QUANTIFYING THE IMPACT OF OVERRIDE BEHAVIOR ON THE PERFORMANCE OF A SUMMER DIRECT LOAD CONTROL PROGRAM

by

Pamela Jordan Wildstein

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science (Environment and Sustainability) in the University of Michigan April 2022

Thesis Committee:
Assistant Professor Michael Craig, Chair
Assistant Professor Parth Vaishnav
Abstract

Demand response (DR) programs represent an important tool for mitigating electric grid reliability risks posed by extreme weather, heat events, and increased electrification. Residential direct load control (DLC) programs with behavioral elements, such as an override option, are often used during the summer to reduce load from air conditioners (AC). Although override options make DLC more attractive to consumers, the high likelihood of an imperfect response can significantly reduce the effectiveness of the program overall. Thus, utilities must understand the impact of these behavioral elements on load-shaving capabilities. To fill this gap, we design a regression-based, deterministic model to predict the behavior of thermostats participating in ecobee’s Donate Your Data initiative under various temperature conditions. The model is then used to quantify the impact of the override option on the performance of 403 ecobee thermostats participating in Southern California Edison’s 2019 Summer Smart Energy Program. Our analysis estimates that although the group’s participation in the DLC program led to a 41% reduction in AC demand, 48% of potential load reduction was lost to exercise of the override option. Similarly, most of the so-called DR events followed a pattern of near-perfect participation for a short duration of time preceded by a steady increase of overrides as the event progressed. DLC programs with an override option are effective at reducing demand in the aggregate, but designers must consider the savings lost to behavior.

Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1841052. We gratefully acknowledge the support and contribution of ecobee and ecobee customers to this research.

1 Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.
Table of Contents

1. Introduction ................................................................................................................. 4

2. Methods ....................................................................................................................... 6
   a. Data .......................................................................................................................... 6
      i. Southern California Edison Smart Energy Program ........................................... 6
      ii. Ecobee Donate Your Data Initiative ................................................................. 6
      iii. Outdoor Temperature ....................................................................................... 9
   b. Overall Model ....................................................................................................... 9
      i. Regression to Predict the Indoor Temperature ............................................... 10
      ii. Threshold Estimation and Control Logic of Thermostat to Predict State of AC .......................................................................................................................... 10
      iii. Indoor Temperature Regression Training ...................................................... 11
      iv. Model Implementation ..................................................................................... 11
   c. Performance Metrics ........................................................................................... 12

3. Results ....................................................................................................................... 12
   a. AC Behavior Model Quality .............................................................................. 12
   b. Override Analysis ............................................................................................... 14
   c. Impact of Behavior on AC Runtime Savings ................................................. 16
   d. Single Event Behavior Analysis ..................................................................... 19

4. Discussion ............................................................................................................... 21

5. Bibliography ............................................................................................................ 24
1. Introduction

Extreme weather events pose a major threat to the electric grid’s ability to match electricity demand with supply. Recent blackouts in Texas[1], [2] and California[3], [4] point to the susceptibility of many types of power system assets, with the impact of heat waves emerging as a particular concern. Extreme heat can reduce the performance of grid infrastructure [5], while high air temperatures may lead to increased load through greater electricity demand from air conditioning (AC)[6]. Decarbonization pathways in the power system also rely on load flexibility for managing variability in wind and solar power and increased electrification of demand [7], [8].

Demand response (DR) programs, or the change in consumption of electricity in response to a price or incentive signal [9, p. 2], [10], are one strategy for procuring capacity during extreme weather events[11]. Instead of dispatching new generation to meet demand, grid operators can utilize behavior changes as a resource to nudge load in the direction necessary to best manage the system. In this sense, well defined DR programs have the potential to provide reliable capacity to the grid and reshape the role of the consumer within the system.

