Random Walk to Savings: A New Modeling Approach Using a Random Coefficients Model and AMI Data

Stephen Grover, Evergreen Economics, Portland, OR
Ted Helvoigt, Evergreen Economics, Portland, OR
Sarah Monohon, Evergreen Economics, Portland, OR
John Cornwell, Evergreen Economics, Portland, OR

Abstract

This paper presents a new energy savings estimation approach, one that provides accurate impact estimates by taking full advantage of hourly AMI data. This approach differs from traditional methods in that it automatically develops a large number of customer-specific regressions covering a wider range of customer types, weather conditions and time periods. The approach uses a type of hierarchical linear model—the random coefficients model—that allows savings estimates to be tailored more closely to individual customer characteristics. This is accomplished by first grouping customer consumption data into different categories based on energy use and weather conditions. Separate models are then estimated for each usage/weather category, which allows for separate load shape predictions for very specific customer types.

The random coefficients model specification was tested using data from two HVAC efficiency programs in California. Using participant and AMI data from both of these programs, average daily load shapes were calculated for specific day types (weekday, weekend, seasonal) and used to estimate program impacts. When estimated load shapes were compared against a holdout sample of customers, the random coefficients model performed extremely well; load shape estimates were within 1 percent of the holdout sample. Energy savings estimates for these programs ranged from 4 to 7 percent of annual energy use, which was consistent with expectations. Besides producing accurate impact estimates, the automated categorization and modeling processes allow for separate savings estimates and load shapes to be developed easily for a variety of situations.

Introduction

As electric utilities transition to advanced metering infrastructure (AMI), a greater amount and richer source of consumption data are becoming available to evaluators. A single customer's metered data at one-hour intervals translate to over 700 data points per month, providing an opportunity for evaluators to better understand the impact that energy efficiency programs (and other factors) have on energy consumption during specific hours of the day, rather than a daily average derived from monthly data. 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 allows 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. Most of the literature to date has focused on using monthly consumption data, as these are typically all that have been available for estimating impacts at the program level. See the California Evaluation Framework (CPUC 2006) and the Uniform Methods Project (Agnew & Goldberg 2013) for a summary of the more traditional methods using monthly data. Other studies (particularly those in demand response programs) have utilized AMI data to estimate load shapes and demand impacts (Nexant 2015, for example), but these models have typically been developed manually for each specific situation and therefore have not been practical for addressing a large number of customer types and time periods. Other works such as Hsiao et al. (1989) provided an early application of the random coefficients model to energy efficiency, while Granderson et al. (2015) have begun to look at developing AMI regression models in a more systematic fashion, but none of these past studies have presented a method for efficiently developing a large number of models that are tailored to a wide range of customer types and time periods that take full advantage of the information contained in the AMI data.

To explore how AMI data could be used in billing regression models, the California investor-owned utilities[1] (IOUs) contracted with Evergreen Economics and SBW Consulting to conduct exploratory research, and a portion of these research results is presented in this paper. During the course of this research, it became apparent that an innovative new analysis method—the random coefficients model—has the potential to be a groundbreaking impact evaluation approach that fully utilizes the benefits of the more granular AMI data. As discussed in the remainder of this paper, 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.

Analysis Methods

This paper presents the results of a random coefficients model approach that utilizes data from the following sources:

- **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. Program data include household and program participation information including the home climate zone and date of participation in the program.
- **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 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.

Each customer dataset was combined with weather data obtained from the National Oceanic and Atmospheric Administration (NOAA) to develop datasets with both energy consumption and weather data. Weather station data were selected based on proximity to each home’s zip code, matching climate zone, and availability of complete hourly data. The selection process resulted in hourly data for 95.5 percent of hourly observations. Additional analysis was performed to identify unreasonably high or low

---

1 The California investor-owned utilities include Pacific Gas and Electric (PG&E), Southern California Edison (SCE), San Diego Gas & Electric (SDG&E), and Southern California Gas Company (SoCalGas).

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.

