471	418	369	325	388	260	303	306	273	219	287	246	227	209	185	176	121	6,061
	11	121	85	107	119	89	82	62	93	56	75	60	57	54	73	60	56	56	37	35	23	1,400
11	2	4,198	3,760	3,337	3,146	2,777	2,846	2,681	2,681	2,541	2,223	2,336	2,233	2,064	1,849	1,821	1,756	1,415	1,655	1,355	907	47,581
	5	922	735	818	828	717	716	559	666	525	573	528	500	490	532	497	454	406	381	314	247	11,408
	8	88	58	92	85	83	69	58	74	89	61	62	48	38	58	50	44	44	38	27	22	1,138
14	2	3,369	2,914	2,693	2,573	2,462	2,422	2,245	2,095	2,037	1,848	1,887	1,871	1,838	1,651	1,588	1,474	1,301	1,324	1,143	825	39,560
	5	200	150	209	203	188	157	131	174	101	152	148	116	102	134	126	107	108	94	71	62	2,733
17	2	3,090	2,470	2,272	2,071	1,997	2,034	1,796	1,673	1,673	1,574	1,540	1,468	1,521	1,354	1,253	1,238	1,139	1,089	918	681	32,851
20	2	1,642	1,130	1,034	841	919	947	809	656	738	742	690	652	695	602	518	536	552	483	415	333	14,934
23	2	1,076	509	486	414	520	479	372	299	344	394	378	271	330	311	261	294	334	236	261	208	7,777
26	2	4,839	1,456	1,078	814	1,013	833	738	525	634	1,184	991	384	306	527	409	555	954	505	908	517	19,170
Total		47,668	38,417	34,773	33,181	31,306	30,117	28,253	27,782	25,815	24,374	24,201	23,276	22,093	20,784	20,007	19,171	16,884	16,462	14,665	10,552	509,781

Table 11: Number of Home-Days in Each Bin (Annual - Weekends)

Day Type / Household Group		Weekends																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
CDD	HDD																					
2	2	77	33	44	28	27	31	17	29	19	28	33	12	9	16	12	22	13	17	28	13	508
	5	508	495	496	434	387	407	358	406	336	297	345	289	247	239	229	252	148	232	178	127	6,410
	8	1,688	1,340	1,190	1,171	1,072	1,064	988	1,021	937	855	881	805	752	714	702	678	570	572	522	366	17,888
	11	1,161	1,188	1,062	985	945	966	841	820	808	729	702	726	739	626	646	595	496	536	406	351	15,328
	14	1,213	1,171	1,026	1,016	966	894	876	855	777	700	647	729	730	673	640	573	523	457	401	295	15,162
	17	473	483	411	424	411	329	363	358	313	265	248	315	296	267	262	241	215	182	156	119	6,131
	20	400	323	317	310	301	257	238	238	202	225	186	207	215	212	176	172	172	131	105	82	4,469
	23	196	222	170	167	155	135	143	123	132	106	91	130	139	107	107	100	86	77	67	53	2,506
	26	210	196	171	174	178	114	160	134	111	109	91	139	116	117	87	94	89	61	50	35	2,436
5	2	553	521	483	442	393	406	380	414	369	330	355	318	266	251	264	262	160	224	198	143	6,732
	5	1,452	1,263	992	1,044	964	929	931	875	913	697	749	761	724	618	677	618	512	503	503	350	16,075
	8	775	659	597	523	547	562	504	446	468	417	437	408	424	362	357	327	303	303	253	197	8,869
	11	223	196	231	212	196	183	166	187	125	161	160	136	108	131	126	115	103	104	77	61	3,001
	14	160	114	147	160	125	109	84	128	74	103	84	76	71	99	81	77	76	51	49	32	1,900
	17	40	28	35	39	29	26	19	31	18	25	18	17	17	24	19	19	18	11	10	7	450
8	2	1,131	1,093	1,008	969	847	851	834	865	771	647	737	678	588	543	546	512	368	511	414	263	14,176
	5	783	690	666	687	609	600	548	592	504	467	478	460	423	433	422	386	321	318	278	197	9,862
	8	211	170	202	202	167	156	136	173	109	143	134	124	104	129	112	95	96	84	63	44	2,654
11	2	1,766	1,429	1,225	1,186	1,083	1,016	1,053	987	931	865	859	875	761	711	682	654	561	573	507	339	18,063
	5	284	212	272	268	238	213	188	225	146	188	169	160	132	174	148	137	128	113	80	66	3,541
	8	[illegible]	[illegible]	7	4	6	4	4	6	[illegible]	4	9	[illegible]	[illegible]	[illegible]	5	[illegible]	3	6	[illegible]	3	61
14	2	1,124	974	869	806	765	795	696	651	668	590	595	586	592	504	510	497	427	440	387	293	12,769
	5	120	89	109	125	99	77	68	101	57	80	69	64	54	77	62	63	59	39	34	23	1,469
17	2	1,116	869	812	762	719	742	622	597	592	571	549	518	554	502	469	452	415	397	347	264	11,869
20	2	726	654	595	505	545	564	505	399	451	397	399	412	455	358	324	307	300	287	225	178	8,586
23	2	764	282	269	206	265	206	192	154	150	232	200	128	113	146	103	129	189	109	137	89	4,063
26	2	1,714	498	362	279	352	287	260	176	225	409	351	127	102	180	141	189	328	173	329	185	6,667
Total		18,868	15,192	13,768	13,128	12,391	11,923	11,174	10,991	10,206	9,640	9,576	9,200	8,731	8,213	7,909	7,566	6,679	6,511	5,804	4,175	201,645

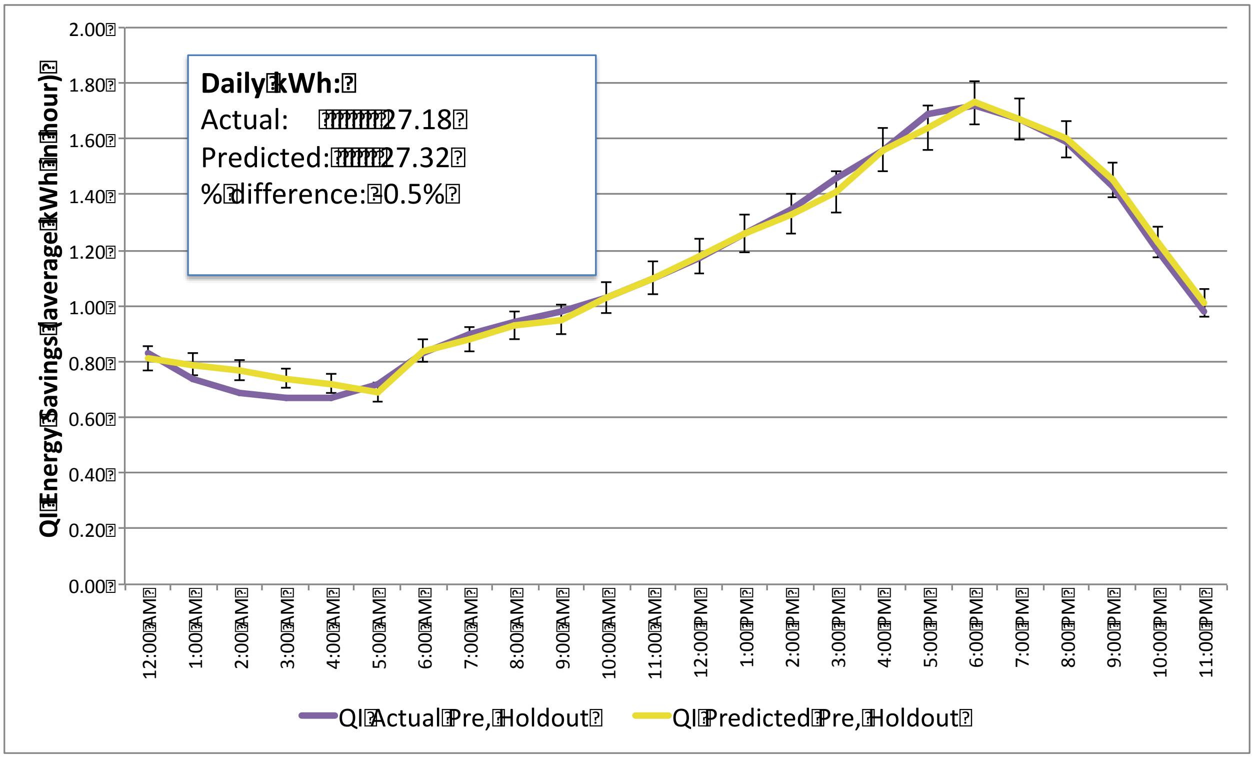
For the savings estimation stage, we recommend using one full year of post-period observations for each household. This ensures that the annual savings estimate is based on all four seasons and a wide range of daily temperatures. If we have a full year of post-period observations, we can calculate savings by comparing our predictions for the post-period (based on a pre-period model results applied to post-period weather) to the actual post-period observations.

At the time of our analysis, SCE's QI program data did not have a full year of post-period data available for the majority of homes in the sample. Rather than base our program-level savings estimate on a small group of homes, we relied on all observed post-period consumption ("actuals") and constructed a post-period model to impute the consumption for the remaining days in that year ("predicted"). Hence, in this case we are comparing our predictions for the post-period (based on a pre-period model and post-period weather) to the post-period, which is made up of all actual post-period observations available and imputed values for the days that are unavailable. These imputed values are generated using a post-period model that is based on the existing post-period data and post-period weather.[28]

In order to test the reliability of our annual model, we randomly selected 30 percent of the homes as a holdout sample and then modeled the remaining 70 percent of the homes. Figure 15 shows the comparison of the predicted pre-period load shape from the model (yellow) with the actual pre-period load shape for the holdout group (purple). As shown below, the model does a good job of predicting energy use in the holdout group in the pre-period, with a difference between estimated and actual usage of about 1 percent over 24 hours.[29] The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate.

[28] Another option for generating annualized savings estimates without a full year of post-period data is to use TMY3 weather.
[29] We used a subsample of the SCE QI data to test the effect of using multiple holdout groups. Using data from just summer weekdays, we ran multiple models to compare the results against six randomly selected holdout groups. Due to the significant computational resources needed to re-run these models using different comparison groups, the comparison test was limited to six randomly selected comparison groups comprising 30 percent of the available sample. The model highlighted in this report and a prediction accuracy of 99.1 percent, compared with an average prediction accuracy of 99.4 percent using all six comparison groups. Although this was a small sample, we took these results as evidence that our selection of a single comparison group would not materially affect the estimation results. The implications of using a larger sample of comparison groups will be explored in the next phase of the AMI analysis.

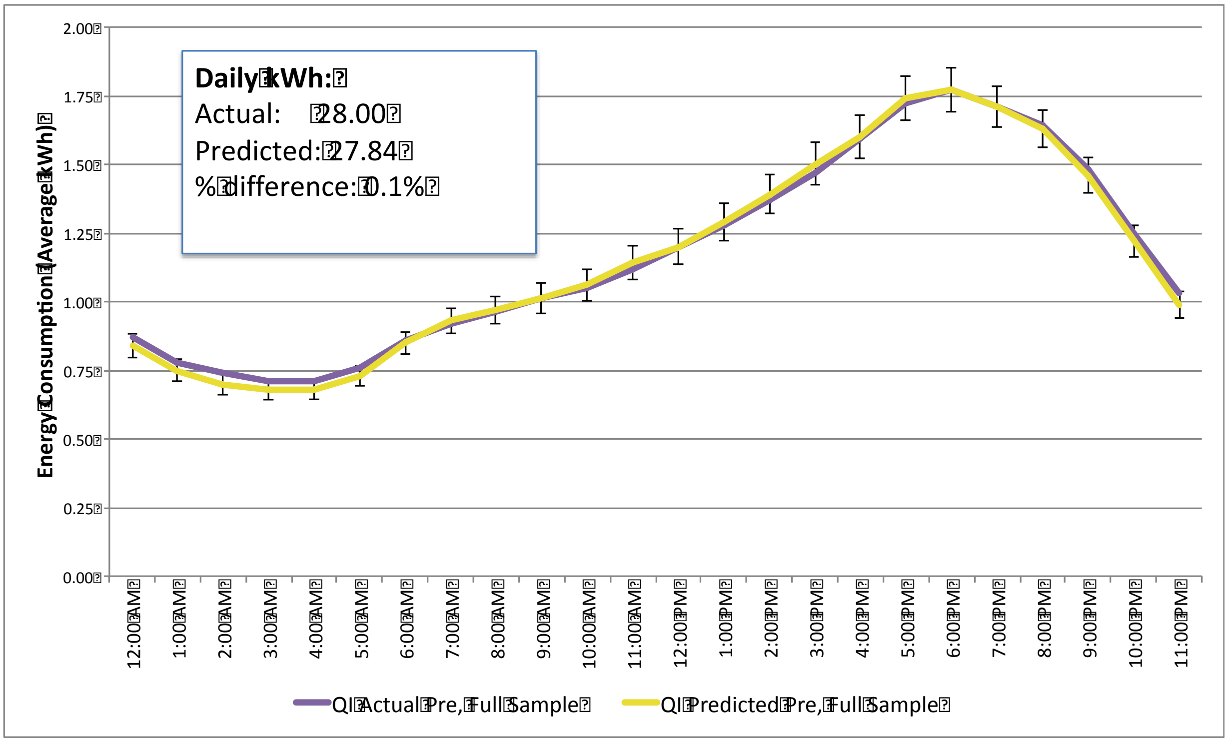
Figure 15: SCE QI Program Annual Predictions versus Actual of Holdout Homes, Pre-Period



RMSE = 0.034, MSE = 0.001

Figure 16 presents the comparison of the pre-installation predicted load shape from the model with the actual pre-installation load shape for all bins combined, including all 2,039 households in 9 different climate zones (i.e., no holdout group). The modeled pre-installation period load shape (yellow) aligns even better with the actual pre-installation load shape (purple), with a difference of less than 1 percent over 24 hours. The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate.

Figure 16: SCE QI Program Annual Predictions versus Actual of Full Sample, Pre-Period



RMSE = 0.023, MSE = 0.001

Next, we constructed a separate post-period model to impute missing observations in the post-period, a step necessary for creating annualized savings estimates without a full year of post-period data. This model is based on post-period data and post-period weather; it is not the pre-period model projected onto post-period weather. In order to test the reliability of this post-period annual model, we randomly selected 30 percent of the homes as a holdout sample, and then modeled the remaining 70 percent of the homes. Figure 17 shows the comparison of the predicted load shapes from the model (yellow) with the actual load shapes for the holdout group (purple) in the post-period. The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. As with the other holdout sample analyses, the model does a good job of predicting energy use in the holdout group in the post-period, with a difference between estimated and actual usage of less than 1 percent over 24 hours.

Figure 17: SCE QI Program Annual Predictions versus Actual of Holdout Homes, Post-Period

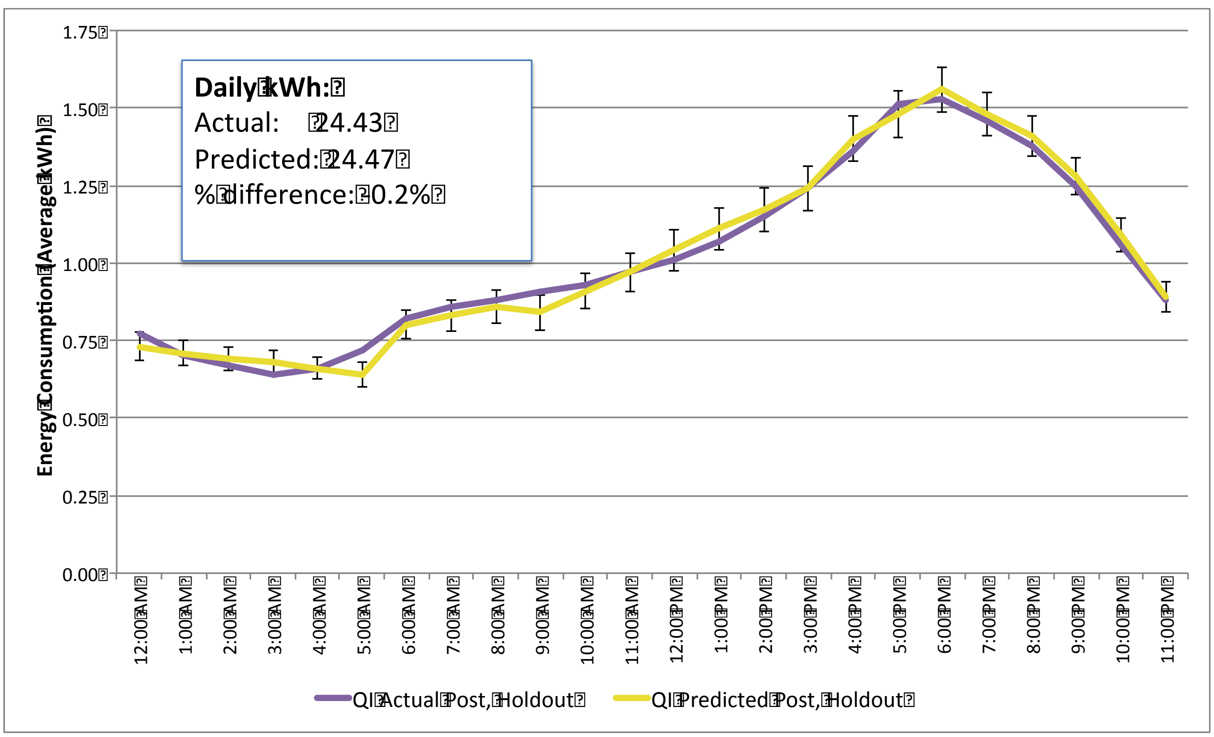


Figure 18 presents the comparison of the post-installation predicted load shape from the model with the actual post-installation load shape for all bins combined, with all 2,039 households in 9 different climate zones (i.e., no holdout group). The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. The full sample's modeled post-installation period load shape (yellow) aligns very closely with the actual post-installation load shape (purple), with a difference of less than 1 percent over 24 hours.

Figure 18: SCE QI Annual Predictions versus Actual of Full Sample, Post-Period

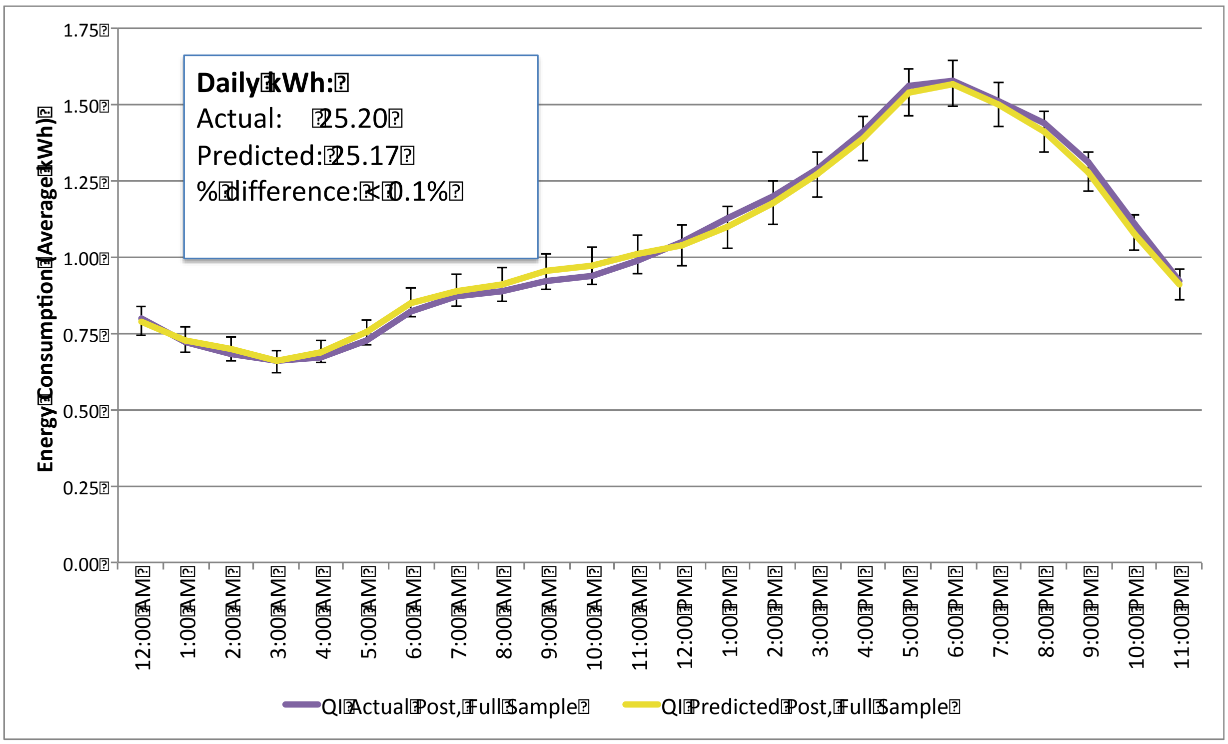
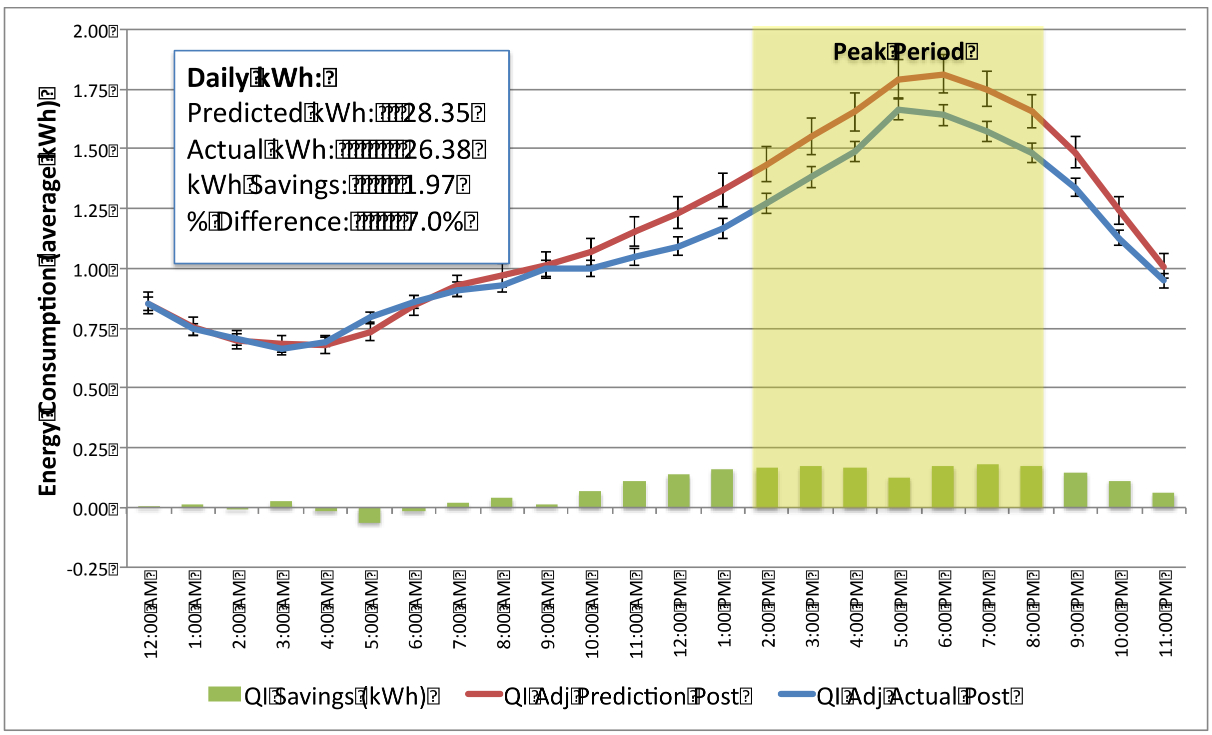


Figure 19 compares the pre-period predicted load shape (red line) with the post-period predicted load shape (blue), averaged across all households. Whenever the post-period load shape falls below the pre-period load shape, this indicates that savings were realized during that hour (green bars). After adjusting for the error in the models, based on the sample of homes used, the modeling approach finds approximately 7 percent annual savings attributable to the HVAC installed through the SCE QI Program.[30] Note also that this approach finds that the majority of savings is realized during the later part of the day including during the peak hour periods of between 2:00 p.m. and 8:00 p.m.,[31] highlighted in yellow. The 95 percent confidence is shown for each estimate, and the error of the hourly predictions is greatest during the late afternoon and early evening, and smallest during the early hours of the morning.

