# Cutting Peak Demand – Two Competing Paths and Their Effectiveness

*Tony Larson, National Grid, Waltham, MA*
*Mona Chandra, National Grid, Waltham, MA*
*Kathleen Ward, Navigant, Boulder, CO*
*Debbie Brannan, Navigant, Boulder, CO*
*Steven Tobias, Navigant, Burlington, MA*

## ABSTRACT

Thermostats are the focal point of many residential energy management programs, and as thermostat vendors have added greater intelligence through software and hardware innovations, "smart" thermostat technology has become increasingly popular. National Grid is tapping the latest evolution of thermostats to offer a Demand Response (DR) program to owners of one of three thermostat types. The program design offers two paths for participation (based on thermostat technology) with varying incentive levels and event attributes (frequency), providing insight into the effectiveness of the program design. Typically, program savings would be measured using interval meter billing data, however, penetration of residential interval meters is extremely limited in the utility's service territory. As a result, the impact evaluation relies exclusively on thermostat runtime data to measure savings. This paper has broad applicability to future DR program designs as it will assess opt-out rates for two program designs, DR event savings, the impact of pre-cooling and post-event recovery, and details an evaluation approach to evaluate DR program savings without interval data.

## Background

National Grid offers a diverse, yet complementary set of demonstration projects targeted to reduce peak demand and inform the design of future DR programs in Massachusetts and in Rhode Island. The Residential Wi-Fi Thermostat DR program was first offered in Massachusetts and Rhode Island in 2016 and reached over 1,400 customers who enrolled over 2,000 thermostats. The demonstration project was designed to test controllable thermostats as a DR technology (testing various thermostat models from multiple thermostat vendors), as well as customer acceptance of the DR program offerings (testing two program platforms that offer different incentive structures, event frequencies, and event durations in Massachusetts).

The Residential Wi-Fi Thermostat DR program includes two program offerings described in Table 1: ConnectedSolutions (CS) and Rush Hour Rewards (RHR). Each offering varies in thermostat model (ecobee, Honeywell, and Nest), DR event attributes (frequency, duration), incentive mechanism (participation requirement), set back strategy (2°F vs. 3°F), pre-cooling, and event dispatch criteria[1] (day-ahead locational marginal price (LMP)). In 2016, the program largely relied upon a Bring Your Own Thermostat (BYOT) approach. In 2017, the program plans to increasingly target customers who install a

[1] The 2016 DR season began June 1, 2016 and ended September 30, 2016. DR events can be called on non-holiday weekdays for both CS and RHR. CS has no limitations to the number of events called per week; RHR could have at most three events per week. CS events could range from 2-4 hours and RHR events were 4 hours. CS allowed up to 160 event hours; RHR allowed up to 60 event hours. CS events were called when the day-ahead weighted average LMP exceeded $49 per MWh for two or more hours. RHR events were called when the day-ahead weighted average LMP exceeded $62 per MWh for four hours. If the LMP criteria was met for more than four hours, the event was centered around the highest LMP hour.

Wi-Fi thermostat as part of the Home Energy Services (HES) program. National Grid is expecting to more than double the size of the program in 2017.

**Table 1.** Program Design

| Category | ConnectedSolutions | Rush Hour Rewards |
|---|---|---|
| State | Massachusetts and Rhode Island | Massachusetts |
| Types of thermostats | ecobee, Honeywell | Nest |
| Total program duration | 108 hours | 52 hours |
| Event duration | 2-4 hours | 4 hours |
| Advance notification | Day of, >2 hour<br>(customer notified) | Day of, >2 hour<br>(customer notified) |
| DR event opt-out option<br>(before event, during event) | No, Yes (ecobee)<br>Yes, Yes (Honeywell) | Yes, Yes |
| Intended DR set point range | +/- 2°F | +/- 3°F |
| Pre-cooling | No (ecobee), Yes (Honeywell) | Yes |
| Customer incentives per thermostat, for up to three thermostats per National Grid account[2] | BYOT:<br>$25 for sign up<br>$25/year if complete >75% of events<br>HES:<br>Free thermostat and installation<br>$25/year if complete >75% of events | BYOT:<br>$40 for sign up; no event requirement |
| Participant delivery channels | BYOT and HES | Nest |
| Price criteria (weighted average of day-ahead LMP) | $49 per MWh | $62 per MWh |
| Number of days meeting dispatch criteria | 38 days | 18 days |

*Source*: National Grid

CS had 29 events called during the 2016 DR season (Figure 1). Start times ranged from 1:00 p.m. to 4:00 p.m. with event durations ranging from 2 to 4 hours, most events began at 2:00 p.m. and lasted 4 hours. Average temperatures varied throughout the DR season, ranging from 69°F on July 8, 2016 to 89°F on July 22, 2016. The average temperature humidity index (THI)[3] ranged from 67°F to 81°F.