Modeling Approach

The standard approach in regression analysis is to focus on the average response of the population of interest. With the most basic billing regression specification, the model produces a regression line that represents the average energy use or savings across all customers included in the sample. Variations on the standard linear regression such as the fixed effects model can be developed to produce separate savings estimates for sub-groups of customers. With the availability of AMI data, however, modeling approaches need to be adapted to account for the additional information as well as the sheer volume of data. In theory, the traditional methods such as the fixed effects model can be adapted for use with hourly AMI data. However, the number of different coefficients (or separate models) required to capture the variations in energy use across hours, days and seasons—in addition to accounting for variation across customers—requires significant amounts of computer processing time, along with a separate process for efficiently evaluating the performance of different model specifications.

Rather than attempting to adapt a fixed effects model for use with a high volume of AMI data, this paper explores a different hierarchical modeling approach that first categorizes the data into manageable sub-groups and then uses regression analysis to estimate energy use within each group. The term "random coefficients model" refers to a type of linear hierarchical model that provides a distribution of model parameters across customer types and weather conditions.2 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.

To incorporate both types of variability and estimate customer load shapes, the random coefficients model utilizes a multi-stage process described below. Note that variations in this approach, such as developing different categories for the data bins and/or exploring different setpoint temperatures for the weather variables, were not attempted in this initial study but are planned for future research phases.

Figure 1 summarizes the random coefficients modeling process and how it is used to estimate load shapes and program savings, with additional detail following the chart.

2 See Snijders and Bosker for a more detailed explanation of the random coefficients model, and a more general discussion of the different types of hierarchical linear regression models.

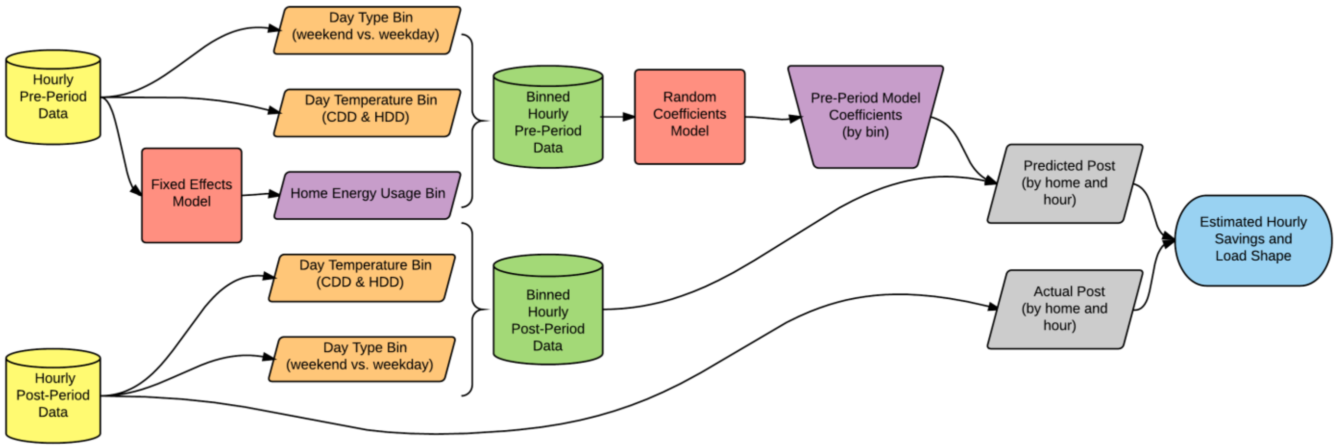

Figure 1. Summary of Random Coefficients Model Savings Estimation Approach

First Stage – Binning Process. In the first stage of the modeling approach, a fixed effects regression model is used to create estimates of daily baseload electricity use for each home, controlling for outside air temperature. The fixed effects model specification is as follows:

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

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 the fixed effects model is the estimation of a specific constant α_i, for every customer site. This constant varies by customer site and accounts for time-invariant effects on electricity consumption over the year. 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). Homes are then ranked in ascending order of baseload energy use and assigned to one of 20 “home groups” based on each home’s weather normalized home usage, prior to program participation. In this way, homes with similar energy consumption were grouped together. Each home group represents about 5 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.