DR can be split into two categories: price-based programs (PBP) and incentive-based programs (IBP) [10]. PBP relies on dynamic pricing to nudge customers in shifting and shaving load. The emphasis is on encouraging certain behaviors of at specific times to better suit overall grid stability. In contrast, classical IBP uses bill credits or discounted rates to compensate participants [10]. Summer direct load control (DLC) is a popular, classical IBP scheme directed at residential and small business customers to shave peak load during the hottest days of the year[10]. Program participants cede control of an appliance during a small number of event days to reduce electricity usage [12]. In the case of AC units, which are the most common form of DLC programs, the utility either adds equipment to the thermostat or communicates with a smart thermostat to remotely increase the setpoint or limit the AC’s cycling runtime [12], [13]. Using DLC for AC can be especially effective at shaving load because the program is automated and does not rely on a customer’s willingness to respond to real-time price signals [12]. DR may be rooted in alterations to consumer behavior, but the design of DLC offers a certainty in load reduction that is not inherent to other programs.

Previous literature has revealed that consumers hesitate to enter residential DR and DLC schemes. Parrish et al. (2019) [14] finds that DR models may overestimate the value that the resources can provide to the grid due to improbable assumptions on the extent to which customers will participate in programs. Specifically, customers do not always act predictably and there is no guarantee that they will accept automation [13]–[15]. In their paper, Parrish et al. (2019) [14] uses the U.S. Electric Power Research Institute’s definitions to frame the ways in which modelers can overestimate the impact of DR programs: the number of participants that enroll (participation), how they respond to signals or perform in the program (performance), and if they choose to continue their enrollment in the program (persistence). Parrish et al. (2020)[16], Stenner et al. (2017)[13], Sintov and Schultz (2017)[17], Fell et al. (2014)[15], and Annala et al. (2012)[18] all attribute these participation, performance, and persistence failures partially to distrust of the utility and the overall loss of control. Sintov and Schultz (2017)[17] specifically report that customers value the agency in manually adjusting their thermostats above the ease of autonomous operation. Similarly, Fell et al. (2014)[15] finds that customers fear that
enrollment in a DLC program could lead to a lack of electricity when they need it most with a “Big Brother” figure dictating their usage.

One method for easing these anxieties is to introduce event “opt-out” and “override” features into the DLC program design, both of which reduce the risk of participation and reintroduce a sense of control. Xu et al. (2018) confirms the importance of these participant controls in DLC programs and finds that the existence of an override option was more effective at recruiting potential customers than increasing financial incentives. This inclusion of an override option adds a behavioral element back into what is designed to be a certain, stable reduction in load and can create substantial variability in the performance of participants.

Empirical literature on the potential role of overrides in undermining the value of DLC programs is limited. Newsham and Bowker (2010) note that although residential AC DLC programs are effective at reducing peak load when events are called, the positive relationship between the likelihood of thermostats initiating an override and the duration of an event can pose a challenge. To quantify override frequency, drivers, and consequences on electricity demand, Sarran et al. (2021) analyzes 2019 ecobee thermostat data from roughly 6,000 DLC participants. The paper concludes that there is a direct relationship between the magnitude of the override rate and the missed savings for the utility. In some instances, a significant number of overrides lead to the savings from an event being cancelled out completely due to the increased AC runtime when the setpoint was lowered.

While Sarran et al. (2021) provides an important initial contribution to empirically quantifying the effect of overrides on DLC programs, it suffers from three significant shortcomings. First, Sarran et al. (2021) quantifies the effect of overrides by comparing consumption between households rather than comparing consumption within a household using counterfactual consumption profiles. This approach ignores numerous factors separate from overrides that could drive differences in household electricity consumption, including the building envelope, occupancy, occupant behavior, and location. Second, the paper is concerned with understanding the override behavior of a large sample of thermostats across a broad geographic area. While this is effective at establishing a general understanding of thermostat override behavior, it does not establish insights into the performance of a specific utility’s program, inclusive of any nuances. Finally, Sarran et al. (2021) considers a program design with a pre-cooling option. Not all DLC programs use extra electricity to cool the home before the event begins, which can lead to drastic differences in performance between designs. There is still a need to quantify the impact of behavior within programs that do not choose to pre-cool homes. Overall, these research gaps mean that utilities that choose to include an override option in their program design must understand how the addition of this behavioral element will impact the load shaving capabilities of the program.

To address these gaps, we develop and apply a new method for quantifying the performance of a DLC program with an override option using a within-household experimental design. In contrast to past literature, we analyze a single utility’s program, thus bounding the scope to a single climate, and seek to quantify the impacts of override behavior by combing a regression-based model with the control logic of the thermostats. We then apply this method to a set of DYD thermostats participating in Southern California Edison’s (SCE) summer DLC program, Smart
Energy Program (SEP), for the summer of 2019. Unlike Sarran et al. (2021)[20], the program we chose to analyze does not offer a pre-cooling option.