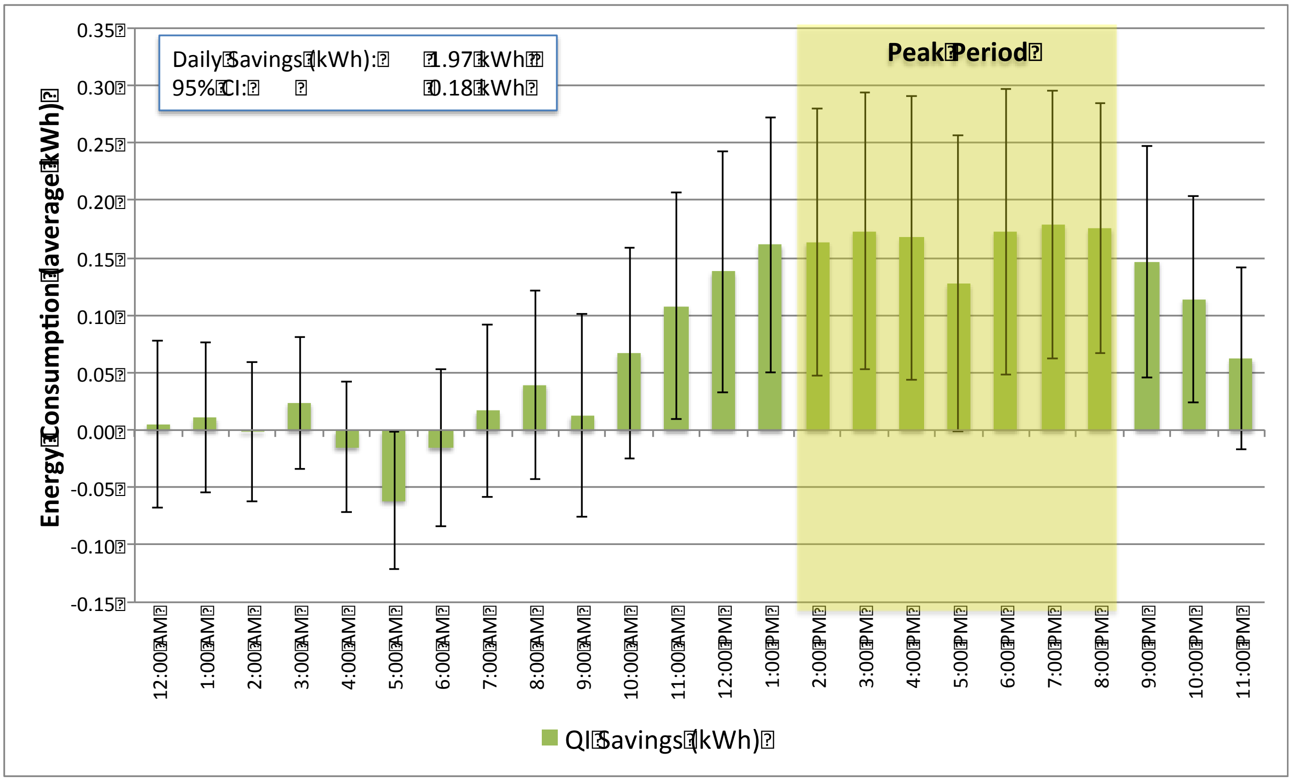
Figure 19: SCE QI Overall Annual Post-Period Model, Includes All Months and Day Types



[30] It is not possible to determine how much of the estimated savings come from the quality installation practices versus the new HVAC equipment. If we needed to separate these impacts, we could compare these results to a control group of customers who replaced their HVAC system but did not use a program contractor for the installation.
[31] We use the residential peak period of 2:00 p.m. to 8:00 p.m., as defined for SCE's residential Time-Of-Use rate plan: https://www.sce.com/wps/portal/home/residential/rates/residential-plan

Figure 20 shows the annual hourly kWh savings estimates along with the 95 percent confidence interval for each hour. We found statistically significant hourly savings from 11:00 a.m. to 4:00 p.m. as well as 6:00 p.m. to 10:00 p.m. Some early morning hours had increases in usage (i.e., negative savings), but none of these were statistically significant.

Figure 20: SCE QI Overall Annual Savings, Includes All Months and Day Types



The SCE QI annual model includes a day-type binning component allowing us to evaluate energy savings for weekdays versus weekends. Figure 21 and Figure 22 compare the pre-period predicted load shape (red line) with the post-period predicted load shape (blue), averaged across all households for weekdays and weekends respectively. The modeling approach finds slightly higher savings on weekends (7.26%) versus weekdays (6.98%) however the differences in hourly savings are not statistically significant.

Figure 21: SCE QI Overall Annual Post-Period Model, All Months; Weekdays

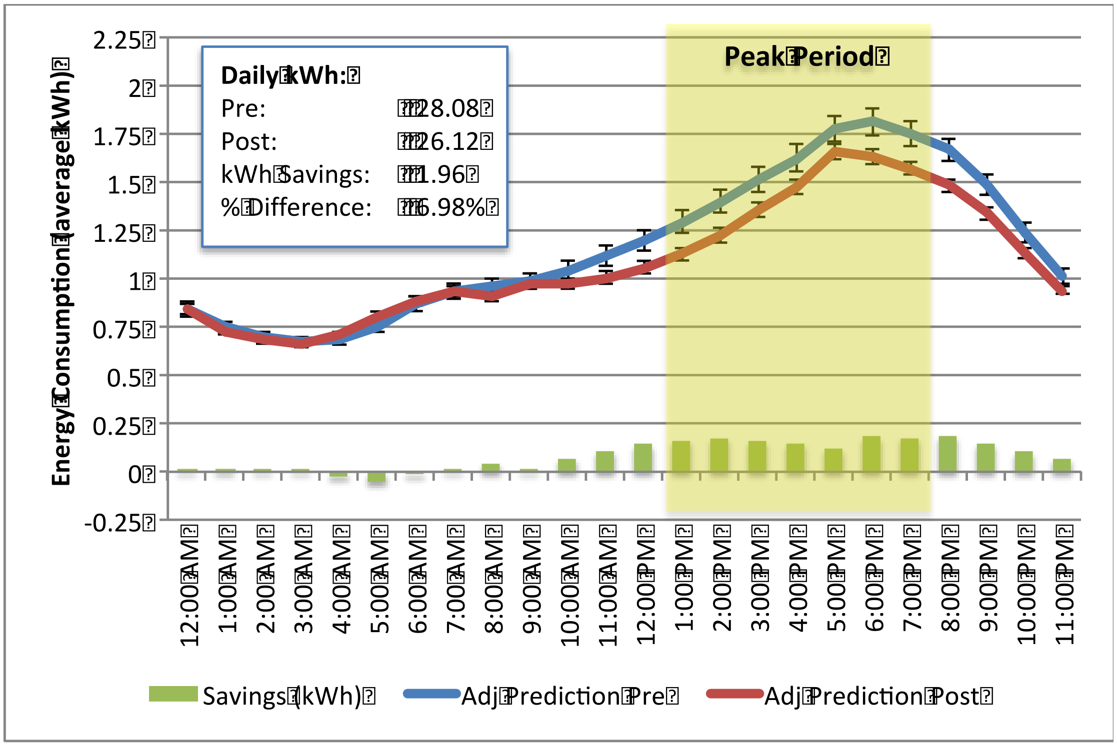


Figure 22: SCE QI Overall Annual Post-Period Model, All Months; Weekends

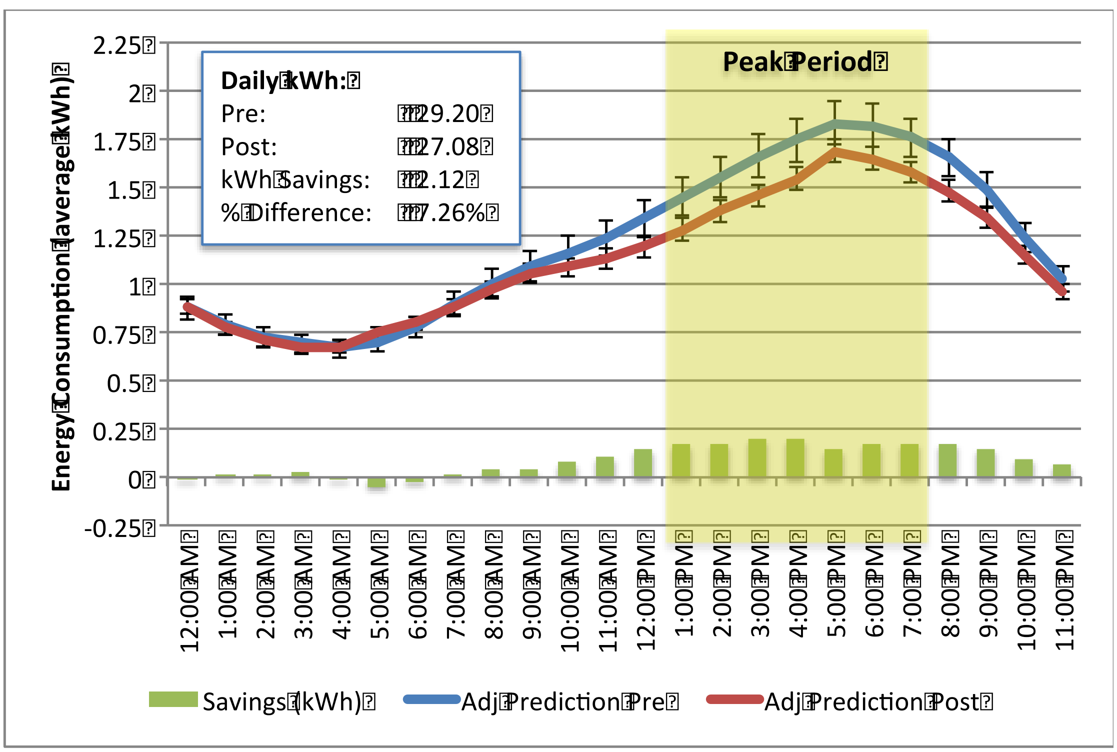


Table 12 and Table 13 provide the average daily savings estimate for each group and bin in the SCE QI annual model on weekdays and weekends respectively. The columns show households grouped by their weather normalized energy usage in the pre-period for each home (highest users on the right) and the rows show days grouped by the temperature via cooling degree-days (hottest days on the bottom). Each cell shows the estimated program savings (kWh per day) for a specific home-day bin. We automatically color-coded the cells with the highest kWh savings in dark blue and the lowest kWh savings in dark red; colorless cells fall in the middle of this spectrum. In general in the annual model we see increased savings as temperatures increase with some deviation from this trend in specific bins.

Table 12: Program Savings (kWh per day) by Bin (Annual – Weekday)

Day Type		Weekdays																			
Household Group		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CDD	HDD																				
2	2	0.9	0.9	-0.4	1.8	-1.1	-1.3	-2.2	-0.8	3.9	1.5	-2.8	-0.2	-0.3	-0.2	-7.0	-4.3	-7.9	-1.4	6.0	-12.5
	5	0.3	0.8	-0.1	0.4	-0.7	0.2	-0.9	0.0	0.4	0.2	-0.3	0.4	0.0	-0.2	-1.9	-4.2	-2.1	-1.7	0.7	-5.6
	8	1.1	1.1	0.7	0.4	-0.7	0.0	-1.2	-0.2	-0.2	-0.6	-0.6	-0.9	0.4	-0.4	-0.9	-3.6	-1.4	-1.8	-1.7	-7.7
	11	1.0	1.1	0.6	0.4	-0.7	-0.1	-0.8	-0.5	-0.5	-0.3	-0.3	-1.2	0.2	-0.2	-0.6	-2.7	-1.4	-1.5	-0.4	-4.8
	14	1.2	0.8	0.8	0.1	-0.1	-0.3	-1.5	-1.2	-0.9	-0.6	-0.3	-1.1	0.0	-1.3	-0.4	-3.6	-0.7	-1.2	0.3	-6.0
	17	1.5	2.3	1.5	1.1	1.1	0.9	0.0	-0.9	0.1	1.7	0.2	-0.2	1.1	-0.8	2.5	-1.9	0.0	1.3	1.1	-2.5
	20	1.1	2.0	3.1	0.8	0.2	-0.1	-0.3	-1.5	0.3	0.7	0.2	0.5	2.1	-0.4	1.7	-3.8	0.1	-2.2	3.8	-14.6
	23	1.7	1.1	1.2	1.0	0.6	0.0	-1.3	-1.6	-0.9	1.6	-0.2	-0.9	1.5	-0.6	1.5	-3.7	1.0	-0.7	-3.7	-11.4
	26	1.2	0.9	1.1	-0.1	-1.5	0.8	-0.6	-2.7	-2.0	1.7	-0.4	-0.3	0.1	-2.3	0.8	-6.3	2.5	3.9	3.5	-6.2
5	2	1.0	1.4	1.0	1.3	0.4	0.8	0.3	0.4	1.2	0.4	1.1	1.0	1.5	1.6	-3.0	-0.9	-1.1	1.3	-4.0	-2.6
	5	0.8	1.3	1.0	0.7	0.1	0.4	-0.4	0.0	0.7	0.1	1.0	0.1	1.9	0.6	-1.3	-0.7	-1.1	1.7	0.2	-1.6
	8	0.8	1.6	1.4	0.9	0.2	0.8	-0.5	0.5	0.3	-0.4	0.3	-0.3	-0.1	0.5	0.1	-0.9	-1.2	-1.1	2.4	-6.0
	11	0.1	0.0	1.1	0.0	0.6	0.6	-1.0	0.3	-0.7	-0.7	-1.4	-1.6	-1.5	-0.6	-0.2	-3.6	-1.9	-3.2	1.5	-13.4
	14	2.1	1.8	3.4	1.1	2.7	2.8	0.4	1.2	-0.2	0.7	0.9	0.7	1.5	2.4	1.4	-1.2	-2.0	2.0	4.1	2.4
	17	0.0	0.0	7.5	0.0	0.0	0.0	-0.5	0.0	0.0	0.0	-4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	2	1.0	1.1	1.5	1.5	0.7	0.9	-0.1	-0.1	0.7	0.6	2.1	2.0	2.6	2.1	-1.6	0.8	0.0	4.5	-1.8	3.5
	5	0.8	1.5	1.2	0.1	-1.0	0.6	-0.8	0.1	1.3	-0.3	0.7	-0.4	1.2	0.6	-1.4	0.6	0.4	1.0	2.3	5.2
	8	0.6	0.5	1.5	-0.1	-0.6	-0.6	-1.4	-0.1	0.1	-0.2	-0.4	-1.0	-1.2	-0.6	-0.6	-2.9	-0.3	0.2	1.0	-4.3
	11	0.0	-0.9	0.4	-2.6	-1.6	-2.1	-4.9	-2.7	-5.9	-2.2	-2.5	-11.6	-8.3	0.6	-4.0	-10.1	-7.6	-9.7	1.1	-9.6
11	2	0.9	2.1	2.5	2.1	1.4	2.7	1.8	3.0	2.8	3.4	4.0	3.3	3.7	3.3	2.5	4.6	4.6	6.8	3.1	9.1
	5	1.0	1.5	1.5	1.5	0.7	1.5	0.5	0.7	1.4	2.1	1.5	0.3	1.6	2.9	0.9	3.8	3.7	1.1	5.0	8.8
	8	-1.0	-1.3	2.1	-0.4	-1.2	-1.3	-3.2	-0.6	1.1	-3.5	-6.2	-1.1	-7.3	-1.7	-6.2	-5.7	-3.7	0.5	0.0	-10.7
14	2	1.3	2.4	2.8	3.6	3.0	3.0	2.8	4.4	5.2	5.3	6.0	6.2	6.3	5.6	4.6	7.7	8.3	10.0	8.6	16.7
	5	1.0	0.6	0.8	-1.3	-2.1	-2.4	0.3	0.6	3.9	0.2	-2.1	0.8	1.5	-0.3	-4.7	4.3	3.7	0.5	5.8	2.9
17	2	1.8	3.8	3.4	4.9	4.9	4.5	4.1	6.1	7.2	7.7	7.7	8.0	7.6	7.8	7.5	10.4	11.7	12.8	14.0	24.4
	5	-13.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	2	3.2	5.8	4.2	7.4	7.6	7.4	6.5	9.5	11.8	10.9	11.0	10.1	11.3	11.8	13.9	14.6	14.8	17.0	20.7	34.1
23	2	4.7	7.1	6.4	8.8	8.1	7.0	7.0	8.9	10.3	11.7	16.4	8.6	11.2	12.9	10.7	17.6	18.0	19.3	24.0	31.6
26	2	7.3	5.4	5.9	14.5	1.9	14.6	14.6	3.5	8.2	10.2	17.8	7.7	14.1	5.4	2.8	14.7	18.1	33.6	12.7	-23.7

Table 13: Program Savings (kWh per day) by Bin (Annual – Weekend)