[2] CS: A $25 sign-up incentive is provided for each thermostat enrolled in the program, up to three thermostats per National Grid account. The enrollment incentive is provided once per thermostat. A $25 participation incentive is provided annually for up to three thermostats, assuming each thermostat participates in more than 75% of DR events each year.
RHR: A $40 sign-up incentive is provided for each thermostat enrolled in the program, up to three thermostats per National Grid account. The enrollment incentive is provided once per thermostat. After the first season, the annual incentive is $25 for each thermostat.

[3] The temperature humidity index (THI) is a weather variable that measures the combined effects of temperature (dry bulb) and relative humidity. The THI calculation used in Figure 4 comes from PJM: *PJM Manual 19: Load Forecasting and Analysis*, Effective Date: June 2016, http://www.pjm.com/planning/resource-adequacy-planning/~/media/documents/manuals/m19.ashx

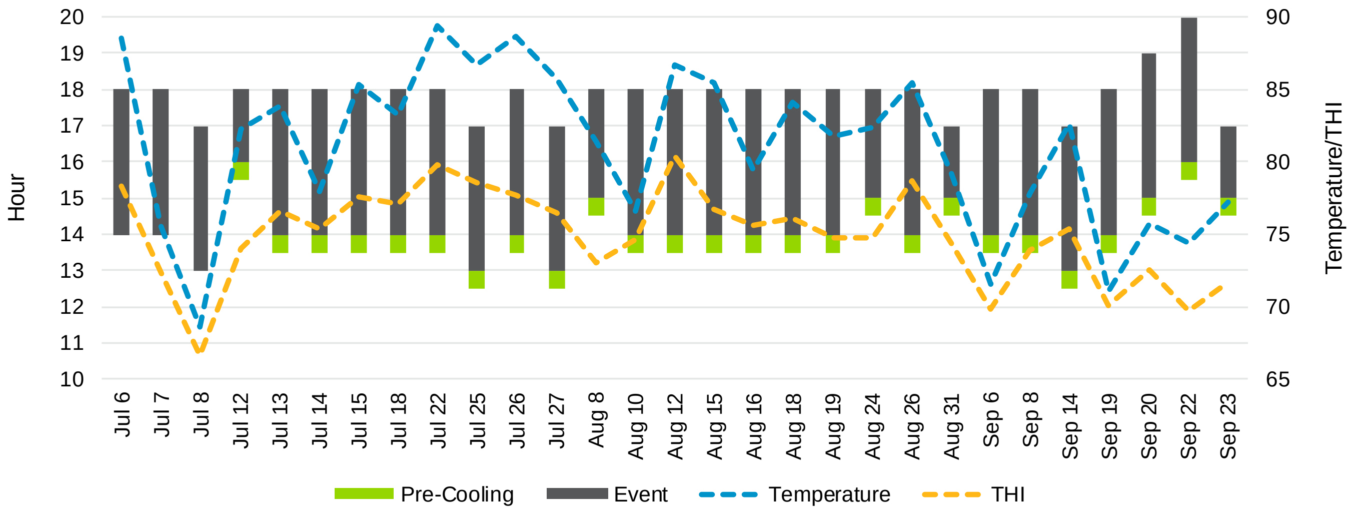


**Figure 1.** Timing of CS events with average temperature and THI.

A total of 13 events were called for RHR during the 2016 DR season (Figure 2). Most events began at 2:00 p.m., although start times ranged from 1:00 p.m. to 4:00 p.m. All events lasted 4 hours. Average temperatures during events were typically higher for RHR than for CS. Temperatures ranged from 74°F on July 7, 2016 to 92°F on August 11, 2016. Average THI ranged from 72°F to 83°F.