2. Methods
To quantify the impact of override behavior on the value of a DLC program, we identify, analyze, and model the performance of 403 DYD dataset[21] thermostats participating in SCE’s summer 2019 SEP. A model is developed to estimate the number of seconds a home’s heating, ventilation, and air conditioning (HVAC) system will run in a five-minute interval, given the external temperature and the thermostat setpoint. The model includes two components: (1) a multiple linear regression to predict the indoor temperature of the home given HVAC operations and (2) a thermostat control algorithm for HVAC operations. The model is then run on each home twice to predict both the expected operation without enrollment in SEP and a perfect response with no overrides. Comparing these results with the actual operation produces an estimate for the impact of the program both with and without the override option.

a. Data
We combine data from the SCE’s SEP event history[22], ecobee’s DYD initiative [21], and the National Oceanic and Atmospheric Administration’s (NOAA) Local Climatological Data (LCD) tool[23] to form a comprehensive picture of the sample groups’ homes and HVAC systems during the DR events.

i. Southern California Edison Smart Energy Program
Southern California Edison (SCE) is an investor-owned utility that serves the bulk of southern California and offers many DR programs for its customers. Most relevant to this work is its summer cooling DLC program, Smart Energy Program (SEP), in which enrolled homes can receive a $75 rebate for their thermostat and up to $40 in bill credits yearly for participating in DR events[24]. When an event is called, a signal is sent to the smart thermostat to raise the cooling setpoint by a few degrees. Homes are not automatically pre-cooled in anticipation of this change, but participants are given a notice up to three houses in advance. If desired, they can pre-cool the home themselves[25]. Events last no longer than four hours and participants are alerted before events are called[25].

SEP events were called 22 times in the summer of 2019, between July 24th and October 22nd. Events occurred between 1 P.M. and 9 P.M. Events were predominately called during peak hours. 55% began at 6 P.M. and 60% concluded at 8 P.M. One event lasted ten minutes, 11 events lasted two hours, one event lasted three hours, and five events lasted four hours.

ii. Ecobee Donate Your Data Initiative
To quantify the impact of override behavior on the performance of the SEP, we analyze a subset of ecobee Inc.’s Donate Your Data (DYD) dataset that were identified as participants. Ecobee is a leading smart thermostat and home company. The DYD initiative allows ecobee customers to make their anonymized thermostat profile available to researchers. Data is recorded in five-minute intervals and includes the data fields listed in Table 1. Ecobee thermostats allow users to wirelessly connect multiple temperature sensors to the thermostat. Users can select one or more sensors, and the average of the readings of these sensors provides the thermostat with a “temperature control.” We assume that this “temperature control” represents the indoor
temperature. Ecobee smart thermostats control the home or building’s HVAC system based off a schedule of cooling and heating setpoints. If the calculated indoor temperature rises above the sum of the cooling setpoint and a user-defined threshold, the compression cooling system will begin to cool the home until it reaches the setpoint. A thermostat user can manually change the setpoint at any time, initiating a so-called “hold” calendar event. If the user manually lowers the setpoint that has been automatically raised as part of a demand response event, this is called an “override.” This paper is mainly concerned with quantifying the impacts of this override behavior. Figure 1 depicts the average number of seconds a thermostat’s AC system ran throughout the day of July 26, 2019, relative to the average outdoor temperature.