Day Type		Weekends																			
Household Group		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CDD	HDD																				
2	2	1.6	0.6	-1.5	0.2	-1.3	-0.5	-1.9	2.0	4.7	0.7	-1.0	2.7	0.9	-0.2	-2.0	-0.2	0.4	15.2	12.9	-12.6
	5	0.8	1.2	0.5	0.4	-0.8	0.5	-0.4	0.2	0.2	0.6	0.3	-0.8	0.2	-0.6	-0.9	-4.2	-2.4	1.6	5.0	-6.2
	8	1.2	1.8	0.6	0.9	-0.3	0.8	-0.4	0.0	0.1	0.3	0.2	-0.9	1.5	-0.1	-0.3	-3.5	-0.4	-0.9	0.4	-6.7
	11	1.2	0.9	0.4	-0.2	-0.2	0.8	-1.4	-0.5	-1.0	0.4	-0.1	-1.4	-0.3	-0.4	0.4	-2.4	-0.4	-2.8	0.2	-2.3
	14	1.6	0.8	0.4	0.5	-0.2	0.9	-1.0	-1.6	-0.6	0.2	-0.7	-1.7	0.4	-0.3	0.6	-2.6	0.5	-2.1	1.9	-1.3
	17	1.6	2.3	1.8	1.5	0.1	1.2	0.3	-0.3	-0.3	0.6	-0.2	-0.7	0.5	-1.1	2.3	-3.7	-0.8	0.9	1.8	-8.0
	20	1.2	2.4	1.5	2.0	0.0	0.8	-0.5	0.5	0.6	1.1	1.2	0.2	2.2	0.0	2.2	-4.1	0.2	0.4	8.1	-8.8
	23	2.5	2.5	4.0	1.7	0.3	2.5	1.0	-0.9	1.0	3.8	0.7	1.6	1.8	-2.9	2.0	-0.3	6.9	6.8	-5.1	-6.4
	26	2.3	3.4	3.0	3.1	2.9	4.5	1.8	1.5	1.7	4.4	0.2	1.8	4.2	1.5	1.7	-1.3	2.9	3.9	7.8	1.1
5	2	1.5	1.2	1.6	1.9	1.4	2.9	1.8	2.2	2.0	2.7	1.5	1.5	4.4	3.0	-0.6	0.5	2.1	1.8	1.8	0.4
	5	1.3	1.6	0.6	1.1	0.1	1.0	-0.5	0.3	-0.2	-0.2	0.5	-0.8	1.8	0.3	-2.7	-1.6	-1.0	0.1	-1.0	-2.9
	8	0.9	1.6	0.8	0.3	-0.8	1.1	-0.9	0.1	1.9	-0.3	-0.3	-1.0	0.9	0.9	0.4	-0.8	-1.8	-0.4	3.0	-6.7
	11	1.4	2.0	2.6	0.4	1.7	0.7	-1.1	-0.5	-1.4	-0.7	-0.7	-0.3	0.5	-0.7	0.4	-1.4	-1.9	-1.5	2.9	-11.5
	14	1.3	0.7	1.7	-0.5	1.4	1.1	-0.2	0.0	-0.3	-0.6	-0.8	-0.9	-1.1	0.9	1.3	-4.2	-2.5	0.8	3.9	9.9
	17	1.7	0.0	7.7	-0.4	0.0	0.0	0.0	0.0	0.2	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	2	1.2	1.9	2.7	2.9	1.4	2.9	1.9	2.0	2.9	2.3	3.2	3.3	4.0	3.2	2.6	2.1	2.7	4.7	3.4	4.1
	5	1.1	2.0	1.9	0.6	-0.5	1.6	0.1	-0.5	0.8	0.4	0.8	-1.1	1.0	-0.8	-1.1	0.3	0.5	1.1	2.9	-1.7
	8	0.5	0.0	0.8	-1.0	-0.3	0.7	-1.5	1.5	0.0	-0.3	-1.2	-2.8	-1.2	-0.7	-1.7	-0.8	2.6	1.9	-1.1	-9.5
	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	2	1.2	1.8	2.2	2.1	1.5	2.0	0.1	2.1	0.8	2.4	3.3	1.6	3.4	3.8	1.1	2.2	3.9	8.2	1.6	11.3
	5	0.4	0.7	2.8	1.3	-1.3	-1.7	-2.3	0.5	0.3	-0.7	-0.5	-1.8	0.8	3.0	-2.4	4.4	2.8	2.7	3.2	8.5
	8	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	2	1.3	2.7	3.4	3.4	2.8	3.8	4.3	4.9	2.9	5.9	7.2	4.2	4.4	4.9	6.7	7.9	8.5	8.8	8.0	19.0
	5	-0.8	-2.8	0.2	-3.5	-4.2	-2.5	1.9	0.2	4.3	2.8	-2.5	-0.6	1.7	-3.3	-5.5	5.9	10.3	-1.3	11.1	4.7
17	2	2.6	5.0	6.3	6.3	6.8	9.0	6.2	8.6	9.8	10.7	11.2	10.1	12.0	10.8	9.3	10.3	14.2	14.0	19.0	25.3
	5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	2	3.4	5.9	5.5	8.8	9.1	9.0	7.1	9.6	10.3	11.0	11.6	11.6	10.8	10.9	11.5	15.2	15.1	14.1	27.1	38.9
23	2	4.5	9.1	4.2	9.6	4.9	11.0	6.9	6.6	9.9	11.5	12.5	10.2	10.1	6.7	5.1	14.7	15.2	23.3	16.8	10.6
26	2	9.7	8.5	6.8	12.5	2.4	18.1	7.2	-1.0	9.6	10.9	20.1	5.7	13.0	12.7	3.3	18.3	17.1	34.4	16.3	-24.0

The model results were also compiled to show load shapes and savings by season, and each of the four seasons has a unique range of temperatures. Consequently, a household's total kWh consumption and load shape will change as the need for heating and cooling changes.

Figure 23 shows our annual model's predicted load shape before (red) and after (blue) the households participated in the SCE QI program by season.[32] Most of the SCE QI program savings occurred in the summer, which had an average daily savings of 5.3 kWh or 12.8 percent. Fall and spring had the next highest savings with 1.8 and 1.2 kWh respectively, corresponding to 7.7 percent and 4.7 percent of the average daily kWh usage. The summer load shape from the annual model is very similar to the summer weekday model presented in the previous section.[33] None of the differences in the models' predictions of hourly or total daily consumption in the post-period are statistically significant. Despite the variation in load shapes across seasons, the random coefficients model is able to produce very accurate estimates in a variety of conditions.

[32] Note that these are not separate models; they are all based on the bin-level output produced by the annual model. We defined these seasons as: summer (July-Sept), fall (Oct-Nov), winter (Dec-Feb), spring (Mar-June).
[33] The annual model summer load shape includes both weekends and weekdays, explaining some of the difference between the summer weekday model load shape and the summer load shape developed from the annual model.

Figure 23: SCE QI Annual Model Results by Season

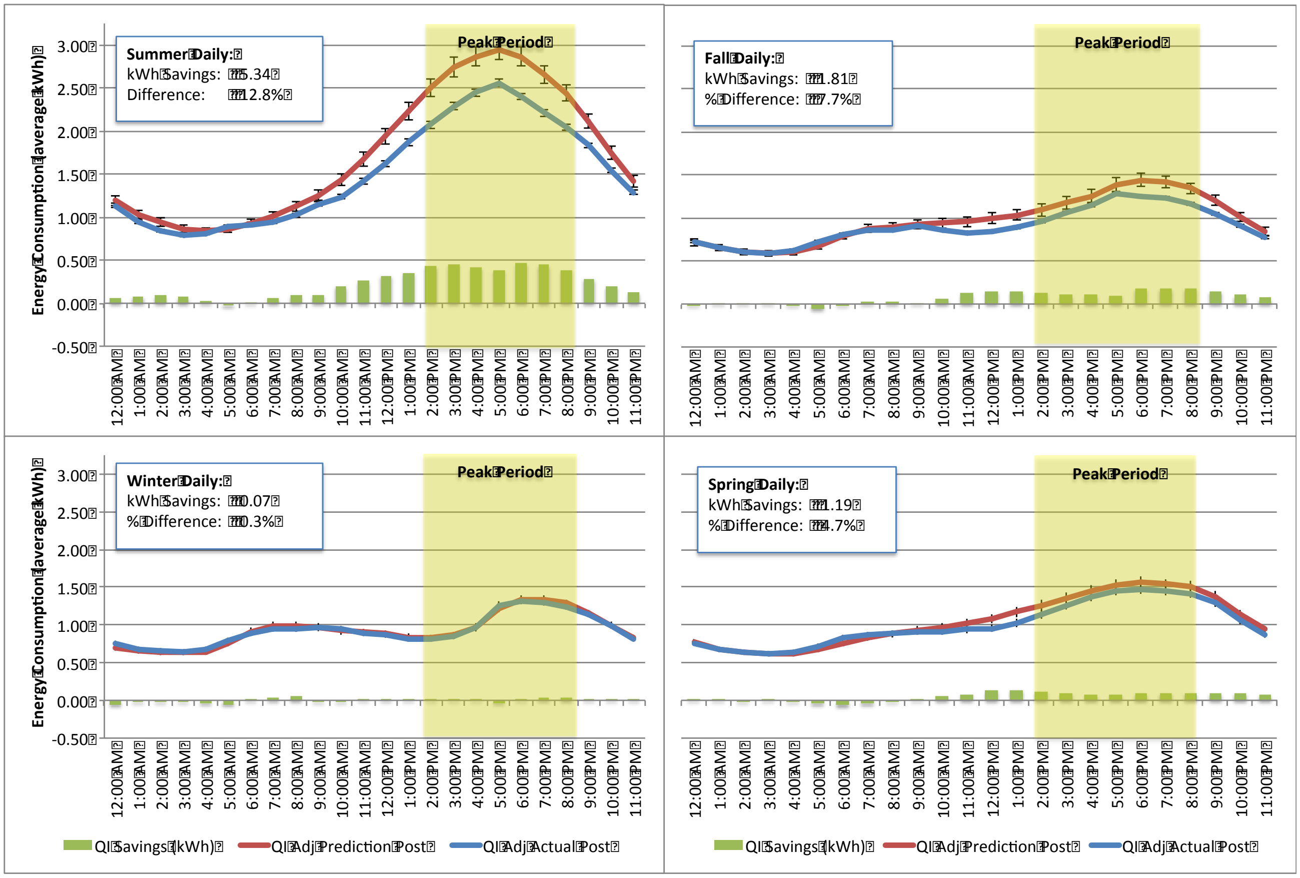
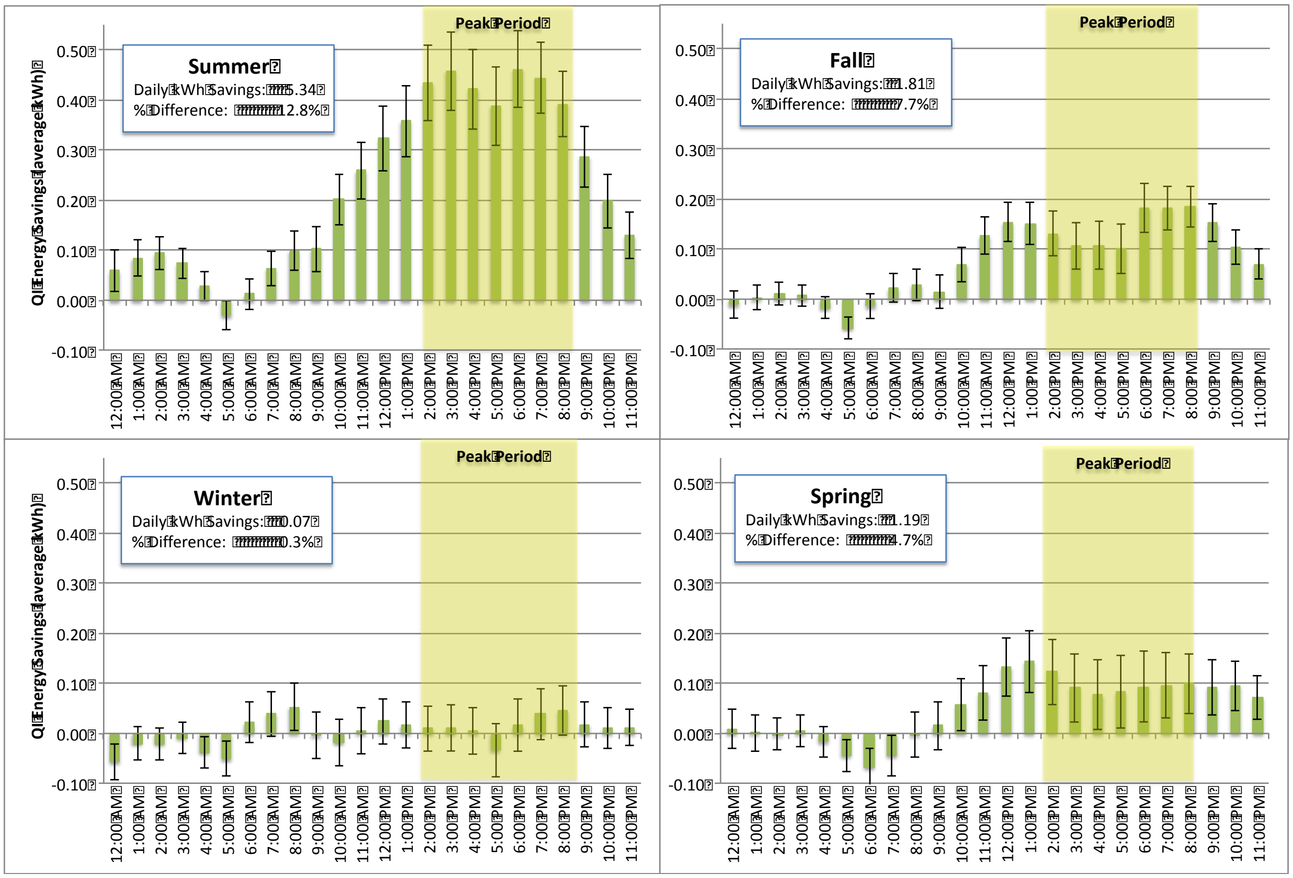


Figure 24 shows the hourly kWh savings estimates from the annual model for each season, with bars representing a 95 percent confidence interval around each estimate. The largest savings in summer months occurred during peak hours, while the savings in fall and spring occurred in the afternoon or evening. Nearly all savings in the winter months are either insignificant or negative (i.e., the households increased their usage).

Figure 24: SCE QI Annual Model Savings Estimates by Season



We have provided additional results from this annual model in the report appendix. These additional results include a table with the results of the pre-period holdout sample and post-period savings estimates for each of these seasons and day types (weekdays and weekends); modeled hourly load shapes and savings estimates for selected groups (household usage, CDD, and HDD groups), as well as a table with daily savings estimates for every bin in the annual model.

2.2.3 PG&E Quality Maintenance Program Results

In addition to SCE's QI Program, we also tested the random coefficients model using data from PG&E's Quality Maintenance Program. Each home in the PG&E QM Program had an existing HVAC system[34] repaired as part of a three-year service contract with a contractor who has received additional training through the program. During each visit, the contractor conducts a full ACCA Standard 4 HVAC System Assessment and then performs any required maintenance. Examples of these maintenance activities include airflow correction, blower motor retrofits, and refrigeration system assessment with savings. These activities should improve cooling delivery (from reduced runtime and/or power draw) and thereby improve efficiency.

Note that for the QM program, the existing conditions in the pre-participation period is the appropriate baseline and therefore no additional adjustments are needed to the baseline to calculate program impacts. The savings estimates would benefit from utilizing a comparison group, however, which we were not able to explore in this analysis due to the data limitations discussed previously. Incorporating data from a comparison group into the random coefficients model is something that will be explored in the next phase of this analysis.

PG&E QM Program Model – Weekdays, Summer Only

The modeling steps for the QM model are the same described previously for the QI program. For the PG&E QM model, the analysis sample includes all participating homes with AMI data for the 2012-2014 cooling seasons (the summer months of July–September). All homes in our sample participated between December 2012 and May 2014 and had non-zero *ex ante* savings listed in the program documentation (i.e. tracking data).[35] Since the HVAC equipment repaired through the QM Program were all listed as air conditioners, we expected the majority program savings to occur during the cooling season. This model uses only weekday data to avoid possible differences in energy usage between weekends and weekdays while using a simplified binning procedure. The resulting dataset includes 1,166 homes dispersed across four climate zones.

Since this model only includes weekdays during summer months, the home-days are assigned to two-dimensional home-day bins that do not include bins for HDD or day type. For this model, we used 20 home groups and 25 CDD groups, resulting in 500 home-day bins. We assigned each rounded CDD to its own CDD group but capped the CDD at 25, including all days with CDD greater than 25 to CDD group 25. This was done to limit the total number of bins

[34] Eligible homes must have a central forced air conditioner or heat pump and be a single-family home or duplex.

[35] We received AMI data for these PG&E households from 2012 through the beginning of 2015, including both a full pre- and post-period cooling season for each household in the sample. Thus, we did not have to exclude any homes due to missing data. Some QM participants did not require any adjustments (i.e. tests revealed system did not need any maintenance), these participants were excluded from our analysis because they did not involve one of the following activities: airflow correction, blower motor retrofit, and/or refrigeration system assessment with savings.

and thereby reduce processing time, but for program evaluations we suggest binning up to the true maximum CDD in the data. In order to isolate days with expected cooling, we removed all days with a CDD of zero.

Table 14: Summary of PG&E QM Program Summer Weekday Binning

Group	Description	Number of Groups
Homes	Usage – weather normalized annual energy usage grouped by percentile, with 1/20th of the total kWh assigned to each group in order from smallest to largest	20
Days	CDD – average of CDH rounded up to a whole number, assigned one CDD per group from 1-25 with all days higher than CDD 25 put into the last group	25
Total	Home-Days	500

Table 6 shows the count of home-days in the post-period assigned to each bin for the PG&E QM summer model. As with previous tables, the cells are automatically color-coded with the highest count in dark blue and the lowest count in dark red, white cells fall somewhere in the middle of this spectrum. Similarly to the SCE QI summer model we see more mid-temperature days with CDDs ranging from 7 to 18 than especially high or low temperature days within each of the household groups. We also see more home days in the home groups at the lower end of the usage spectrum. This is because each home group represents about 5 percent of total baseline electricity usage for the homes in our sample. Because of this, the number of homes in each home group varies, with more homes in the lower home groups and fewer homes in the higher home groups. However, the amount of daily kWh each home group represents is approximately the same.

Table 15: Number of Home-Days in Each Bin

Cooling Degree-Day Groups	Household Groups 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
1	3	9	13	8	10	5	7	9	3	7	3	4	4	5	2	4	4	0	7	0	107
2	12	45	61	40	42	25	27	40	13	34	13	15	19	22	11	19	20	5	32	2	497
3	12	41	49	36	25	23	29	36	18	35	15	21	15	21	15	20	21	8	27	6	473
4	50	100	118	87	67	58	69	80	34	78	41	46	43	51	39	47	53	25	66	12	1,164
5	56	109	106	94	65	68	122	108	96	102	72	122	70	88	86	79	90	68	67	53	1,721
6	83	119	126	109	103	78	114	92	79	83	76	94	67	104	81	74	69	61	70	44	1,726
7	147	240	231	207	174	165	224	213	177	191	144	240	131	184	172	150	152	127	134	95	3,498
8	255	310	295	293	261	219	250	193	174	211	196	211	175	208	190	155	149	145	143	84	4,117
9	235	322	314	310	262	232	250	207	182	237	168	253	170	165	180	153	142	105	132	79	4,098
10	348	444	348	358	334	291	291	216	245	309	269	281	232	233	233	193	194	172	152	107	5,250
11	488	517	411	423	327	371	360	258	326	350	300	298	283	247	231	185	231	190	158	82	6,036
12	254	299	279	254	220	205	230	197	179	217	175	218	171	169	164	147	143	133	120	79	3,853
13	386	370	317	299	250	256	261	204	209	257	222	220	205	218	177	159	170	155	140	83	4,558
14	439	457	347	380	324	309	324	271	313	315	310	301	290	281	279	202	249	239	166	150	5,946
15	344	315	251	267	234	217	242	178	204	229	223	207	205	198	192	178	165	190	116	113	4,268
16	359	316	243	233	223	218	233	157	185	214	208	170	189	190	165	147	154	170	117	89	3,980
17	414	340	281	277	215	235	243	181	215	196	205	172	202	181	140	117	147	127	101	45	4,034
18	210	169	151	146	96	120	135	97	123	118	113	90	113	86	76	66	79	100	50	36	2,174
19	292	183	176	155	109	147	151	118	154	145	120	127	135	107	104	88	116	107	56	57	2,647
20	86	58	59	45	36	41	46	28	40	49	37	42	37	35	31	33	27	33	14	15	792
21	107	51	43	29	27	28	49	30	43	33	25	34	39	21	29	29	33	42	10	28	730
22	261	117	103	83	87	78	87	84	82	82	83	78	74	79	73	61	72	75	53	69	1,781
23	87	34	42	19	22	27	35	30	24	29	13	32	17	14	19	17	12	29	13	26	541
24	161	75	69	56	49	57	52	52	44	40	39	35	44	33	30	39	33	32	30	33	1,003
25	2,229	348	541	225	254	251	167	498	242	160	200	59	184	71	138	130	234	170	341	490	6,932
Total	7,318	5,388	4,974	4,433	3,816	3,724	3,998	3,577	3,404	3,721	3,270	3,370	3,114	3,011	2,857	2,492	2,759	2,508	2,315	1,877	71,926

As before, to test the reliability of our annual model we randomly selected 30 percent of the homes as a holdout sample and then modeled the remaining 70 percent of the homes. Figure 25 shows the comparison of the predicted pre-period load shape from the model (yellow) with the actual pre-period load shape for the holdout group (purple). The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. The holdout analysis yields results similar to the QI model, with a difference between estimated and actual usage of less than 1 percent over 24 hours.

