Thriller in Asilomar: Battle of the Smart Thermostats

Noah Lieb and Scott Dimetrosky, Apex Analytics, Dan Rubado, Energy Trust of Oregon

ABSTRACT

Momentum continues to build for home automation products offering robust gas and electric savings. According to Navigant Research, global revenue for communicating and smart thermostats is expected to grow from $146.9 million to $2.3 billion over the next decade (Navigant, 2014). A growing number of program administrators are either adding or looking to add these smart thermostat measures to their residential efficiency portfolio, with the most significant of the recent announcements being Commonwealth Edison's "million thermostats by 2020" initiative.

As one of the first program administrators to offer advanced smart thermostats, the Energy Trust of Oregon recently completed their second advanced smart thermostats pilot. The most recent pilot (Apex Analytics, 2015) involved installing the Nest and Honeywell Lyric thermostats in gas furnace heated single-family homes. Apex Analytics, in partnership with Energy Trust, evaluated the pilot, which involved staff and participant surveys coupled with a billing analysis. The evaluation focused on estimating gas savings, customer interaction, usage, satisfaction, and the thermostats' control of the comfort of participant's homes.

This paper covers the methods used and results of the recent pilot evaluation, including the lessons learned regarding study design and the pilot implementation. Ultimately, the Nest thermostat offered easier installation with less complications, higher satisfaction ratings, and statistically higher gas energy savings (6.0% gas heat load savings) relative to the Honeywell Lyric (4.9% increase in gas heat load).

Background

In 2013, Energy Trust of Oregon (Energy Trust) launched a pilot, the Nest Thermostat Heat Pump Control Pilot, to study the electric energy savings impacts of installing a smart thermostat in lieu of heat pump controls in heat-pump heated homes. The Nest Thermostat Heat Pump Control Pilot was a successful undertaking (Apex Analytics, 2014), with high participant satisfaction and robust energy savings. In 2014, Energy Trust initiated a "Smart Thermostat Pilot" to continue testing smart thermostats and explore the potential for a new cost-effective gas savings measure. This Pilot focused on the Honeywell Lyric and the Nest Thermostat, two smart thermostats in the market. Both thermostats claimed to offer simple user interfaces with advanced features to save energy. Features included automated and occupancy-based temperature management and various remote control options. Both products were available at retail stores for approximately $250 (at the time of the launch of this Pilot).

The primary goals of the evaluation were to:

- Quantify the annual natural gas savings that result from installing smart thermostats in single family homes heated with a gas furnace.
- Identify variations in savings between participants based on demographic and household characteristics and any differences in savings between the two thermostats.

- Obtain feedback from program staff and participants to understand thermostat installation issues, how well the thermostats worked, and what kinds of operational issues were encountered.
- Understand participants' interaction and satisfaction with the thermostats.

The Energy Trust executed the pilot under its Existing Homes program, which purchased all of the thermostats up-front, maintaining inventory control for the Pilot by accurately recording product serial numbers. Energy Trust offered participants discounted smart thermostats for $219 per unit, made available through a bulk-purchase order. This study required participants to self-install (either on their own or through a contractor of their choice) their thermostat, connect it via Wi-Fi to the internet and link it to their online Nest or Honeywell account, and then forward the account verification email to Energy Trust for participant verification and rebate processing. Participants received a $200 rebate for their thermostat.

The Pilot ran from the fall of 2014 through the spring of 2015, covering one entire heating season. Participants were recruited primarily through a collaborative marketing effort with Northwest Natural Gas Company (NW Natural), the primary natural gas utility associated with the Energy Trust pilot. Based on eligibility criteria provided by the program, NW Natural randomly selected and contacted a sample of 22,000 customers who paid their bills online, had a gas account for at least a year and had a winter gas usage at least twice that of summer months.

The recruitment email sent by NW Natural directed interested candidates to complete an online survey to determine if they qualified to participate in the study. The program performed additional recruiting among Energy Trust employees and program management contractors to increase the number of participants. Candidates who met the criteria received a follow-up qualification email from Energy Trust containing information and directions on how to purchase the thermostat. Candidates whose answers indicated they did not meet one or more of the eligibility criteria received a customized email informing them of the reason they did not qualify.