<table>
<thead>
<tr>
<th>Data Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Province/State</td>
<td>The state in the United States or the Canadian province the thermostat is located in.</td>
</tr>
<tr>
<td>City</td>
<td>The city the thermostat is located in.</td>
</tr>
<tr>
<td>Time</td>
<td>The date and time of the reading. It is taken in five-minute intervals.</td>
</tr>
<tr>
<td>HVAC Mode</td>
<td>Indicates which system(s) the thermostat is set to use. Options include Heat, Cool, Auto, and Off.</td>
</tr>
<tr>
<td>Calendar Event</td>
<td>Any modifier to the thermostat’s pre-set schedule. Indicates demand response program participation.</td>
</tr>
<tr>
<td>Temperature Control</td>
<td>The average indoor temperature based on relevant sensors.</td>
</tr>
<tr>
<td>Cooling Setpoint</td>
<td>Indoor temperature at which the thermostat will begin cooling the home.</td>
</tr>
<tr>
<td>Fan</td>
<td>Run time in seconds for the fan. Recorded in periods of 15 seconds.</td>
</tr>
<tr>
<td>Compression Cooling System 1</td>
<td>Run time in seconds for the first cooling system. Recorded in periods of 15 seconds.</td>
</tr>
<tr>
<td>Compression Cooling System 2</td>
<td>Run time in seconds for the second cooling system. Recorded in periods of 15 seconds.</td>
</tr>
</tbody>
</table>

*Table 1: Data fields found in the DYD dataset that are used as part of this study’s methodology with a short description.*
To be considered a member of the sample group, thermostats must be located within SCE’s territory and show evidence of participation in at least one SEP DR event. Participation was determined through a DR event in the calendar event column and the HVAC system settings. The DYD dataset does not include a standard calendar event for intervals in which the thermostat is participating in a DR event. Instead, we investigated the calendar events present in the dataset during the intervals of called events, ground truthing the events with thermostat behavior, and verification with ecobee research staff[26] to determine the likely calendar event for the SEP program. For HVAC settings, a thermostat could only be considered a participant if the system was set to “Auto” or “Cool” for at least one interval during the event. Ecobee thermostats include four options for HVAC system behavior: “auto,” “cool,” “heat,” and “off.” The “auto” and “cool” settings imply that the AC can be used to cool the home if the indoor temperature surpasses the cooling setpoint. In contrast, the “heat” or “off” options prevent any cooling from occurring. In those cases, the relationship between the sum of the indoor temperature and cooling threshold, and the cooling setpoint does not hold, regardless of if the cooling setpoint had been surpassed. If the HVAC setting is set to “heat” or “off,” then we cannot assume that there would have been savings from raising the cooling setpoint through a DLC event because the AC would have never been on, regardless of utility intervention. The final sample consisted of 403 thermostats after excluding thermostats that did not have any DR calendar events or that had been set to “heat” or “off” during events.

Figure 2 illustrates the frequency at which the 403 thermostats participated in demand response events and the number of participants per event. On average, a home participated in 13 out of the 22 called events with a mode of 16. Events had on average 239 participants, with 50% of events consisting of at least 309 participants. The largest event had 362 participants, or 90% of the sample.
iii. Outdoor Temperature

Outdoor temperatures were taken from the hourly dry bulb temperature column of the National Oceanic and Atmospheric Administration’s (NOAA) NOAA’s Local Climatological Data (LCD) tool[23]. The database consists of summaries for approximately 950 U.S. Automated Surface Observing System stations[23]. Each thermostat was paired with the station closest to its indicated city with a data coverage of 100% for the year 2019.

Average outdoor temperatures during events from 73°F to 85°F, with a mean of 79°F. 75% of events had an average temperature higher than 77°F and 25% had an average temperature higher than 82°F.

b. Overall Model

In this paper, we develop a model consisting of a multiple linear regression to estimate the indoor temperature and the control logic of the thermostat to ascertain if the cooling system was on or off for a five-minute interval. Figure 3 illustrates this design.

Figure 2: (Left) A histogram of the distribution of the number of events thermostats in the sample participated in. (Right) A bar graph showing the number of sample thermostats that participated in each event.

Figure 3: The workflow of the model. Circles are the independent variables of the model, red rounded squares represent the indoor temperature regression and control logic of the thermostat, and the squares depict the results of the model. Blue
variables are lags of the indoor temperature, yellow variables are lags of the outdoor temperature, purple elements are the state of the cooling system, and gray elements are the cooling setpoint and estimated thermostat threshold assumed from the data.

