

Findings from the Opower/Honeywell Smart Thermostat Field Assessment

ET Project Number: ET11PGE3074



Meet your new Thermostat.

Manage your temperature settings with the FREE iPhone app and convenient remote-control thermostat.

Now that's smart!



Prepared For: Lucy Morris
Pacific Gas and Electric Company

Prepared By: Candice Churchwell and Michael Sullivan
Nexant, Inc.
101 Montgomery St., 15th Floor
San Francisco, CA 94104

Issued: July 24, 2014

© Copyright 2014 Pacific Gas and Electric Company. All rights reserved.



**Pacific Gas and
Electric Company®**

ACKNOWLEDGEMENTS

Pacific Gas and Electric Company's Emerging Technologies Program is responsible for this project. It was developed as part of Pacific Gas and Electric Company's Emerging Technology under internal project number ET11PGE3074. Candice Churchwell and Michael Sullivan of Nexant, Inc. conducted this technology evaluation for Pacific Gas and Electric Company with overall guidance and management from Michael Seelig and Lucy Morris. For more information on this project, contact LLAA@pge.com.

LEGAL NOTICE

This report was prepared for Pacific Gas and Electric Company for use by its employees and agents. Neither Pacific Gas and Electric Company nor any of its employees and agents:

- (1) makes any written or oral warranty, expressed or implied, including, but not limited to those concerning merchantability or fitness for a particular purpose;
- (2) assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, process, method, or policy contained herein; or
- (3) represents that its use would not infringe any privately owned rights, including, but not limited to, patents, trademarks, or copyrights.

ABBREVIATIONS AND ACRONYMS

HVAC	Heating, ventilation and air conditioning
PC	Personal computer
RCT	Randomized control trial

CONTENTS

EXECUTIVE SUMMARY	1
INTRODUCTION	4
BACKGROUND	6
EMERGING TECHNOLOGY/PRODUCT	7
ASSESSMENT OBJECTIVES	8
TECHNOLOGY/PRODUCT EVALUATION	9
TECHNICAL APPROACH/TEST METHODOLOGY	13
RESULTS	17
Energy Savings Estimates	17
Participant Surveys	19
Overall Project Experience	20
Opinions on the Energy Coaching Messages	21
User Experiences with the Smart Thermostat System	24
Reducing Energy Usage	26
Connectivity Problems and Customer Service	28
EVALUATIONS	29
RECOMMENDATIONS	31
APPENDIX	32
Fixed Effects Model Estimation	32
Lagged Dependent Variables Model Estimation	33
Consequences of Model Misspecification	34
What about Randomization?	35
REFERENCES	37

EXECUTIVE SUMMARY

This Emerging Technologies report documents the results of the Opower/Honeywell Smart Thermostat field assessment. The assessment was launched with two waves of recruitment in mid-to-late 2012. Under a recruit-and-deny research design, a retail intercept recruitment campaign resulted in 695 volunteer participants that were randomly assigned to a control group and 693 volunteers were randomly assigned to the treatment group. The assessment's treatment consisted of a professionally-installed Honeywell programmable communicating thermostat, or smart thermostat, that connects to the internet through the customer's home Wi-Fi network. The smart thermostat's connectivity enables the customer to control their heating, ventilation, and air conditioning (HVAC) system from any location using an iPhone or other smartphone that uses the Android operating system. The application, or app, running on the smartphone that enables remote HVAC control is provided by Opower. The Opower application features behavioral messaging designed to nudge the user towards making more efficient HVAC usage choices.

PROJECT GOAL

The objective of this project was to obtain a quantitative understanding of how customers interact with smart home technologies such as the Opower/Honeywell smart thermostat solution. Residential HVAC energy usage is a significant portion of overall household energy consumption and presents a prime end use to target for energy efficiency gains. A behavioral energy efficiency program centered on smart thermostats can be used to increase the percentage of thermostats that have heating and/or cooling schedules; increase the depth of setback temperatures; expand the length of setback periods; set a more efficient temperature during times when the customer is at home and active; and for when the customer is at home and asleep. This assessment in particular focused on estimating the energy savings that would result from encouraging more of these efficient HVAC usage behaviors. Another important objective for this project was to understand how customers engage with the smart thermostat system and to learn about their preferences for – and attitudes towards – the technology. Together, the energy savings estimates and research on customer engagement provided the basis for an overall assessment of the potential value of communications-enabled thermostats that are coupled with behavioral messaging strategies.

PROJECT DESCRIPTION

The approach for this technology assessment was that of a randomized control trial (RCT) in which the primary data of interest was the energy savings induced by the thermostat in the treatment group as compared with the control group. Also of interest were learnings on the customer experience with the technology. Data on the customer experience was gathered through two online surveys. The first survey was administered within two months of the thermostat's installation. The second survey was administered online at the end of the assessment in February 2014 after participants had at least one year of experience with the thermostat.

PROJECT FINDINGS/RESULTS

Hourly electric and daily gas interval data for nearly all assessment participants were made available for estimating the electricity and natural gas savings due to the smart thermostat. This dataset began in July 2011 and ended in February 2014. The panel regression estimates of energy savings took into account the differences in energy consumption patterns between the treatment and control group that existed prior to the onset of the treatment. The RCT research design for this project was implemented towards the end of eliminating such pre-treatment differences; whether these differences occurred as a result of a failure of the randomization protocol or simply arose by chance is unknown. Once these pre-existing differences were taken into account, no significant electricity or natural gas energy savings were found at the 95% confidence level.

Data gathered from the end-of-project survey revealed that even though respondents described the smartphone app as a highly valued function of the smart thermostat system, the importance of the usability of the thermostat wall unit itself remained very high. Over half of all open-ended commentary on the wall unit detailed specific complaints and areas for improvement. The website portal was not highly utilized – only 47% reported using it.

Participants were asked how much they like the energy coaching messages and how helpful they are. On a scale of 1 to 10, where 1 means "Not at all helpful/Do not like at all," and 10 means, "Very helpful/Like very much," customers gave the messaging mildly positive likability/helpfulness scores of 6.3 and 6.8, respectively. The most common type of commentary on the messaging was positive, but half of the praise comments were qualified to discount the value of the messages to the customer's decision making process.

Participants did take advantage of the smart thermostat system's capability for controlling the HVAC system remotely, reporting that they valued the ability to change thermostat settings from many different locations such as work, the car, on vacation, or even simply from bed.

More than half of survey respondents reported that they think they reduced their energy use as a result of having the smart thermostat system. The smartphone app was ranked the highest as the component that contributed most to their perceived savings.

Half of respondents reported having problems with connecting to, or controlling the thermostat, remotely. Some customers stated that the connection problems improved over time, or specifically, after the Z-Wave thermostat was changed out for the Wi-Fi thermostat.

Less than half, 41%, of respondents reported contacting customer service for assistance with their smart thermostat system. On average, these customers gave a moderately high rating of 7.76 (on a scale of 1 to 10 where 1 means "very bad" and 10 means "very good") for their customer service experiences.

PROJECT RECOMMENDATIONS

Future efforts to study and evaluate the next generation and iterations of this technology should focus on the following:

- Improving the connectivity between the thermostat and the internet;
- Ensuring that customers understand the basic operating functions of the thermostat so that it is allowed to run as a programmed thermostat, rather than simply being turned on and off like a manual thermostat;

- Future investigations into behavioral messaging-enabled thermostat systems should include a control group that does not receive the messaging so that the effects of the messaging itself can be evaluated; and
- Similarly, future tests of behavioral-messaging thermostat products should incorporate test-and-learn messaging approaches towards the end of more thermostats being used to their full energy-efficiency potential.

INTRODUCTION

The California Energy Commission (CEC) reports that heating, ventilation, and air conditioning (HVAC) usage accounts for 9% of household electricity consumption and 37% of household natural gas energy consumption in California. These findings are from the 2010 Residential Appliance Saturation Survey (RASS), which also shows that HVAC usage is in fact the fifth highest contributor to household electricity consumption and the second highest contributor to natural gas energy consumption in households across California (Palmgren, et al. 2010). Figure 1 presents shares of electricity and natural gas energy consumption for various household appliances as estimated by the 2010 RASS and illustrates that only lighting, refrigeration, electronics, pools, and spas have greater shares of the average California household's electricity consumption. Only water heating represents a larger share of household natural gas usage than HVAC. By accounting for such a significant portion of overall household energy usage, HVACs offer a prime opportunity for reducing overall household energy consumption through energy efficiency programs.

FIGURE 1. APPLIANCE SHARES OF HOUSEHOLD ENERGY USAGE

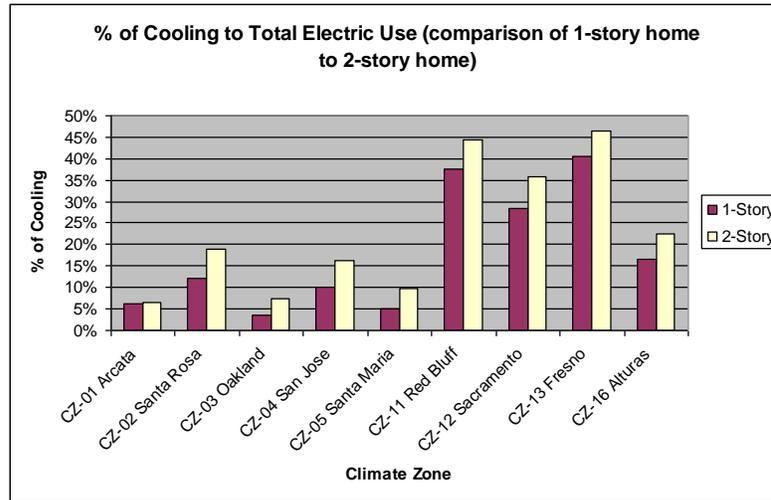


*Note: An estimate of 1,200 kWh per household (20% of the total consumption) has been designated as interior lighting and was shifted from the *Miscellaneous* category to *Lighting* category where it is combined with exterior lighting consumption. This number comes from other lighting studies* that are better able to pinpoint this estimate than a conditional demand model as was used for the RASS.

Source: 2010 California Residential Appliance Saturation Survey

But the share of household energy usage attributed to HVAC is weather-sensitive, varying widely across regions that experience hotter or cooler weather. Figure 2 shows the HVAC share of household electricity consumption for one- and two-story homes across several PG&E climate zones. Using building simulation models, PG&E estimates the HVAC share of household electricity usage is actually 30% or more in several of PG&E's climate zones, ranging as high as 45% for two-story homes in certain climate zones. As a result, the potential for HVAC energy efficiency savings is even higher in many regions of PG&E's service territory than across the state on average.