Implementation staff controlled for product selection bias by randomly assigning qualifying candidates into one of two treatment groups based on the two thermostats involved in this study. Those in the Nest group were provided a link to purchase a Nest in their qualification email, while candidates in the Lyric group received a link to purchase a Lyric. In addition to the treatment groups, Energy Trust also identified two additional groups to be used for the billing analysis: a comparison group and later, to help verify the robustness of the model, an intention-to-treat group. The intention-to-treat group comprised both the treatment group and those participants that had successfully signed up for the pilot but did not end up purchasing a thermostat. NW Natural provided approximately 1,000 randomly selected customers' information to the Energy Trust Evaluation Team to serve as a comparison group for the billing analysis. These customers met the same pre-screening criteria as those customers who were contacted. Customers in the comparison group were not contacted. Though a total of 400 participants were targeted (200 Nest, 200 Lyric), the final thermostat household sample contained 212 Nest and 171 Lyric homes. The final billing analysis, after attrition, included 153 Nest and 127 Lyric participant homes. In addition, there were 592 Nest and 580 Lyric intention-to-treat homes, 1,816 Nest and 1781 Lyric matched comparison homes, and 978 randomized comparison homes available for analysis. Additional details regarding the methods and analysis are included below.

Study Methodology

There were three primary components associated with the Pilot evaluation: staff interviews, participant surveys, and a billing analysis. Staff interviews were conducted with the goal of collecting insight and feedback from those staff members most familiar with the Pilot and to supplement the program summary report compiled by the program management contractor, CLEAResult. Interviews were held with four members of CLEAResult, and one was held with a member of the Energy Trust team.

There were two separate participant surveys administered to the entire population of Pilot participants, one in January 2015 and a second one at the end of the heating season in May 2015, but only to those who had completed the first survey. The initial survey focused on customer motivations for participating in the pilot, installation and setup of the device, attitudes about the device, valued features, home comfort, use of the device, commitment to saving energy, and satisfaction with the pilot. A core set of questions remained consistent in both surveys to gauge whether participants changed their opinions of the devices during the heating season. In addition, the surveys explored participant characteristics and behaviors that might be related to the amount of energy savings and which features of the device might also affect energy savings. A web-based survey was deemed the best approach because all participants had Internet access (a requirement for the thermostat installation), and were assumed to have some degree of familiarity with technology due to the high-tech nature of the advanced smart thermostats.

For the billing analysis, the Pilot was set up similar to a randomized encouragement design (RED) study with a second stage of randomization where treatment group homes were assigned to one of the two thermostats, once they had opted in to the study (see Figure **1**). The standard approach for determining the effect of an intervention in a RED study is to compare the entire randomized treatment group (known as the intention-to-treat group) with the entire randomized comparison group. This preserves the initial randomization and helps control for self-selection bias. However, for this study, the typical intention-to-treat analysis could not be performed on the entire randomized treatment group, prior to opting-in, due to the very low response rate to the recruitment emails (8%) and the fact they were not randomized into specific thermostat groups. The Evaluation Team focused instead on the second stage of randomization, where treatment group households that expressed interest in the Pilot were assigned to the Nest or Lyric thermostat group. These two groups of respondents became the intention-to-treat groups for the study. The Evaluation Team acknowledges that a self-selection bias could have been introduced here, but any bias should be the same for participants assigned to both thermostat groups.

An additional problem with the study was that the conversion rate from treatment homes that expressed interest in the Pilot to those that actually received a thermostat was quite low (24%). Even using the reduced intention-to-treat group, this level of attrition significantly limited our ability to observe an effect. So, the Evaluation Team separately analyzed the homes that received a thermostat against the comparison group. Unfortunately, if a strong self-selection bias was present, the comparison homes would not properly represent homes that received a thermostat. To investigate and address this potential source of bias, the Evaluation Team also synthesized a matched comparison group for each group of thermostat recipients. Matched comparison groups were created by randomly selecting 200 residential gas customers from each decile of pre-Pilot raw annual gas usage, based on the distributions of annual usage for

thermostat recipient homes. This resulted in two comparison groups of 2,000 homes, with gas usage distributions that closely matched either the Nest or Lyric recipient homes. These matched comparison groups allowed us to conduct a quasi-experimental analysis that the Evaluation Team compared to the other analyses. The end result was that the Evaluation Team did three different comparisons for each thermostat:

- the thermostat recipient group versus the randomized comparison group,
- the intention-to-treat group versus the randomized comparison group,
- the thermostat recipient group versus the matched comparison group

Given the limitations of the design and implementation of the Pilot, the Evaluation Team believes this combination of analyses provides the best possible estimates of gas usage impacts. A more detailed chart demonstrating this two-stage randomization is included in Figure **1** below.