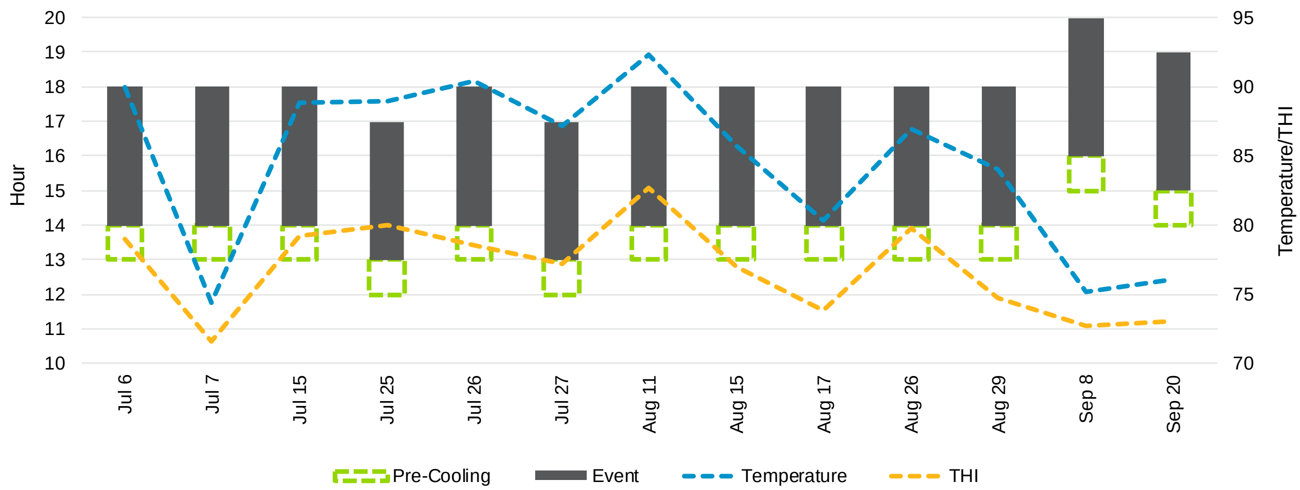


**Figure 2.** Timing of RHR events with average temperature and THI.

As of September 30, 2016, there were 1,492 customers who enrolled 2,065 thermostats in the residential DR program in Massachusetts and Rhode Island.

## Methodology

There are two elements of the impact evaluation that make this program evaluation unique, described in further detail below:

(1) Lack of interval meter billing data. Typically, program savings would be measured using interval meter billing data, however, penetration of residential interval meters is extremely limited in National Grid's service territory. As a result, this study used metering

data to estimate the relationship between runtime and power, applying this relationship to the population of DR program participants.

(2) Randomized Events. The proposed methodology relied upon designated "non-event" days to predict baseline usage; these days met the same day-ahead LMP pricing criteria as event days, but were randomly designated as a "non-event" day (i.e., a DR event was not initiated). However, because event dispatch criteria was based on day-ahead LMP, the temperature on event days varied considerably (from 70°F to 90°F) resulting in an unbalanced temperature distribution across event and designated "non-event" days.

**Converting Thermostat Runtime to Power**

In the absence of interval meter billing data, this study relied on thermostat runtime data to estimate demand impacts. Thermostat runtime could serve as a proxy for power, however, there are limitations with this approach. Most notably, each hour of air conditioning (AC) runtime does not correspond to the same amount of power. Differences in power may be due to outdoor temperature, humidity, operating conditions, etc. As a result, we used a small sample of metering data from the 2017 Massachusetts Baseline Study (nine sites in total) to model the relationship between air conditioning runtime and power, accounting for size and efficiency of the AC unit and temperature.

Figure 3 presents the variation of metered power consumption by AC runtime, showing that power may vary by as much as +/- 10% and confirming the relationship between runtime and power is not constant.

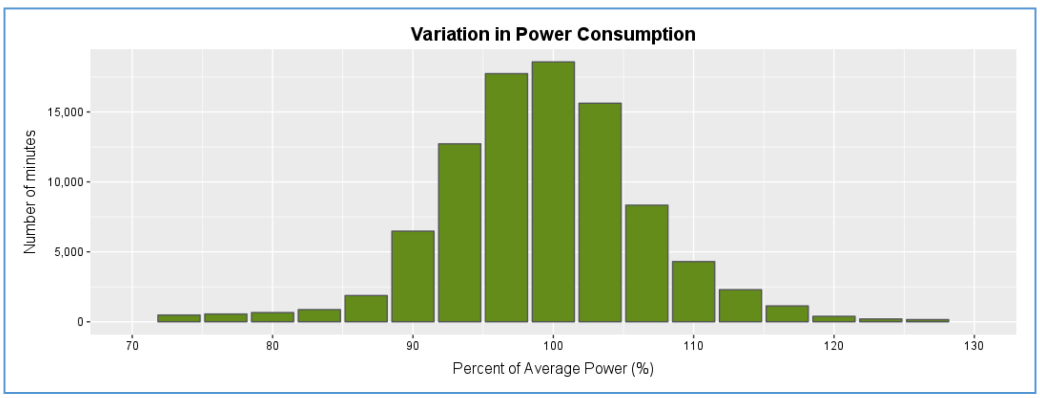


**Figure 3.** Variation in power consumption for metered ACs.

Equation 1 describes the model estimated.

$$kW_{it} = Intercept + \alpha_i + \beta_1 * \frac{Runtime * Btu/h}{EER * 1000} + \beta_2 * CDH + \beta_3 * CDH * \frac{Runtime * Btu/h}{EER * 1000}$$

**Equation 1.** Conversion of Runtime to kW

In Figure 4 we plot predicted power as estimated by Equation 1 against actual power. The model performs well in predicting power as a function of AC runtime, size and efficiency of the air conditioning unit, and weather (multiple $R^2$ is 0.99).