FIGURE 2. SHARE OF HOUSEHOLD ELECTRICITY USAGE BY CLIMATE ZONE AND NUMBER OF STORIES



Source: eQuest – Building Energy Use and Cost Analysis Software, developed by James J. Hirsch & Associates (JJH), version 3.64 was the latest release, <http://www.doe2.com/>

Long-established energy efficiency programs, such as rebates on new HVAC system purchases, encourage investment in more energy efficient technology. The goal of these programs is to reduce household energy consumption by introducing new HVAC units that use less energy into homes at faster adoption rates than would otherwise occur. But another approach is to reduce HVAC energy use through behavior change, that is, to change the way households use their HVAC systems so as to affect a reduction in HVAC energy usage. At the same time, the market for home automation platforms and products that can help customers manage, control, and optimize their household’s energy use is growing. This burgeoning customer interest in home automation, combined with increased utility experience with and confidence in the potential for behavioral energy efficiency programs, may be combined to become a promising venue for PG&E to engage with customers and help them optimize their HVAC energy use.

BACKGROUND

It is estimated that 33% of PG&E customers have programmable thermostats for air conditioning and 47% for heating, however, research also shows that the majority of customers do not know how to properly set their thermostats (Palmgren, et al. 2010). The benefits of using a programmable thermostat rely upon running heating and air conditioning according to a schedule, but approximately 50% of programmable thermostats have been found in some surveys to be set to long-term hold and 20% do not even have the correct time programmed (Meier, 2010). Energy Star, in fact, found that homes with programmable thermostats were using more energy than homes with manual thermostats; and as a result, terminated programmable thermostats from the Energy Star program (Meier, 2010).

Behavioral-orientated approaches to changing or influencing the way customers use electricity and natural gas include pricing, feedback, and control technologies. These programs are now playing a growing role in utility demand-side management (DSM) program portfolios; PG&E's rollout of Opower Home Energy Reports that provide comparative home energy analysis reports is a prime example. These reports have been shown to successfully motivate customers to become more efficient users of energy to the tune of 0.9% to 1.1% for electricity and 0.4% to 0.9% for natural gas (Perry and Woehleke, 2012). While these percentage energy savings may seem diminutive, aggregated across 550,000 customers, they add up to nearly 50 GWh and 1,500,000 therms during the period of August 2011 through December 2012 (Perry and Woehleke, 2012).

The approach used by the Opower Home Energy Reports seeks to motivate customers to lower their energy usage through periodic reports sent to the customer by U.S. Mail or email. The reports compare the customer's monthly electric and gas usage to an average of similar homes' usage as well as to an average of the most efficient 20% of similar homes' usage. The comparisons are known as "neighbor comparisons" and are based on a variety of customer characteristics such as location, home square footage, presence of a pool, and type and number of HVAC units. The general hypothesis about the energy savings mechanism behind the Home Energy Reports is that the neighbor comparisons provide a social motivation to customers to adjust their energy usage habits in multiple ways such as turning off lights or adjusting the thermostat a few degrees.

A behavioral strategy for lowering household HVAC energy consumption could seek to change the way that customers use the programmable thermostats and improve upon the energy savings achieved by programmable thermostats. A behavioral approach can make it easier for customers to use programmable thermostats as they are intended to encourage customers to reduce their HVAC energy consumption. Delivering feedback about HVAC usage behaviors can be used to increase the percentage of thermostats that have a heating or cooling schedule, increase the depth of setback temperatures, expand the length of setback periods, and set a more efficient temperature both when the customer is at home and active and at home and asleep. Towards these ends, Opower and Honeywell have produced a smart thermostat system that enhances a programmable thermostat with behavioral messaging. Further, this system also enables customers to control their thermostat from their mobile phone, personal computer (PC), or tablet; which also meets increasing customer demand for home automation technologies.

EMERGING TECHNOLOGY/PRODUCT

The Opower/Honeywell smart thermostat solution is a combination of a Honeywell internet-enabled thermostat with Opower behavioral software that features normative messaging and recommended thermostat settings through a smartphone application (also referred to as an app). The thermostat can be programmed either through the smartphone app, through a website on the customer's PC, a tablet app, or on the wall unit itself. Opower's normative messaging is designed to keep the customer better informed about their heating and cooling choices and the cost ramifications of those choices. The solution is designed to use consistent encouragement and feedback to motivate customers to reduce their energy consumption and, in turn, make it easy for customers to act on their motivation.

The thermostat wall unit in the home functions as a standard programmable thermostat, and allows the user to set wake/leave/return/sleep or wake/sleep schedules and temperature settings for each day of the week. The innovation of the solution lies in the simplicity of the process used to program the thermostat through the app or portal, and in the interaction with the Opower messaging. The messaging is designed to nudge customers towards using less energy through their HVAC usage by making suggestions. For example, the suggestions alert customers to opportunities for saving energy and money if they were to choose a more efficient thermostat setpoint.

This thermostat system would replace either a manual thermostat or a conventional programmable thermostat already installed in the home. The primary market barrier is that the customer must have broadband internet in order to use this product. As recently as 2009, approximately 35% of households in the United States do not have broadband internet service (NTIA 2010). An additional market barrier is price: home automation technologies in general are currently offered at a premium cost to the incumbent technologies they are designed to replace. Prices for home automation technology such as the Opower/Honeywell smart thermostat system are expected to decrease with greater market penetration.

ASSESSMENT OBJECTIVES

The objective of this Emerging Technology Project is to obtain a quantitative understanding of how customers interact with smart home technologies. This assessment focuses on energy savings as well as customer engagement with different types of messaging and user interfaces (with the smart phone, web portal, or at the thermostat wall unit). Energy savings estimates and findings on customer engagement, preferences, and attitudes provide the basis for an overall assessment of energy savings potential for communications-enabled thermostats coupled with behavioral messaging such as the technology tested here.

TECHNOLOGY/PRODUCT EVALUATION

The Opower/Honeywell smart thermostat system tested here is comprised of three primary customer-facing components: the thermostat wall unit, a website portal, and a smartphone and tablet app.

The thermostat wall unit is wired directly to the customer's HVAC system. The thermostat tested here is the Honeywell Z-Wave Touchscreen thermostat (model number TH8320ZW1000). The thermostat communicates with the website portal and the smartphone app (hosted by Opower) via the internet, but it cannot connect directly to the internet by itself. A Z-Wave gateway, the Mi Casa Verde gateway, is installed in conjunction with the thermostat so that it can communicate with the internet via the Z-Wave connection to the gateway.

Midway through this assessment, a number of participants began to experience connectivity issues with the original wall unit and gateway hardware combination. All assessment participants were contacted by Honeywell to make an appointment for the original Z-wave thermostat and gateway to be removed and replaced with another Honeywell communicating thermostat (TH8320ZW1000). This model of communicating thermostat connects directly with the customer's Wi-Fi router for access to the internet; a gateway device is not used in conjunction with the new thermostat.

Once connected to the internet, the customer can remotely program and control their thermostat using either a website portal or a smartphone app. The Program feature in the app and web portal allow the user to indicate their typical occupancy habits (at home all day, leave during the day, unpredictable) for each day of the week. They can then indicate the timing of those habits, and save the desired temperature setting for the times that they are home, away, and sleeping.

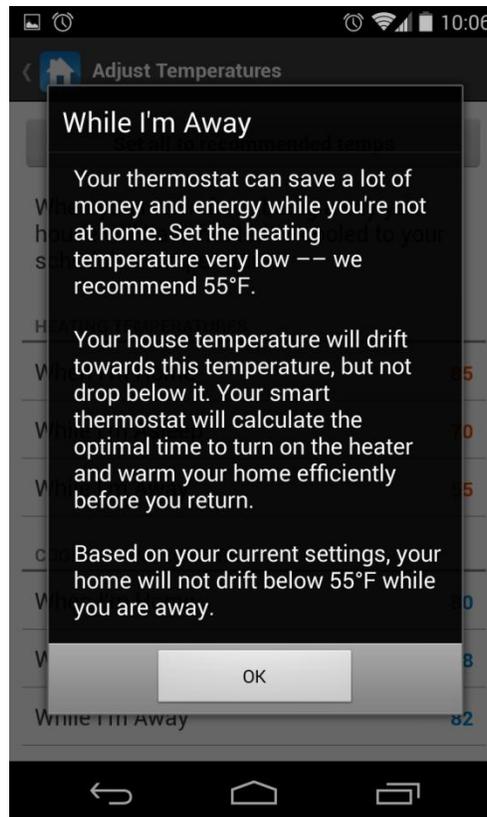
FIGURE 3. OCCUPANCY HABITS AND SCHEDULING

The screenshot shows a mobile application interface for programming a thermostat. At the top, the status bar shows 'AT&T' service, signal strength, Wi-Fi, and the time '11:48 AM'. Below the status bar is a dark header with a 'Program' button and the text 'Mondays'. The main content area is divided into two sections: 'On a typical Monday:' and 'My Monday schedule:'. The first section has three radio button options: 'I leave during the day.' (which is selected with a blue checkmark), 'I'm usually at home.', and 'It's unpredictable.'. The second section lists four events with their times and a right-pointing arrow for each: 'Wake up: 06:00 AM >', 'Leave home: 07:45 AM >', 'Return home: 05:00 PM >', and 'Bedtime: 10:00 PM >'.

Controlling the thermostat using the app or website portal can be done from any location where the device has internet access. Logging into the website portal from a computer at work, using the smartphone app in a coffee shop, or even simply using the smartphone app in any room in the home are all examples of locations and ways the customer can control their thermostat.

During the course of a customer using the website portal or app to make temporary or permanent changes to the thermostat's schedule for heating or cooling, Opower provides feedback messaging to the user, which is designed to nudge customers towards making energy efficient choices. For example, in setting the desired temperature for the home during the times of day when the customer is at home but asleep, customers are prompted by an Opower message to consider setting a lower temperature (for heating) because the thermostat is programmed to bring the house to the desired "at home and active" temperature prior to the time they wake up. The feedback also informs the customer how much money they would save by choosing the recommended "at home but asleep" temperature setting. Similar prompts appear when customers schedule temperatures for times when they are normally away from the home. Figure 4 shows an example of this type of message for the temperature setting for when the customer is away from home.

FIGURE 4. MESSAGING ENCOURAGING A LOWER SETPOINT WHILE AWAY FROM HOME



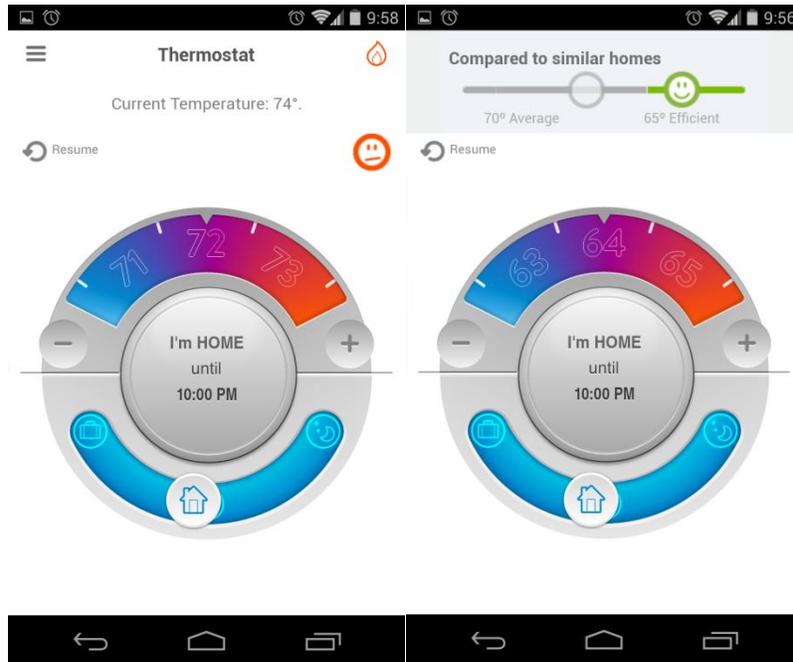
Customers can also use the app to manually override their saved schedule or temperature setting. For example, a customer might indicate that they are home on a day they are usually scheduled to be away, or that they want the current temperature to be cooler or hotter than the saved setpoint for that day/time and occupancy mode. If a customer chooses to set a more efficient temperature than is in their saved schedule, they are asked if they would like to adopt the more efficient temperature as their new saved temperature for that occupancy mode. For example, the message may read, "Would you like to save 79° as your usual temperature for Home? Not saving this temperature could cost you an extra \$20 this year."