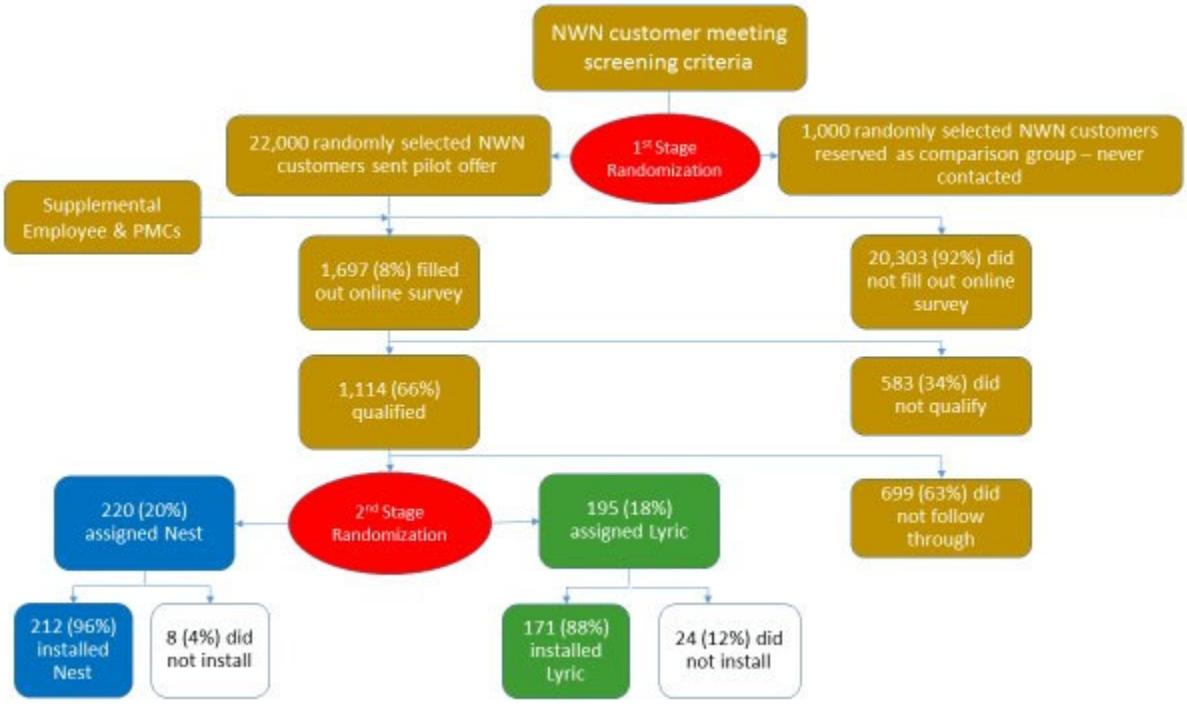

Figure 1. Pilot Two-Stage Randomization Results

A billing analysis was performed to estimate the annual heating season gas savings associated with the two thermostats tested in the Pilot[1]. Gas billing data were analyzed for the year prior to the Pilot (September 2013 to November 2014), and the post-installation study period was defined as January 25, 2015 through October 2015. The Evaluation Team constructed a panel dataset of study participants and comparison homes by merging Pilot data, response data from participant surveys, daily weather data from the National Climatic Data Center, and monthly gas usage data. As described above, the Evaluation Team set up three comparisons

[1] Though there may be electric cooling savings attributable to these thermostats as well, this study focused exclusively on the potential gas savings benefits of these thermostats

between the different study-groups for each thermostat. A series of regression models were constructed using variables for weather, study group, and study period to predict the average daily gas usage for each billing period. Average daily temperature was used to calculate heating degree-days (HDDs) for each billing period, using reference temperatures ranging from 40 to 75 degrees Fahrenheit. The average daily HDDs were computed for each billing period, so that the units were directly comparable to the average daily gas use. The evaluation team created a study period flag to indicate observations occurring in the pre-Pilot vs. post-installation period. A study group flag was created to indicate whether homes were in the participant or comparison group.