Figure 25: PG&E QM Program Summer Predictions versus Actual of Holdout Homes, Pre-Period

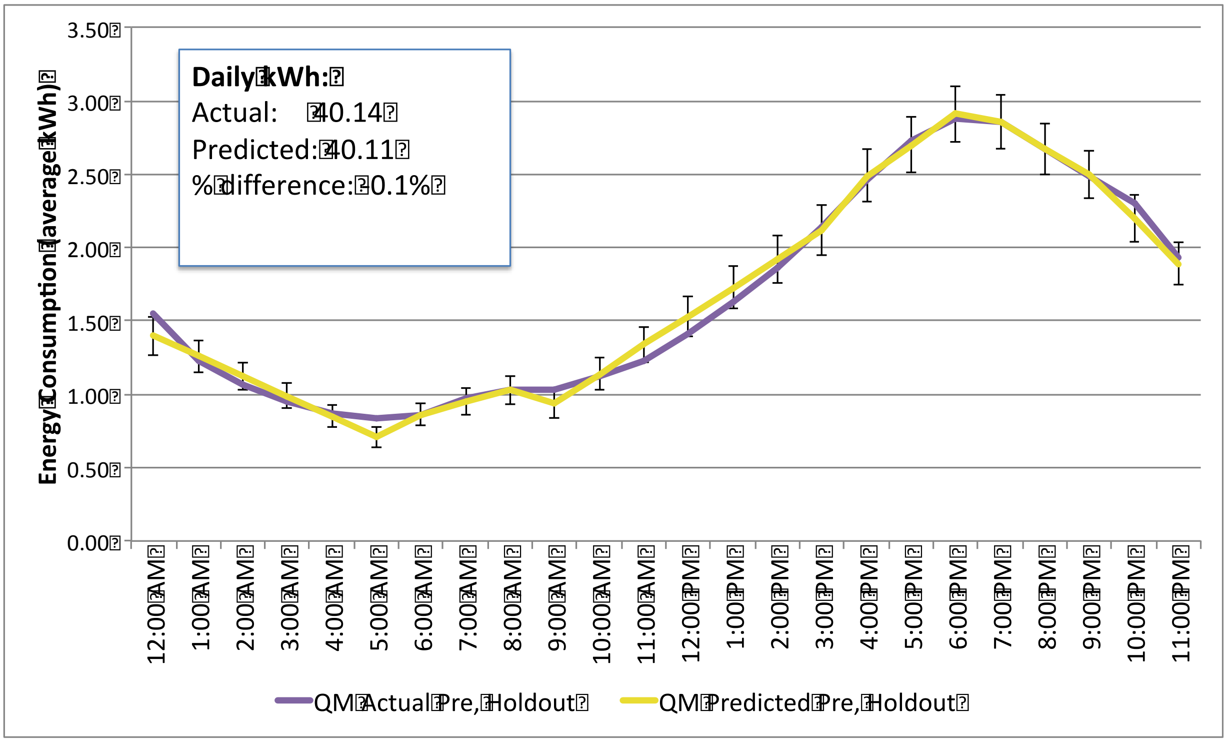
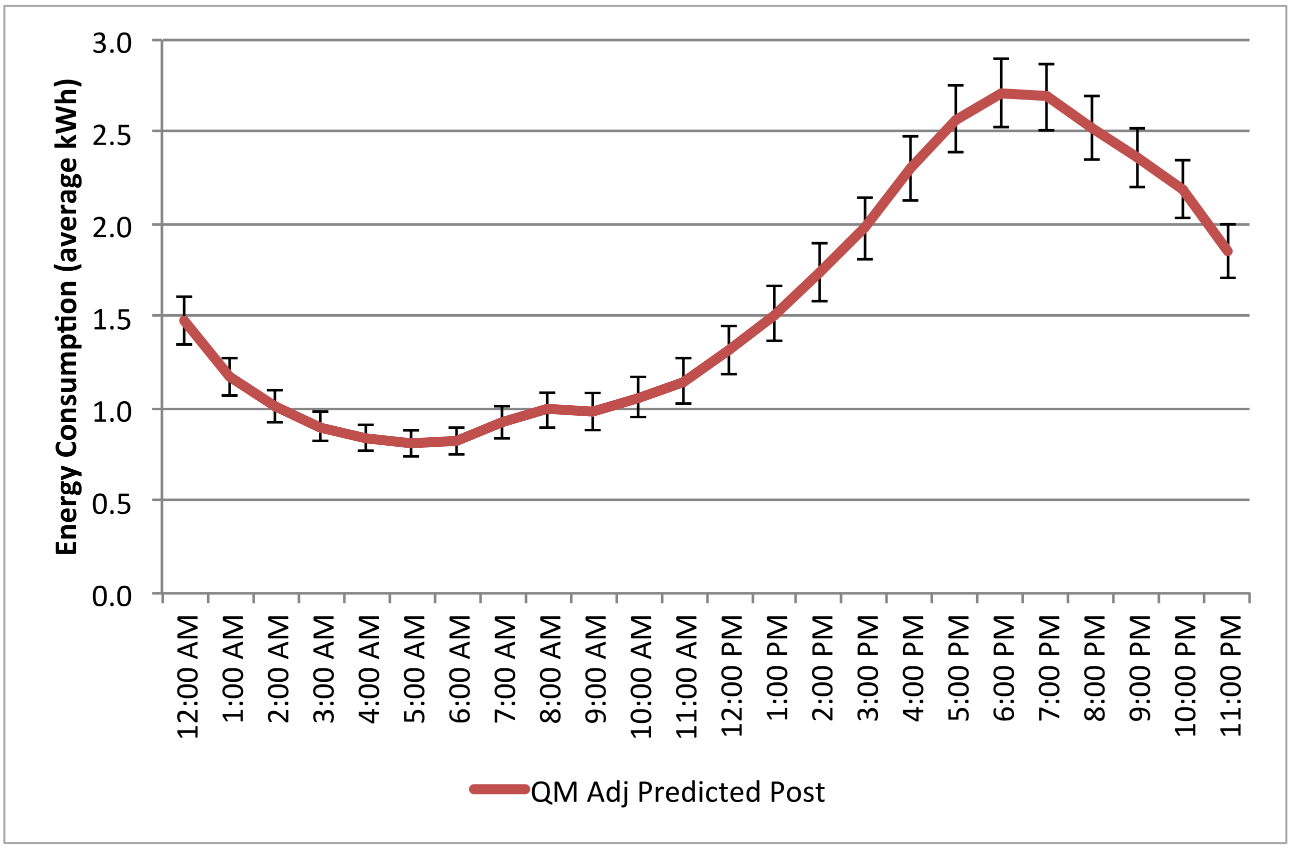


Figure 26 below shows the adjusted model prediction of post-period consumption for all households that participated in the PG&E QM Program.[36] This prediction is based on the pre-period consumption model and post-period weather data; it represents the expected load shape for these households in absence of PG&E QM Program participation. The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. The errors of the hourly estimates are smallest in the early hours of the morning and are widest during the peak hours.

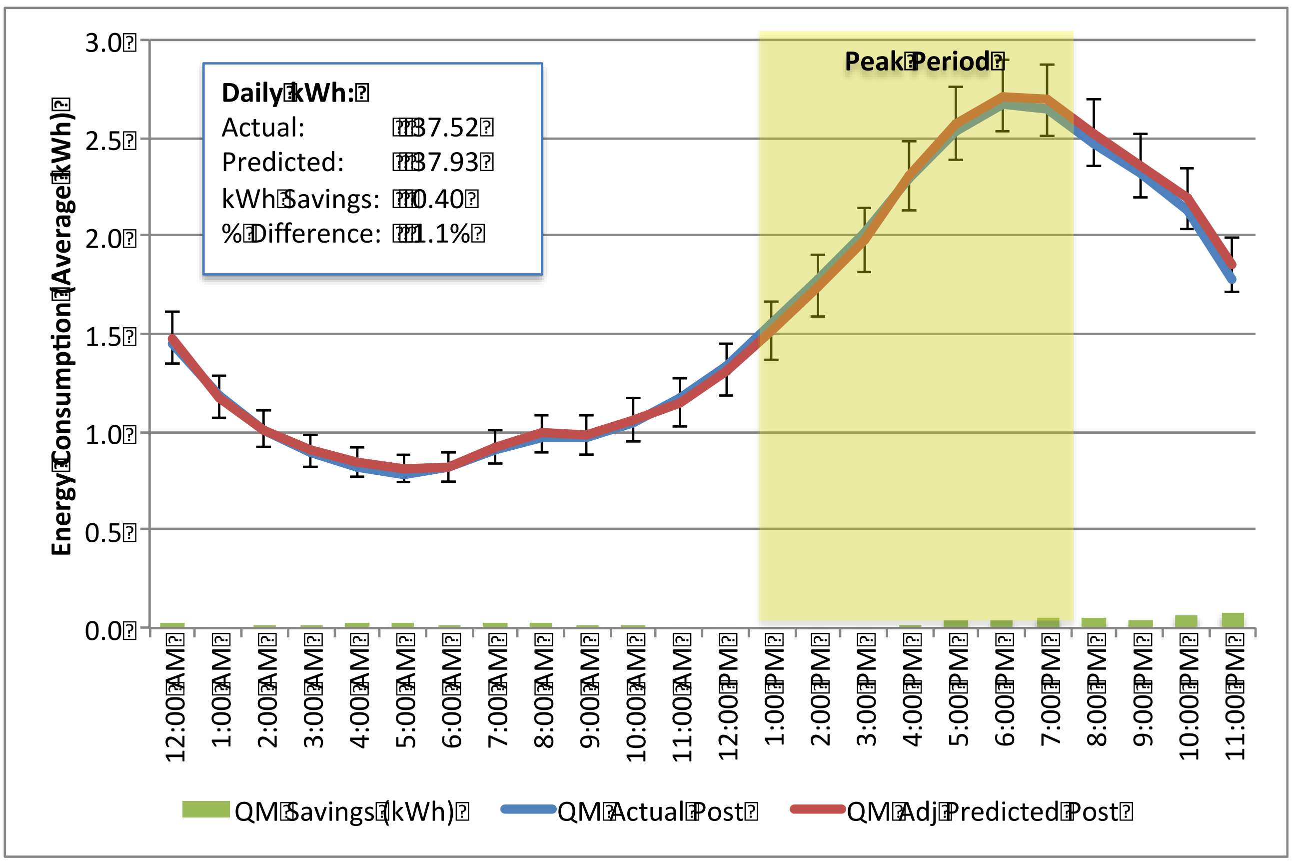
Figure 26: PG&E QM Summer Predictions of Post-Period with Error Bars



[36] The same adjustment methodology was used to correct for model bias using the difference between the model and the holdout sample as was demonstrated with the QI summer weekday model in Table 7.

Figure 27, below compares the post-period predicted load shape (red line) with the actual post-period load shape across all households (blue). Whenever the actual post-period load shape falls below the predicted post-period load shape, this indicates that savings were realized during that hour (green bars). After adjusting for error in the model using the same method described for the QI model, the modeling approach finds approximately 1 percent savings during summer months attributable to PG&E QM Program. Unlike SCE's QI Program, homes that participated in QM did not install new equipment, so all observed savings could theoretically be attributed to the PG&E QM Program. Note also that this approach finds the majority of savings is realized during the later part of the day including during the peak hour periods between 1:00 p.m. and 7:00 p.m.,[37] highlighted in yellow.

Figure 27: PG&E QM Program Summer Predictions versus Actual, 2014 Post-Period



[37] We use the residential peak period of 1:00 p.m. to 7:00 p.m., as defined for PG&E's residential Time-Of-Use rate plan (E-6). http://www.pge.com/nots/rates/tariffs/ResTOUCurrent.xls

Figure 28 shows the summer weekday hourly kWh savings estimates from the previous figure with bars depicting 95 percent confidence intervals around each estimate. None of the increases or decreases in hourly energy consumption are statistically significant, however, when aggregated to the daily level the savings are statistically significant at the 95 % confidence level.

Figure 28: PG&E QM Summer Hourly Savings Estimates with Error Bars

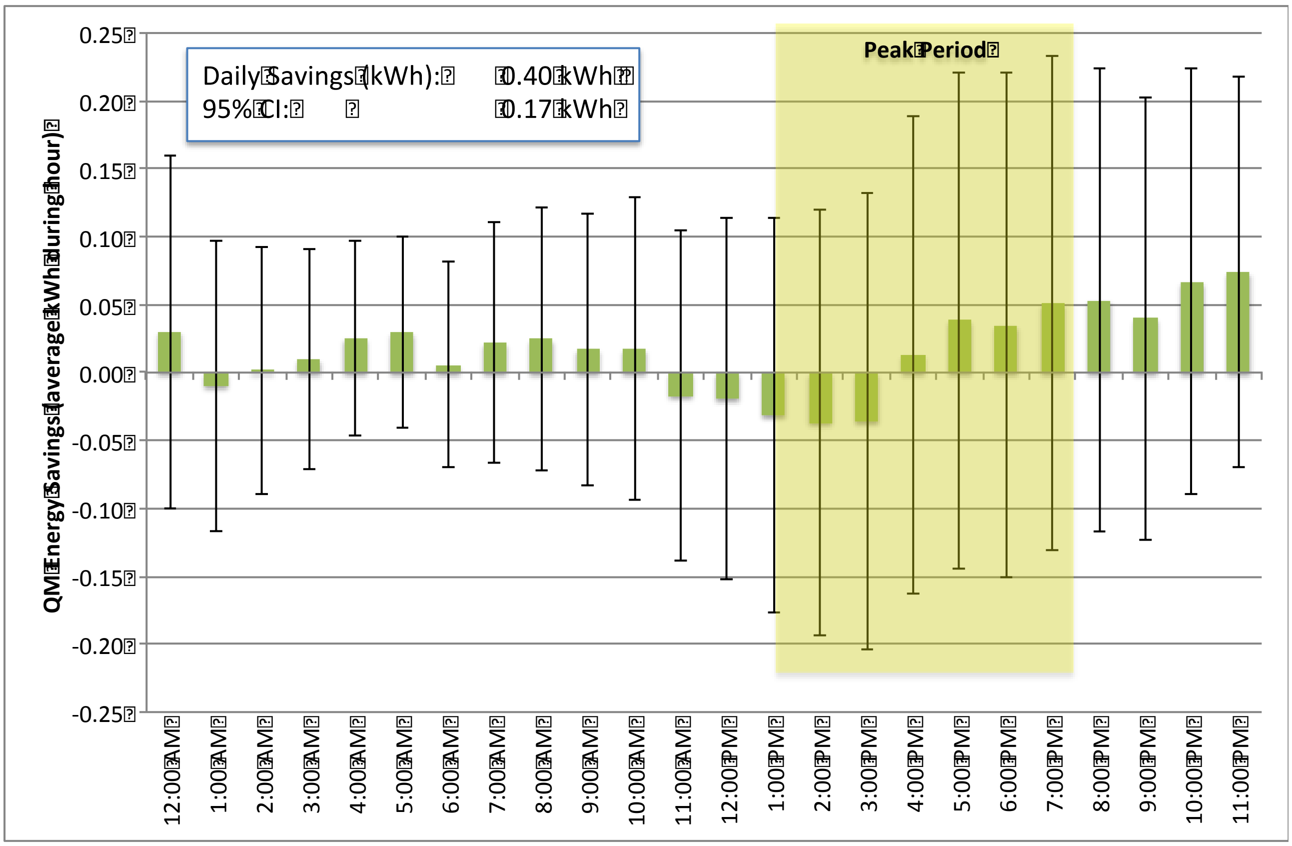


Table 16 shows the average daily savings estimate for each group and bin in the PG&E QM Summer weekday model. Similar to previous tables, the columns show households grouped by their weather normalized energy usage in the pre-period for each home (highest users on the right) and the rows show days grouped by the temperature via cooling degree-days (hottest days on the bottom). Each cell shows the estimated program savings (kWh per day) for a specific home-day bin. The cells are color-coded with the highest kWh savings in dark blue and the lowest kWh savings in dark red; colorless cells fall in the middle of this spectrum. Within each household group, there are home-days from a wide range of temperatures, each with their own savings estimate. Similarly, each group of days with similar temperatures (i.e., CDD) includes home-days from a range of households (i.e., high, mid, and low users), which experience a wide range of daily kWh savings. For the PG&E QM summer model, we see a similar trend to the SCE QI model with savings trends upward as temperature and household energy increase, although some specific bins deviate from this trend, for example home-group 14[38].

Table 16: Program Savings (kWh per day) by Bin

		Household Groups																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Cooling Degree-Day Groups	1	0.0	0.0	11.4	0.0	-1.9	0.0	1.7	6.0	7.4	0.0	0.0	0.0	0.0	11.0	0.0	-12.3	0.0	0.0	0.0	0.0	7.6
	2	-5.6	-1.6	-0.7	5.2	-3.7	1.4	-3.9	-1.0	3.1	2.1	3.7	0.0	0.9	5.0	1.9	8.4	-2.4	0.0	0.0	-0.2	5.8
	3	-0.3	-2.3	0.3	1.4	-0.4	4.8	-1.2	8.7	8.3	2.1	-4.1	-3.9	-4.6	7.2	1.1	1.6	-3.8	17.8	-20.1	4.2	1.0
	4	-3.0	-0.8	-1.3	2.2	2.2	3.6	1.5	4.2	1.1	-0.5	1.4	8.6	-2.0	4.6	3.1	4.6	11.7	3.8	12.2	4.0	1.8
	5	-0.8	-0.1	-2.1	3.3	1.6	3.7	0.5	-1.2	-1.9	-0.9	1.9	2.7	-2.6	1.7	-7.3	3.9	-6.0	-2.9	-26.7	-17.2	-1.0
	6	-4.3	-1.5	-1.8	0.5	0.6	1.8	-0.5	-2.5	0.5	-1.7	0.4	1.8	-3.5	4.5	-0.8	8.9	2.1	0.0	-0.4	1.4	-0.1
	7	-4.3	-1.4	-2.4	-1.8	-1.5	2.1	1.4	0.2	2.8	1.6	0.7	1.5	2.0	1.1	0.5	13.6	8.6	0.0	4.6	0.0	1.3
	8	-0.9	-1.1	-1.7	0.5	0.4	1.4	3.8	-2.7	1.4	0.2	-1.0	-0.1	-1.4	1.1	1.1	9.6	2.8	-2.1	-5.7	-13.8	-0.1
	9	-3.5	-3.5	-2.7	-1.1	0.5	0.8	1.3	-1.9	1.1	0.3	2.0	-0.3	-0.2	5.8	-0.2	6.7	2.5	11.0	8.0	3.2	0.0
	10	-3.0	-4.0	-2.2	-0.3	-0.1	0.1	4.5	-0.6	3.6	2.0	1.2	-2.5	1.2	1.8	3.5	6.6	3.9	11.4	4.8	-10.0	0.5
	11	-2.4	-3.6	-2.6	-0.7	-1.5	2.2	0.3	-3.2	2.5	1.1	2.4	-1.9	-0.1	-0.3	3.4	6.7	6.5	19.8	2.0	7.2	0.3
	12	-2.8	-3.4	-0.8	-1.9	-3.7	1.8	2.0	-2.7	-0.1	0.9	-3.8	0.5	-1.9	-0.4	4.3	8.4	3.4	-0.3	-0.7	-7.7	-0.9
	13	-3.6	-3.7	-4.4	-1.7	-2.9	0.5	2.0	-2.6	3.3	0.7	0.7	2.5	0.9	-2.1	5.0	5.0	7.6	4.8	3.4	2.7	-0.2
	14	-1.8	-1.8	-0.5	-0.4	-2.2	2.3	0.7	2.2	5.6	0.1	4.6	1.4	-1.5	-0.7	7.1	2.6	6.6	1.3	4.4	1.9	0.7
	15	-3.0	-1.7	-3.1	-0.9	-1.3	-0.5	0.9	-1.3	2.2	1.0	0.1	3.7	-1.9	-4.1	3.2	3.2	3.5	0.2	4.2	-4.4	-0.8
	16	-2.0	-3.3	-1.2	-0.3	-0.6	0.8	0.7	-1.8	0.1	1.5	3.4	2.2	-0.3	-2.0	4.4	2.4	3.4	2.4	3.7	-1.9	-0.5
	17	-1.8	-4.0	-4.6	0.4	-2.6	-0.4	0.2	-3.4	2.6	1.6	5.3	5.5	4.3	-1.0	5.6	3.7	6.2	5.0	3.8	6.1	0.1
	18	-2.2	-2.8	-3.3	0.2	-1.9	-2.9	-0.2	-3.7	2.1	0.6	5.7	-0.1	1.4	-0.2	3.4	1.3	5.8	4.4	3.1	4.6	-0.6
	19	-0.8	-2.6	-2.7	-1.6	-1.3	1.8	-2.2	-4.1	3.0	1.4	5.5	-0.2	2.8	2.0	7.6	2.8	4.1	3.2	7.2	12.1	0.4
	20	-0.7	-2.3	-5.1	0.1	-3.5	-1.1	-2.1	1.4	5.4	2.6	1.4	2.9	1.5	-4.7	8.8	4.1	2.4	6.3	6.0	11.6	0.3
	21	-1.0	-0.7	-2.0	-0.2	-0.9	0.9	-1.3	-0.6	3.1	3.2	4.1	-1.6	5.1	-0.2	11.3	0.5	2.7	9.5	6.1	20.8	1.4
	22	-0.6	-0.2	-4.1	1.5	-0.9	1.3	-3.2	1.2	0.7	4.0	4.4	6.0	1.3	3.9	6.7	-5.2	9.7	3.7	7.8	18.7	1.2
	23	-6.2	-4.9	-6.4	-3.9	-3.5	0.5	-1.3	0.1	-6.2	0.8	0.7	-3.5	1.0	-0.5	-1.3	0.3	3.7	1.9	14.6	11.2	-1.9
	24	-2.9	-2.5	-1.6	-4.6	-3.4	5.4	2.3	0.4	-3.7	4.4	1.8	-3.4	4.3	1.6	2.1	1.8	11.8	7.8	0.8	27.5	1.1
	25	0.0	-0.4	-0.8	0.5	-0.8	4.5	2.4	2.3	-0.6	5.1	-3.5	1.9	5.8	12.7	9.3	9.0	10.3	9.6	15.0	28.7	3.6
Total		-1.9	-2.4	-2.6	-0.5	-1.5	0.7	0.5	-1.2	1.9	1.5	1.5	1.1	0.7	0.9	3.9	5.1	5.4	3.3	4.8	9.2	0.4

Annual PG&E QM Program Model

In addition to the summer peak day model, we also developed an annual QM model that could be used to estimate yearly program impacts. For the annual model, the sample includes all homes that participated in the PG&E QM program between December 2012 and May 14 that had non-zero *ex ante* savings listed in the program documentation (i.e. tracking data). Unlike the summer weekday model, the annual model uses all months and day types (i.e., weekdays and weekends) with a more complex binning procedure. The resulting dataset includes 1,216 homes dispersed across four different climate zones.

Since this model includes all seasons and day types, we binned the home-days to four-dimensional bins. Specifically, we used 20 home groups, 9 CDD groups, 9 HDD groups, and 2 day type groups resulting in 3,240 possible home-day bins. We assigned each day to a CDD group and an HDD group that included a range of three degree-days each, up to a maximum of 26. Using multiple degree-days per group and setting a maximum limits the total number of bins and thereby reduces processing time.

---

[38] As with the overall hourly savings in Figure 28, many of these average daily savings values will not be statistically significant (particularly some bin-level savings with very few home-days). Thus, individual bins with extreme values (e.g. household group 19 with CDD 5) should not be a cause for alarm.