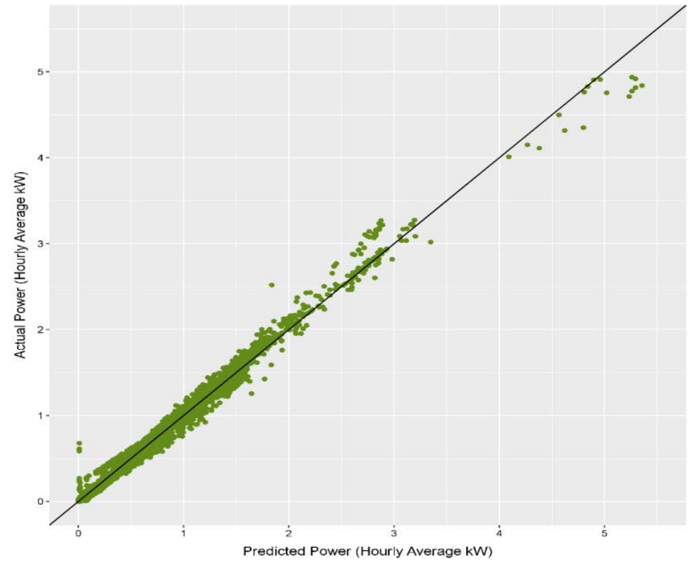


**Figure 4.** Predicted kW vs. actual kW for metered ACs.

Using the coefficient estimates, we converted thermostat runtime to power. As central AC nameplate was not collected for DR participants, we assume average size (2.6 tons) and efficiency (11 Energy Efficiency Ratio) based on the Central AC Digital Check-Up/Tune-Up measure in the 2015 Massachusetts Technical Reference Manual. For example, for a 15-minute interval with 100% runtime at 80°F, the estimated power is 2.6 kW.

While the results of the conversion are promising, there are several limitations to this approach. First, thermostat runtime and AC runtime are not perfectly aligned. There is a delay between the thermostat signaling cooling and the AC unit turning on. A similar delay exists when the thermostat signals the end of a cooling cycle. This delay varies across thermostats and AC systems and was not accounted for in the conversion as thermostat data was not available for the metering sample.

Second, this conversion assumes an AC system size and efficiency based on the Massachusetts TRM. Since no information was collected from CS and RHR program participants, there is no way to verify the accuracy of these assumptions. Given that homes with smart thermostats tend to be larger than the average home, the assumed size and efficiency could be too low. This would mean the conversion is underestimating power and biasing down program impacts.

Finally, this conversion does not account for the effect of DR programs on other end uses. When an event starts and the thermostat sets back, household occupants could be turning on AC window units or fans. This increase in energy use would not be reflected in thermostat runtime. Alternatively, occupants may choose to leave the house during an event, turning off lights and other appliances. Again, thermostat runtime would not reveal these decreases in energy use. By relying solely on thermostat runtime to evaluate demand reductions, secondary impacts from other end uses are excluded from the estimation.

In 2017, a small metering study of DR participants will be conducted which will begin to address each of these three limitations. In addition, a larger sample from the 2017 Massachusetts Baseline Study will be available to re-estimate the AC runtime to power conversion.

## Savings Estimation

The proposed methodology relied on a regression-based within-subject baseline where non-event days serve as the baseline (or counterfactual) for observed impacts on event days. A limitation of the standard within-subject approach is that event days are, in general, systematically different from non-event days, suggesting a baseline based on all non-event days may be biased (Spurlock et al. 2016). To address this potential source of bias, National Grid developed a protocol for randomly assigning days meeting the criteria for calling an event to be designated as an "event" day or a "non-event" day. Designated "non-event" days are identical to "event" days in the criteria required to call a DR event, but the DR thermostat setback algorithm is not initiated.

There were two limitations of this approach: (1) National Grid's event criteria is based on day-ahead LMP pricing. As a result, the temperature on event days varied considerably from (from 70°F to 90°F) resulting in an unbalanced temperature distribution across event and designated "non-event" days. "Non-event" days were, on average, hotter than event days resulting in an inaccurate baseline. To mitigate this imbalance, all other non-event days between June 1 and September 30, 2016 were included. (2) The random protocol resulted in one program having a designated "non-event" day on the ISO-NE summer peak. In 2017, the pilot program will be use an experimental design addressing both limitations and potential bias.