The Evaluation Team created multi-level linear mixed-effects models for both Nest and Lyric participants. The multi-level model was used to account for the longitudinal nature of the data, where gas usage observations were made repeatedly on each home over time. Using random effect terms in the model, an intercept and HDD coefficient were fitted to each home separately, in each study period, and then pooled across the sample using fixed effect terms. Average daily gas usage was modeled as a function of average daily HDDs, the study period, and the study group. Interaction terms between all three variables were added to model the effect of installing a smart thermostat on gas usage. The following equation describes the linear mixed-effects model that was used.

$$Usage_{ij} = \beta_0 + \beta_1 HDD_{ij} + \beta_2 Group_{ij} + \beta_3 Post_{ij} + \beta_4 Group_{ij} * Post_{ij} + \beta_5 Group_{ij} * HDD_{ij} + \beta_6 Post_{ij} * HDD_{ij} + \beta_7 Group_{ij} * Post_j * HDD_{ij} + u_{0i} + u_{1i} HDD_{ij} + u_{2i} HDD_{ij} * Post_{ij} + \epsilon_{ij}$$

Where:

$Usage_{ij}$ = the average daily gas usage for home *i* during billing period *j*,

β = regression coefficients for each variable in the model (indexed from 0 to 7),

β_0 = fixed intercept for all homes,

HDD_{ij} = heating degree-days for home *i* during billing period *j*,

$Group_{ij}$ {0,1} = dummy variable where 1 indicates that home *i* is part of the Nest or Lyric participant group, which is static across all *i* billing periods,

$Post_{ij}$ {0,1} = dummy variable where 1 indicates that billing period *j* for home *i* is in the post-installation period,

u_{0i} = random intercept for home *i* that is independent from ϵ_{ij},

u_{1i} = random slope coefficient of HDD for home *i* that is independent from ϵ_{ij},

u_{2i} = random coefficient of the interaction between HDD and installation period for home *i* that is independent from ϵ_{ij},

ϵ_{ij} = model error for home *i* for billing period *j*.

The model provided two key parameter estimates for computing energy savings: the interaction term coefficients β_4 and β_7. Together, these coefficients described the mean difference in the change in consumption between the participant and comparison groups from the pre- to post-installation periods for a given number of HDDs. A linear combination of these coefficients, plus the long-run annual HDDs, was used to compute the weather-normalized

annual gas savings in therms per home for each thermostat. The Evaluation Team also computed the pre-Pilot annual gas usage and heating usage for the treatment group to calculate the gas savings as percentages.

$$\text{Annual Savings} = 365 * \beta_4 + LRHDD * \beta_7$$
$$\text{Normalized Annual Usage} = 365 * (\beta_0 + \beta_2) + LRHDD * (\beta_1 + \beta_5)$$
$$\text{Normalized Heating Usage} = LRHDD * (\beta_1 + \beta_5)$$

Where:

β = regression coefficients for each variable in the model (indexed from 0 to 7),
LRHDD = long-run annual HDDs, derived from TMY3 data.

For the intention-to-treat savings analysis, an additional step was necessary to compute gas savings attributable to the installation of one of the two thermostat models. The intention-to-treat savings estimates were divided by an adjustment factor (the proportion of homes where a thermostat was installed – 26% for Nest and 21% for Lyric) to obtain the portion of savings due to the treatment, known as the Local Average Treatment Effect (LATE). Unfortunately, due to the small proportion of homes in the intention-to-treat group that received thermostats, the observed effect sizes were small with a relatively large amount of error.

$$\text{LATE Annual Savings} = Savings / \left(\frac{\#\ of\ Thermostat\ Installations}{\#\ of\ Intention-to-Treat\ Homes} \right)$$

In addition to the overall savings for each thermostat, the Evaluation Team was interested to see if there were differences in savings between subgroups of participant homes for each thermostat. Most of the data used to define subgroups was self-reported, obtained through the participant intake survey or follow-up surveys. The Evaluation Team assessed differences in gas savings between categories of participants for a variety of factors including pre-pilot annual gas usage, housing characteristics, heating equipment, demographics, and experience with the thermostats; this allowed the Team to analyze whether gas savings varied within the sample, and to identify factors that might be driving those differences. Although the estimates were not precise, they allowed the Team to coarsely assess whether there were any large differences in gas savings between categories of participants.