Additionally, the app provides push notifications that can be sent five minutes before the heating or cooling system turns on to begin to transition between "Away" to "At Home" temperature settings. This notification allows users to set a later time for the system to activate if they are not going to be home soon. Push notifications are also available for fan usage, warning customers that their fan has been running for either 2, 8, or 24 hours. The notification suggests switching the fan setting to "Auto." There is also a feature that notifies customers via the app when someone has adjusted the thermostat when the thermostat had been set to "Away."

The Opower app additionally provides messaging that compares programmed setpoints to those of similar homes. One mechanism is an always-on indicator that provides the comparative status of the current setpoint to setpoints used at similar homes. Figure 5 shows an example of the comparison indicator in the case of an inefficient setpoint, followed

by another example of the app's indicator when moving the setpoint to a more efficient temperature.

FIGURE 5. COMPARATIVE MESSAGING INDICATOR FOR CURRENT SETPOINT



TECHNICAL APPROACH/TEST METHODOLOGY

The technical approach for this technology assessment was that of a randomized control trial (RCT) in which the primary data of interest were the energy savings induced by the thermostat in the treatment group as compared to the control group. Also of interest were several aspects of the customer experience with the technology, which include:

- How much participants liked the thermostat device and how useful they found it;
- How much they liked the smartphone app and website portal and how useful they found them;
- Whether participants liked the messages that were sent to them through the smartphone app; and
- Whether participants thought that the technology would lead them to save energy.

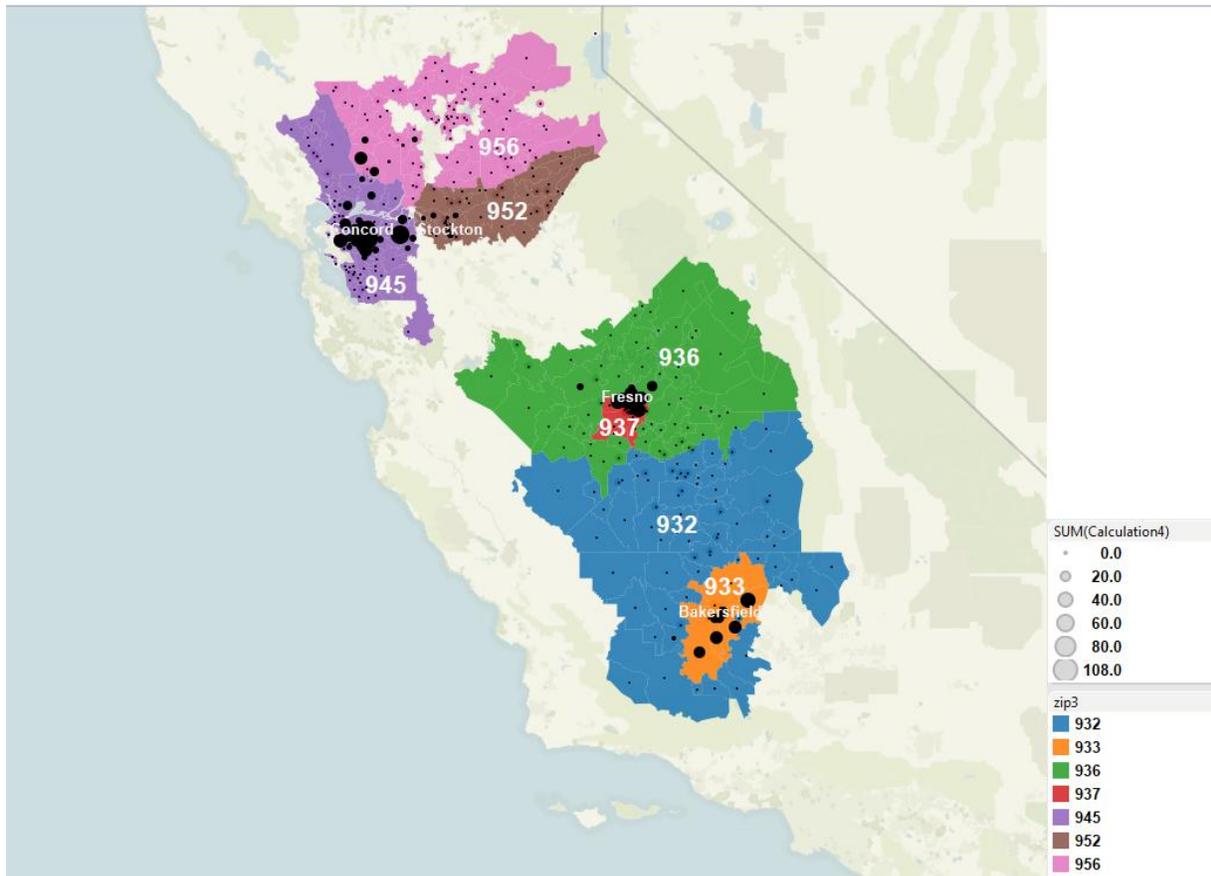
Data on these customer experience research questions were gathered using two surveys.

Prior to launching the recruitment and installation phase of the RCT, an alpha test of the Opower/Honeywell smart thermostat was conducted in order to test functional readiness of the thermostat and behavioral software solution in addition to assessing customer satisfaction with the installation process. All alpha test participants were PG&E employees; a total of 17 employee volunteers participated. Feedback from the participants was gathered through two surveys and one focus group and were consolidated for reporting purposes. Recommendations for addressing problems prior to the full RCT launch included:

- Improve follow-up communication from Opower or Honeywell to the customer when the thermostat or app doesn't work; and
- Test the smartphone app control as a part of the thermostat installation process; and
- Fully verify that the app controls the thermostat correctly prior to completing the installation.

The RCT assessment and evaluation of the Opower/Honeywell smart thermostat solution is made possible by the voluntary participation of PG&E customers in two geographically clustered areas of PG&E's service territory. Recruitment for the two areas was conducted in two waves. The East Bay/Northern Central Valley cluster of participants was recruited from ZIP codes starting with 945, 952, and 956 during the period of July 19, 2012 through September 30, 2012. Smart thermostat installations for customers assigned to the treatment group began in July and were completed in October 2012. The Southern Central Valley cluster of participants was recruited from ZIP codes beginning with digits 932, 933, 936, and 937. Recruitment and installations were completed in these ZIP codes between December 2012 and February 2013. Figure 6 shows that participants are clustered near the cities of Concord, Stockton, Fresno, and Bakersfield. These areas chosen for project eligibility were based on the expectation they would have a high number of eligible and interested customers who could realize measurable energy savings from using the smart thermostat system because of hotter climates and higher likelihood of AC usage.

FIGURE 6. GEOGRAPHIC DISTRIBUTION OF TREATMENT AND CONTROL CUSTOMERS



Other criteria that PG&E customers had to satisfy in order to take part in the assessment included:

- Had to receive electric and gas service from PG&E;
- Reside in select ZIP codes;
- Have operational central heating and air conditioning;
- Have a smartphone;¹
- Own a single-family home, townhouse, or condo;
- Have only one thermostat in the home; and
- Have high speed internet service at their home and an available port on their router.

The RCT framework of this research effort, coupled with fairly restrictive participation requirements of having a smartphone and high speed internet service at the home, necessitated a recruit and deny approach to finding study participants.

¹ The first wave of recruited customers specifically had to have an Apple iPhone. Eligibility was expanded to customers who have other smartphones that run the Android operating system for the second recruitment wave.

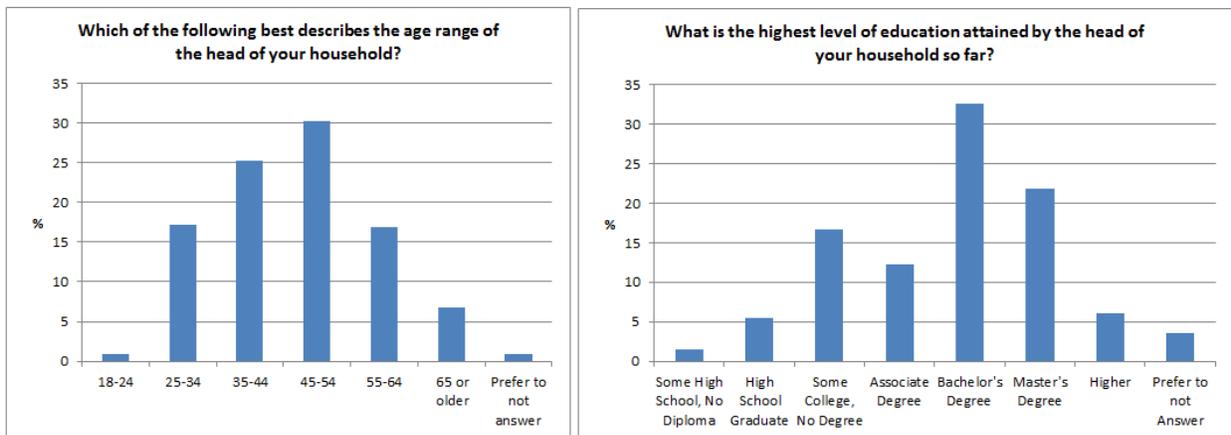
The recruit and deny process was facilitated by a subcontractor to Honeywell, GoGorilla Media. GoGorilla staged a series of retail intercept events outside retail stores, malls, festivals, and farmers' markets in the areas targeted for recruitment. The intercept events entailed GoGorilla staff's presence at a promotional display for the thermostat where they would invite people to talk to them about it and then attempt to persuade customers to sign up for the assessment. Customers who agreed to sign up entered their personal information into an onsite computer, including their PG&E account information. After a set of qualifying questions on the computer to determine eligibility, in addition to questions about their home thermostat settings, the customer was randomly assigned to either the treatment or control group. All participants had equal odds to be assigned to the treatment or control groups. Customers assigned to the treatment group then made an appointment to have a Honeywell technician come to their home and install the device.