The analysis estimated both the average treatment effect (ATE), in which impacts are estimated for all enrolled devices—full participant, opt out, and nonparticipants, including devices with connectivity issues—and the treatment effect of the treated (TOT), in which impacts are estimated for all full participant-enrolled devices only. These two estimates provide an upper and lower bound for the expected savings if opt out rates and connectivity issues are minimized.

Formally, the model used to predict baseline usage is:

$$\widehat{\boldsymbol{kW_{\imath t}}} = \boldsymbol{\alpha_i} + \boldsymbol{\lambda_t} + \boldsymbol{\beta_{1t} Pre_{it}} + \boldsymbol{\beta_{2t} Event_{it}} + \boldsymbol{\beta_{3t} Post_{it}} + \boldsymbol{\gamma Weather_{it}} + \boldsymbol{\varepsilon_{it}}$$

**Equation 2.** Impact model specification

Where:

$\widehat{kW_{\imath t}}$ = The estimated kW for device *i* in time period *t*.

$\alpha_i$ = A device-specific fixed effect.

$\lambda_t$ = A time fixed effect.

$Pre_{it}$ = A dummy variable equal to 1 if time period *t* for device *i* falls in the pre-event period, and 0 otherwise.

$Event_{it}$ = A dummy variable equal to 1 if time period *t* for device *i* falls in the event period, and 0 otherwise.

$Post_{it}$ = A dummy variable equal to 1 if time period *t* for device *i* falls in the post-event period, and 0 otherwise.

$Weather_{it}$ = A set of weather variables specific to device *i* and time period *t*.

## Results

Figure 5 presents the ATE and TOT per device. The ATE per device, which included all enrolled devices regardless of participation status, reflects actual impacts, acknowledging a portion of devices will opt out or experience connectivity issues. Average demand impacts varied by thermostat model, with the Honeywell and Nest thermostats achieving average demand reductions of 0.39 kW per event and the ecobee thermostat achieving 0.31 kW per event.

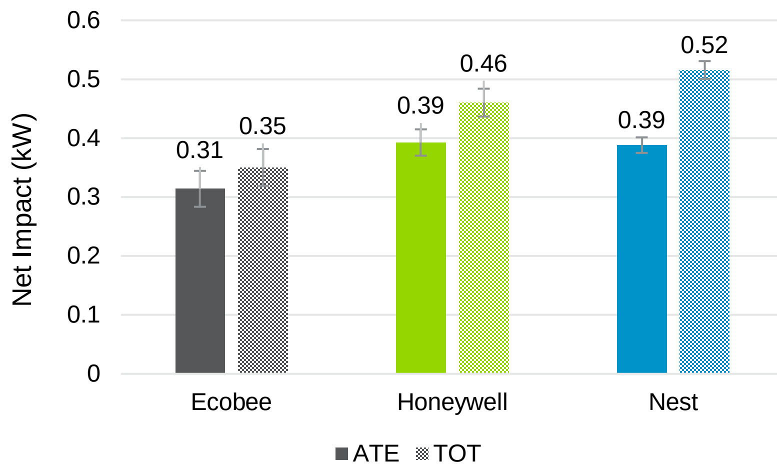


**Figure 5**. Average ATE vs. TOT per device for all events.

The Honeywell thermostat experienced relatively high rates of connectivity issues, while the Nest experienced relatively high opt-out rates. As a result, Navigant also estimated the TOT per device, which only included full participants (i.e., excluding participants with connectivity issues or who opted out). Figure 5 shows the higher TOT average impacts. The Honeywell and ecobee thermostats are both part of the CS program; as a result, the remaining differences in TOT average impacts were likely due to differences in precooling or customer attributes. Differences between the Nest and the ecobee and Honeywell thermostats may be due to differences in program design (set back strategies, event temperature, event duration, etc.).

**Average Hourly Impacts**

The previous section presented average impacts—i.e., the average of hourly impacts over the duration of the event. In this section, we present the hourly impacts for both the event period and for the precooling and recovery periods. The precooling period included the 3 hours preceding an event, while the recovery period covered the 3 hours immediately following an event. Average hourly impacts are displayed in Figure 6 and average impacts in each period in Table 2. Of note, the largest impacts are observed during the final precooling interval (for the Honeywell and Nest, which have precooling) and the first intervals during the event and recovery periods.[4] Impacts during the event are largest during the first hour and steadily degrade. This is a common feature of thermostat DR programs where indoor air temperatures increase throughout the duration of the event, resulting in increased cooling loads, and as some customers choose to opt-out during the event. As shown in Table 2, over the pre-cooling, event and recovery period, each thermostat-type showed a decrease in energy use.