Findings

The findings presented here are ordered chronologically and align with how participants experienced the Pilot: the early stage includes participant recruitment and installation; the middle stage includes participant experiences with the thermostats, including usage, satisfaction, and feedback on the thermostats; and the final stage, after the first heating season in which the thermostats were installed, includes determining the gas savings associated with thermostats.

Finding 1: Recruitment – The self-installation model proved to be a low-cost delivery approach but this model may have led to substantial attrition among interested and qualified customers.

Staff concluded that the recruiting and targeting of customers was considerably improved from the approach used for the Nest Heat Pump Control Pilot (Energy Trust directly installed the thermostats). Acquiring approximately 400 participants in less than two months, with minimal cost to Energy Trust, proved the ease and success of the self-install pilot model. Participants that required support successfully received assistance from the manufacturers either via phone or website rather than having to rely on Energy Trust or CLEAResult staff for guidance.

The most serious recruiting challenge, however, arose at the gap between qualifying and purchasing participants: only 35% of candidates who completed the intake survey and qualified for the study actually purchased a thermostat, despite being offered a $250 thermostat for only $19. Staff speculated that the large drop– between those who completed the survey and qualified to participate, to those who actually purchased a thermostat– was likely due to the perceived technical difficulty of self-installation of the thermostats (and potential need of contractor-based installation) while another barrier may have been the $219 upfront cost of the thermostat in advance of receiving the rebate.

Finding 2: Installation – Thermostat installation was faster, easier, and received higher satisfaction ratings for Nest participants compared to Lyric participants.

Nest participants were able to install the thermostat in less time and with less difficulty than the Lyric participants. Respondents reported that the average installation time for the Nest was less than an hour (51 minutes) whereas the Lyric took one hour and 13 minutes – a difference of only 22 minutes, but about 40% longer. Only 4% of Nest participants believed initial setup and configuration was either difficult or very difficult, compared to 17% of Lyric participants. An even higher percentage of Lyric users indicated experiencing installation issues (37%) – over three times that of the Nest user base (10%).

Participant satisfaction with installing the thermostat was highly dependent on the device: 90% of participants indicated a satisfaction rating of either a 4 or 5 out of 5 for the overall installation process for the Nest, while only 63% of the Lyric participants provided an equivalent satisfaction rating for their installation process.

Finding 3: Satisfaction – Nest users reported much higher rates of satisfaction with the user interface, scheduling, and overall thermostat compared to the Lyric users.

The vast majority (95%) of Nest participants rated the overall user interface either easy or very easy to use, while only 70% of Lyric participants gave the same rating for their thermostat. Scheduling proved to be the most difficult aspect of the Lyric, with over 20% of participants indicating this was somewhat or very difficult, whereas only 3% of Nest participants reported experiencing the same difficulty. Lyric participants also experienced considerably more non-installation-based issues: 50% of first-survey and 27% of second-survey respondents reporting additional issues with their Lyric thermostat; whereas 16% of first-survey and 7% of second-survey respondents experienced Nest-related post-installation issues.

In terms of overall satisfaction, Nest users gave considerably higher satisfaction ratings relative to the Lyric: over 65% of Nest users rated the thermostat a 5 out of 5, whereas only 24% of Lyric users rated the thermostat a 5 out of 5. In addition, although participants were committed to retaining their thermostats for the duration of the Pilot, if given the option to return their units, only 10% of Nest participants would have returned the unit whereas over three times as many Lyric participants (34%) would have returned the thermostat.

Participant dissatisfaction followed the same trend: while 6% of Lyric users were dissatisfied in the first survey, over 13% of users were dissatisfied by the second. For the Nest, less than 2% of first survey users were dissatisfied with the device, and this declined to less than 1% by the second survey. Furthermore, 50% of Lyric respondents in the first survey, and 27% in the second survey, reported having additional issues with their thermostat (mostly general operational issues, scheduling adjustments, Wi-Fi connectivity, and occupancy detection among others). Nest recipients had a considerably lower level of post-installation issues than Lyric recipients, with only 16% of Nest respondents in the first survey, and 7% in the second survey reporting additional issues (mostly general operational issues and Wi-Fi connectivity).

Home comfort was one area where both participants were in agreement, the two thermostat populations did not show a statistically significant difference in home comfort: over half of both Lyric and Nest survey respondents in both the first and second survey described the temperature of their home to be either "somewhat more comfortable" or "much more comfortable" after installing the thermostat.