A total of 693 customers were assigned to the treatment group and made appointments for installation of the smart thermostat. Of these, 505 had successful installations; the remaining 188 customers had unsuccessful, cancelled, or no-show installation appointments.

Treatment customers with successful installations filled out a short survey on the characteristics of their household upon completion of the installation. Completed surveys were received for 426 of 505 completed installations, a completion rate of 84%. These surveys revealed that smart thermostats were mostly (70%) installed in single-story homes with an average of 3.5 bedrooms. On average, 2.3 adults reside in the home along with 1.1 children.

Figure 7 presents the distributions of the age and educational attainment for the head of household for customers in the treatment group with successful smart thermostat installations.

FIGURE 7. AGE AND EDUCATIONAL ATTAINMENT OF TREATMENT CUSTOMERS WITH SUCCESSFUL INSTALLATIONS



Midway through the data collection period of the study, a number of project participants experienced connectivity issues between the thermostat and the website portal and smartphone app through the Z-Wave portal. While not every participant's thermostat system had this problem, all project participants were contacted by Honeywell and given the option to change out the thermostat at no charge. The equipment change involved the removal of the Z-Wave gateway from the system and replacement of the Honeywell Z-Wave Touchscreen thermostat with the Z-wave (TH8320ZW1000) that connects to the internet via the customer's home Wi-Fi network. Not all customers responded to the offer for the

equipment change out, and some customers refused it. Others reported moving out of the premise where the original thermostat was installed. In all, 423 of the 505 original thermostats were changed out for Wi-Fi-enabled thermostats.

RESULTS

ENERGY SAVINGS ESTIMATES

The data available for analysis consists of hourly electric and daily gas interval data for nearly all customers participating in the assessment. The data begins in July 2011 and ends in February 2014. Since all PCTs were installed by the end of February 2013, at least one full year of post-treatment interval data is available for all participants.

Gas and electric energy savings were estimated using a panel regression, specified with both fixed and time effect variables, and with error estimates clustered at the customer level. More formally, the model is specified below. The model variables are defined in Table 1 for the case of estimating electric energy savings. The model for estimating gas energy savings is the same, except the regress and $kW_{i,t}$ is replaced with $therms_{i,t}$ (monthly natural gas consumption of participant i during month t). The full context and rationale for the choice of this model specification, including consideration of other model specifications, is provided in Appendix A.

$$kW_{i,t} = b \cdot (T_i \cdot TreatmentPeriod_t) + c \cdot CDH_t + d \cdot (T_i \cdot CDH_t) + e \cdot HDH_t + f \cdot (T_i \cdot HDH_t) + g \cdot u_t + v_i + \epsilon_{it}$$

TABLE 1: DEFINITION OF MODEL VARIABLES

Variable	Definition
$kW_{i,t}$	average daily electric consumption of participant i during month t
b	estimated treatment effect
c, e	estimated effect of weather on treatment and control group
d, f	estimated effect of weather specific to treatment group, allowing the weather response to differ between treatment and control groups
g	estimated monthly time effect on treatment and control group
T_i	indicator of whether or not the participant is assigned to the treatment condition
$TreatmentPeriod_t$	indicator of whether or not the treatment was in effect during month t
u_t	Time effects for each month that control for unobserved factors that are common to all treatment and control customers but unique to month t
v_i	Participants' fixed effects that control for unobserved factors that are time-invariant and unique to each customer. Fixed effects do not control for fixed characteristics such as air conditioning that interact with time-varying factors such as weather
$\epsilon_{i,t}$	error for each participant and month

While this assessment was designed as a randomized controlled trial, approximately 30% of the participants assigned to the treatment condition did not complete the PCT installation process. With such a large percentage of the treatment group failing to receive the treatment, it is necessary to interpret the experiment as a randomized encouragement design (RED). Estimation of the impact of the treatment in a RED design is a two-step

process. In the first step, the difference between the treatment group and the control group is found. This step includes all treatment and control customers in the analysis regardless of whether or not the PCT equipment was successfully installed. The difference between treated customers and customers who were not treated is then found in a second step in which the intent to treat (ITT) difference is inflated by dividing it by the success rate for treatment – in this case about 70%. This second estimate is referred to as the effect on treated (EOT) effect. Tables 2 and 3 show Nexant's ITT savings estimates, where positive values indicate energy savings. The ITT estimate for electric energy savings is about 0.18 kWh per day, representing 0.7% electricity savings. This estimate for electric energy savings is not significant at the 95% confidence level. The estimate for gas energy savings is negative, 0.02 therms per day, representing a negative 1.4% savings. This estimate of gas energy savings is also not statistically different from zero at the 95% confidence level.

TABLE 2. ELECTRIC ENERGY SAVINGS ESTIMATED BY FIXED EFFECTS PLUS WEATHER MODEL – ITT

Regression	Absolute Daily Impact (kWh)	Percent Impact	Standard Error	95% Conf. Lower Bound	95% Conf. Upper Bound
FE with weather & weather interacted with treatment	0.18	0.7%	1.1%	-1.5%	2.9%

TABLE 3. GAS ENERGY SAVINGS ESTIMATED BY FIXED EFFECTS PLUS WEATHER MODEL – ITT

Regression	Absolute Daily Impact (therms)	Percent Impact	Standard Error	95% Conf. Lower Bound	95% Conf. Upper Bound
FE with weather & weather interacted with treatment	-0.02	-1.4%	0.9%	-3.2%	0.3%

Tables 4 and 5 present the effect-on-treated estimates of electric and gas energy savings. The intent-to-treat estimates and the effect-on-treated estimates are presented because they are both of interest for evaluating the outcome of the assessment and for planning future programs or pilots. Like the ITT estimates on which they are based, these estimates of the effect-on-treated participants are not statistically different than zero at the 95% confidence level.

TABLE 4. ELECTRIC ENERGY SAVINGS ESTIMATED BY FIXED EFFECTS PLUS WEATHER MODEL – EFFECT-ON-TREATED

Regression	Absolute Daily Impact (kWh)	Percent Impact	Standard Error	95% Conf. Lower Bound	95% Conf. Upper Bound
FE with weather & weather interacted with treatment	0.25	1.0%	1.1%	-1.3%	3.2%

TABLE 5. GAS ENERGY SAVINGS ESTIMATED BY FIXED EFFECTS PLUS WEATHER MODEL – EFFECT-ON-TREATED

Regression	Absolute Daily Impact (therms)	Percent Impact	Standard Error	95% Conf. Lower Bound	95% Conf. Upper Bound
FE with weather & weather interacted with treatment	-0.03	-2.0%	0.9%	-3.7%	-0.3%

PARTICIPANT SURVEYS

An important objective of this Emerging Technologies project is to provide insights into residential customer preferences and attitudes toward enabling technologies. Interval data made available through each participant's SmartMeter was the primary source of data to support the estimation of energy savings due to this technology. The primary source of information used for learning about the project participants engagement with this smart thermostat solution are datasets collected through two surveys administered to the project participants. None of the customers assigned to the control group were surveyed, and only those customers assigned to the treatment group who had successful thermostat installations were surveyed.

The first survey was administered to project participants within a month or two of the thermostat's installation. For the first group of participants in the northern central valley, this survey was in-field during the month of November 2012. The second group of participants in the southern central valley received survey invitations in the month of March 2013. The survey was accessed on the web by navigating to a URL that was provided to each participant in their survey invitation; all survey invitations and follow-up reminders were delivered by email. The northern central valley group's response rate was 52% while 40% of the southern central valley participants completed the survey; the overall response rate to the first survey was 47%. The first survey included 38 questions, some of which were open-ended and optional; the average completion time of the first survey was 15 minutes and 15 seconds.²

The second survey was administered at the conclusion of the assessment in February 2014. Like the first survey, the second survey was administered online and announced by email. The second survey was completed by 242 treatment customers, representing a 48% completion rate. The second survey was slightly shorter, with 31 questions and, on average, was completed in 16 minutes and 30 seconds.³

The first survey inquired about the participants' enrollment experience, initial impressions of the thermostat wall unit, app, website portal, and of the behavioral messaging. The survey also asked about participants' opinions on the usability of the smart thermostat system and whether or not they thought that it was helping them use less energy. Participants' experiences with assessment-related customer service and opinions of educational materials were also areas of inquiry in the first survey.

The second survey, administered at the close of the assessment, dropped or de-emphasized the enrollment, customer service, and educational topics that were included in the first survey. Questions on opinions of each component (wall unit, app, website portal, and behavioral messaging) were revisited and expanded. More questions about the participants' opinions on the usability of the system, in addition to questions about how the participants were using the system, were added. Questions about whether participants believe the system helps them use less energy were included again and expanded upon.

Summarized results of the first survey were presented in an earlier Emerging Technologies project report (Perry and Oh, 2012). The remainder of this section presents the findings of the second survey, including comparisons of results with the first survey where applicable.

² Five customers' survey completion times were removed from this average. The survey instrument recorded completion times in excess of three hours for these five customers.

³ Eight customers' survey completion times were removed from this average using the same filter that was used in the first survey.

Survey results are presented below in five topical areas: overall assessment experience; opinions of the energy coaching messages; user experiences with the smart thermostat system; reducing energy usage; and customer service.

OVERALL ASSESSMENT EXPERIENCE

Participants were asked first whether they have used all three components of the smart thermostat system that was made available to them: the thermostat wall unit; the smartphone app; and the website portal. Nearly all of the second survey's 242 respondents, 98%, reported ever using the thermostat wall unit and the smartphone app, but only 47% reported ever using the website portal.

Respondents were asked to score how much they liked each of the system's components on a score of 1 to 10, where 1 means "do not like at all" and 10 means "like very much." The average "liked" score of 8.4 was the highest for the app, followed by an average score of 7.7 for the wall unit and an average score of 7.6 for the portal. When comparing the answers of the 161 customers who took both surveys and accepted the Wi-Fi thermostat upgrade, the scores were slightly different. On the first survey, the customers rated the wall unit, app, and portal 8.6, 8.6, and 7.8, respectively. These same customers rated all of the three components lower in the second survey, at 7.7, 8.4, and 7.6. The only statistically significant drop was the drop in the likeability score of the thermostat wall unit. It is impossible to conclusively determine whether the customers' reported satisfaction with the wall unit declined as a result of the installation of the new version or whether the overall experience of the wall unit was tainted by the performance of the malfunctioning first version.

Respondents were then asked to articulate the rationale behind their likability scores. Negative commentary had the highest profile among comments about the thermostat wall unit; around half of responses were negative. Commonly cited issues with the wall unit were:

- The thermostat's programming aims to heat or cool the house to a certain temperature by a certain time of day, or not to begin heating or cooling at that time. Some customers complained of their heater coming on too early and waking them up;
- The wall unit can only be programmed via the app or portal and not directly at the unit itself;
- The new thermostat has no clock;
- The new thermostat has a backlight that never turns off;
- Display is difficult to read or see;
- Touch screen is not responsive enough;
- Buttons are too small;
- Dependency on the home's Wi-Fi network; and
- Changes to the Wi-Fi network resulted in the need to call customer service to reconnect the thermostat.