i. Regression to Predict the Indoor Temperature
To estimate the indoor temperature of the home for a five-minute interval, a multiple linear regression is designed and trained. Since so few homes have a second compression cooling system, the compression cooling system 1 and compression cooling system 2 columns were summed to make a single compression cooling column with the number of seconds both systems ran. Equation 1 illustrates the regression where $y_t$ is the indoor temperature for time $t$, $\beta_s$ are coefficients, $x_{1,(t)}$ is the indoor temperature indoor temperature lagged by one five-minute period, $x_{1,(t-2)}$ is the indoor temperature indoor temperature lagged by two five-minute periods, $x_{2,t}$ is the outdoor temperature, $x_{2,(t-1)}$ is the outdoor temperature lagged by one five-minute period, $x_{2,(t-2)}$ is the outdoor temperature lagged by two five-minute periods, $x_3$ is a binary 0 or 1 representing if the cooling system is on or off, and $\epsilon$ is the error term.

\[
y_t = \beta_0 + \beta_1 x_{1,(t-1)} + \beta_2 x_{1,(t-2)} + \beta_3 x_{2,t} + \beta_4 x_{2,(t-1)} + \beta_5 x_{2,(t-2)} + \beta_6 x_3,(t-1) + \epsilon
\]

Equation 1: A multiple linear regression that uses lags of the indoor temperature, outdoor temperature, and state of the cooling system to predict the indoor temperature of the home.

We performed a Durbin-Watson test and concluded that there was strong evidence of autocorrelation among the regressors. Strong collinearity was also established between the outdoor temperature and indoor temperature.

ii. Threshold Estimation and Control Logic of Thermostat to Predict State of AC
To determine if the cooling system is on or off for each period, the indoor temperature result from the regression is compared to the threshold setting of the thermostat and current cooling setpoint. Figure 4 illustrates the control logic of the thermostat. The threshold for each home was estimated by taking the median of the difference between the indoor temperature and cooling setpoint for intervals where the cooling system switched from off to on and the HVAC Mode had been set to “auto” or “cool” for at least one hour. This value was then rounded to the nearest 0.5 degrees Fahrenheit, starting at the lowest possible setting and default, 0.5°F [27]. This mimics human behavior and ensures that no thermostat in the sample is estimated to have a threshold below the minimum.
iii. Indoor Temperature Regression Training

To account for any differences between the homes in the sample that are not fully captured by the DYD dataset, the regression is trained separately for each thermostat. For training, the original dataset is filtered in many ways. First, the trainer dataset only considers the summer, June to October, during intervals where the outdoor temperature is more than or equal to the minimum outdoor temperature that occurred during the DR events the thermostat participated in. This is done to simulate moments when a DR event is most likely to be called. Second, the HVAC Mode must have been set to “cool” or “auto” for at least an hour. If the mode is “off” or “heat” the cooling system will never turn on, so these periods are irrelevant to our analysis. Third, the calendar event cannot be set to the smart recovery, smart away, smart home, or Homekit hold features. These algorithms preemptively cool the home in anticipation of changes of schedule transitions, and the relationship is unlikely to be easily modeled. Finally, both the fan and compression cooling system must be on or off. Moments where the fan is cycling air, but the compression cooling system is off are not applicable to savings from this DLC program design because the SEP is only intended to target the AC system. The fan and compression cooling system are considered to be off if they do not run at all during the interval, and on if they are recorded at 15 or fewer seconds. Because of the nature of the lags, each interval’s regression is dependent on the results of the prediction for the time step before it. Therefore, the regression must be fit for each time step sequentially when testing overall accuracy.

iv. Model Implementation

To predict the AC runtime of events for both perfect participation and no participation, the trained indoor temperature regression and control logic of the thermostat were used. The SEP is designed to raise the cooling setpoint of the home by a predetermined number of degrees at the start of the event and hold it constant throughout the duration of the event or until an override is initiated. Thus, the cooling setpoint for perfect participation is estimated to be the setpoint recorded in the first five-minute interval, before any override can occur. The setpoint ten minutes before the event began is used for modeling the AC runtime if the thermostat had not been enrolled in the SEP.
Each thermostat was only evaluated for the events that it participated in. If the thermostat’s calendar event column did not include a DR calendar event during the first five minutes of the event, then it was not considered a participant. If an interval did not have an HVAC mode of “auto” or “cool,” it was not evaluated as part of this analysis because the system was turned off.

c. Performance Metrics
To describe the operation of the sample group thermostats, three metrics were used: the percent change in demand from the recorded participation in the event, the percent change in demand from the modeled perfect performance in the event, and the percent of AC runtime savings lost to the override option. The first two metrics rely on the modeled AC runtime of the thermostats if they had not participated in the program or specific event under consideration. The third, in contrast, considers the difference between actual participation and the predicted performance with not override option.