Finding 4: Thermostat Use – Nest users were more likely to utilize the occupancy detection features and less likely to override the unit.

One of the primary energy-saving features of the smart thermostats is occupancy detection. For the Nest, this feature is called "Auto-Away," which minimizes heating and cooling when the device determines no one is home based on occupancy sensors (motion sensors). For the Lyric, this feature is called "geofencing," which is dependent on the GPS location of the smartphone that is matched with the thermostat. When the Nest thermostat is installed, the Auto-Away feature should be preset as enabled, whereas for the Lyric, the geofencing is not enabled by default, and the user is required to enable the geofencing during initial setup.

Nest users overwhelmingly left Auto-Away enabled: 88% of first- and second-survey respondents reported that they left this feature enabled. The Lyric respondents were not as likely to have enabled geofencing: only 57% of Lyric users had enabled this feature by the first survey, and slightly less (50%) had this feature enabled by the second survey.

In addition, Nest participants showed a 60% relative decline in daily adjustments between the first and second survey. Lyric participants only showed a 35% relative decline in daily adjustments, indicating participants continued to rely on manual adjustments. The fact that a significant proportion of Lyric participants continued to make frequent adjustments (daily or weekly) shows that participants still were unable to rely on the thermostats to perform one of their primary functions – to automate home heating and cooling.

Finding 5: Energy Savings – The energy savings, the most important feature associated with this study, proved to be the most notable difference between the two thermostats: Nest participants showed decreased gas consumption while Lyric participants showed increased gas consumption.

The final thermostat recipient sample contained, after attrition, 153 Nest and 127 Lyric participant homes. In addition, there were 592 Nest and 580 Lyric intention-to-treat homes, 1,816 Nest and 1781 Lyric matched comparison homes, and 978 randomized comparison homes available for analysis. In terms of pre-Pilot raw annual gas usage, the participant and comparison group homes were very similar. Lyric recipient homes appeared to have slightly higher baseline

gas usage versus to the randomized comparison group, but this difference was not statistically significant. Less than 10% of both Nest and Lyric thermostats were installed in the homes of Energy Trust employees and contractors.

Based on the best fit linear mixed effects model, the average annual gas savings for Nest recipients was estimated at 34 therms (90% CI: 13, 55), which was statistically significant. In contrast, Lyric recipients experienced an estimated 29 therm increase (90% CI: -58, -7) in annual gas usage, on average. This estimate was also statistically significant at the 90% confidence level. Although the savings estimates for both thermostats had relatively large standard errors (> 30% of point estimates) and wide confidence intervals, they provide a clear indication that Nest participants significantly reduced their annual gas usage, while Lyric participants increased it. The difference in estimated gas savings between the two thermostats was statistically significant.

Gas savings were similar when modelling the intention-to-treat homes against the original randomized comparison group. After savings were calculated for the entire intention-to-treat group, using the best fit linear mixed effects model, an adjustment factor was applied to estimate the LATE savings. Using the intention-to-treat analysis with the LATE adjustment had the disadvantage of introducing a large amount of error to the savings estimate, based on the noise from the large proportion of homes that did not receive a thermostat. The annual gas savings for Nest recipients was estimated at 40 therms (90% CI: -21, 100), on average, but this result was not statistically significant at the 90% confidence level. Lyric recipients experienced an estimated increase in annual gas use of 55 therms (90% CI: -130, -20), on average, but this was not statistically significant either. Although the savings estimates for both thermostats had very large relative standard errors and wide confidence intervals, they provide some indication that Nest participants reduced their annual gas usage, while Lyric participants increased their gas usage.

As an alternative to using the original Pilot randomization scheme, a quasi-experimental analysis was used to compare thermostat recipients with matched comparison groups. The matched comparison group was created by randomly selecting residential gas accounts with a similar distribution of pre-Pilot annual gas use as the thermostat recipients. The average annual weather normalized gas savings for Nest recipients was estimated at 34 therms (90% CI: 15, 53), which was statistically significant. The Evaluation Team estimated that Lyric recipients experienced a 24 therm increase (90% CI: -58, -7) in annual gas usage, on average. This increase in usage was statistically significant at the 90% confidence level. These results are very similar to the estimates above, using the original randomized comparison group. These results clearly indicate that Nest recipients significantly reduced their annual gas usage, while Lyric recipients increased it. The difference in estimated gas savings between the two thermostats was statistically significant.