[4] For certain ecobee devices on certain events, the start time of the event began 15 minutes prior to the scheduled start time (e.g., the DR algorithm initiated at 1:45 p.m. rather than 2:00 p.m.). This led to the decrease in demand observed immediately prior to the event start and likely contributed to the lower average impacts for the ecobee thermostat.

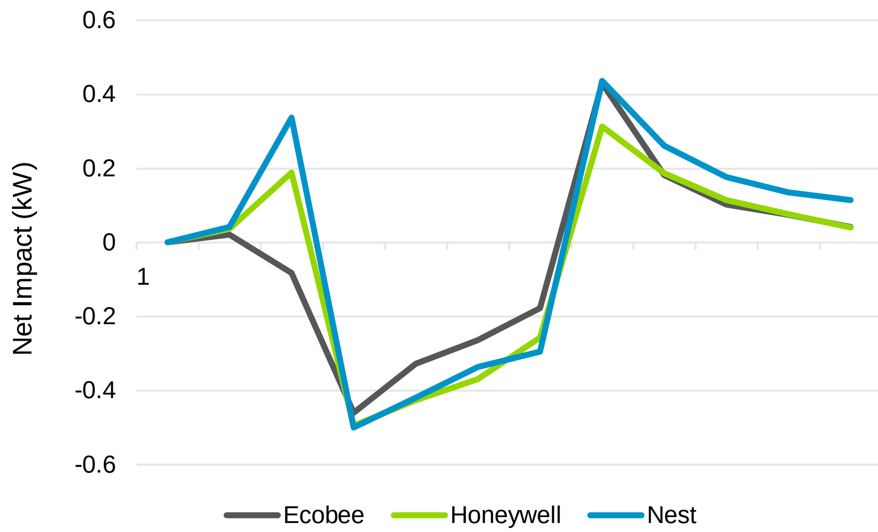


**Figure 6.** Average hourly impacts for all events (ATE).

**Table 2.** Summary of Average Impacts (ATE) by Period

| | Ecobee | Honeywell | Nest |
|---|---|---|---|
| Pre-Cooling (kW) | -0.02 | 0.08 | 0.13 |
| Event (kW) | -0.31 | -0.39 | -0.39 |
| Recovery (kW) | 0.17 | 0.15 | 0.23 |
| Energy (kWh) | -0.82 | -0.90 | -0.8 |

**Event-Specific Impacts**

Figure 7 displays the average impacts (ATE) by event for each thermostat type, along with the average temperature. Note that the average temperature varies by thermostat type—a result of the different geographic distributions of enrolled thermostats. As is evident in the figure, average impacts were correlated with temperature, with impacts ranging from 0.05 kW to 0.2 kW on the coolest event days to 0.6 kW to 0.7 kW on the hottest event days.