The percent change in demand from participation in the event, inclusive of the override option, is calculated using Equation 2:

\[
\text{change in demand from DR event} = \frac{\text{actual AC runtime} - \text{expected AC runtime}}{\text{expected AC runtime}} \times 100\% \quad (\text{Eqn. 2}).
\]

The percent decrease in demand associated with a perfect response to the event, and no usage of the override option, was calculated using Equation 3:

\[
\text{change in demand from perfect participation} = \left(\frac{\text{perfect AC runtime} - \text{expected AC runtime}}{\text{expected AC runtime}}\right) \times 100\% \quad (\text{Eqn. 3})
\]

The percent of AC runtime savings lost due to participant exercise of the override option was calculated using Equation 4:

\[
\text{change in demand reduction due to override behavior} = \frac{\text{perfect AC runtime} - \text{actual AC runtime}}{\text{perfect AC runtime}} \times 100\% \quad (\text{Eqn. 4})
\]

3. Results

a. AC Behavior Model Quality
To assess the performance of our model, we conduct ten-fold cross validation is preformed to ascertain the coefficient of determination (R²) and root mean square error (RMSE) goodness of fit estimates for each thermostat’s individual regression. The R² of the regression is calculated using Equation 5, where RSS is the sum of squares of the residuals and TSS is the total sum of squares, and the RMSE is calculated is calculated through Equation 6, where \( r \) is the residuals and \( N \) is the number of observations.

For the average home, 98% of the variance in the indoor temperature is explained by the variables included in the regression. 75% of the thermostat’s had an R² above 98%, while 25% was calculated to be 100%. The average RMSE of the indoor temperature regressions was 0.19 °F. 75% of the sample had a RMSE below 0.21°F and 25% below 0.13°F.
The accuracy score, sensitivity, and precision metrics are used to evaluate the overall model’s ability to predict the state of the HVAC system for a given interval. Figure 5 presents a histogram of both the accuracy of the overall model and individual events. The average thermostat’s model accurately predicts if the AC unit is on or off 77% of the time, with 75% of the sample having an accuracy score above 69% and 25% above 84%. The model preforms particularly well for demand response events. These are the moments this analysis is primarily concerned with. The left graph of Figure 5 depicts a histogram of the accuracy scores for demand response event the thermostats in the sample participated in. The mean accuracy of prediction during demand response events is 87%, with 75% of DR events having an accuracy above 77% and 50% of events having an accuracy of 100%. We reason that this high accuracy for demand response events can be attributed to the larger difference between the cooling setpoint and indoor temperature. This makes it more likely that the deterministic HVAC relationship will predict correctly, even in the event of an error from the indoor temperature regression.

Figure 6 presents a scatterplot of each thermostat’s prediction sensitivity and precision. The sensitivity metric considers the number of times the HVAC was observed to be on and predicted to be on. Similarly, the precision metric is the fraction of instances where the HVAC was correctly predicted to be on. Sensitivity and precision are widely spread with a positive correlation. The mean sensitivity and precision are 50% and 59%, respectively. The 25th percentiles of the sample for sensitivity and precision are 37% and 48%. The 75th percentiles of the sample for sensitivity and precision are 61% and 70%.
b. Override Analysis
We analyze the occurrences of overrides based on the duration of events, indoor temperature, and discomfort. Override rates varied from 0% to 29% with the average event seeing an override rate of 14%. Figure 7 illustrates the relationship between the rate of override and the duration of events. In general, longer events are associated with a higher override rate.

Figure 7: A scatter plot showing the relationship between the rate override and the duration of events in minutes. The data’s fitted, linear regression is included.

Figure 8 presents the distribution of the time of override relative to the duration of the event. For one-hour events, the average thermostat that initiated an override did so 24 minutes into the event and 75% of overrides were initiated before the halfway point of the event. For two-hour
events, the most frequent duration used by SCE, the average thermostat overrode the event 52 minutes into the event and 75% of overrides were initiated before 70 minutes into the event. For the single three-hour event dispatched in 2019, the average thermostat overrode the event 80 minutes into the event and 75% of overrides were initiated before 120 minutes into the event. The average thermostat that overrode a four-hour event did so 101 minutes in and 75% of overrides were initiated 155 minutes into the event. Overall, the mean time of override varies little between events when understood as the percent of the event they participated in. The average override occurred 40% of the way through a one-hour event, 43% for a two-hour event, 44% for a three-hour event, and 42% for a four-hour event. On average, and across all event durations, the average override was initiated 43% of the way into the event.