Table 1. Summary Results of the three billing analysis groups

Billing Analysis Group	Thermostat	Annual Therm Savings	90% Conf. Interval	p-value	% Savings	% Heating Savings
Treatment vs Randomized Comparison Group	Nest	34	13, 55	0.018*	4.50%	6.00%
	Lyric	-29	-55, -3	0.071*	-3.70%	-4.90%
Intention-to-Treat vs Randomized Comparison Group	Nest	40	-21, 100	0.254	4.90%	6.60%
	Lyric	-55	-130, 20	0.209	-7.00%	-9.20%
Treatment vs Matched Comparison Group	Nest	34	15, 53	0.012*	4.40%	5.90%
	Lyric	-24	-48, -1	0.090*	-3.10%	-3.90%

Results from the subgroup analysis showed some counterintuitive results, though not all results were statistically significant. First and foremost, the findings determined that there was a negligible effect on overall savings from participants recruited from Energy Trust employees and contractors and from those who removed their thermostats mid-Pilot – verifying the absence of potential bias due to the supplemental sample. The first subgroup analysis compared groups of homes by annual gas usage. Findings showed lower usage homes realized larger savings than higher usage homes, though the differences for Nest recipients were not statistically significant and could be due to random variability in the sample while Lyric participants saw large, significant differences between the lowest and highest usage categories. It is not clear what is driving this trend.

Another factor of interest was whether participants had a secondary heating system installed. In homes with secondary heating systems, such as gas fireplaces, if the secondary system is not centrally controlled, there will be lower savings potential for the thermostat. For Nest participants, homes with no secondary system appeared to have higher gas savings, although the difference was not significant. For Lyric homes, there did not appear to be any difference in savings.

The primary energy saving strategy for smart thermostats in gas-heated homes is to increase the number of temperature setbacks and better match heating with occupant schedules. This strategy depends on the level of control and number of setbacks that homes have in place before installing a smart thermostat. To assess the impact of the pre-Pilot control strategy on gas savings, the Evaluation Team compared homes where the prior thermostat was reported to be programmed against homes where the prior thermostat was either manual or not programmed with any setbacks. Unfortunately, the sample sizes of manual and non-programmed prior thermostats were small, so the results are not reliable. However, it appears that Nest recipients who replaced non-programmed or manual thermostats realized higher gas savings than those who replaced thermostats with programmed setbacks. Although indicative, these differences were not statistically significant. For Lyric homes, there did not appear to be any differences in savings.

The number of temperature setbacks achieved by both the Nest and Lyric thermostats depends partly on how well the occupancy detection features work in the field. From The Nest Heat Pump Control Pilot, the Evaluation Team learned that a small portion of participants disabled the occupancy detection feature because it did not work well for their home. To analyze the impact of occupancy detection on savings, the Evaluation Team compared participants who disabled occupancy detection versus those who did not. The number of Nest recipients who reported disabling the auto away feature was very small, so the results for that group are not reliable. For Lyric recipients, there were a larger number of participants who reported that they were not successful in setting up and enabling the geo-fencing feature. For both Nest and Lyric homes, participants who kept occupancy detection enabled appeared to have higher gas savings, although these differences were not significant.

Summary and Conclusions

The Energy Trust Smart Thermostat Pilot built on the successes experienced during the previous Nest Thermostat Heat Pump Control Pilot and in many ways improved upon the previous Pilot. Improvements included the increased speed of Pilot inception through implementation, greater planning and filtering for participant recruitment, and a new delivery approach that placed most of the installation and processing requirements on the participants (in order to reduce cost and demonstrate what a larger program delivery scheme would likely entail). The Pilot did suffer some setbacks, including low enrollment and low purchase uptake for qualifying participants, and there was also a vocal minority of participants who were very dissatisfied with their thermostat, almost entirely comprised of Lyric owners.