Next, every day of the study period is characterized (binned) in terms of the weather and day type. The weather groups are created by calculating cooling degree hours (CDH) for each hourly observation using a base temperature of 65 degrees Fahrenheit, and then taking 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, rounded up to the nearest integer. For models covering the heating season, this process is repeated to assign days to heating degree day (HDD) groups, again using a base temperature of 65 degrees Fahrenheit.3 There were a total of 25 weather groups for both CDD and HDD, and categorizing days using

3 An alternative model using 75 degrees Fahrenheit to define cooling days was explored in this analysis but did not have a significant effect on the estimation results. The next phase of this research will explore using different base temperatures for defining both cooling and heating degree days.

outdoor temperature in this manner explicitly incorporates temperature into our modeling approach. To reflect possible differences in energy usage between weekends and weekdays, home-days are also binned based on the day type. Weekends were assigned to day type group 1 and weekdays were assigned to day type group 0.

Lastly, all groups are combined 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) based on the average daily weather value described above, and day type group (weekday versus weekend). 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.4
- This process models home-days rather than households, which allows for individual days with missing observations to be excluded from the analysis sample.5 For example, specific days with less than a complete 24 hours of hourly data can be removed rather than limiting the analysis to homes with flawless data throughout the study period.
- 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), or at the program-level (i.e., incorporating all bins), without the need to run additional models.

Second Stage – Random Coefficients Model. For the next stage, 70 percent of sample homes were randomly selected for use in the random coefficients regression model to develop predicted hourly load shapes for each home-day bin using pre-period consumption data. The remaining 30 percent of homes were set aside as a holdout sample to test the performance of the predicted load shape. In this way, the predictive power of the model is tested against data that were not used to develop the model. For both the QI and QM datasets, the model was able to estimate load shapes within 1 percent accuracy for the holdout samples in both cases. In this application, the 1 percent was used as an adjustment factor for the load shape forecasts to improve accuracy.

The average hourly kWh value was computed 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.

The random coefficients model specification is used, as it allows for the daily load shape (i.e., hourly kWh usage) of each home-day bin to be estimated while accounting for covariance with other home-day bin load shapes. Unlike a typical fixed effects regression, which produces a single set of

---

4 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.

5 For both the QM and QI customer data, less than 1 percent of the observations were missing and so the likelihood of bias due to missing data is small.

coefficients and household-specific constants, the random coefficients model produces a vector of regression coefficients for each home-day bin.

The final random coefficients model specification is as follows:[6]

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

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*.

$\varepsilon_{i,t}$ = Random error, assumed normally distributed.

The coefficient estimates from the random coefficients model in each bin are then used to estimate consumption and eventually energy savings, as explained in the third and final stage of the modeling process.

Third Stage – Savings Estimation. The third and final stage requires that load shapes be calculated for the post-period using the results of the random coefficients model. To accomplish this, the post-period data are subjected to the same binning process as was used with the pre-period data. Each individual home remains in the same weather-normalized usage group that it was 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. Next, each day is assigned to a weather group (by CDD and/or HDD) and day type group (weekday or weekend).

After assigning each home-day in the post-period to a home-day bin, the predicted hourly pre-period kWh values for each home-day bin are combined with the random coefficients regression model results. This process gives us a predicted estimate of each household's consumption during each hour of the post-period if it had not participated in the program.

Once the forecasts of post-period usage are created (based on the pre-period consumption model and post-period weather data), they are compared 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. This calculation could also be generalized to use long-term average weather data (rather than the specific post-period data used here) to create load shapes and subsequently energy savings estimates that assume more typical weather conditions.

Note that this process attributes the entire difference in actual versus predicted usage to the program intervention, which may or may not be appropriate depending on the program or market context.

---

6 Different model specifications at this stage could be used, including a model that includes separate coefficients for each hour of the day. Given that the predictive power was quite good with this initial specification, we did not explore possible alternatives.

It also assumes that the existing site conditions are an appropriate baseline—any adjustments to account for a standard practice or market baseline would need to be made outside the random coefficients model.7

Analysis Results

Figure 2 presents the impact results from the model using the SCE QI Program data. This graph 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). The random coefficients modeling approach results in approximately 7 percent annual savings attributable to the HVAC installed through the SCE QI Program.8 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 p.m. and 8 p.m., highlighted in yellow. The 95 percent confidence interval 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.