On the other hand, many of the comments about the thermostat were positive and they were often simply stated as "easy to use" or "very convenient." A few customers in fact said the always-on backlight was a nice feature for their household.

Respondents generally praised the smartphone app as easy to use, convenient, simple, or user-friendly. But some complaints did surface, including:

- Loss of connectivity with the thermostat or making changes on the app but the HVAC system not responding;
- The need to enter a password too frequently; and
- Vacation mode was lost with the second thermostat.

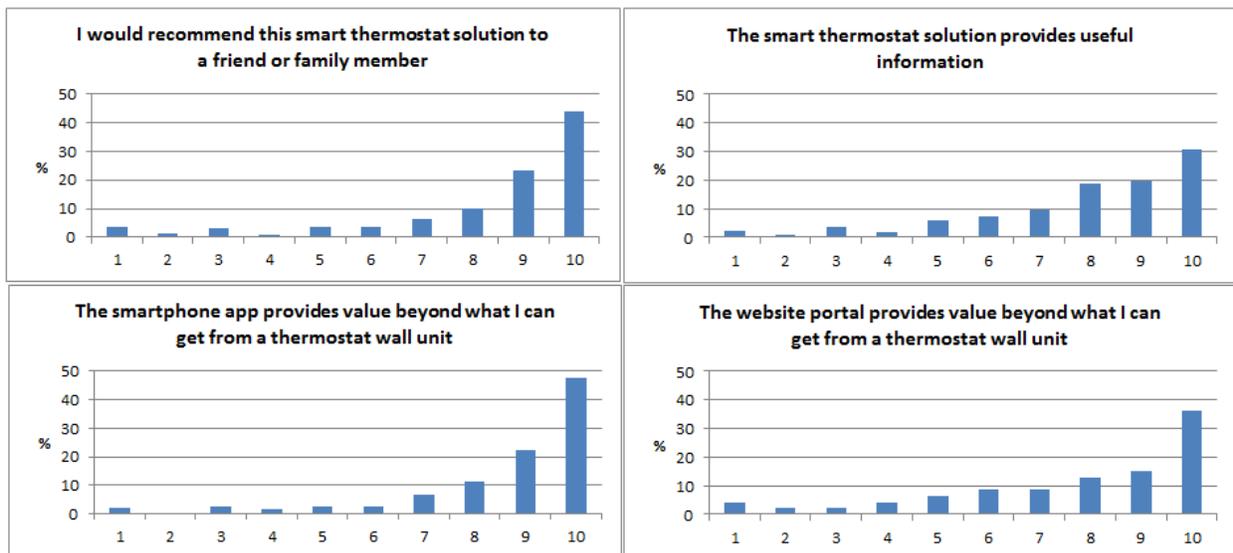
A limited number of users also stated that they wished they could program more scenarios than HOME, SLEEP, and AWAY.

Most of the participants who used the portal gave the portal positive feedback, but they reported that they preferred using the app. Some respondents said they lost the URL and a few others mentioned that they used a website as an alternative way to access the thermostat when app connectivity was down.

In general, the components of the smart thermostat system tested here moderately met participants' expectations: the average agreement score (on a scale of 1 to 10 where 1 means "strongly disagree" and 10 means "strongly agree") for the statement that a system component has met their expectations was 7.6 for the wall unit; 8.2 for the app; and 7.9 for the portal. The percentage of customers reporting agreement in the "top two boxes" is 48%, 63%, and 57% for the wall unit, app, and portal, respectively.

Agreement with other statements describing the value and usefulness of the smart thermostat follow a similar distribution along the scale of ratings. Figure 8 shows the distribution of agreement scores for four questions related to the perceived value and usefulness of the technology tested in this trial.

FIGURE 8 . AGREEMENT SCORES WITH SURVEY QUESTIONS PERTAINING TO VALUE AND USEFULNESS, WHERE 1 MEANS STRONG DISAGREEMENT AND 10 MEANS STRONG AGREEMENT



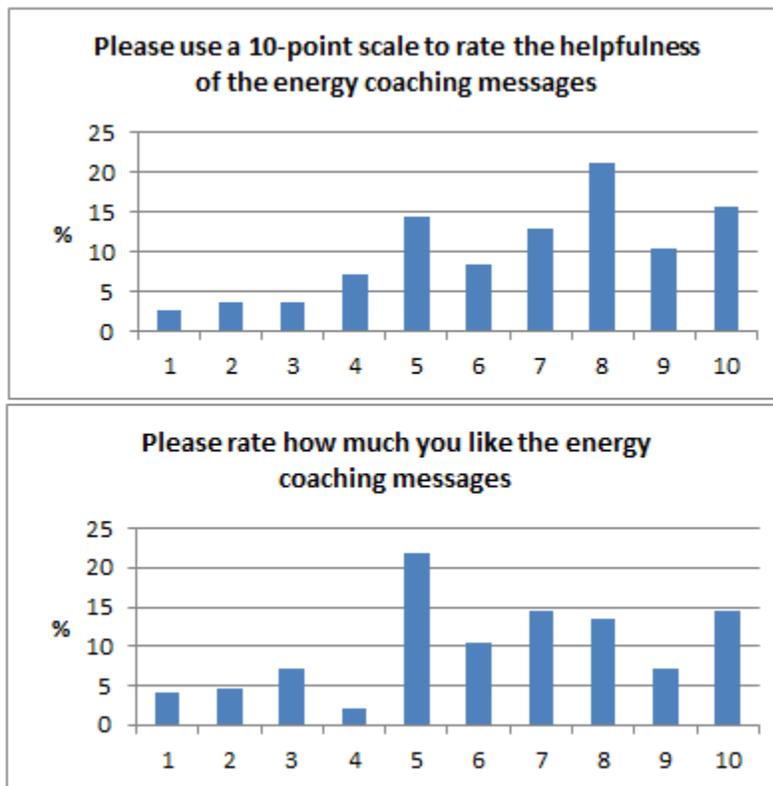
OPINIONS ON THE ENERGY COACHING MESSAGES

A primary component of the smart thermostat system tested in this assessment were Opower's energy coaching messages that appears when thermostat setpoints were made using the app and website portal. These messages are intended to guide or encourage

users to making thermostat setpoint choices that resulted in lower HVAC energy consumption than what they otherwise would have used without the coaching.

The second survey asked for participants' opinions specifically about the energy coaching messages. A large majority (81%) of respondents recalled seeing the messages; these customers were asked to rate on a scale of 1 to 10 how helpful they thought the messages were and how much they liked them (where 1 means "not at all" and 10 means "very"). Respondents on average gave mildly positive scores for the helpfulness and likeability of the coaching. Average scores of 6.8 and 6.3 were given for helpfulness and likability, respectively; Figure 9 presents the distribution of these scores. Responses indicate that customers were more ambivalent about how much they like the messaging than they were about how helpful they were: 15% of customers gave a neutral score of 5 for the helpfulness of the coaching and 22% of customers gave them a neutral score of 5 for whether they liked it or not.

FIGURE 9. HELPFULNESS AND LIKEABILITY OF THE ENERGY COACHING MESSAGES, , WHERE 1 MEANS "NOT HELPFUL AT ALL/DO NOT LIKE AT ALL" AND 10 MEANS "VERY HELPFUL/LIKE VERY MUCH/"



Open-ended comments were solicited from respondents about their scores for how helpful the coaching was for them and for how much they liked the energy coaching. The most common type of commentary about the messages was positive in nature, but approximately half of the statements of praise or appreciation for the coaching were qualified in the same breath that mitigate the coaching's value to the respondent. This style of commentary commonly went along the following lines:

- "As much as they are informative, they aren't enough to convince me to change my routine."

- "Somewhat helpful, but doesn't really influence my decision as to what temperature to keep my house at."
- "I like the idea of knowing the suggested optimal temperature of my house, but if I feel too hot or cold that is not going to change my mind."

Negative opinions or opinions stating that the messages were not needed were the next most common response. Responses were varied but some examples reflecting typical sentiments of this type include:

- "Do not really pay attention to them."
- "The messages become annoying, as if there is no setting (other than OFF) that will make the program happy!"
- "It doesn't really tell me anything I don't already know."
- "It is condescending."
- "My house always feels cold and they just make me feel guilty that I don't have it set lower."

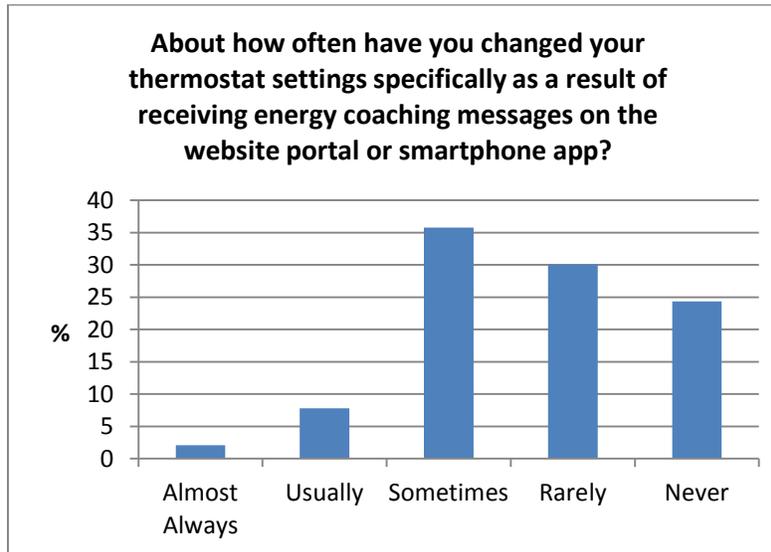
The type of comment that was least given, reflected unqualified adherence to the coaching's guidance or a middle-of-the-road acceptance, along the lines of, "the information was good, sometimes I took the advice of the coaching." A few examples of this rarer type of comment were:

- "I usually go with the message in the hopes of saving money and energy."
- "Helpful. They at least keep you aware of what others are doing around you and sometimes you'll dial it down a notch."
- "I guess I like the suggestions but sometimes I want it warmer or colder."

The survey separately asked for comments about how much the customer thought the messages were helpful and how much they liked the messages. Most respondents did not distinguish between the two concepts in their comments. At the prompt for comments on how much the respondent liked the energy coaching, many responses referred back to their preceding comments on the helpfulness of the coaching; others repeated their comments. Probing the distinction between the behavioral messaging being liked and perceived as helpful would need to occur in a focus group setting where a more in-depth conversation about the respondent's thinking is possible.

Finally, the survey also inquired about how often participants changed their thermostat settings specifically as a result of receiving the energy coaching; about 35% of respondents reported that they "sometimes" do so. The distribution of responses to this question are shown in Figure 10. This question offers an interesting point of reference to the open-ended comments solicited just before it in the survey: 55% reported here that they "rarely" or "never" change settings as a result of the energy coaching, echoing the responses in the earlier open-ended question indicating that the messaging was not very influential for many participants.