The billing analysis results showed that the Nest thermostat was associated with significant energy savings during Energy Trust's Pilot. It produced about 6% heating load savings, on average, in gas-heated homes. On the other hand, the Honeywell Lyric thermostat was associated with significant increases in energy use during the Pilot. The Lyric added about 5% to heating loads, on average, in gas-heated homes. The difference in realized energy savings between the two thermostats was unambiguous and statistically significant. The findings from this billing analysis support the expansion of the smart thermostat incentive that Energy Trust currently offers to include homes with gas and electric forced air furnaces provided the reported gas savings more than offsets the costs for Energy Trust cost effectiveness screening. The findings do not support extending the Energy Trust incentive to include the Lyric thermostat. However, it is worth noting that the Lyric was at a much earlier phase in product development during the Pilot, and feedback from the program has allowed Honeywell to make improvements to the next generation of thermostats. Further testing of future versions of the Lyric and other smart thermostats may reveal energy savings for additional products.

For the Smart Thermostat Pilot, potential candidates were pre-screened to qualify based on stringent criteria, and were further screened for compatibility with an additional enrollment survey. Ultimately this led to a recruitment sample that closely reflects the likely populations of candidates that may participate in the future. Additionally, with an almost completely participant-driven application and installation delivery approach, this Pilot was able to replicate the issues faced by a more realistic self-install program structure.

Unfortunately, this early stage in the Pilot was also faced with some challenges. With well over 1,000 participants completing the in-take survey and qualifying to participate in the Pilot, only one-third of the potential candidates chose to purchase the thermostat and actively participate in the Pilot. This was a lower-than anticipated uptake for the measure and resulted in additional sample being drawn from a pool of Energy Trust staff and contractors to achieve the targeted 400 installs. Even with the increased sample, there was a significant number of terminated participants, either due to returns (close to 10% of units returned), product failures, or equipment incompatibility. Overall, the self-install approach is definitely a viable model, though caution should be made with respect to the specific products that should be offered through the program.

The evaluation team believes a brief online survey administered to the qualified population that chose not to participate would help understand the low-uptake. This will help with future program planning by understanding the obstacles to engage potential successful participants. For example, if self-installation is a barrier, Program Administrators can provide a list of qualified contractors to help with installation. In addition, program administrators considering a smart thermostat measure should reach out to the manufacturers to assess a measure installation verification system that ensures participant privacy while also ensuring rebated products are not resold into the open market. Finally, any future smart thermostat program with a similar delivery approach may want to consider the 10% return rate until further research can be conducted to determine the underlying causes of the returns.

Based on the findings from staff interviews, the two participant surveys, and direct communication from several participants providing feedback on the Pilot, participants overwhelmingly preferred the Nest thermostat. Satisfaction levels with installation, user interface, scheduling adjustments, and the overall user experience were significantly higher for Nest relative to the Lyric participants. Also, considering that over three times as many Lyric participants would have returned the thermostat if given the chance, suggests that the Lyric thermostat would benefit from significant design changes. Since Lyric participants were considerably less likely to have enabled the primary energy-saving function (geofencing) relative to Nest users (for whom the feature was enabled by default at installation), the prospects for the current version of the Lyric to demonstrate energy savings are very questionable, as confirmed by this study.

Based on participant experience alone, outside of any potential energy savings, the team cannot recommend the current version of the Lyric for any future incentive offerings. Similar to the previous Nest Heat Pump Control Pilot, the Nest thermostat user experience was considerably more positive, and the team can more easily recommend the Nest for future program offerings, should the associated gas savings, balanced against the costs, offer a cost-effective solution for Program Administrators gas measures.

The findings also point to using additional judiciousness when Program Administrators are considering offering brand-new, un-vetted technologies. Any new measure that has not been market-vetted (including third-party testing and independent reviews) and where feasible should receive in-house testing (by either Program Administrator or implementation staff) for several

weeks to understand whether it is truly a viable candidate for rolling out as a Pilot. The strongest argument for this is to avoid what could be considerable negative reactions and fallout from participants unhappy with the measure.

References

Apex Analytics and Energy Trust of Oregon. 2014. Energy Trust of Oregon Nest Thermostat Heat Pump Control Pilot Evaluation.
http://energytrust.org/library/reports/Nest_Pilot_Study_Evaluation_wSR.pdf

Apex Analytics and Energy Trust of Oregon. 2016. Energy Trust of Oregon Smart Thermostat Pilot Evaluation.
http://energytrust.org/library/reports/Nest_Pilot_Study_Evaluation_wSR.pdf

Navigant Research. 2014. Smart Thermostats: Communicating Thermostats, Smart Thermostats, and Associated Software and Services: Global Market Analysis and Forecasts.
https://www.navigantresearch.com/research/smart-thermostats