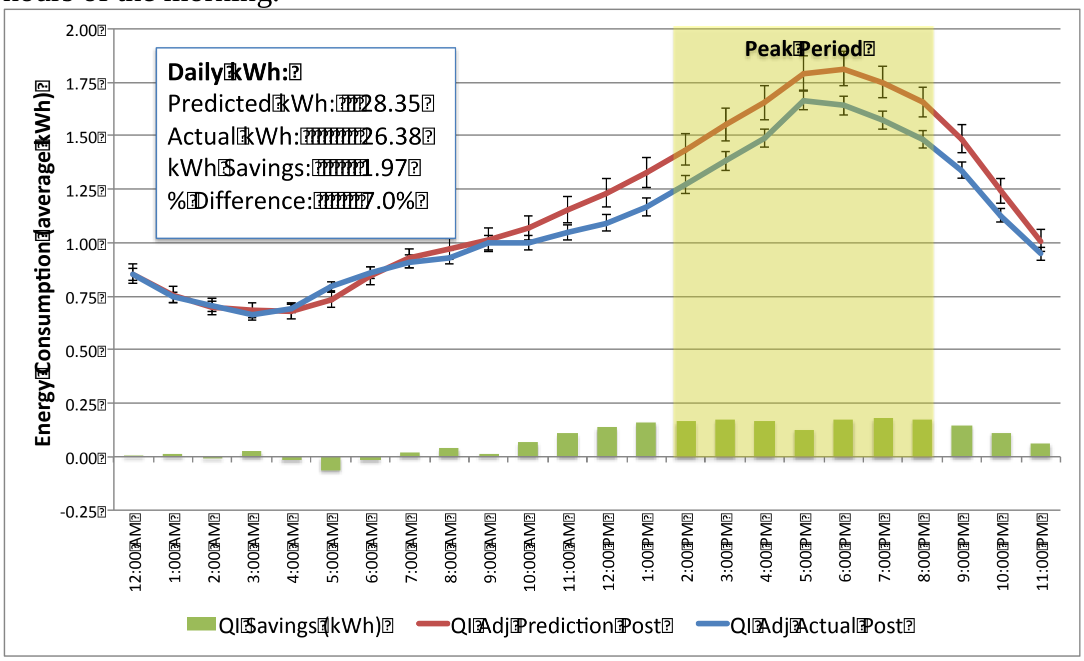

Figure 2. SCE QI Overall Annual Post-Period Model, Includes All Months and Day Types

Figure 3 shows the average annual hourly kWh savings estimates along with the 95 percent confidence interval for each hour. Statistically significant hourly savings occurred from 11 a.m. to 4 p.m. as well as from 6 p.m. to 10 p.m. Some early morning hours had increases in usage (i.e., negative savings), but none of these were statistically significant.

---

7 Both these issues are also present with the traditional billing regression methods and are not unique to the random coefficients model.

8 It is not possible to determine how much of the estimated savings come from the quality installation practices versus the new HVAC equipment from the data available for this study. In order to separate these impacts, the model would need to include a control group sample of customers who replaced their HVAC system but did not use a program contractor for the installation.