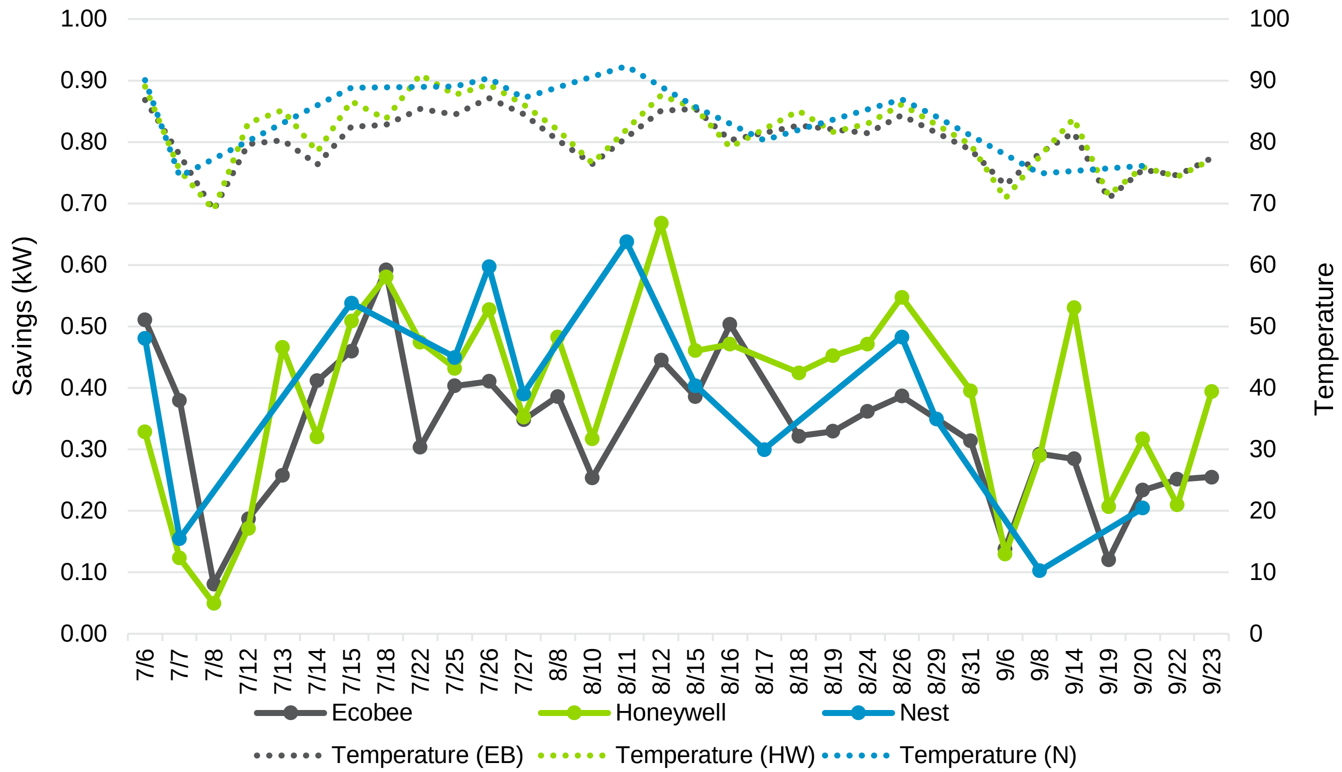


**Figure 7.** Average demand savings (ATE) by event..

### Weather Analysis

Given some events were called when outdoor air temperature was relatively mild, this study estimated the average impacts per event (ATE) for events with an average temperature above and below the temperature threshold of 80°F. As shown in Figure 8, events above 80°F had significantly higher savings impacts as compared to events on sub-80°F days.

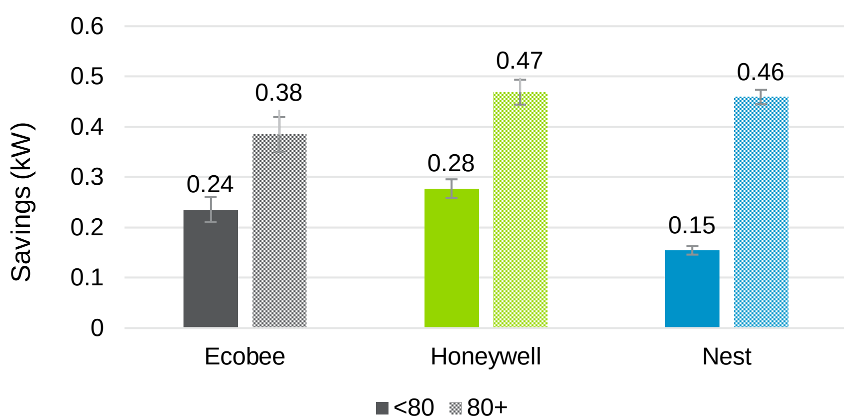


**Figure 8.** Event average demand savings (ATE) by temperature threshold.

Figure 9 presents average impacts (ATE and TOT) for events above 80°F. Compared with Figure 5-7, average impacts are notably higher, ranging from 0.42 kW to 0.59 kW.

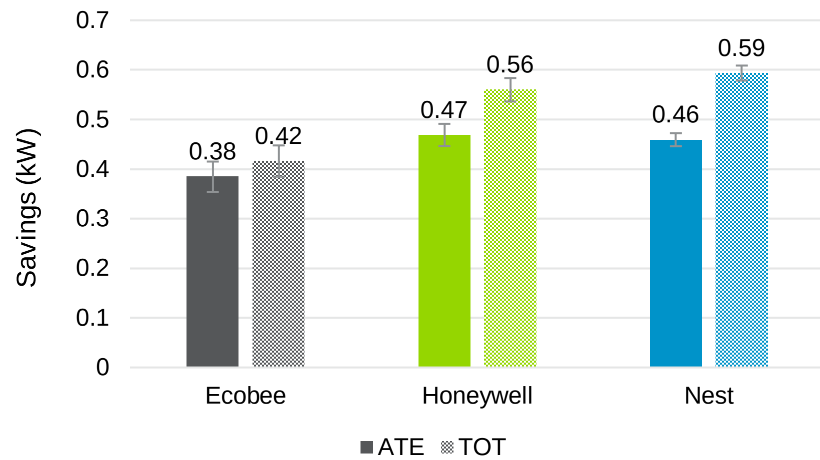


**Figure 9.** Event average demand savings ATE vs. TOT per device for events above 80°F.

**Customer-Level Impacts**

The results presented thus far represent device-level impacts. However, program participants had, on average, 1.4 thermostats. To better assess the DR associated with a participating customer, Navigant re-analyzed the thermostat telemetry data, aggregating it to the customer level. Figure 10 presents a comparison of the average device-level and customer-level impacts (ATE) per event. Customer-level average impacts (0.46 kW to 0.54 kW) are larger than device-level average impacts (0.31 kW to 0.39 kW).[5]