FIGURE 10. FREQUENCY OF CHANGING SETTINGS IN RESPONSE TO ENERGY COACHING



USER EXPERIENCES WITH THE SMART THERMOSTAT SYSTEM

The smart thermostat tested in this assessment has another very important feature designed to help participants reduce their HVAC-related energy consumption: controlling the thermostat remotely from the thermostat itself. This feature was available through both the website portal and the smartphone app, but early questions in the survey revealed that the website was little used by participants of this assessment. The app and portal allowed the customer to control the thermostat from any location where their PC, tablet, or smartphone could access the internet. Customers reported in this survey that they did take advantage of this feature, changing their thermostat settings from many locations like work, the car, on vacation, or even simply from bed.

The second survey asked a series of questions about how often different kinds of thermostat adjustments were made and where the adjustment was made – at the wall unit, with the app, or on the website portal:

- Changing the current temperature setpoint;
- Changing the current occupancy modes between Home, Away, and Asleep occupancy modes;
- Changing the scheduled temperatures for Home, Away, and Asleep occupancy modes;
- Changing the scheduled times for the Home, Away, and Asleep occupancy modes;
- Changing the system mode between Heat, Cool, or Off; and
- Changing the fan setting between Auto, Circulate, or On.

The most frequently made system adjustment is changing the current temperature setpoint – 53% of respondents reported changing the setpoint at least a few times a week. Respondents were about equally likely to report changing the setpoint using the wall unit or the app.

While 20% of customers said they never changed the occupancy mode between Home, Away, and Asleep, 64% said they did so a few times a month, monthly, or less often than

monthly. Nearly three-quarters of respondents also reported changing the saved temperatures for Home, Away, and Asleep occupancy modes a few times a month, monthly, or less often than monthly. Similarly, 75% reported changing the scheduled times for occupancy modes at that frequency. For all three of these activities, between 78% and 86% of respondents reported that they usually did them using the app.

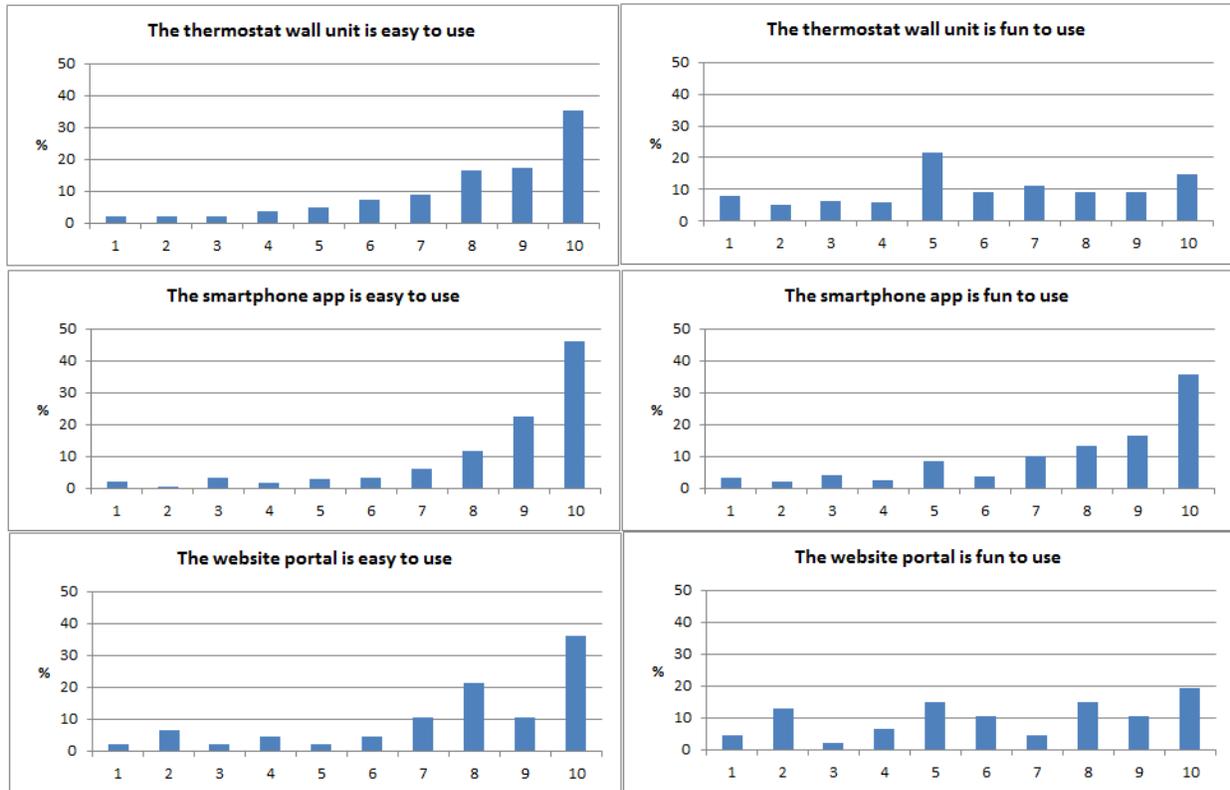
Thermostat activities that are reported to occur with even less frequency were changing the system mode between Heat, Cool, or Off and changing the fan setting between Auto, Circulate, or On. 58% say they never changed their fan setting. 56% said they changed the system mode less often than monthly – this activity was just as likely to be reported to be done with the wall unit as with the app.

The survey asked about agreement with the statements that each element of the smart thermostat system was easy or fun to use, where 1 means “strongly disagree” and 10 means “strongly agree.”

The app was rated the easiest to use, with an average agreement score of 8.5, while the thermostat unit was rated 8.0 in ease-of-use. Only 47 respondents reported using the portal, and these customers rated it 7.8 with respect to agreement with the statement that the portal was easy to use.

The app was reported to be more fun to use than the thermostat, and by a greater margin than ease-of-use: the app received an average “fun” score of 7.8 and the thermostat received a fairly low average score of 6.0. In fact, 25% of respondents indicated that they disagreed with the statement that the thermostat was fun to use. The 47 respondents who used the web portal gave it a fun score of 6.4 on average, which is also fairly low. Figure 11 presents the distributions of the responses to these questions and clearly indicates that the app was viewed as the easiest and most fun to use component of the system.

FIGURE 11. OPINIONS ON AGREEMENT WITH STATEMENTS ABOUT USING EACH SMART THERMOSTAT SYSTEM COMPONENT, WHERE 1 MEANS STRONG DISAGREEMENT AND 10 MEANS STRONG AGREEMENT

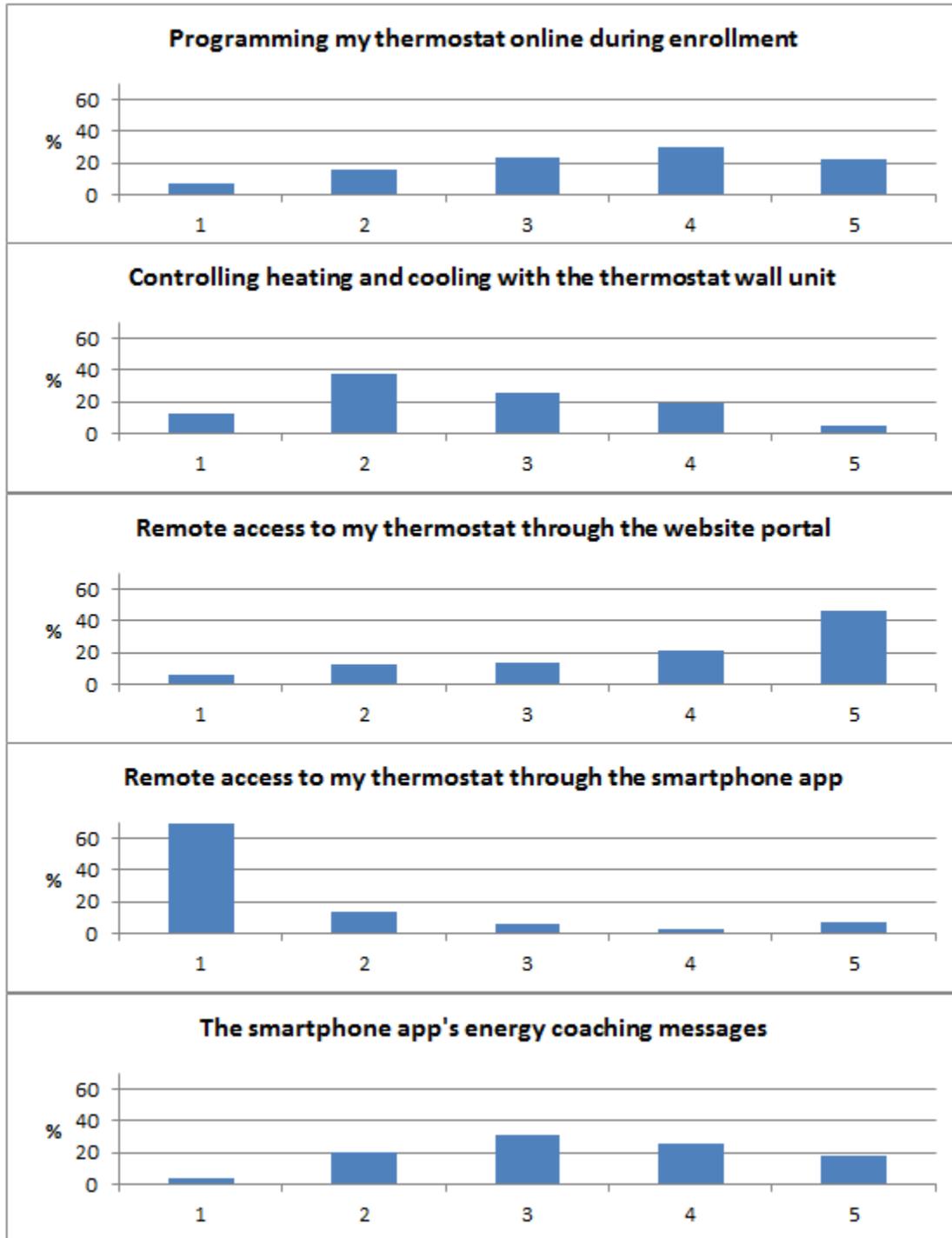


REDUCING ENERGY USAGE

More than half, 65%, of respondents reported that they thought they reduced their energy use as a result of using the smart thermostat system. The survey further asked those that thought the smart thermostat did help them save energy why that was the case. Most comments cited the system’s ability to control the HVAC system away from the house, allowing them to turn the HVAC system off remotely when they had forgotten prior to leaving the house, or turning it on and off related to being away for vacation. A few respondents reported liking the ability to adjust thermostat settings without getting out of bed. Other comments mentioned appreciation of the energy coaching’s intent to help save energy and to give the customer greater energy awareness. Other comments simply stated that having greater control of the system led to energy savings.