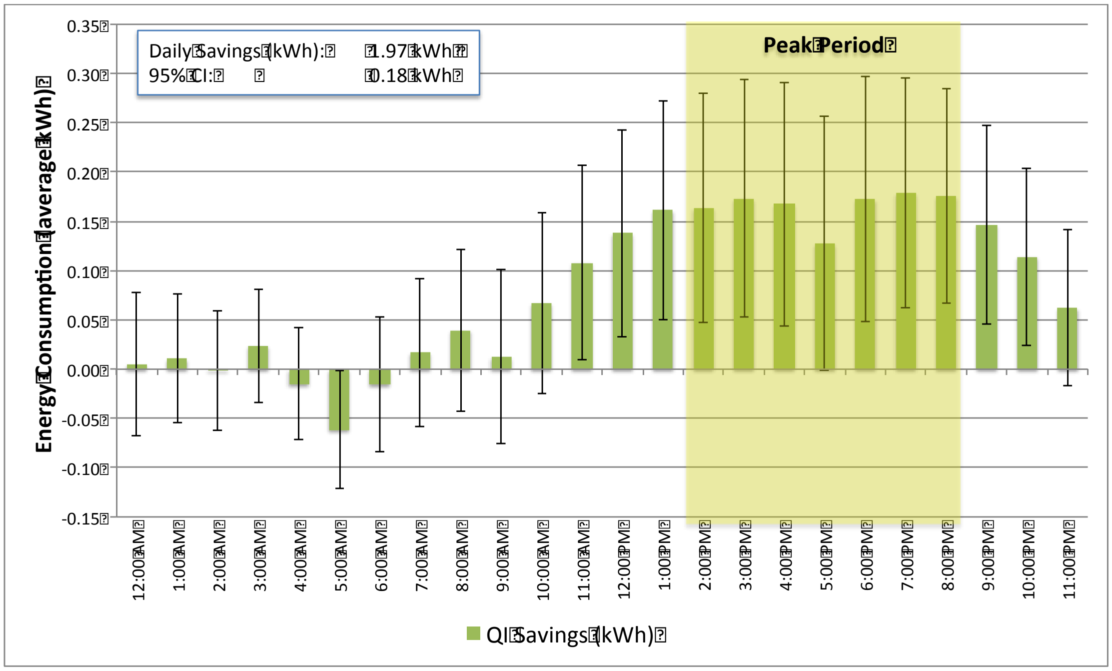

Figure 3. SCE QI Overall Annual Savings, Includes All Months and Day Types

Figure 4 shows the model's predicted savings by season.9 Note that this is an important feature of the random coefficients approach. Once the individual models are estimated, they can be aggregated easily to show savings for different sub-categories of interest. As shown in the graph, 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 kWh 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. Despite the variation in load shapes across seasons, the random coefficients model is able to produce very accurate estimates in a variety of conditions. This ability for a single modeling process to match automatically a range of different load shapes is a key benefit of the random coefficients modeling approach.

9 Seasons are defined as: summer (July-September), fall (October-November), winter (December-February), spring (March-June).