Table 17: Summary of PG&E QM Program Annual Binning

Group	Description	Number of Groups
Homes	Usage – weather normalized annual energy usage grouped by percentile, with 1/20th of the total assigned to each group in order of smallest to largest	20
Days	CDD – average of CDH rounded up to a whole number, assigned three CDDs per group from 0-26 with all days higher than CDD 26 put into the last group	9
	HDD – average of HDH rounded up to a whole number, assigned three HDDs per group from 0-26 with all days higher than HDD 26 put into the last group	9
	Day Type – flag for weekends that separates them from weekdays	2
Total[39]	Home-Day Bins	3,240

Table 18 and Table 19 present the count of home-days in the post-period for the PG&E QM annual model on weekdays and weekends respectively. As with previous tables, these tables show the actual distribution of participant households and the weather they experienced in the post period. In the annual model day-types are binned by combinations of both CDD and HDD, and the table is labeled with the upper limit of each day-type (e.g. the day type bin CDD 2 includes all days with CDD between 0 and 2). We see more moderate days with CDD or HDD ranging from 6 to 17 than especially high or low temperature days within each of the household groups. Again, there are more home days in the home groups at the lower end of the usage spectrum, because each home group represents about 5 percent of total baseline electricity usage for all the homes in our sample.

[39] Some bins have zero home-days. This is expected as certain combinations of groups are not present in the data, In particular combinations of HDD and CDD groups because there were no days with extreme temperature ranges. For example, the data did not include any days with a temperature range from 40˚F-90˚F, so there are no home-days assigned to both CDD 25 and HDD 25.. Our final pre-period model includes 990 bins.

Table 18: PG&E QM Annual Model: Number of Home-Days in Each Bin (Annual - Weekdays)

	Day Type	Weekdays																				
	Household Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
CDD	HDD																					
2	2	102	94	94	71	59	66	76	62	67	55	55	54	49	47	49	44	50	39	25	27	1,185
	5	1,943	1,668	1,423	1,414	1,297	1,290	1,194	1,193	1,111	1,004	1,086	998	1,009	940	940	847	881	805	745	644	22,432
	8	4,750	4,083	3,989	3,596	3,390	3,292	3,281	2,908	2,939	2,669	2,545	2,494	2,488	2,347	2,311	2,104	2,154	1,945	1,672	1,216	56,173
	11	5,814	5,066	5,042	4,471	4,228	4,096	4,113	3,671	3,679	3,307	3,153	3,063	3,092	2,924	2,954	2,613	2,732	2,390	2,021	1,433	69,862
	14	5,195	4,442	4,283	3,849	3,614	3,561	3,483	3,164	3,140	2,842	2,896	2,645	2,658	2,529	2,488	2,210	2,403	2,046	1,829	1,273	60,550
	17	4,866	4,176	4,263	3,776	3,700	3,529	3,406	3,149	3,029	2,903	2,672	2,536	2,635	2,529	2,445	2,261	2,261	1,950	1,781	1,143	59,010
	20	3,378	2,963	3,141	2,735	2,704	2,610	2,590	2,223	2,345	2,053	1,919	1,836	1,826	1,835	1,798	1,581	1,627	1,402	1,197	728	42,491
	23	1,983	1,673	1,726	1,539	1,482	1,430	1,417	1,214	1,269	1,118	1,065	1,036	1,022	1,002	975	858	912	784	647	412	23,564
	26	1,787	1,535	1,709	1,454	1,433	1,347	1,388	1,138	1,221	1,096	959	940	924	957	900	793	834	704	572	321	22,012
5	2	839	696	675	646	597	541	564	476	502	448	422	432	426	388	392	358	377	349	285	279	9,692
	5	4,243	3,747	3,627	3,315	3,102	2,964	2,984	2,641	2,670	2,460	2,312	2,315	2,285	2,085	2,119	1,950	1,981	1,790	1,529	1,118	51,237
	8	2,567	2,368	2,528	2,138	2,109	2,006	2,118	1,722	1,899	1,706	1,455	1,499	1,449	1,466	1,432	1,277	1,264	1,167	931	597	33,698
	11	285	289	310	275	295	237	275	198	232	170	176	163	162	164	171	129	160	126	100	58	3,975
	14	12	14	6	11	6	5	6	8	9	3	10	7	6	4	8	4	8	7	3	5	142
	17	6	3	3		3	3															18
8	2	3,071	2,587	2,598	2,272	2,183	2,169	2,117	1,815	1,882	1,741	1,677	1,575	1,567	1,485	1,476	1,348	1,398	1,218	1,017	713	35,909
	5	3,064	2,708	2,957	2,584	2,580	2,461	2,498	2,040	2,248	2,081	1,726	1,799	1,765	1,764	1,714	1,548	1,467	1,431	1,164	709	40,308
	8	123	127	148	116	143	115	133	94	101	85	82	65	77	82	74	62	65	49	47	22	1,810
	11	6	3	3		3	3															18
11	2	4,542	3,930	4,222	3,429	3,343	3,434	3,415	2,801	2,950	2,760	2,517	2,416	2,314	2,401	2,269	2,059	2,126	1,829	1,585	985	55,327
	5	1,045	952	1,344	995	1,051	1,067	1,127	760	1,017	831	629	686	616	723	678	604	539	503	335	157	15,659
	8	9	4	4	5	9	1	6							3	5		3				59
14	2	4,860	4,249	4,219	3,675	3,572	3,410	3,378	3,029	2,841	2,768	2,680	2,553	2,451	2,443	2,313	2,135	2,293	1,875	1,824	1,100	57,668
	5	131	111	153	130	144	176	166	80	167	92	88	80	75	103	98	64	65	76	41	20	2,060
17	2	4,781	4,141	3,658	3,479	3,292	3,032	2,903	2,972	2,510	2,549	2,657	2,479	2,365	2,277	2,220	1,976	2,201	1,852	1,916	1,229	54,489
20	2	3,882	3,405	2,773	2,796	2,582	2,305	2,196	2,423	1,927	1,993	2,125	2,029	1,975	1,776	1,801	1,586	1,787	1,524	1,631	1,101	43,617
23	2	2,528	2,208	1,715	1,797	1,649	1,461	1,368	1,578	1,197	1,247	1,379	1,286	1,267	1,141	1,153	999	1,160	980	1,083	722	27,918
26	2	2,219	1,993	1,515	1,605	1,451	1,275	1,194	1,416	1,064	1,120	1,215	1,180	1,153	1,007	1,064	905	1,023	886	989	684	24,958
Total		68,031	59,235	58,128	52,173	50,021	47,896	47,396	42,775	42,016	39,101	37,500	36,166	35,656	34,422	33,847	30,315	31,771	27,727	24,969	16,696	815,841

Table 19: PG&E QM Annual Model: Number of Home-Days in Each Bin (Annual - Weekends)

	Day Type	Weekends																				
	Household Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
CDD	HDD																					
2	2									3						3						6
	5	477	397	344	327	303	303	282	268	260	241	265	238	235	219	210	200	203	183	169	163	5,287
	8	▯,276	▯,073	▯,118	988	920	916	945	741	869	701	692	666	651	631	627	547	578	518	385	300	15,142
	11	2,359	2,070	▯,925	▯,751	▯,618	▯,575	▯,596	▯,408	▯,411	1,287	1,247	1,215	1,217	1,107	1,127	998	1,094	952	837	606	27,400
	14	2,041	▯,780	▯,713	▯,558	▯,497	▯,464	▯,389	▯,345	▯,260	1,187	1,175	1,070	1,098	1,065	1,047	947	953	835	767	519	24,710
	17	▯,546	▯,277	▯,317	▯,233	▯,152	▯,127	▯,109	939	▯,014	786	852	781	815	762	759	647	727	606	502	329	18,280
	20	▯,804	▯,545	▯,568	▯,361	▯,371	▯,312	▯,222	▯,189	▯,103	1,117	1,000	972	961	972	919	859	830	726	686	435	21,952
	23	784	707	880	660	663	629	691	524	596	529	433	433	427	432	418	380	399	322	232	135	10,274
	26	531	448	495	431	414	390	417	323	366	319	268	271	266	274	256	230	231	204	158	84	6,376
5	2	359	321	313	274	255	235	248	220	220	200	194	197	188	168	164	159	171	143	130	114	4,273
	5	▯,938	▯,706	▯,585	▯,499	▯,418	▯,324	▯,294	▯,238	▯,154	1,121	1,048	1,061	1,056	984	973	914	895	821	741	538	23,308
	8	▯,309	▯,183	▯,232	▯,052	▯,040	980	▯,029	846	902	837	715	734	718	702	710	621	630	579	472	319	16,610
	11	212	200	205	173	181	189	181	159	171	133	139	112	115	132	127	102	105	96	82	42	2,856
	14	3						8														11
8	2	▯,027	876	756	747	680	682	615	624	583	541	561	534	542	473	498	435	476	426	388	298	11,762
	5	▯,221	▯,078	▯,093	▯,047	994	897	955	777	844	772	650	705	703	650	652	586	570	573	473	304	15,544
	8	153	162	200	139	171	165	169	130	145	136	107	99	94	115	107	94	85	76	59	29	2,435
	11	17	6	7	6	6	12	8		10		6										78
11	2	▯,604	▯,380	▯,665	▯,230	▯,262	▯,354	▯,321	▯,008	▯,141	1,066	907	875	812	909	854	781	775	639	516	295	20,394
	5	367	323	443	351	357	369	385	251	350	268	207	232	223	243	230	202	183	183	120	62	5,349
	8					7																7
14	2	▯,702	▯,450	▯,563	▯,270	▯,270	▯,210	▯,212	▯,059	▯,023	993	965	878	834	878	796	741	811	636	573	321	20,185
	5	47	34	43	51	53	61	62		72		35	25	25	33	35		17	30	13		636
17	2	▯,875	▯,653	▯,491	▯,379	▯,302	▯,227	▯,203	▯,182	▯,041	1,031	1,063	984	945	905	891	798	877	740	741	484	21,812
20	2	▯,772	▯,556	▯,321	▯,289	▯,196	▯,084	▯,060	▯,113	904	907	981	915	897	825	812	730	812	671	718	468	20,031
23	2	▯,432	▯,254	980	▯,030	941	813	774	875	665	727	746	749	725	653	657	564	648	574	642	434	15,883
26	2	▯,006	894	680	725	664	588	540	641	481	516	552	530	520	463	481	412	465	409	451	304	11,322
Total		26,862	23,373	22,937	20,571	19,735	18,906	18,715	16,860	16,588	15,415	14,808	14,276	14,067	13,595	13,353	11,947	12,535	10,942	9,855	6,583	321,923

Since all of the households participated in the PG&E QM program during 2012-2013, we used the full year of 2014 as the post-period for all savings estimations. This ensures that the annual savings estimate is based on all four seasons and a wide range of daily temperatures.

To test the reliability of the annual QM model, we randomly selected 30 percent of the homes as a holdout sample and modeled the remaining 70 percent of the homes. Figure 29 shows the comparison of the predicted load shape from the model (yellow) with the actual load shape for the 30 percent holdout group (purple). The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. As with the previous models, the annual QM model is able to produce very accurate predictions of energy use for the holdout sample, with a difference between estimated and actual usage of less than 1 percent over 24 hours.

Figure 29: PG&E QM Program Annual Predictions versus Actual of Holdout Homes, Pre-Period

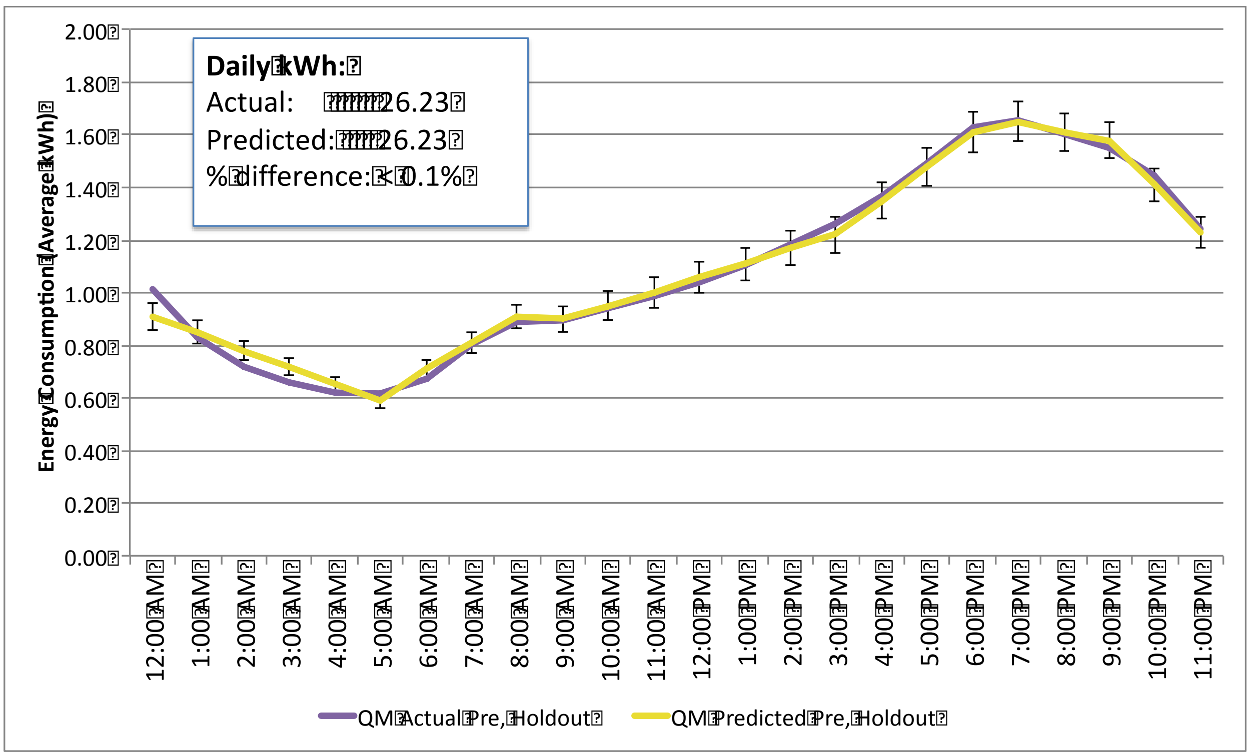


Figure 30 presents the comparison of the pre-installation predicted load shape from the model with the actual pre-installation load shape for all bins combined, with all 1,166 households in four different climate zones (i.e., no holdout group). The error of each hourly consumption prediction is depicted with a 95 percent confidence interval shown as bars around each estimate. The modeled pre-installation period load shape (yellow) aligns very closely with actual pre-installation load shape (purple), with a difference of about 0.1 percent over 24 hours.

Figure 30: PG&E QM Annual Predictions versus Actual of Full Sample, Pre-Period

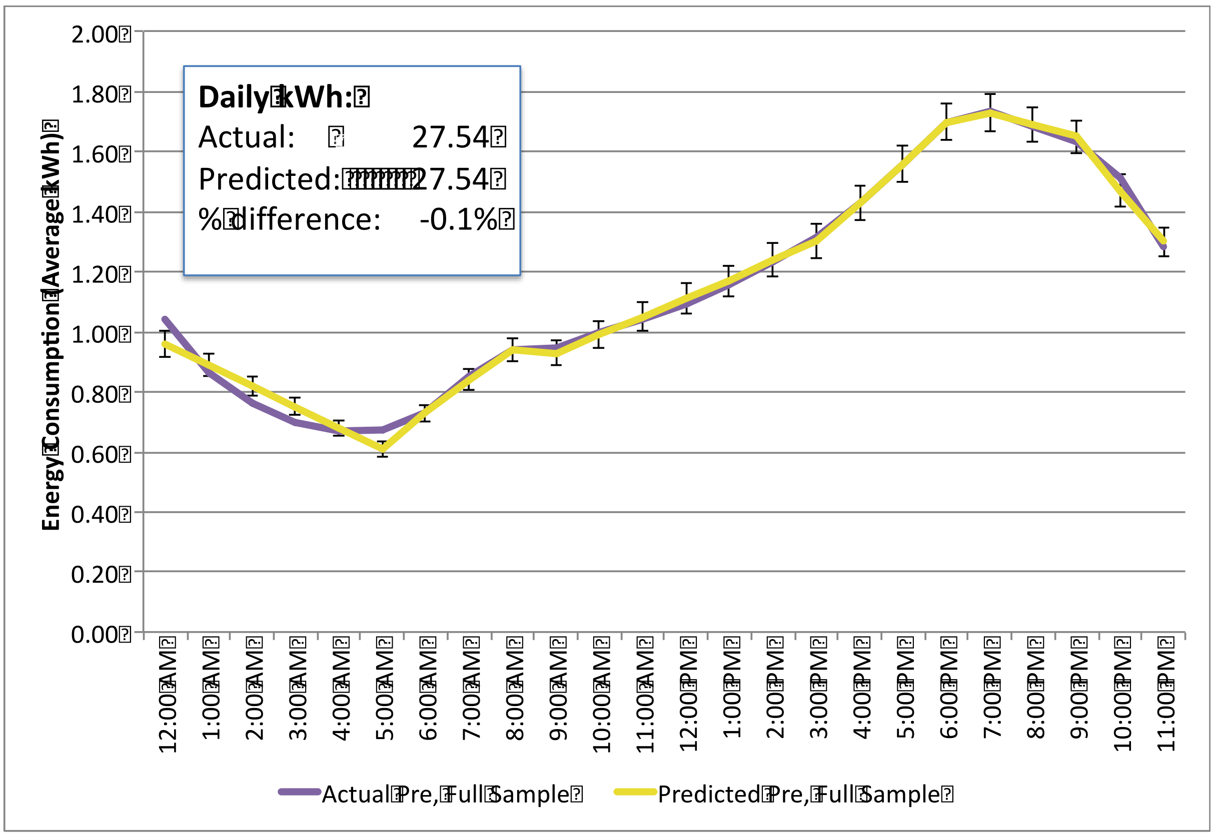
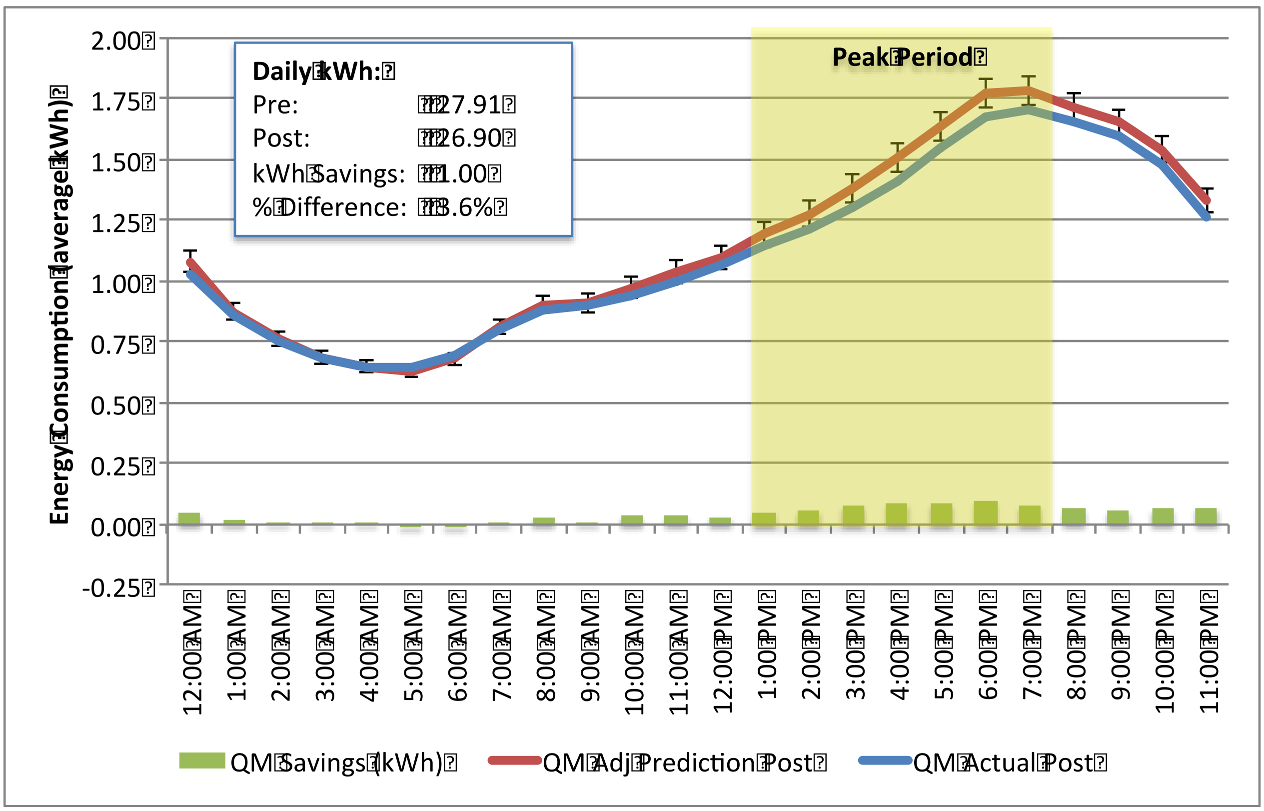


Figure 31 below compares the pre-period predicted load shape (red) with the post-period actual load shape (blue) averaged across all households. Whenever the post-period load shape falls below the pre-period load shape, this indicates that savings were realized during that hour (green bars). After adjusting for the error in the model, based on the sample of homes used, the modeling approach finds approximately 3.6 percent annual savings attributable to the QM program. As before, the largest impacts are realized during the later part of the day, including during the peak period between 1:00 p.m. and 7:00 p.m.,[40] highlighted in yellow. The error of the hourly consumption predictions is shown using a 95 percent confidence interval depicted with bars around each estimate. The error bands are tightest in the morning from midnight to 7:00 a.m. and widest during the peak hours from 2:00 p.m. to 11:00 p.m.