Figure 8: A histogram for each event duration illustrating the minutes into the event overrides occurred.

Figure 9 illustrates the range of indoor temperatures at which an override was initiated across all events to give insight into what program participants consider comfortable. The average indoor temperature at the time of override was 79°F with 75% of overrides occurring at a temperature above 77°F and 25% occurring at a temperature above 80°F. In other words, the data indicates that people may begin to feel discomfort at around 77°F.
Figure 9: A histogram of recorded indoor temperatures at the moment of override.

Figure 10 depicts how large the difference between the indoor temperature and the preferred temperature of the home needs to be before participants initiate an override. We consider the preferred temperature of the home to be the cooling setpoint ten minutes before the start of the event. The average difference between the indoor and preferred temperatures was 2°F at the time of override. 75% of overrides were initiated at a difference above 1°F and 25% above 4°F. As shown by Figure 10, the distribution is relatively consistent across event durations. Therefore, regardless of the duration of the event, program participants tended to become uncomfortable enough to override the event between about a 1°F and 4°F difference.

Figure 10: A histogram of the difference between the recorded indoor temperature and the preferred setpoint (the cooling setpoint in the five-minute interval before the start of the event) at the time of override. Different colors indicate different event durations in minutes.

c. Impact of Behavior on AC Runtime Savings
Analysis of estimated participation with no override option and expected AC behavior without any participation concludes that the DLC program as designed is associated with a 41% savings
in AC runtime across all events. A perfect response from participants without usage of the override option is predicted to have led to a 60% reduction in AC runtime. Thus, override usage behavior is associated with a 48% difference in demand reduction, or 19 percentage point decrease.

Figure 11 shows a breakdown of the aggregated results for all events by month and Figure 8 shows a breakdown by hour. Results varied by month. The majority of events and abated seconds of AC runtime occurred August and September, giving them the most potential for load reduction. August had a greater change in demand from event participation at a 46% decrease and a greater change in demand from perfect participation 66% decrease, relative to September’s 38% and 59% decrease, but was impacted more by the override behavior. August is calculated to have a 58% decrease in demand reduction due to override behavior and September is calculated to have a 51% decrease. July had the largest percent reduction in demand from event participation, 47%, and potential percent decrease in demand from perfect participation, 70%, but a 73% percent decrease in demand reduction due to override behavior, making the month’s savings the most impacted by override behavior. October had the lowest percent decrease in demand from event participation at 24%, the lowest percent decrease in demand from perfect participation at 26%, and the lowest percent reduction in demand due to override behavior at 2%.
Figure 11: (Top) a breakdown of the three metrics of analysis by month. (Bottom) The seconds of AC runtime for each month broken down by no event override, override, and no event participation scenarios.

Figure 12 illustrates how results vary by hour. Trends tend to follow if the hour is considered on-peak, in this case 5 P.M to 8 P.M., or off peak. The on-peak hours are defined by both a higher magnitude of seconds of AC runtime and a higher impact of override behavior on demand reduction. The average percent decrease in demand reduction due to override behavior for off-peak hours is 7%, compared to on-peak’s average reduction of 59%. This presents a challenge for program designers. The goal is to reduce load during peak hours, or the times when participants are most likely to override events.
d. Single Event Behavior Analysis