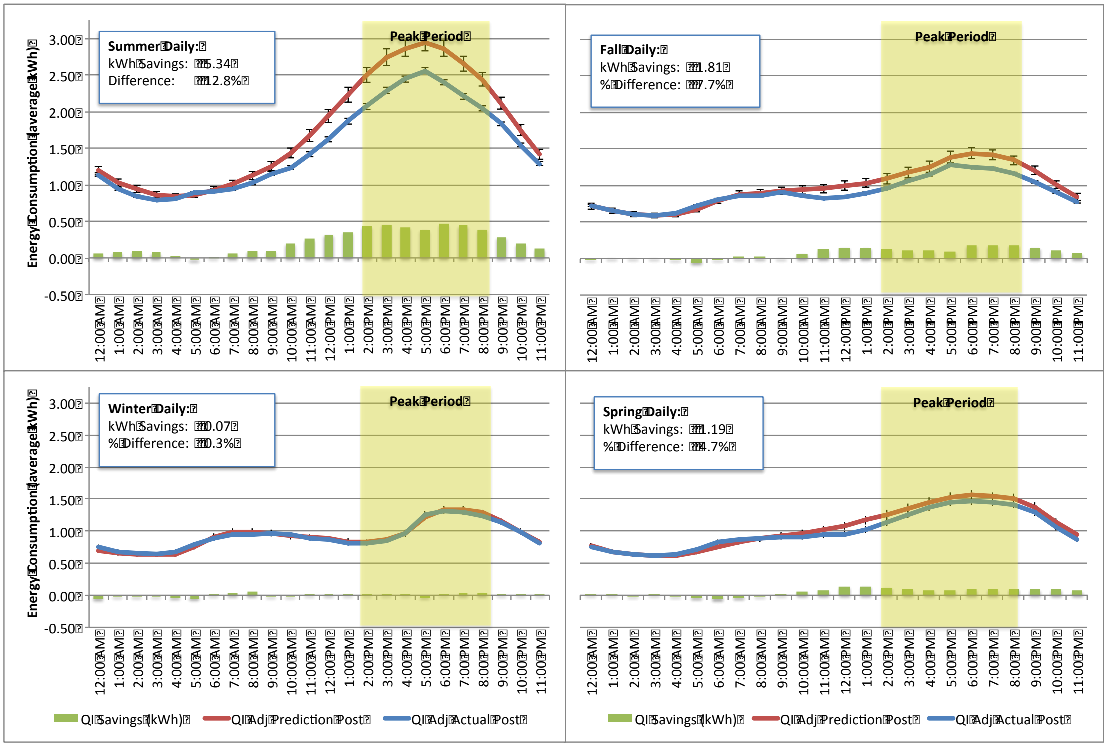

Figure 4. SCE QI Annual Model Results by Season

A similar set of results was produced using data from participants in the PG&E QM Program. Figure 5 compares the pre-period predicted load shape (red) with the post-period actual load shape (blue) averaged across all participating 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 with the QI Program, the largest QM impacts are realized during the later part of the day, including during the peak period between 1 p.m. and 7 p.m. highlighted in yellow.10 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 a.m. and widest during the peak hours from 2 p.m. to 11 p.m.

10 The residential peak period of 1 p.m. to 7 p.m. is used, as defined for PG&E's residential Time-Of-Use rate plan (E-6). http://www.pge.com/nots/rates/tariffs/ResTOUCurrent.xls

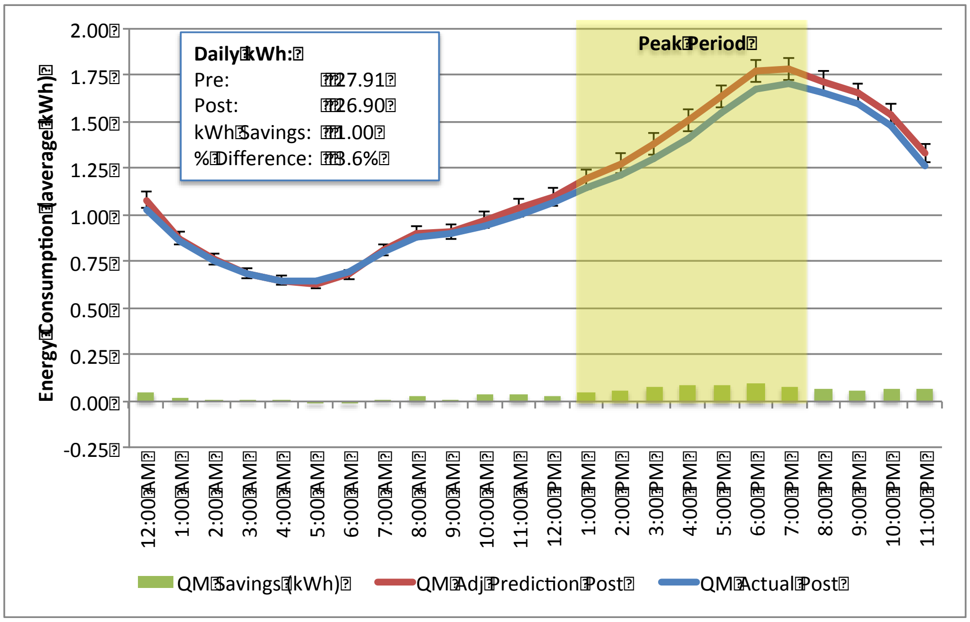

Figure 5. PG&E QM Program Overall Annual Post-Period Model, All Months and Day Types

Figure 6 shows the annual hourly kWh savings estimates from the previous figure with bars depicting 95 percent confidence intervals around each estimate. Statistically significant hourly savings were found from 1 p.m. to 11 p.m. and included the peak period. None of the increases in usage (i.e., negative savings) were significant.