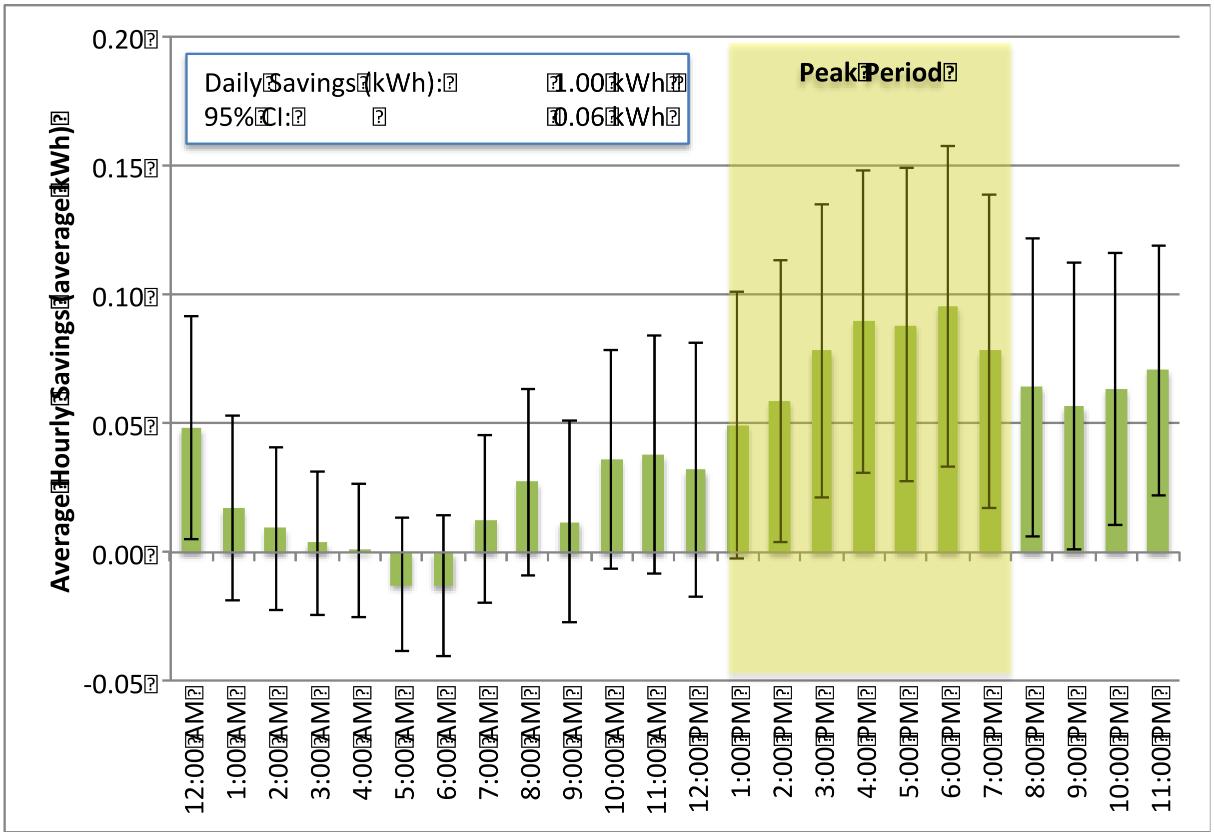
Figure 31: PG&E QM Program Overall Annual Post-Period Model, All Months and Day Types



[40] We use the residential peak period of 1:00 p.m. to 7:00 p.m., as defined for PG&E's residential Time-Of-Use rate plan (E-6). http://www.pge.com/nots/rates/tariffs/ResTOUCurrent.xls

Figure 32 shows the annual hourly kWh savings estimates from the previous figure with bars depicting 95 percent confidence intervals around each estimate. We found statistically significant hourly savings during the peak hours from 14 through 23. As with SCE, none of the increases in usage (i.e., negative savings) were significant.

Figure 32: PG&E QM Program Annual Hourly Savings Estimates with Error Bars



The PG&E QM annual model includes a day-type binning component allowing us to evaluate energy savings for weekdays versus weekends. Figure 33 and Figure 34 compare the pre-period predicted load shape (red) with the post-period predicted load shape (blue), averaged across all households for weekdays and weekends respectively. The modeling approach finds slightly higher savings on weekdays (3.7%) versus weekends (3.5%) however the differences in hourly savings are not statistically significant.

Figure 33: PG&E QM Overall Annual Post-Period Model, All Months; Weekdays

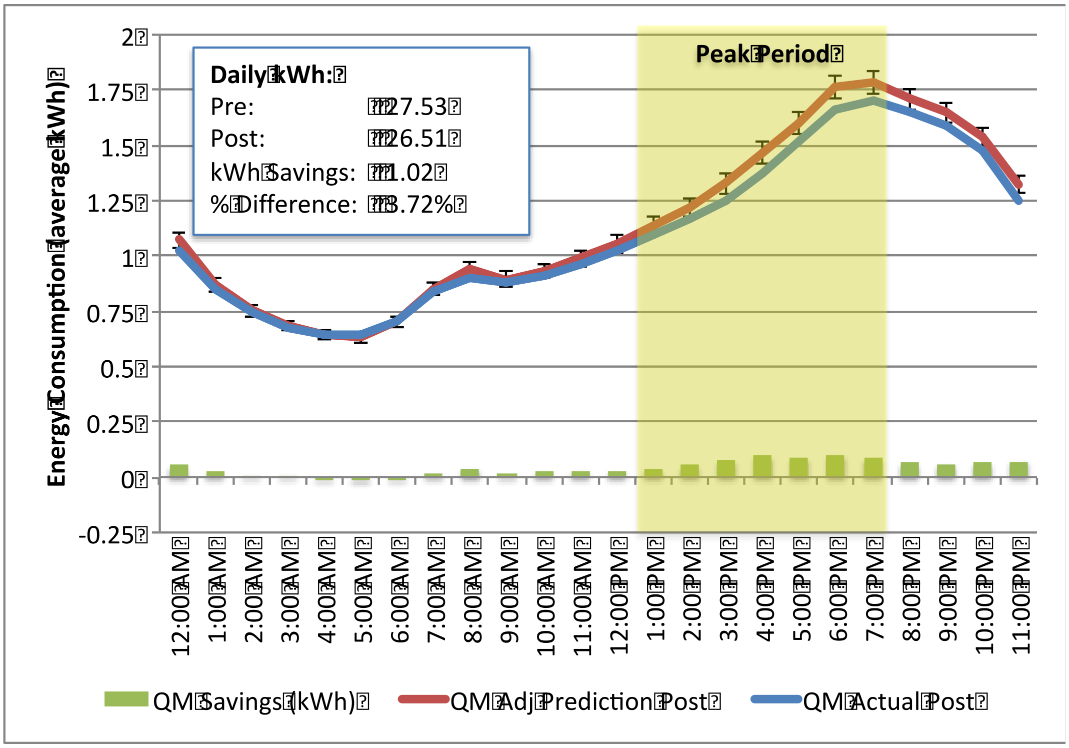


Figure 34: PG&E QM Overall Annual Post-Period Model, All Months; Weekends

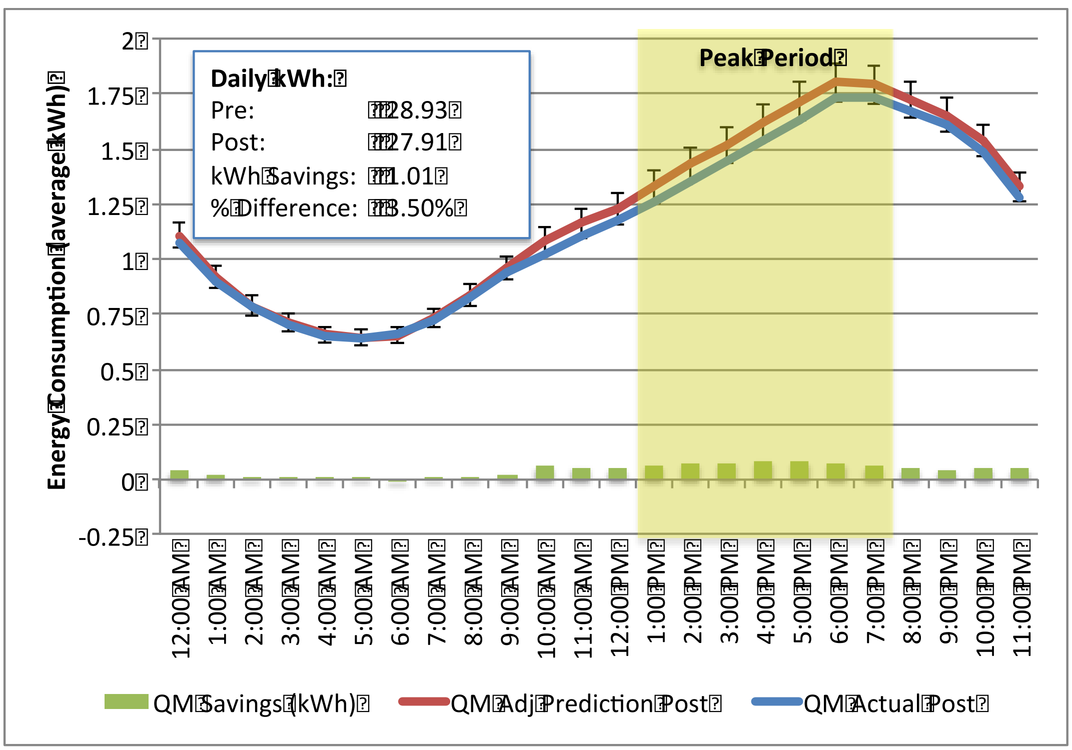


Table 20 and Table 21 provide the average daily savings estimate for each group and bin in the PG&E QM annual model on weekdays and weekends respectively. The columns show households grouped by their weather normalized energy usage in the pre-period for each home (highest users on the right) and the rows show days grouped by the temperature via cooling degree-days (hottest days on the bottom). Each cell shows the estimated program savings (kWh per day) for a specific home-day bin. We automatically color-coded the cells with the highest kWh savings in dark blue and the lowest kWh savings in dark red; colorless cells fall in the middle of this spectrum. In general in the annual model we see increased savings as temperatures increase and weather normalized consumption increase with some deviation from this trend in specific bins.

Table 20: PG&E QM Annual Model: Program Savings (kWh per day) by Bin (Annual – Weekday)

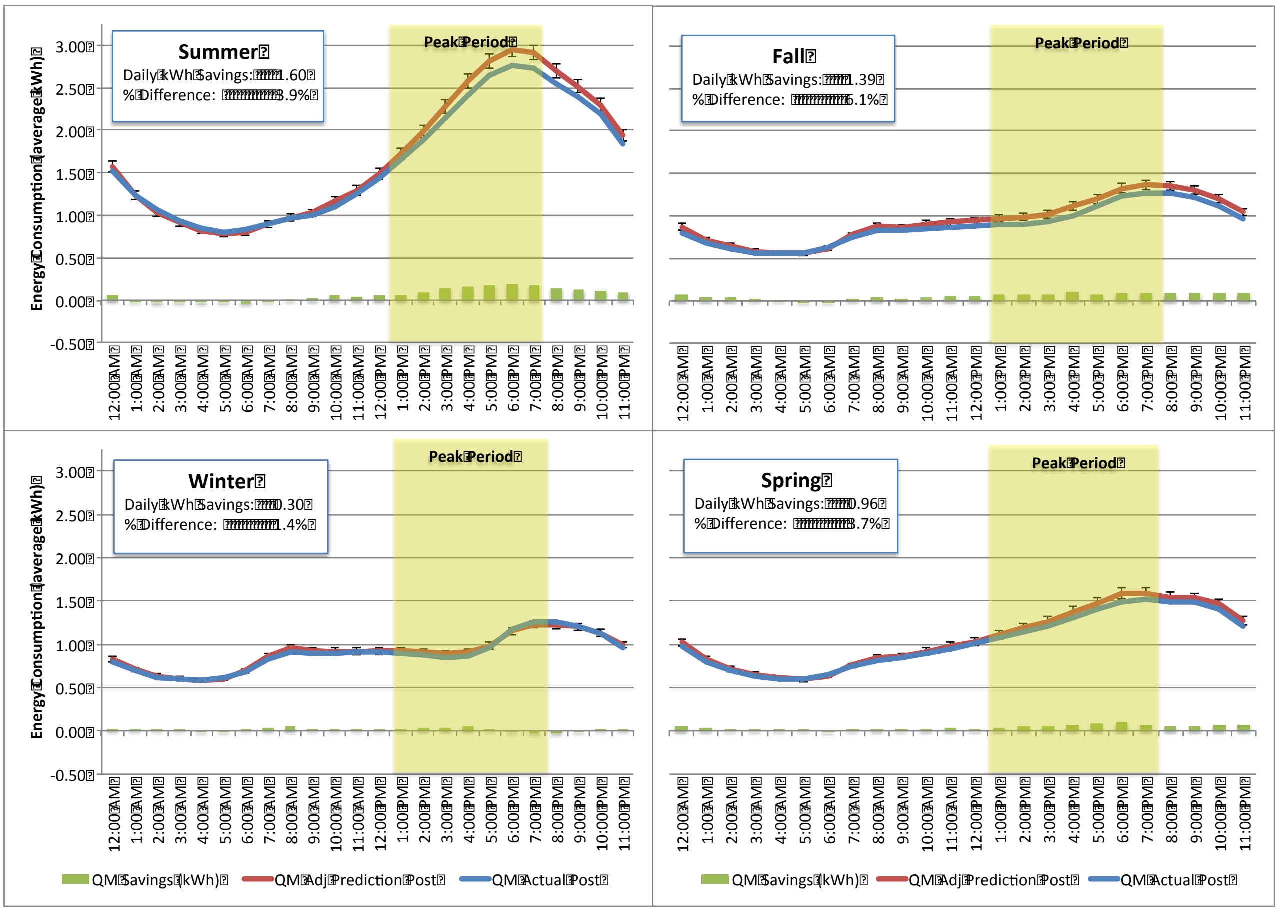
Day Type		Weekdays																			
Household Group		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CDD	HDD																				
2	2	0.9	0.9	-0.4	1.8	-1.1	-1.3	-2.2	-0.8	3.9	1.5	-2.8	-0.2	-0.3	-0.2	-7.0	-4.3	-7.9	-1.4	6.0	-12.5
	5	0.3	0.8	-0.1	0.4	-0.7	0.2	-0.9	0.0	0.4	0.2	-0.3	0.4	0.0	-0.2	-1.9	-4.2	-2.1	-1.7	0.7	-5.6
	8	1.1	1.1	0.7	0.4	-0.7	0.0	-1.2	-0.2	-0.2	-0.6	-0.6	-0.9	0.4	-0.4	-0.9	-3.6	-1.4	-1.8	-1.7	-7.7
	11	1.0	1.1	0.6	0.4	-0.7	-0.1	-0.8	-0.5	-0.5	-0.3	-0.3	-1.2	0.2	-0.2	-0.6	-2.7	-1.4	-1.5	-0.4	-4.8
	14	1.2	0.8	0.8	0.1	-0.1	-0.3	-1.5	-1.2	-0.9	-0.6	-0.3	-1.1	0.0	-1.3	-0.4	-3.6	-0.7	-1.2	0.3	-6.0
	17	1.5	2.3	1.5	1.1	1.1	0.9	0.0	-0.9	0.1	1.7	0.2	-0.2	1.1	-0.8	2.5	-1.9	0.0	1.3	1.1	-2.5
	20	1.1	2.0	3.1	0.8	0.2	-0.1	-0.3	-1.5	0.3	0.7	0.2	0.5	2.1	-0.4	1.7	-3.8	0.1	-2.2	3.8	-14.6
	23	1.7	1.1	1.2	1.0	0.6	0.0	-1.3	-1.6	-0.9	1.6	-0.2	-0.9	1.5	-0.6	1.5	-3.7	1.0	-0.7	-3.7	-11.4
	26	1.2	0.9	1.1	-0.1	-1.5	0.8	-0.6	-2.7	-2.0	1.7	-0.4	-0.3	0.1	-2.3	0.8	-6.3	2.5	3.9	3.5	-6.2
5	2	1.0	1.4	1.0	1.3	0.4	0.8	0.3	0.4	1.2	0.4	1.1	1.0	1.5	1.6	-3.0	-0.9	-1.1	1.3	-4.0	-2.6
	5	0.8	1.3	1.0	0.7	0.1	0.4	-0.4	0.0	0.7	0.1	1.0	0.1	1.9	0.6	-1.3	-0.7	-1.1	1.7	0.2	-1.6
	8	0.8	1.6	1.4	0.9	0.2	0.8	-0.5	0.5	0.3	-0.4	0.3	-0.3	-0.1	0.5	0.1	-0.9	-1.2	-1.1	2.4	-6.0
	11	0.1	0.0	1.1	0.0	0.6	0.6	-1.0	0.3	-0.7	-0.7	-1.4	-1.6	-1.5	-0.6	-0.2	-3.6	-1.9	-3.2	1.5	-13.4
	14	2.1	1.8	3.4	1.1	2.7	2.8	0.4	1.2	-0.2	0.7	0.9	0.7	1.5	2.4	1.4	-1.2	-2.0	2.0	4.1	2.4
	17	0.0	0.0	7.5	0.0	0.0	0.0	-0.5	0.0	0.0	0.0	-4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	2	1.0	1.1	1.5	1.5	0.7	0.9	-0.1	-0.1	0.7	0.6	2.1	2.0	2.6	2.1	-1.6	0.8	0.0	4.5	-1.8	3.5
	5	0.8	1.5	1.2	0.1	-1.0	0.6	-0.8	0.1	1.3	-0.3	0.7	-0.4	1.2	0.6	-1.4	0.6	0.4	1.0	2.3	5.2
	8	0.6	0.5	1.5	-0.1	-0.6	-0.6	-1.4	-0.1	0.1	-0.2	-0.4	-1.0	-1.2	-0.6	-0.6	-2.9	-0.3	0.2	1.0	-4.3
	11	0.0	-0.9	0.4	-2.6	-1.6	-2.1	-4.9	-2.7	-5.9	-2.2	-2.5	-11.6	-8.3	0.6	-4.0	-10.1	-7.6	-9.7	1.1	-9.6
11	2	0.9	2.1	2.5	2.1	1.4	2.7	1.8	3.0	2.8	3.4	4.0	3.3	3.7	3.3	2.5	4.6	4.6	6.8	3.1	9.1
	5	1.0	1.5	1.5	1.5	0.7	1.5	0.5	0.7	1.4	2.1	1.5	0.3	1.6	2.9	0.9	3.8	3.7	1.1	5.0	8.8
	8	-1.0	-1.3	2.1	-0.4	-1.2	-1.3	-3.2	-0.6	1.1	-3.5	-6.2	-1.1	-7.3	-1.7	-6.2	-5.7	-3.7	0.5	0.0	-10.7
14	2	1.3	2.4	2.8	3.6	3.0	3.0	2.8	4.4	5.2	5.3	6.0	6.2	6.3	5.6	4.6	7.7	8.3	10.0	8.6	16.7
	5	1.0	0.6	0.8	-1.3	-2.1	-2.4	0.3	0.6	3.9	0.2	-2.1	0.8	1.5	-0.3	-4.7	4.3	3.7	0.5	5.8	2.9
17	2	1.8	3.8	3.4	4.9	4.9	4.5	4.1	6.1	7.2	7.7	7.7	8.0	7.6	7.8	7.5	10.4	11.7	12.8	14.0	24.4
	5	-13.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	2	3.2	5.8	4.2	7.4	7.6	7.4	6.5	9.5	11.8	10.9	11.0	10.1	11.3	11.8	13.9	14.6	14.8	17.0	20.7	34.1
23	2	4.7	7.1	6.4	8.8	8.1	7.0	7.0	8.9	10.3	11.7	16.4	8.6	11.2	12.9	10.7	17.6	18.0	19.3	24.0	31.6
26	2	7.3	5.4	5.9	14.5	1.9	14.6	14.6	3.5	8.2	10.2	17.8	7.7	14.1	5.4	2.8	14.7	18.1	33.6	12.7	-23.7

Table 21: PG&E QM Annual Model: Program Savings (kWh per day) by Bin (Annual – Weekend)