Single events generally followed the same trends regarding behavior when the results for AC runtime of all participating thermostats are aggregated. Figure 13 is an example of a shorter two-hour event on July 24th and Figure 14 is an example of a longer four-hour event from September 5th. Events start out with a low percent decrease in demand reduction due to override behavior and almost equivalent values for actual participation and perfect participation. After the first few five-minute intervals, participants begin overriding the DLC settings, so there is a steady decline in the actual percent decrease in load from participation and steady increase in percent decrease in demand reduction due to override behavior. In the shorter two-hour events, the load reduction lost to override behavior even out during the last hour. This contrasts with the four-hour events where events often end with an increase in load relative to if there had been no participation.
Figure 13: (Top) a breakdown of the three metrics of analysis by five-minute interval for a two hour event on July 24th. (Bottom) The seconds of AC runtime broken down by no event override, override, and no event participation scenarios at five-minute intervals for a two hour event on July 24th
4. Discussion
To better understand the impact of behavior on the energy savings from a DLC program, we used household-level five-minute ecobee thermostat data to quantify the impact of participant-initiated overrides on the 2019 Southern California residential Summer Energy Program. We find that overrides have a significant impact on the demand reduction from the called DR events. Although the program is associated with a 41% reduction in AC runtime, participant exercise of the override option led to a 49% decrease in demand reduction. We estimate that there would have been a 60% reduction in AC runtime if no thermostats had overridden the DLC program’s settings. Events saw, on average, a 14% override rate, with a positive correlation between the duration of the event and the percentage off participants that overrode the utilities settings. The average override occurred 43% of the way into the event, regardless of duration, and the mean indoor temperature at the moment of override was 79°F. When compared to the preferred temperature of the home, the average override was initiated when the indoor temperature was 2°F higher. These results indicate that program participants may only be able to tolerate a few
degrees of discomfort before making the decision to cool their home. Similarly, participants respond better and are less impacted by behavior during off-peak hours relative to peak hours. This may be because people are less likely to be home at those times. Finally, the events themselves generally follow similar patterns with near perfect participation for the first few five-minute intervals, followed by a steady increase in the percent of load reduction lost to override behavior. In shorter events, this loss tends to even out near the end of the period, but in longer four-hour events the impact is predicted to have caused an increase in AC runtime in final intervals. This may be because the HVAC system needs to overcompensate to cool homes that overrode the DLC cooling settings later in the event, and thus were much warmer than they would have been otherwise.

The results of this research produces many recommendations and insights for DLC program designers. First, to minimize overrides an ensure a more stable group to solicit AC runtime reduction from, demand response events should be kept to shorter durations. Shorter events offer more stability and a higher certainty in load reduction, both in the lower rate of overrides and the percentage decrease in load from the event. Overrides during longer, four-hour events in particular were shown to run the risk of leading to an increase in demand near the end of the event. Second, DLC program operators should favor beginning events at the moments in which they require the largest decrease in load and should expect that the bulk of overrides will occur about 40-50% of the way through the event, regardless of its duration. This is especially important if the event is called between the 5 P.M. and 9 P.M., for participants are more likely to exercise the override option during these peak hours. Finally, smart thermostats such as ecobee should continue to be used as a tool to better manage the comfort settings of homes. An understanding of how much of an increase in indoor temperature participants are willing to endure can be leveraged to decrease overrides and ensure a higher certainty in load reductions.

This study has several limitations that can be expanded on by future research. First, the accuracy, sensitivity, and precision of the model point to the increasing complexity of the algorithms used to operate smart thermostats. Although the high accuracy in predicting DLC events are likely the result of the clear differences between the indoor temperatures and cooling setpoints, there is more ambiguity at other times. “Smart” settings are built to anticipate the needs of thermostat owners, predict their lifestyle patterns, and create more options for consumers to control their homes. This information is not necessarily intuitive or available in the DYD dataset. Similarly, there are extraneous variables occurring in the home that researchers will never be able to build into generalized models. More complex models may be able to better predict how the thermostat would have operated without the DLC settings. Second, future studies should continue to investigate at what point the usage of behavioral elements such as override features lead to diminishing returns. We did not calculate opt-out rates, but the results of Sarran et al. (2021)[20] point to concerns surrounding program design flexibility. Is it better to have a small participant group that responds predictably to events, or a larger group with both more potential for load shaving and a higher likelihood of overriding? Third, it should be noted that the thermostat sample considered in this analysis is a rather narrow group that owns a smart thermostat, is willing to participate in a DLC program, and voluntarily donated their personal data to researchers. While this is not representative of the general population, it does provide insights to human behavior and a potential type of group that can successfully shave load through this program design. Future research should consider other sample groups and types of DR programs.
Utilities may find greater success in creating a more diverse array of programs built for different types of customers.
5. Bibliography