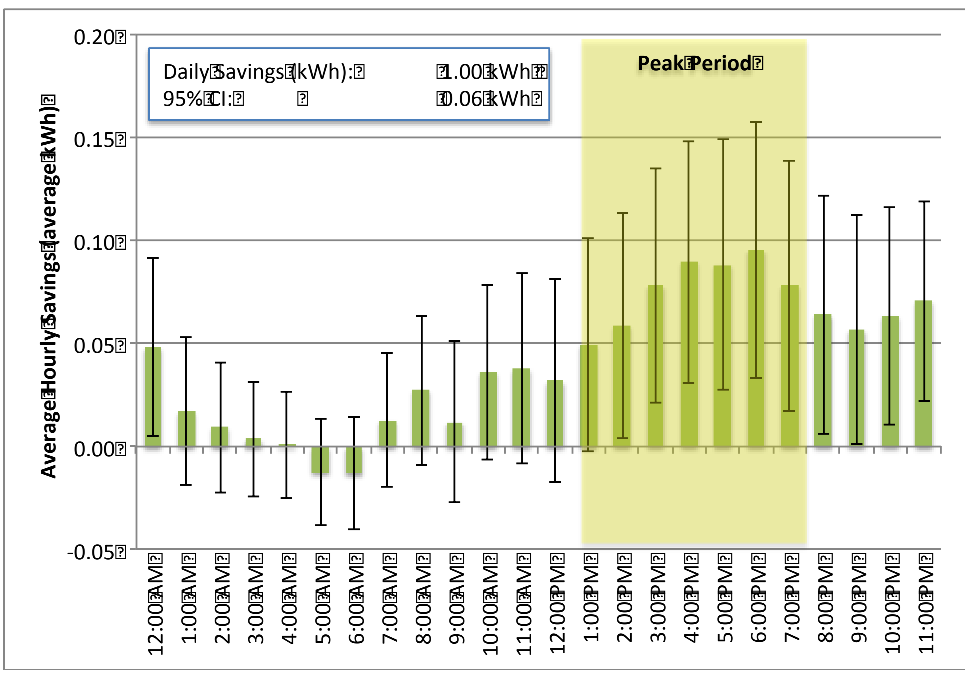

Figure 6. PG&E QM Program Annual Hourly Savings Estimates with Error Bars

Figure 7 shows the predicted load shape (red) and the actual consumption (blue) by season for the QM Program participants. 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 as expressed 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.