Day Type		Weekends																			
Household Group		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CDD	HDD																				
2	2	1.6	0.6	-1.5	0.2	-1.3	-0.5	-1.9	2.0	4.7	0.7	-1.0	2.7	0.9	-0.2	-2.0	-0.2	0.4	15.2	12.9	-12.6
	5	0.8	1.2	0.5	0.4	-0.8	0.5	-0.4	0.2	0.2	0.6	0.3	-0.8	0.2	-0.6	-0.9	-4.2	-2.4	1.6	5.0	-6.2
	8	1.2	1.8	0.6	0.9	-0.3	0.8	-0.4	0.0	0.1	0.3	0.2	-0.9	1.5	-0.1	-0.3	-3.5	-0.4	-0.9	0.4	-6.7
	11	1.2	0.9	0.4	-0.2	-0.2	0.8	-1.4	-0.5	-1.0	0.4	-0.1	-1.4	-0.3	-0.4	0.4	-2.4	-0.4	-2.8	0.2	-2.3
	14	1.6	0.8	0.4	0.5	-0.2	0.9	-1.0	-1.6	-0.6	0.2	-0.7	-1.7	0.4	-0.3	0.6	-2.6	0.5	-2.1	1.9	-1.3
	17	1.6	2.3	1.8	1.5	0.1	1.2	0.3	-0.3	-0.3	0.6	-0.2	-0.7	0.5	-1.1	2.3	-3.7	-0.8	0.9	1.8	-8.0
	20	1.2	2.4	1.5	2.0	0.0	0.8	-0.5	0.5	0.6	1.1	1.2	0.2	2.2	0.0	2.2	-4.1	0.2	0.4	8.1	-8.8
	23	2.5	2.5	4.0	1.7	0.3	2.5	1.0	-0.9	1.0	3.8	0.7	1.6	1.8	-2.9	2.0	-0.3	6.9	6.8	-5.1	-6.4
	26	2.3	3.4	3.0	3.1	2.9	4.5	1.8	1.5	1.7	4.4	0.2	1.8	4.2	1.5	1.7	-1.3	2.9	3.9	7.8	1.1
5	2	1.5	1.2	1.6	1.9	1.4	2.9	1.8	2.2	2.0	2.7	1.5	1.5	4.4	3.0	-0.6	0.5	2.1	1.8	1.8	0.4
	5	1.3	1.6	0.6	1.1	0.1	1.0	-0.5	0.3	-0.2	-0.2	0.5	-0.8	1.8	0.3	-2.7	-1.6	-1.0	0.1	-1.0	-2.9
	8	0.9	1.6	0.8	0.3	-0.8	1.1	-0.9	0.1	1.9	-0.3	-0.3	-1.0	0.9	0.9	0.4	-0.8	-1.8	-0.4	3.0	-6.7
	11	1.4	2.0	2.6	0.4	1.7	0.7	-1.1	-0.5	-1.4	-0.7	-0.7	-0.3	0.5	-0.7	0.4	-1.4	-1.9	-1.5	2.9	-11.5
	14	1.3	0.7	1.7	-0.5	1.4	1.1	-0.2	0.0	-0.3	-0.6	-0.8	-0.9	-1.1	0.9	1.3	-4.2	-2.5	0.8	3.9	9.9
	17	1.7	0.0	7.7	-0.4	0.0	0.0	0.0	0.0	0.2	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	2	1.2	1.9	2.7	2.9	1.4	2.9	1.9	2.0	2.9	2.3	3.2	3.3	4.0	3.2	2.6	2.1	2.7	4.7	3.4	4.1
	5	1.1	2.0	1.9	0.6	-0.5	1.6	0.1	-0.5	0.8	0.4	0.8	-1.1	1.0	-0.8	-1.1	0.3	0.5	1.1	2.9	-1.7
	8	0.5	0.0	0.8	-1.0	-0.3	0.7	-1.5	1.5	0.0	-0.3	-1.2	-2.8	-1.2	-0.7	-1.7	-0.8	2.6	1.9	-1.1	-9.5
	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	2	1.2	1.8	2.2	2.1	1.5	2.0	0.1	2.1	0.8	2.4	3.3	1.6	3.4	3.8	1.1	2.2	3.9	8.2	1.6	11.3
	5	0.4	0.7	2.8	1.3	-1.3	-1.7	-2.3	0.5	0.3	-0.7	-0.5	-1.8	0.8	3.0	-2.4	4.4	2.8	2.7	3.2	8.5
	8	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	2	1.3	2.7	3.4	3.4	2.8	3.8	4.3	4.9	2.9	5.9	7.2	4.2	4.4	4.9	6.7	7.9	8.5	8.8	8.0	19.0
	5	-0.8	-2.8	0.2	-3.5	-4.2	-2.5	1.9	0.2	4.3	2.8	-2.5	-0.6	1.7	-3.3	-5.5	5.9	10.3	-1.3	11.1	4.7
17	2	2.6	5.0	6.3	6.3	6.8	9.0	6.2	8.6	9.8	10.7	11.2	10.1	12.0	10.8	9.3	10.3	14.2	14.0	19.0	25.3
	5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	2	3.4	5.9	5.5	8.8	9.1	9.0	7.1	9.6	10.3	11.0	11.6	11.6	10.8	10.9	11.5	15.2	15.1	14.1	27.1	38.9
23	2	4.5	9.1	4.2	9.6	4.9	11.0	6.9	6.6	9.9	11.5	12.5	10.2	10.1	6.7	5.1	14.7	15.2	23.3	16.8	10.6
26	2	9.7	8.5	6.8	12.5	2.4	18.1	7.2	-1.0	9.6	10.9	20.1	5.7	13.0	12.7	3.3	18.3	17.1	34.4	16.3	-24.0

Figure 35 shows our annual model's predicted load shape (red) and what the households actually consumed (blue) after participating in the PG&E QM program, by season. Note that these are not separate models, but rather each season is calculated from the bin-level output from the single annual model. When looking at the savings in kWh, most of the savings occurred in the summer, which had an average daily savings of 1.6 kWh or 3.9 percent. However, when looking at the savings as a proportion of total energy use, most of the savings occurred in the fall, which had an average daily savings of 6.1 percent or 1.4 kWh.

Figure 35: PG&E QM Program Annual Model Results by Season



Note that the summer load shape from the PG&E QM program annual model is similar to the summer weekday model presented earlier, but the savings appear to be larger and occur earlier in the day. While we did not investigate this issue in depth, factors that may be contributing to these differences between the annual and summer only model include the following:

- The summer peak model only includes weekdays, while the annual model includes summer weekdays and weekends. These extra days result in additional savings during the summer.

- CDDs were assigned to 9 bins in the annual model rather than the 25 bins used in summer model. The random coefficients model is creating predictions based on the specific observations in each bin, if the binning procedure is changed then the mix of observations in a bin and the model's predictions for that bin may also change.
- Season was not one of the binning factors. Hence, hot days in spring and fall were included in the same bins as summer days with the same HDD and CDD. Bins containing observations of warm summer days from the summer models will also have similarly warm days from the spring and fall in the annual model. As with a change in binning procedure, this would result in a different specific mix of observations in each bin and thereby cause a change in the model predictions for each bin. We believe the random coefficients model could likely be improved by including seasonal indicators in the binning procedure.

We believe the results from the annual model better represents actual electricity usage because, (1) the results from the annual model are based on many more days of summer-like weather, and, (2) the annual model also considers the effect that any cooler hourly temperatures that may occur in the summer have on electricity usage.

1. The Annual Model Considers More Days of Summer-Like Weather

First, in order to keep the size of the working data file manageable as we developed the random coefficients model, we defined the summer model to include only data for weekdays in July, August, and September. In doing so, we left out of this data set many summer-like days in the shoulder months in which CDD was greater than zero. Our reason for doing this was strictly parsimony. In the early stages of this project, we wanted to keep our approach simple by focusing on the primary cooling months and only for weekdays. Once we were confident that the random coefficients approach was sound, we then expanded the analysis to include data for the entire year.

The difference in size of the working data set was substantial when we moved to the annual model. The annual model not only includes weekends and days of cooler temperature, but also includes many days with higher temperatures. For example, in the summer weekday model, there were on average fewer than 200 observations per hour for each home-day bin. Comparatively, when considering the entire year, but restricting to weekdays with CDD greater than zero, there were nearly 350 observations per hour per home-day bin. The annual model includes all of this additional information likely resulting in estimated load shapes that are more representative of actual hourly electricity usage. In addition, the larger sample (on average) in each bin results in a smaller standard error—all else equal—due to the greater degrees of freedom.[41]

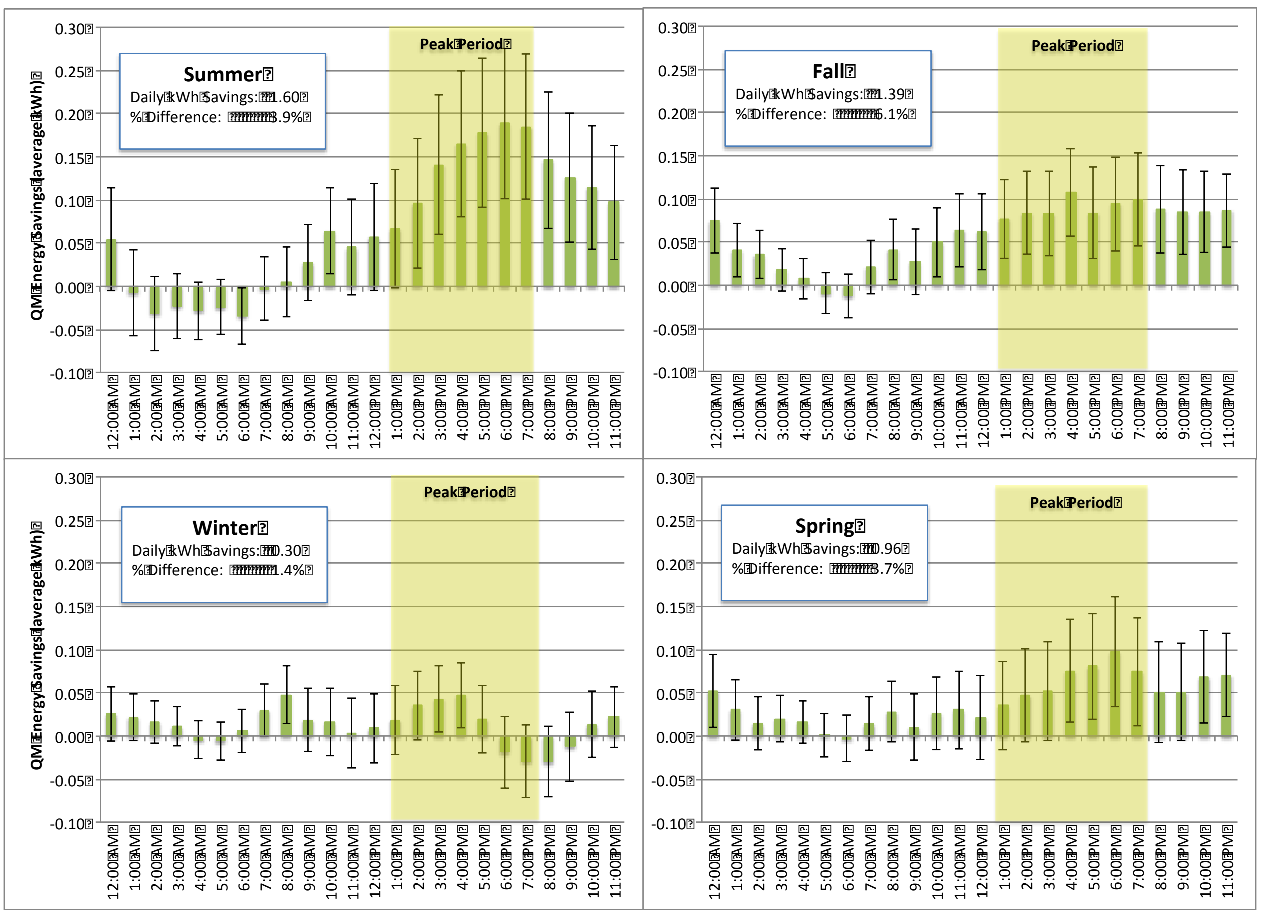
[41] Degrees of freedom represents the number of different ways in which a random variable can vary without violating any constraint placed upon it. For our purposes, we assume degrees of freedom is equal to the number of observations per hour for each home bin minus one.

2. The Annual Model Considers the Effect of Cooler Hourly Temperatures

Due to the parsimony of the summer weekday model, we considered only CDD when considering the temperature attribute of the home-day bin. Because of this, the summer weekday model does not account for hours of cooler temperature that many residential customers take advantage of to cool their home and reduce electricity usage. Comparatively, the annual model explicitly considers both CDD and HDD in the development of each home-day bin. We believe that the inclusion of HDD (representing opportunities for natural cooling) resulted in estimated load shapes that better fit actual summer-time load shapes of program participants.

Figure 36 shows the hourly kWh savings estimates and the 95 percent confidence intervals from the annual model for each season. We also found that nearly all of the savings during and after the peak period (hours 14-23) were statistically significant during summer and fall months.

Figure 36: PG&E QM Annual Savings Estimates by Season



A table with the results of the pre-period holdout sample and post-period for each of these seasons by day type (weekdays versus weekends), as well as charts of the hourly savings and error for each can be found in the report appendix.

3 Model Comparison Summary and Recommendations

The preceding chapter describes how AMI data can be used in several different impact analysis approaches, with details on both the analysis methods and results. As mentioned in the beginning of this report, we also explored several other methods for analyzing AMI data, but these methods were given less attention than the random coefficients model. The other methods include:

- Fixed effects regression model
- PRISM
- ECAM

The results of the AMI analysis using these other methods are included as an appendix to this report. The remainder of this chapter provides a comparison of estimation results across these methods. The chapter concludes with some recommendations for future research.

3.1 Comparison Of Fixed Effects and Random Coefficients Models

An appropriate comparison of the fixed effects model approach with the random coefficients model requires that each model use the same set of homes and billing data. For the purposes of this research, we relaxed a common criterion for the fixed effects models, which is to limit homes in the analysis to those with at least a full year of pre and post installation data. To ensure a direct comparison in the results, we used all homes and all observations used in the random coefficients model to model energy savings with a fixed effects model.

Table 22 compares the average household daily savings estimates produced by the random coefficients model with the daily savings estimates produced by the fixed effects model using monthly data for both the SCE QI and PG&E QM data. On an annual basis, the results of the two modeling approaches are comparable. The SCE QI program savings estimate produced by the random coefficient model of 1.91 kWh/day (7%) falls within the 95 percent confidence interval of the fixed effects savings estimates of 1.89 kWh/day (6.76%).

The PG&E QM program results are also close between the two models, with PG&E QM program savings estimated at 1.00 kWh/day (3.6%) by the random coefficient model and 0.76 kWh/day (3.04%) by the fixed effects model. The random coefficient model result for kWh savings falls outside the 95 percent confidence interval of the fixed effects approach, although the percentage of savings falls within the 95 percent confidence interval.[42]

[42] Note that the denominator when calculating the percentage of savings differs between the two models. The random coefficients model uses the modeled expected daily consumption in the post period in the absence of the HVAC installation as the denominator, whereas the fixed effects model uses the pre-period average actual daily consumption.

The similarity of the annual impact estimates between the two models gives us confidence that the random coefficients model provides reasonable savings estimates, as the estimates are comparable to the industry standard modeling approach, the fixed effects model.

Table 22: Comparison of Random Coefficients and Fixed Effects Models

Model	Random Coefficients Model Savings		Fixed Effects Model Savings	
	Daily kWh	%	Daily kWh (95% CI)	% (95% CI)
SCE QI Annual	1.91 ± 0.18 kWh	7.00% ± 0.60%	1.89 ± 0.29 kWh	6.76% ± 1.05%
PG&E QM Annual	1.00 ± 0.06 kWh	3.60% ± 0.21%	0.76 ± 0.15 kWh	3.04% ± 0.58%

The estimated savings in Table 22 represent the average annual program savings across all households in the program. We did not attempt to develop separate fixed effects models for some of the different subgroups covered by the random coefficients model. While it is possible to develop fixed effects models for some subgroups (months, seasons), other sub-models are not feasible (daily models, weekday vs. weekend models). In the typical fixed effects specification, a single coefficient (or sets of coefficients, depending on the variables used) are applied to all customers to estimate savings, in contrast to the random coefficients model that lets the savings estimates vary by bin. Additional fixed effects models, therefore, need to be developed manually for each subgroup to obtain savings estimates that vary across subgroups. Developing separate models can be cumbersome for more than a few subgroups, and in this area the random coefficients model provides a distinct advantage over the fixed effects model as these models are generated automatically and therefore can be easily developed for a high number of subgroups.

3.2 Comparison of ECAM and the Random Coefficients Model

In order to make an apples-to-apples comparison of the annual random coefficients model and ECAM, we ran each model on the exact same set of homes and observations. We selected a sub-group of PG&E QM program customers residing in CZ12, identified previously as CZ12f. This group includes 193 homes, each with three years of complete data (i.e., no missing observations from 2012-2014).

One important difference between these analyses is that the ECAM analysis used calendar year 2012 to construct the pre-period baseline model for the group of homes and excluded all observations in 2013, as this was the first year when the homes participated in the QM program. This approach was a simple way of avoiding bias due to missing data.

The random coefficients model is not modeling groups of homes, but rather groups of home-days. This more sophisticated approach allows the inclusion of all pre-period days from 2012 and 2013 for all homes in the modeling sample, resulting in one model for each type of home on a day with specific weather conditions (CDD and HDD) and day type (weekend or weekday). Both models used the full calendar year of 2014 as the post-period to ensure that the same days were being included in the calculation of PG&E QM program impacts and load shape for this sample.

Figure 37 shows the results of the ECAM analysis, and Figure 38 shows the results of the random coefficients model using the same 193 homes and days in 2014. In both charts, the blue line represents the average actual consumption during each hour, and the red line represents the model's prediction of the average consumption during each hour. The random coefficients model also depicts the estimated savings in each hour using green bars.

The daily actual kWh, predicted kWh, kWh savings, percent difference, and load shapes of both modeling approaches are very similar. ECAM estimated the savings for these homes was 1.20 kWh per day (4.71%), while the random coefficients model estimated that the savings for these homes was 1.26 kWh per day (4.94%). However, there are some differences in the specific hours when the models suggest most of the savings occurred, with the largest hourly savings occurring at 6:00 a.m. and from 3:00 p.m. to 5:00 p.m. in the ECAM model compared to from 1:00 p.m. to 4:00 p.m. in the random coefficients model.

Figure 37: ECAM Results for CZ12f Sample

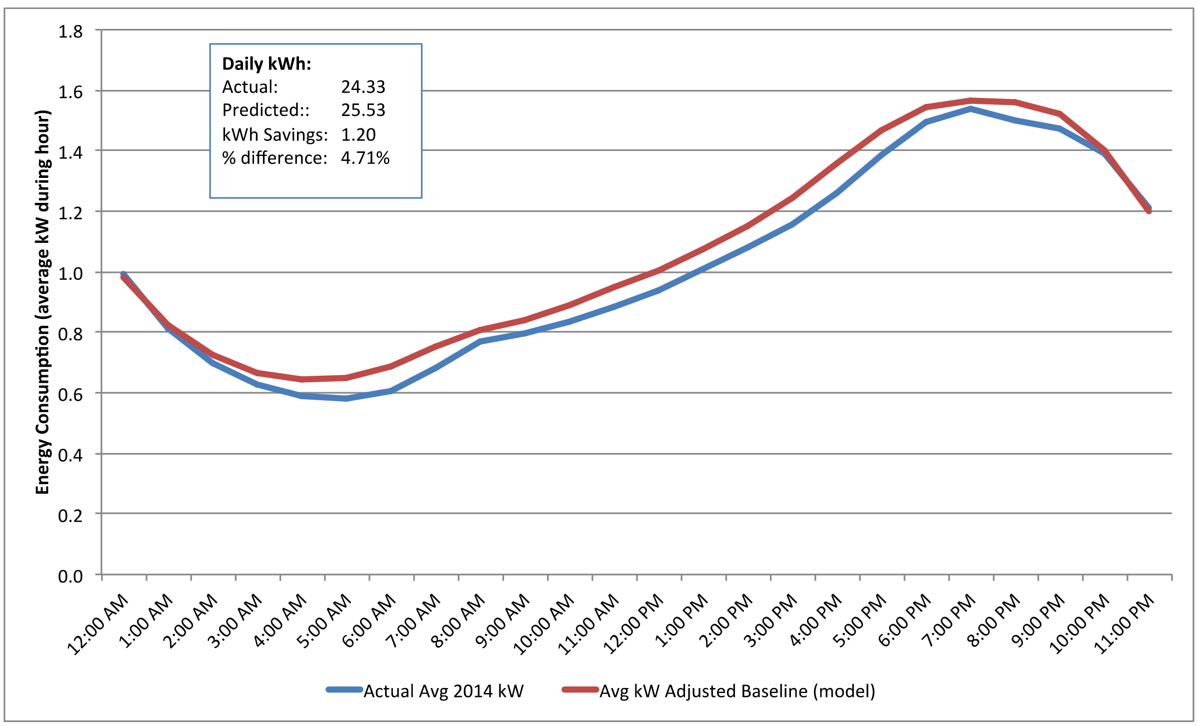
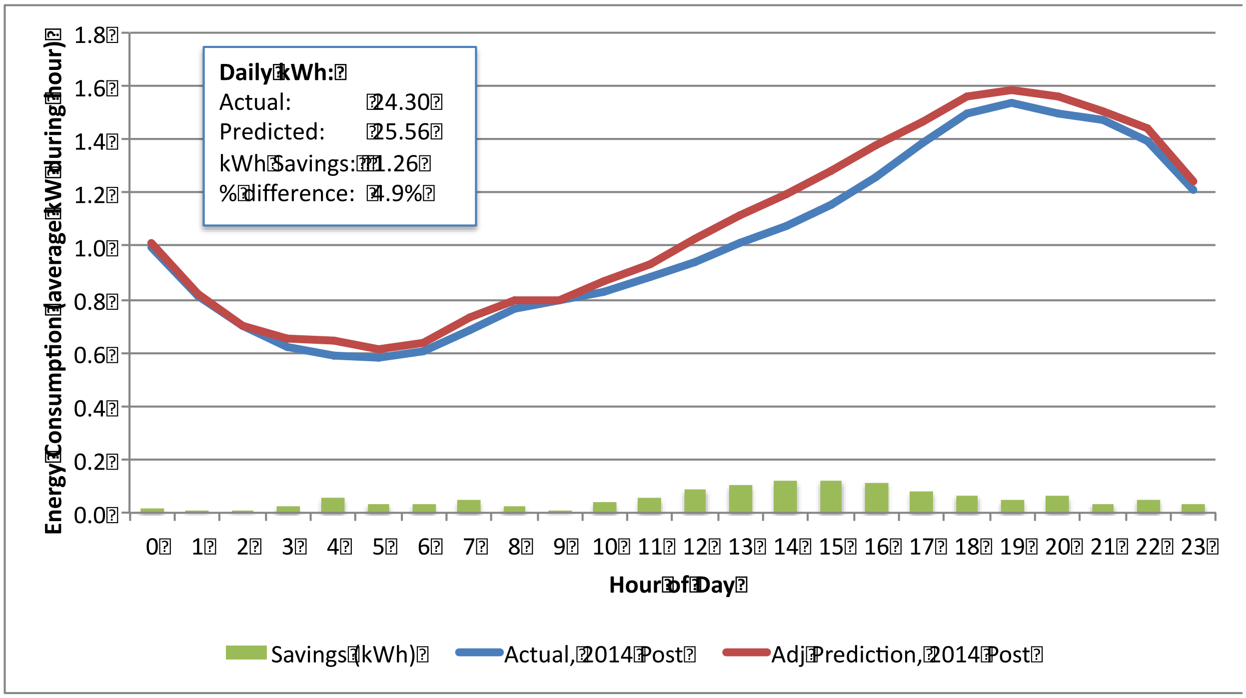


Figure 38: Random Coefficients Model Results for CZ12f Sample

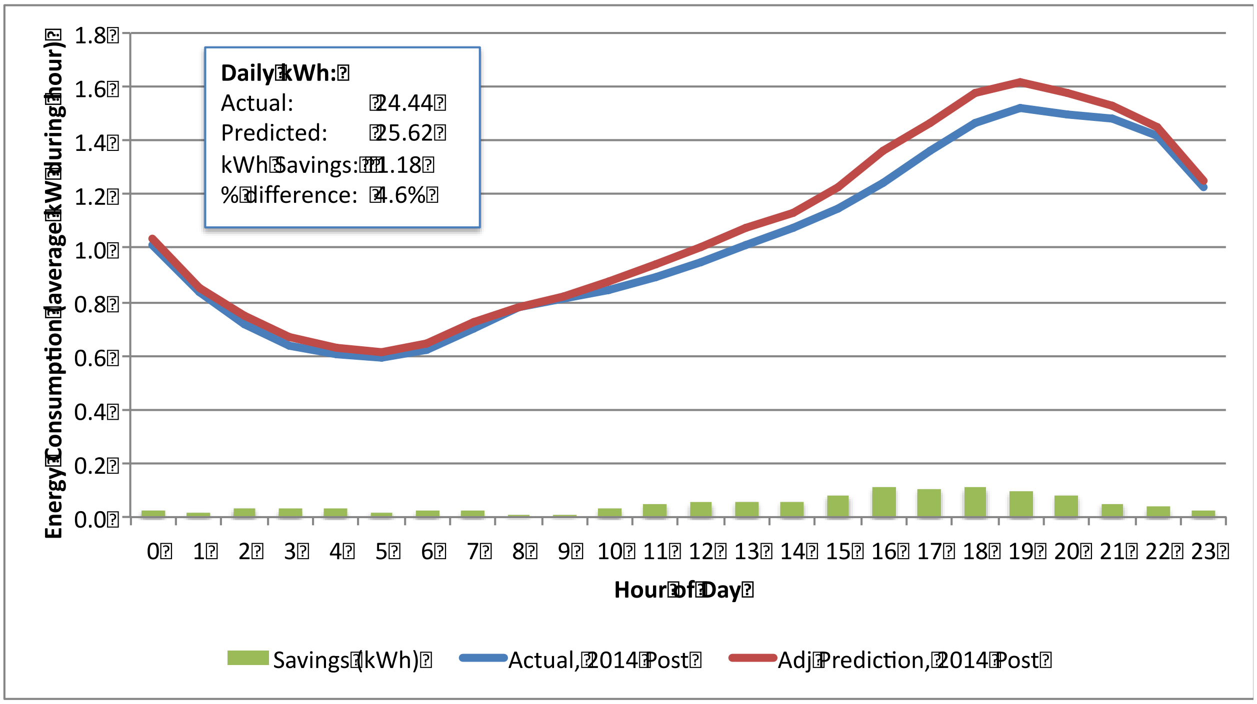


ECAM and the random coefficients model generated very similar estimates of the hourly load, total daily consumption, and kWh savings for this group of homes. In order to make this comparison, we used the same filtered dataset for the random coefficients model that was used by ECAM. This filter removed all homes with one or more missing hourly observations during the study period of 2012-2014 from the sample. This filter excluded 168 households (47%) from the modeling sample.