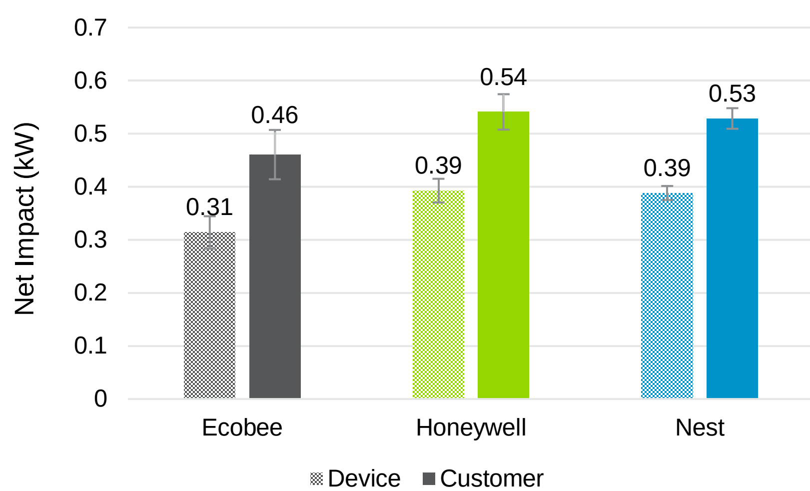


**Figure 10.** ATE comparison: device vs. customer.

Figure 11 presents a comparison of customer-level impacts when including all enrolled customers (ATE) and only full participants (TOT). Customer-level impacts for full participants ranged from 0.51 kW to 0.70 kW.

[5] Navigant analyzed device-level impacts for customers with one, two, and three thermostats to determine whether average impacts vary with the number of thermostats. Results suggested each additional thermostat yields lower impacts, though this difference was not statistically significant.

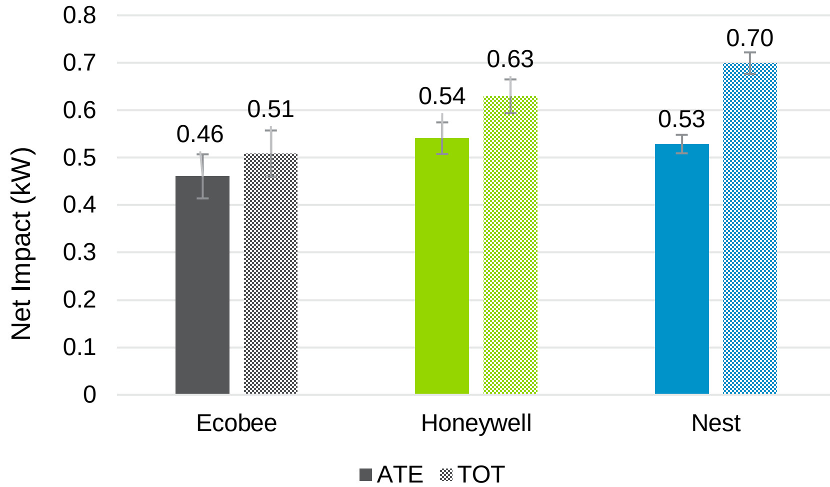


**Figure 11.** Average ATE vs. TOT per customer.

## Conclusion

This study found National Grid's Residential Wi-Fi Thermostat DR program was successful in testing the thermostats as a residential DR technology and customer acceptance of the program offering. The evaluation shows promise for thermostats as a residential DR technology, though important differences exist across different thermostat models. The study revealed important findings regarding how various program design features affect customer acceptance of the DR program offerings and program savings. National Grid is now positioned to leverage the experience of the 2016 program year to further test the technology and the Residential Wi-Fi Thermostat DR program offering in 2017.

## References

Spurlock, A., Baylis, P., Cappers, P, Jin, L. and A. Todd. "Go for the Silver? "Gold Standard" RCT Versus Quasi-Experimental Methods." Presentation at the Behavior, Energy and Climate Change Conference, 2016. Available at: http://beccconference.org/wp-content/uploads/2016/10/Spurlock_presentation.pdf