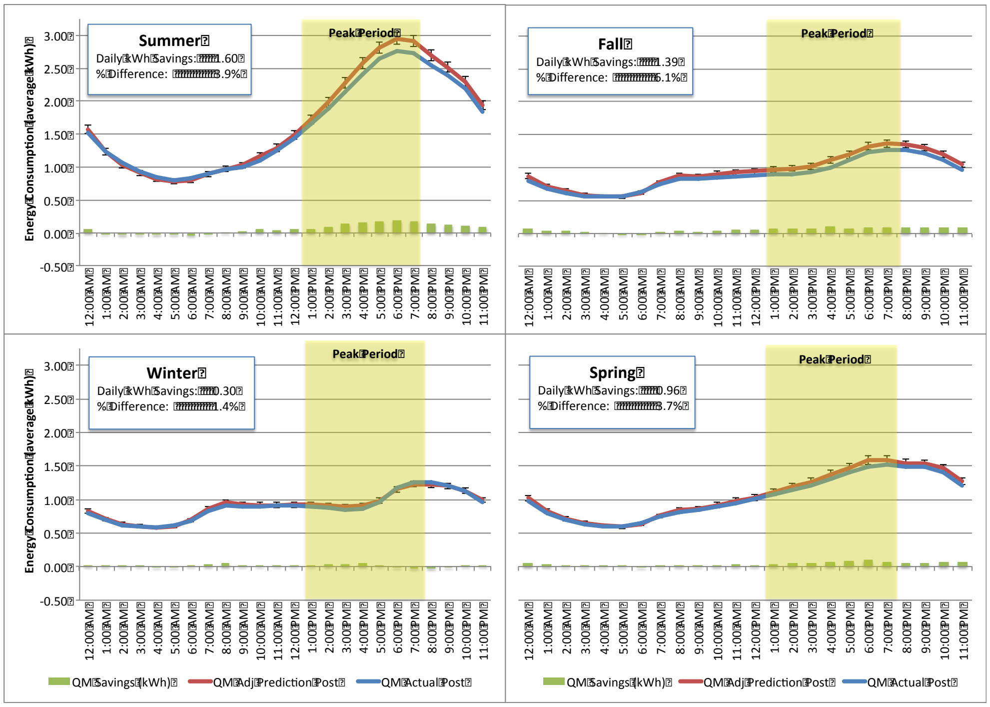

Figure 7. PG&E QM Program Annual Model Results by Season

Summary and Conclusions

The results of this exploratory research demonstrate enormous potential for the random coefficients model and represent a significant and positive departure from current approaches to analyzing AMI data and estimating program impacts. While analytically and conceptually more sophisticated than the traditional billing regression 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 multi-stage approach also provides an efficient means for processing the high volume of AMI data and organizing it for use in the billing regression model in a manner that controls for important sources of variation.

The test application of the random coefficients model using data from both the QI and QM Programs provided some encouraging results. First, the random coefficients model approach was able to produce very accurate predictions of load shapes, generally within 1 percent of actual use for a holdout sample of customers for both programs. Secondly, the random coefficients approach was able to produce savings estimates that were consistent with expectations for both programs. Finally, the predicted load shapes and impact estimates at the daily and seasonal level were also consistent with expectations and, in the case of the seasonal models, were able to account for significant differences in load profiles across periods.

Perhaps the greatest benefit of the random coefficients model is the ability to estimate impacts for different day types and seasons in an efficient manner. If only annual impact estimates are needed, then the fixed effects model is likely sufficient for estimating savings. However, a single annual savings number does not take full advantage of the information provided by the AMI data. The random coefficients approach, in contrast, uses the AMI information to create customer subcategories that help control for significant amounts of variation and ultimately allows for accurate load shape predictions. Once this process is completed, it provides a flexible method that enables different load shapes (and subsequently impact estimates) to be developed easily for a wide range of different time periods.

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 these initial results are very promising, they also suggest that further research in other areas is warranted. Suggestions for the next research phase include the following:

- **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. Using commercial customer AMI data, an initial test of the random coefficients model can be done to determine how well the modeling process can predict commercial load shapes. To estimate energy savings, additional data collection will be needed to identify HVAC installation date and the number of HVAC units at a site. If HVAC end use meter data are available, the method can also be used to test how well the random coefficients model can be used to estimate commercial HVAC loads.
- **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 model should explore the effects of using a comparison group on the load shape forecasts and the impact 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. 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. Variations in the setpoint temperature (currently at 65 degrees Fahrenheit 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 (CPUC 2008), where historical customer billing data are used to forecast energy use during an event period, 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.

References

Agnew, K. and M. Goldberg. 2013. "Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol." *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures (UMP)*.

California Public Utilities Commission (CPUC). 2006. *The California Evaluation Framework.*

California Public Utilities Commission (CPUC). 2008. *Load Impact Estimation for Demand Response: Protocols and Regulatory Guidance.*

Granderson, J, PN Price, D. Jump, N. Addy, and M. Sohn. 2015. "Automated Measurement and Verification: Performance of Public Domain Whole-Building Electric Baseline Models." *Applied Energy* 144: 106-113.

Granderson, J., S. Touzani, C. Custodio, S. Fernandes, M. Sohn, and D. Jump. 2015. *Assessment of Automated Measurement and Verification (M&V) Methods*. Lawrence Berkeley National Laboratory, LBNL#-187225.

Hsiao, C., D. Mountain, M.W. Chan, K.Y. Tsui. 1989. “Modeling Ontario Regional Electricity System Demand Using a Mixed Fixed and Random Coefficients Approach.” In *Regional Science and Urban Economics* Volume 19, Issue 4: 565-587.

Nexant. 2014. *2013 Load Impact Evaluation for Pacific Gas and Electric Company’s SmartAC Program.* Prepared for Pacific Gas and Electric Company.

Snijders, Tom and R. J. Bosker. 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks: Sage Publications.