To see the impact of this filter, we ran the same random coefficients model on the full sample of 358 homes. We excluded three homes with a missing period (e.g., no post-period observations). We also excluded individual home-days with missing observations from the 168 homes with one or more missing hourly kWh consumption value, rather than removing all observations for that home.

Figure 39 compares the pre-period predicted load shape (red line) with the post-period actual load shape (blue) averaged across all households. Whenever the post-period load shape falls below the pre-period load shape, this indicates that savings were realized during that hour (green bars). After adjusting for the error in the model, based on the sample of homes used, the modeling approach finds approximately 4.6 percent annual savings attributable to HVAC maintenance provided through the PG&E QM program. In this example, the total daily savings is quite similar to the sample of 193 homes shown in Figure 38, but the hours when most of the savings occur has shifted from 1:00pm-4:00pm to 4:00pm-7:00pm.

Figure 39: Random Coefficients Model Results for CZ12f Full Sample



As with the fixed effects model, both ECAM and the random coefficients model produced similar annual savings results. Both approaches also provide hourly load shape estimates, but the random coefficients model is able to make use of more data, potentially making the results more representative of all homes treated by the program. ECAM also requires that each model be constructed manually, which further limits its ability to develop separate estimates for different sub-groups of interest.

3.3 Comparison Of PRISM and the Random Coefficients Model

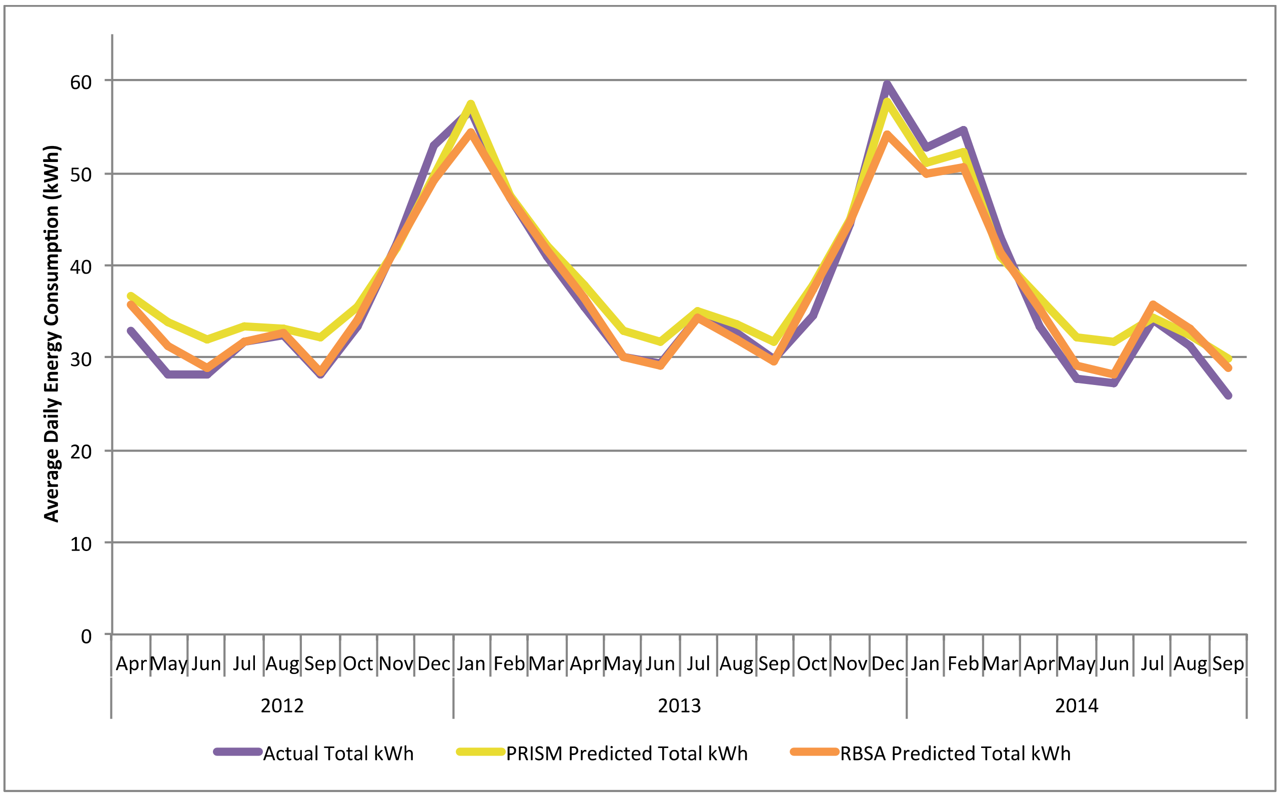
In order to compare directly the random coefficients model to PRISM, we ran both models on the same set of 99 households from NEEA's RBSA dataset.[43] Because the PRISM model was based on daily consumption while the random coefficients model uses hourly data, we aggregated the hourly predictions from the random coefficients model to the daily level.

Figure 40 shows the actual average daily total kWh consumption (purple), PRISM's prediction of the average daily total kWh (yellow), and the random coefficients model prediction of the average daily total kWh (orange) for each month during the test period of April 2012-September 2014. Both models were able to predict consumption reasonably well throughout the study period. The overall percentage difference between the actual and predicted daily

[43] This analysis excludes 4 of the 103 RBSA households due to a high number of missing observations in the metering data.

kWh consumption was 4.0 percent for PRISM and 0.1 percent for the random coefficients model. PRISM's predicted consumption in winter months was more accurate than the random coefficients model, with a percentage difference of -2.6 percent in the PRISM model and -5.7 percent in the random coefficients model. This underestimation during the winter months in the random coefficients model may be caused by the fact that we capped heating degree days at 70 for this example to reduce processing time, resulting in the same model predictions being assigned to home-days with HDD of 70 as HDD of 80. During summer months, the random coefficients model has more accurate predictions of total daily consumption than the PRISM model for this group of homes, with a percentage difference of 5.4 percent in the PRISM model and 2.2 percent in the random coefficients model.

Figure 40: Actual versus Predicted Daily Total Consumption Comparison

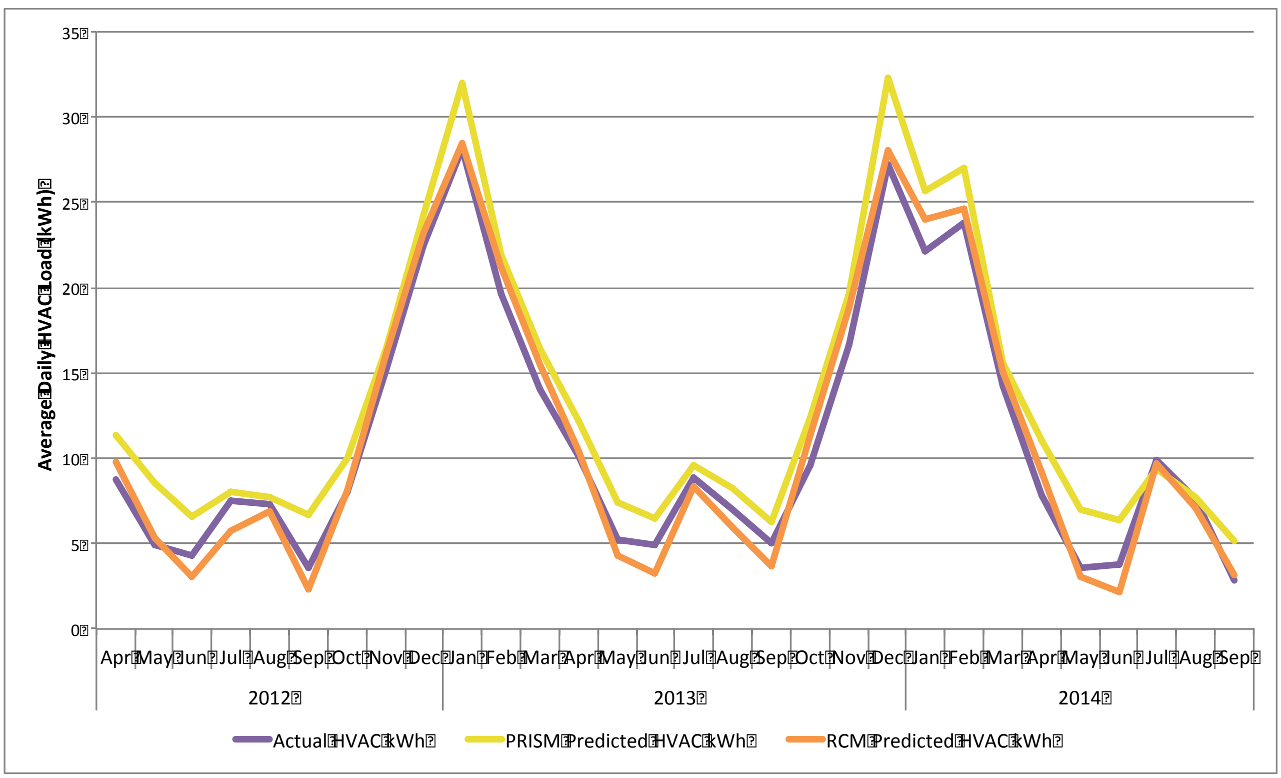


The RBSA database has the advantage of having both whole house and HVAC metering data. This allowed us to estimate a billing regression model and then validate the model's ability to isolate the HVAC consumption with the actual HVAC metering data for that same time period. For both models, we assumed that all predicted weather-dependent consumption (i.e., any consumption over the weather normalized baseline) was caused by operation of HVAC equipment. We extracted this estimated HVAC consumption from the total energy consumption predictions generated by PRISM and the random coefficients model, then compared these to the actual HVAC consumption, as measured by the RBSA submeters.

Figure 41 shows the actual average daily HVAC kWh consumption (purple), PRISM's prediction of the average daily HVAC kWh (yellow), and the random coefficient model's prediction of the average daily HVAC kWh (orange) for each month during the test period of April 2012-September 2014.

The overall percentage difference between the actual and predicted daily HVAC kWh consumption was 19.4 percent for PRISM and 1.2 percent for the random coefficients model. PRISM's model overestimated HVAC consumption during most months, while the random coefficients model slightly overestimated HVAC consumption during winter months and underestimated it during summer months. In the previous figures (total daily consumption) we found that the random coefficients model underestimated total daily kWh consumption during winter months, but here we see that the random coefficients model is quite accurate at predicting the HVAC consumption for these months. This suggests that the random coefficients model underestimation in total usage comes from an underestimation of the consumption in winter months from end uses other than HVAC (e.g., lighting). It may be possible to further improve the accuracy of the random coefficients model by including seasonal indicators in the binning process, but this option was not explored as part of this study.

Figure 41: Actual versus Predicted Daily HVAC Consumption Comparison



In general, this comparison demonstrates that the random coefficients model is able to produce results that are similar to PRISM. Additionally, we found some benefits to using the random coefficients model instead of PRISM to estimate household energy consumption:

1. **Interval Data** - The finest granularity that PRISM can handle is daily observations, while the random coefficients model can easily work with hourly data (it can also be adapted to work for 15-minute or finer intervals).
2. **Controls** - The random coefficients model can control for additional factors when modeling specific days. This version of the random coefficients model bins by day type; other bins of interest could include holidays, weather variables beyond temperature (e.g., rainfall), or period of home vacancy (e.g., tenant turnover). PRISM allows the user to identify estimated meter reads, but all other days with unusual energy usage would have to be dropped from the analysis to be controlled. Given the large number of data points with AMI data, we would expect that the number of estimated meter reads is much less of an issue than with monthly data.
3. **Variation in Observations** - PRISM requires the user to run sets of models with households depending on which weather station they are assigned to. The random coefficients model uses a station identifier to assign weather data to observations, but from then on can look at all observations (from multiple stations) simultaneously. This is particularly important when the data include homes from a large territory that spans many climate zones.

3.4 Comparison Summary

Table 23 provides a summary of the benefits and limitations of these different approaches to modeling AMI data. While all four approaches are capable of modeling daily and monthly intervals, only three of the four (all except PRISM) are capable of modeling hourly or finer intervals.

Based on our analysis, we believe that the random coefficients model provides the most advantages for estimating impacts using AMI data. The random coefficients model has shown itself to be extremely accurate when predictions are compared with a holdout sample. It also has the advantage of being able to automatically generate load shapes (and subsequently impact estimates) across a wide range of subgroups, in contrast to the other methods where separate models need to be developed manually for each group. Since the random coefficients model is a new technique, no existing software or programming text is readily available.

Table 23: Benefits and Limitations of Four AMI Modeling Approaches

Characteristic	Modeling Approaches			
	Random Coefficients	Fixed Effects	ECAM	PRISM
Capable of modeling daily and monthly intervals	Yes	Yes	Yes	Yes
Capable of modeling hourly (or finer intervals)	Yes	Yes	Yes	No
Automatically generates results for different segments (e.g., home types, day types, season)	Yes, this is inherent in the design	No, separate models need to be developed manually	No, separate models need to be developed manually	No, separate models need to be developed manually
Accuracy of estimates (based on analysis presented in this report)	Very accurate, predictions for holdout sample typically within 1%	Accurate, annual results similar to random coefficients model	Accurate, annual results similar to random coefficients model	Less accurate, estimates consistently overestimated HVAC load
Availability of software	Limited, some options available (LimDep, likely R)	Common, many options available (e.g., LimDep, R, SAS)	Free, public-use Excel tool with detailed user guides available	Available for purchase from the developer
Capable of handling large datasets	Yes	Yes	No, Excel has data limits	No

The four AMI modeling approaches each provide advantages in certain situations. Based on our experience using each method with AMI data from the RBSA, QI, and QM programs, we believe these models should be the preferred approach in the following situations:

- **Random Coefficients Model –** When hourly savings estimates and load shapes are desired for multiple groups, at the customer and program level. This could include groups of households (by demographics, regions, equipment type), types of days (hottest summer days, seasons, weekday vs. weekend), or both.
- **Fixed Effects Model –** When only a single, annual program-level savings estimate is needed, with no separate estimates needed by subgroup.
- **Energy Charting and Metrics Tool (ECAM) –** When working with small to mid-sized datasets and segmentation is only needed for a few groups, and/or free software with detailed user guides is desired.
- **Princeton Scorekeeping Method (PRISM) –** When there is no interest in hourly savings or load shape, household-level savings are desired, and homes are suspected to have significantly variable baseload heating/cooling temperature setpoints.

3.5 Recommended Areas for Future Research

We believe that the preliminary research presented in this report demonstrates enormous potential for the random coefficients model and represents a significant and positive departure from current approaches to analyzing AMI data and estimating program impacts. While analytically and conceptually more sophisticated than the fixed-effects model, the additional complexity of the random coefficients model is necessary to take full advantage of AMI data. As utilities continue to migrate their customers to interval meters, we believe it is necessary that evaluators embrace methods of analysis that fully exploit the abundant information contained in AMI data.

The initial analysis results relied on data from residential customers only and examined a handful of scenarios to test the ability of the random coefficients model to simulate customer load shapes and estimate energy savings. Although we believe these initial results are very promising, they also suggest that further research in other areas is warranted. Suggestions for research topics in the next research phase are discussed below.

Commercial customers. A logical next step is to test the random coefficients model on commercial customers. Commercial customers typically will have greater variations in energy use given the wider ranges of end uses, building types, and business activities relative to residential customers. Potential IOU sources of commercial participant data involving HVAC are the Upstream Commercial HVAC Program, the Commercial QM Program and the Commercial QI Program. An initial test of the model can be done by using only the AMI billing data for these commercial customers and testing how well the process outlined in this report can predict load shapes. To estimate energy savings, additional data collection will be needed to identify HVAC installation date (for the Upstream program), the number of HVAC units at a site and the portion that were covered as part of the program. If any of these sites have HVAC meters, then this information can be used to test how well the random coefficients model can estimate the HVAC load for commercial customers.

Customer targeting based on demographics/firmographics. An intriguing area for future research is linking customer characteristics to specific load shapes pattern. Given the binning process, some of the steps needed to establish these links are completed automatically based on the initial bin assignments. Once the load shapes are calculated, additional modeling would focus on what customer characteristics are most correlated with specific load shapes. This could be done through established discrete choice (i.e., logit) modeling techniques based on whatever customer data are available, either through the utility records or other publically available data sources (e.g., US Census, Dun and Bradstreet, InfoUSA). With additional research, the discrete choice modeling component can also be automated to calculate a propensity score for each customer based on their characteristics and estimated load shape. This information can then be used by program managers for recruiting to specific energy efficiency or demand response programs, or for tailoring programs that are more closely matched to specific customer types.

Comparison group. It is often desirable to include an appropriately matched non-participant comparison group in the regression sample to help account for other factors that might be affecting energy use but are not controlled for explicitly in the model. Without a comparison group, the model may erroneously attribute changes in energy use to the program intervention rather than to external factors such as economic conditions that might be affecting energy use throughout the population. Future work with the random coefficient models should explore the effects of using a comparison group on the load shape forecasts and the energy estimates.

Changes to customer binning, setpoint temperature, holdout samples. This initial test of the random coefficients approach only explored a limited number of variations in model parameters, and examining more variations in these areas may yield additional improvements to the approach. As discussed earlier in the report, one variation that should be explored is to expand the binning processes to include a seasonal element, which may help explain the differences observed across the daily and annual models for the summer impact estimates. Additional work should also explore the accuracy of the model using a larger sample of randomly selected holdout groups. Variations in the setpoint temperatures (currently at 65 degrees in the current models) should also be explored to determine if the model results are sensitive to assumptions made regarding this parameter.

Demand response. A logical extension of the random coefficients model is to test it with demand response programs. The basic modeling steps are consistent with the current impact evaluation methods commonly used for demand response programs[44], where historical customer billing data are used to forecast energy use during an event period and then the difference between the observed and predicted consumption during the event is used as the estimate of program impacts. Current methods generally rely on developing these load forecasts manually, and the random coefficients model provides an opportunity for this process to be automated. Additional research comparing the traditional impact methods with the random coefficients model for the demand response programs would be very beneficial and could allow for more accurate models that are tailored more closely to different customer groups.

Load forecasting. While this report explores using the random coefficients model in the context of program evaluation, the ability of the model to forecast load shapes provides an opportunity for broader load forecasting using a wider group of customers. With the binning process, the customers are segmented based on energy use and weather conditions in such a way as to remove a substantial amount of uncertainty from the model. In the applications presented here, this has resulted in very accurate load shape predictions. Future research could expand this to address larger customers groups (the entire residential or commercial

---

[44] See *Load Impact Estimation for Demand Response: Protocols and Regulatory Guidance,* California Public Utilities Commission, April 2008.

population in a geographic area, for example) and determine how well a bottom-up approach using the random coefficients model can produce accurate load forecasts. This could include developing short term forecasts under extreme weather conditions as well as longer term forecasts assuming historical average weather conditions.