To further gauge which components of the smart thermostat solution contributed most to enabling participants to reduce energy usage, the survey also asked the customer to rank each smart thermostat system element with respect to its impact on energy usage. A rank of 1 indicated the greatest impact on energy use and a rank of 5 indicated the least impact. Figure 12 presents the average rankings for each system component. Respondents reported that the web portal had the least impact on energy use and the app is shown to have the most impact on energy use. In fact, nearly 70% of respondents cited the app as having the greatest impact. Using the wall unit was also perceived to have a high impact on energy savings in the opinion of the users. Programming the thermostat online during the enrollment process, and the energy coaching messages, were both viewed as having moderate or low impact on energy savings.

FIGURE 12. RANKING THE IMPACT OF SMART THERMOSTAT ACTIVITIES ON PERCEIVED ENERGY SAVINGS, WHERE 1 INDICATES THE ACTIVITY THAT HAD THE GREATEST IMPACT AND 5 INDICATES THE ACTIVITY THAT HAD THE LEAST IMPACT



CONNECTIVITY PROBLEMS AND CUSTOMER SERVICE

Half of respondents reported having problems connecting with or controlling the thermostat remotely. Nearly all respondents who reported problems offered some description of what was wrong, but relatively few articulated particulars about the connectivity problem they experienced; the most common description was along the lines of "sometimes it (the app) can't connect." When customers did describe what was going wrong, they mostly indicated that the trouble involved their router, and others indicated that the problem involved passwords. Some customers stated that connection problems improved over time or, specifically, after the thermostat was changed out; a small number of customers said that the new Wi-Fi thermostat resulted in worse connectivity than the Z-wave thermostat.

Less than half (41%) of respondents called the customer service phone line for support at any point in the assessment. Most (95%) respondents who said they called customer service gave a description of why. Most common descriptions were:

- Needed help getting the app to connect or how to install it;
- Needed help with issues related to changing their router or Wi-Fi network;
- Problems with the first Z-wave thermostat;
- The thermostat was not operating as expected. One issue that surfaced repeatedly was regarding the Wi-Fi thermostat's operating protocol that brings the home to a desired temperature by a certain time, rather than beginning to cool or heat at that time. This caused heating or cooling to begin earlier than expected and was a surprise to some respondents.

The mean rating for customer service was 7.76 (a 1 through 10 scale where 1 is "very bad" and 10 is "very good." Most customers who rated customer service less than 8 gave a reason. Chief complaints were:

- Too many transfers to different people;
- Problems never got solved; and
- No response from call center staff or non-responsive handling of complaints.

EVALUATIONS

Central to this Emerging Technology assessment is the estimate of energy savings that resulted from customers using the Opower/Honeywell smart thermostat solution. The data gathered during the course of this study do not provide convincing evidence that significant changes in energy consumption occurred due to the thermostat solution. It is possible that a small treatment effect occurred but that the relatively small sample size, and systematic differences between the treatment and control groups, have masked it. Only experiments with larger sample sizes can improve the likelihood of estimating small effects with standard statistical uncertainty (i.e., at the 95% level of confidence).

However, the other objective of this assessment was to obtain an understanding of customer experiences with, and attitudes towards, this emerging home automation/smart home technology. Survey results clearly revealed that customers liked this technology's functionality that allowed them to control their HVAC system from any room of the house, at work or on vacation. This same functionality was offered via a website but was rarely used, which highlights the convenience that apps bring to mobile devices like smartphones.

Further probing on using the app, however, revealed that the technology tested here may not have been "ready for prime time," or at a minimum revealed the weaknesses home automation systems such as the one tested here still face. The novel functionality of remote HVAC control hinges on connecting the thermostat to the internet. Keeping the thermostat connected was the fundamental challenge for this assessment, so much so that the original equipment needed to be changed out for more robust technology. Once the thermostat was changed out to another that connects to the customer's Wi-Fi network, customers vocalized frustration with the thermostat's dependency on the fact that their Wi-Fi must also be in working order and that any changes to that network cascade to affect the thermostat. Another common complaint centered on the further communication hop that the system depends on: the app must be connected to Opower servers via either the cell network or home Wi-Fi. Many customers found that the connectivity of the app was not reliable. Together, the as-intended user experience relies on two communication connections that in most cases customers are neither prepared to, nor interested in, troubleshooting. These are currently the weakest links for home automation products such as the one tested here. Providers will have to improve their customer service to help customers quickly overcome barriers that may arise pertaining to connectivity, in order to keep them satisfied and engaged with the product.

This research also revealed how important the thermostat wall unit still is. Customers were very vocal with their praise and complaints of the wall unit in the survey setting. While new thermostat products featuring attractive design, and endowed with blockbuster marketing efforts such as the Nest, are beginning to woo customers towards wanting a new programmable or smart thermostat, the key for smart thermostats in the setting of the utility energy efficiency portfolio may not necessarily lie in attracting customers to a new thermostat. Instead, keeping customers engaged with the thermostat, and using the thermostat as it is intended, are crucial for realizing energy savings.

Compelling customers to use programmable thermostats in the right way is at the heart of the matter for these devices to deliver energy efficiency savings. Towards this end, this research suggests that the thermostat tested still has room to improve with respect to the size and readability of the screen, the size of the buttons, and connectivity with the internet. Also of note was the perception that the thermostat could not be programmed directly on the unit itself. While the thermostat model did in fact allow for 7-day programming, the

process was complicated and not clearly explained in the User Manual, thus leading many customers to believe it could only be programmed using the app or the website.

The "anticipation" function of this thermostat was also not understood or appreciated by users. Some customers were surprised by it, while others would have preferred that the system simply turn on at a certain time rather than turning on early in order to bring the home to a specific temperature by a certain time. This preference was sometimes noted to be in response to the simple fact that customers didn't like being woken up early when the furnace turned on. One conclusion drawn from these reported experiences is that the participants should have been better educated about how their thermostat works; another is that the anticipation feature should be configurable or optional.

While smart thermostats face the challenges enumerated above, they are nonetheless poised to ride a growing wave of interest in home automation that is largely facilitated by mobile computing through devices like smartphones and tablets. While these products depend on the customer having broadband internet service at the home, the number of customers with this service is growing and interest in home automation products is also growing. HVAC energy consumption remains an opportune end use for energy efficiency programs to target with behavioral programs, and the intersection of more efficient HVAC usage with home automation technology could be a golden opportunity for PG&E to better serve their customers and diversify their energy efficiency portfolio performance.

RECOMMENDATIONS

The findings of this study do not show statistically significant energy savings attributable to the smart thermostat system tested here. However, the survey and usage data collected point to strong customer interest in the feature of remote control of the HVAC system, which fits into the developing and growing market for home automation technologies. As such, smart thermostats have great promise and attraction, but have yet not graduated to the status of tested and reliable providers of savings. Future efforts to study and evaluate the next generation, and iterations of this technology, should focus on the following:

- The connectivity between the thermostat and the internet, and in turn between the internet and the personal electronic device, must become more robust. At a minimum, more resources for customer service must be made available to help customers troubleshoot connection problems.
- The basic operating function of the thermostat program must be better explained to customers. In order to realize the energy efficiency gains from the thermostat, it must be allowed to run as a programmable thermostat, rather than simply being turned on and off like a manual thermostat.
- Future investigations into behavioral messaging-enabled thermostat systems should include a control group that does not receive the messaging so that the effects of the messaging itself can be evaluated.
 - Data should be gathered on the extent to which measurable energy efficiency goals are met: do customers who receive the messaging have a thermostat schedule in place; have deeper setback temperatures; have longer setback periods; and set more efficient set points when home or asleep than customers who do not receive the messaging?
- Similarly, future tests of behavioral messaging thermostat products should incorporate test-and-learn approaches to the effectiveness of the messages, and towards the end of more thermostats being used to their full energy-efficiency potential. This test-and-learn approach should include surveying to ascertain whether the setback suggestions are too deep, or the messaging too frequent, resulting in customers feeling guilty or irritated.

APPENDIX

The analytic approach taken in this evaluation uses panel data to determine energy savings. Panel data is a data structure where multiple observations over time are present for multiple individuals. However, it is possible to use different model specifications for using the panel data to estimate energy savings. The model specified in this evaluation is called a fixed-effects (FE) panel regression model, whereas another modeling possibility is called a post-only model, which is a panel regression model that incorporates lagged dependent variables (LDV) as explanatory variables. Each of these model specifications has merit under the appropriate circumstances, but they are fundamentally different approaches to estimating treatment impacts.

FIXED EFFECTS MODEL ESTIMATION

To understand the differences between these two modeling approaches, begin with a description of fixed-effects model estimation. To make concepts clear, consider a simple panel regression model, shown in Equation 1:

$$y_{it} = \alpha + \beta T_{it} + \delta z_{it} + \varepsilon_{it} \quad (1)$$

where the i subscripts denote individuals and the t subscripts denote time periods. In the context of this assessment, the outcome variable is energy consumption (y_{it}) and the parameter of interest is the coefficient on the treatment variable, β , which represents the impact of the treatment. The term z_{it} represents any control variables that may (or may not) be included and the ε_{it} is an idiosyncratic, white noise error term. In this framework, both pre-treatment and post-treatment time periods are included in the dataset that are analyzed.

The key assumption in the fixed-effects model is that there are time-invariant variables for each customer omitted from Equation 1 that are relevant in explaining energy consumption and are also potentially correlated with the treatment variable (T_{it}). If these fixed-effects are present and indeed correlated with the treatment, then estimating the model using the standard ordinary least squares (OLS) estimator will produce inconsistent estimates of the treatment effect.⁴

To obtain a consistent estimate of β , the FE model modifies Equation 1 so that the intercept (α) becomes unique for each customer:

$$y_{it} = \alpha_i + \beta T_{it} + \delta z_{it} + \varepsilon_{it} \quad (2)$$

The term α_i is called a fixed effect and captures all characteristics for each individual customer related to their energy consumption that do not change over time and are not captured by the z variables that are included in the model. In order to get a consistent estimate of the treatment effect from Equation 2, the fixed effects must first be eliminated. The most common way to eliminate the fixed effects is by mean-differencing the data for

⁴ A consistent estimator has the property that, as the sample size increases, the estimate converges to the true value that we seek an estimate for (i.e., that the estimator converges to a single value and the bias of the estimator converges to zero).

each individual,⁵ which is the approach Nexant employed. Once this is accomplished, the resulting transformed equation can be estimated by OLS to provide a consistent estimate of the treatment effect.

The identifying assumption underlying the FE model was that the average consumption without treatment would be the same for both treatment and control customers once Nexant accounted for the fixed effects and the control variables. The primary advantage of the FE model is that it controlled for all of a customer's time-invariant characteristics that are related to energy consumption, no matter whether they were observed or unobserved. Thus, any idiosyncratic habits or behaviors of a particular consumer did not influence the estimate of the treatment effect as long as they did not change over time. The FE model also produced a consistent estimate of the treatment effect if the fixed effects were not correlated with treatment.

The FE model has its disadvantages as well. While controlling the time-invariant characteristics can be useful, it also means that impacts of any such characteristics cannot be estimated because they were eliminated from the regression equation during estimation. The exact nature of the unobserved variables was also typically somewhat mysterious. This shortcoming does not pertain to the validity of the estimated effect they were trying to measure, rather, it reflects the fact that while the effect of the treatment on the outcome variable becomes known, the identity and specific effects of the unobserved variables on the outcome variable were never brought to light.

Another disadvantage of the FE model is that it did not control for any omitted variables that changed over time.

Finally, because estimating the FE model required mean-differencing the data, the only variation used to estimate the treatment effect was variation in variables over time within particular individuals. Because mean-differencing does not make use of any cross-sectional (also called "between") variation, the estimates from FE was less precise than other estimators.

LAGGED DEPENDENT VARIABLES MODEL ESTIMATION

An alternative to the FE model is to incorporate customer heterogeneity by explicitly including past values of an individual's energy consumption as control variables on the right-hand side of the regression equation. Such a model specification is called a lagged dependent variable (LDV) model. In this case, Nexant modified Equation 1 to something like Equation 3:

$$y_{it} = \alpha + \beta T_{it} + \delta z_{it} + \theta y_{i,t-1} + \varepsilon_{it} \quad (3)$$

⁵ To mean-difference the data, the average values for each variable in the equation were calculated for an individual and these averages were then subtracted from all of the observations for that individual. Since the fixed effects did not change over time, the average value of α_i was equal to each individual α_i and so the difference was equal to zero and the fixed effects disappeared. The transformed version of the model can be written as $\tilde{y}_{it} = \beta \tilde{T}_{it} + \delta \tilde{z}_{it} + \varepsilon_{it}$ where $\tilde{y}_{it} = y_{it} - \bar{y}_i$, $\tilde{T}_{it} = T_{it} - \bar{T}_i$ and $\tilde{z}_{it} = z_{it} - \bar{z}_i$.

In this specification, the intercept remained the same for everyone (no i subscript on the α) and there was an additional term $y_{i,t-1}$ that represented the energy consumption for individual i in a previous period. This is akin to saying that what makes consumers unique was captured entirely by their past levels of consumption. Equation 3 was estimated by pooled OLS, provided that there was no serial correlation in the error term and there was no omitted variables that were correlated with the treatment. The underlying identification assumption was that average consumption without the treatment would be the same for both treatment and control customers, which were conditional on the control variables and their previous usage.

Unlike the FE model, an LDV model explicitly controlled for any measurable customer characteristics that varied over time. Another advantage was that pooled OLS used both variation over time within customers and variation between customers so that the resulting estimates of the treatment effect was relatively more precise than the FE estimate.

The main downside to the LDV model was that if fixed effects did indeed exist, and were correlated with the treatment variable, then the LDV estimate of the treatment effect was inconsistent. Additionally, the LDV model relied on being able to measure the relevant time-varying variables and could not control for unobserved variables like the FE model could.

CONSEQUENCES OF MODEL MISSPECIFICATION

In considering which model would be most appropriate for this assessment, the key question was, what would happen if the wrong model was used to estimate the treatment effect, i.e., what would happen if we used an LDV model when fixed effects were present (or vice versa)? Fortunately, there is a useful discussion of this matter in Angrist and Pischke's "Mostly Harmless Econometrics" text. To summarize, although the FE and LDV models are not nested, they do have a useful bracketing property in that the estimates obtained from each model will bound the true treatment effect (Angrist and Pischke, 2008). In a situation where the treatment effect is negative (energy savings are positive), using the LDV model in the presence of fixed effects will overestimate energy savings if the fixed effects cause the control group to have higher levels of pre-treatment consumption than the treatment group. In the reverse scenario where time-varying omitted variables are important, but the FE model is used, the FE estimate will underestimate the treatment effect. In the case where both fixed effects and lagged dependent variables are appropriate, the true treatment effect will be somewhere between the FE and LDV estimates.

TABLE 6. CONSEQUENCES OF MODEL MISSPECIFICATION

Case	Use FE model	Use LDV model
Fixed effects exist and are correlated with treatment	$\hat{\beta}_{FE}$ provides consistent estimate of β	$\hat{\beta}_{LDV}$ overestimates β
Time-varying omitted variables are correlated with treatment	$\hat{\beta}_{FE}$ underestimates β	$\hat{\beta}_{LDV}$ provides consistent estimate of β

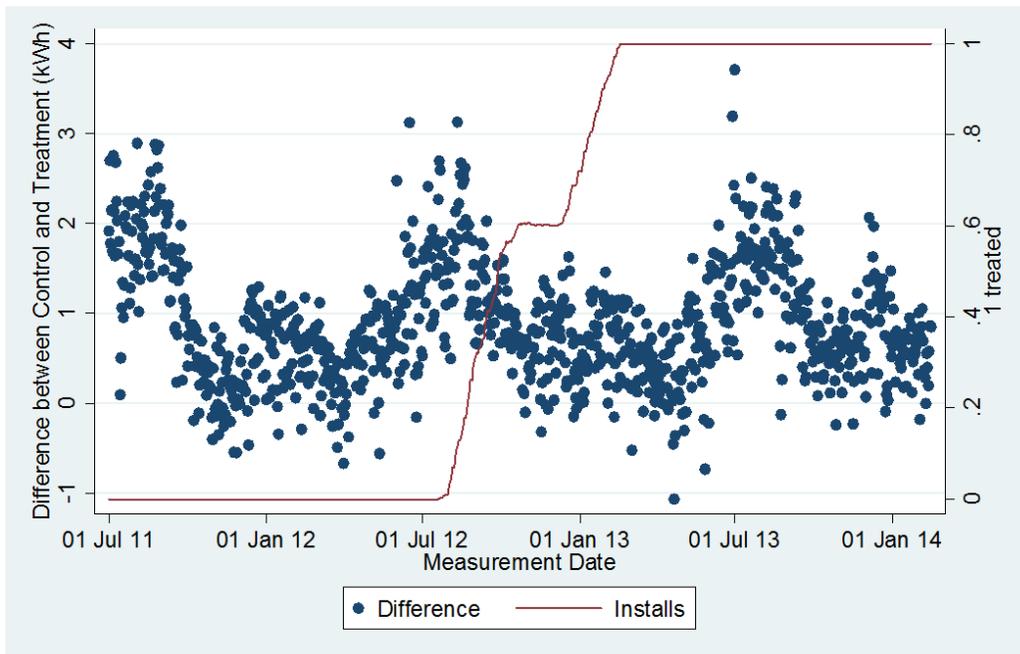
WHAT ABOUT RANDOMIZATION?

In order for misspecification to be a problem, the omitted variables (whether fixed effects or lagged dependent variables) must be correlated with the treatment. If there is no correlation between the omitted variables and treatment, then the error term will remain uncorrelated with the explanatory variables and both models will give consistent estimates of the treatment effect. In theory, this is what we would expect to see in a randomized control trial because randomly assigning the treatment should cause it to be uncorrelated with any other variables that are related to energy consumption. This is why randomization is desired – it removes the need for fancy econometrics by eliminating confounding factors. If this was indeed the case, then we would expect to get consistently estimated results from both models. Adding control variables, fixed effects, or lagged dependent variables, would serve only to increase the precision of the estimates.⁶

Unfortunately, we have evidence that the randomization has been compromised because of a difference in temperature sensitivity between treatment and control customers. This difference is apparent in the pre-treatment period and exists throughout the entire study period. Figure 13 shows the difference between the average daily electricity consumption of the treatment and control groups over time. Each blue dot represents the difference between control and treatment kWh on that day (difference = control - treatment). The red line shows the number of installs over time. Ideally, the blue dots would be centered on zero and would not show any seasonal trends. However, as Figure 16 shows, there is in fact a clear seasonal trend, with significant peaks during the summer months, and smaller peaks during the winter months. These summer peaks indicate that, during the warmer months, the control group tends to have higher electricity consumption than the treatment group.

⁶ In this case, random effects would probably be preferred to fixed effects because it would produce more precise estimates.

FIGURE 13. DAILY DIFFERENCES IN ENERGY CONSUMPTION BETWEEN CONTROL AND TREATMENT GROUPS



The pattern in Figure 16 indicates that there is an omitted variable related to energy consumption that is also correlated with the treatment. One possible explanation for this would be that each customer has a unique preference for thermal comfort and that, on average, control customers prefer to keep their homes cooler in the summer than treatment customers. If this preference is stable over time, as it appears to be, then it is a fixed effect that could be confounding the LDV estimate.

The discussion of FE and LDV models above is predicated on the inclusion of both pre-treatment and post-treatment time periods in the analysis dataset. Under this scenario, the treatment variable that is included is the treatment variable times a dummy variable for the post-treatment period (an interaction term) where the value is 0 for all customers in the pre-treatment period and 1 only for treatment customers in the post-treatment period. With this specification, panel models are effectively performing a difference-in-differences (DD) calculation to estimate the treatment effect. If the impact of the omitted variable is the same throughout the study period (which it appears to be in Figure 16), the DD estimates will not be affected and will provide consistent estimates for the treatment effect.

Estimating an LDV model only using the post-treatment period would not utilize a DD approach. Doing so does not address the omitted variable (weather-sensitivity) that is correlated with the treatment and will subsequently produce biased estimates of the treatment effect. The reason for this is that the omitted variable is related to temperature and will therefore create serial correlation in the error term. OLS does not account for this serial correlation, which is the source of the bias.

REFERENCES

Angrist, J.D. and J. Pischke (2008), "Mostly harmless econometrics: An empiricist's companion". Princeton University Press. See sections 5.3-"Fixed Effects versus Lagged Dependent Variables" and 5.4-Appendix: "More on fixed effects and lagged dependent variables"

Greene, William H. (2003), "Econometric Analysis". 5. Upper Saddle River, NJ: Prentice Hall.

Meier, A. (2010), "How People Actually Use Thermostats: Quantifying Usability in Programmable Thermostats". Delivered at the Behavior, Energy and Climate Change Conference. http://stanford.edu/group/peec/cgi-bin/docs/events/2010/becc/presentations/3D_AlanMeier.pdf

NTIA (2010), Digital Nation: 21st Century America's Progress toward Universal Broadband Internet Access. http://www.ntia.doc.gov/files/ntia/publications/ntia_internet_use_report_feb2010.pdf

Palmgren, C., N. Stevens, M. Goldberg, R. Barnes and K. Rothkin (2010), California Statewide Residential Appliance Saturation Study Final Report Executive Summary. Oakland, CA: KEMA, Inc.

Perry, M. and J. Oh. (2012), Findings from the Initial Installation Phase of the Opower/Honeywell Thermostat Trial. San Francisco, CA: Freeman, Sullivan and Co. http://www.etcc-ca.com/sites/default/files/reports/ET11PGE3073%20Opower%20Honeywell%20Thermostat%20Trial_12.12.12.pdf

Perry, M. and S. Woehleke (2013), Evaluation of Pacific Gas and Electric Company's Home Energy Report Initiative for the 2010-2012 Program. San Francisco, CA: Freeman, Sullivan and Co. http://www.calmac.org/publications/2012_PGE_OPOWER_Home_Energy_Reports_4-25-2013_CALMAC_ID_PGE0329.01.pdf