Smart Tech, Dumb Humans: The Perils of Scaling Household Technologies

Alec Brandon
Johns Hopkins University

Christopher M. Clapp*
University of Chicago

John A. List
University of Chicago & NBER

Robert D. Metcalfe
University of Southern California & NBER

Michael Price
University of Alabama & NBER

November 10, 2021†

Abstract

Smart-home technology has been heralded as an important way to increase energy conservation, but causal evidence remains scarce. We estimate the causal impact of smart thermostats on energy use using data from two novel field experiments in which a random subset of treated households were given a smart thermostat that was installed in their home free of charge. We combine this experimental data with 18 months of high-frequency data on household energy consumption in the form of more than 16 million hourly electricity and daily natural gas observations. In contrast to advertised savings based on engineering models, we find little evidence that smart thermostats have a statistically or economically significant effect on energy use. This result is robust to the inclusion of numerous controls and when the model is estimated on various subsamples of relevance for grid managers and policymakers (e.g., by hour of the day). We explore potential mechanisms using almost four million observations of system events including human interactions with their smart thermostat. Results indicate that user behavior dampens energy savings and explains the discrepancy between estimates from engineering models and those observed in our field experiments.

JEL Codes: D01, Q41, Q48, Q55.

Keywords: Energy Demand, Smart Grid, Smart Technology, Field Experiment.

*Corresponding author: cclapp@uchicago.edu.

†Acknowledgements: Dylan Brewer, Fiona Burlig, Ken Gillingham, Michael Greenstone, Koichiro Ito, David Jimenez-Gomez, Juanna Joensen, David Novgorodsky, Florian Rundhammer, Casey Wichman, and seminar participants at the Advances with Field Experiments Conference, Heartland Environmental and Resource Economics Workshop, Federal Communications Commission, EPIC Faculty Workshop, and NBER Environmental and Energy Economics Summer Institute provided helpful comments and suggestions. Ariel Listo, David Liu, Uditi Karna, and Jack Ogle provided excellent research assistance. We would like to thank Richard Caperton, Augustin Fonts, Ilan Frank, Kevin Hamilton, Nicholas Payton, Jim Kapsis, and many others at Opower for sharing data and offering insights. Opower provided the data analyzed in this paper to the authors under a nondisclosure agreement. The authors and Opower structured the agreement in a way that maintains the authors' independence. In particular, the agreement stipulates that Opower has the right to review the publication prior to public release solely for factual accuracy.

1 Introduction

Households in America spend an average of over $2,200 on energy annually, and residential energy accounts for roughly 20% of the annual carbon dioxide pollution from energy production (Energy Information Administration (EIA), 2018, 2019b).[1] These high private and social costs have led to substantial interest in smart technologies that reduce energy use without reducing consumer utility by increasing efficiency. Given that the largest share of residential energy (almost 40%) goes to heating and cooling the home (EIA, 2019a), smart thermostats are an increasingly popular example of such a technology.

Smart thermostats allow individuals to program temperature setpoint schedules and adjust settings remotely via a smart phone application. While producers of these devices promise consumers substantial savings on their home heating and cooling bills, projected savings are often based on engineering simulations that fail to account for how people actually use their smart thermostats and represent an upper bound on potential savings (ecobee, 2013; Honeywell, 2013). Or they are based on studies that use non-experimental data and have methodological flaws that result in upwardly biased estimates of savings (Nest, 2019). Thus, the true marginal impact of smart thermostats on real world energy usage is uncertain.

In order to determine the causal impact that smart thermostats have on home energy usage, we examine data from two field experiments conducted by Opower and Honeywell in conjunction with Pacific Gas and Electric (PG&E) – the second largest utility in California. For these experiments, the 1,385 households that volunteered to participate in the study were randomized into either a treatment group who received free installation of a Honeywell two-way programmable smart thermostat or a holdout control group that did not receive a smart thermostat.[2] We evaluate the effect of the smart thermostat on energy consumption using high-frequency records of household energy consumption over an 18-month period that includes more than 16 million hourly electricity use records and almost 700 thousand daily observations of natural gas consumption. We estimate this effect using a difference-in-differences instrumental variables (DDIV) specification where we use random assignment to treatment as an instrument for the endogenous decision to install a smart thermostat.

Across numerous specifications, we find that smart thermostats have neither a statistically nor economically significant effect on energy use. This result is robust to the inclusion of controls for weather conditions and a battery of household, location, and time effects. Our midpoint estimate

[1] For perspective on total costs, American households consume 20.8 quadrillion British thermal units (BTU) of energy per year, and energy production results in the emission of over five billion metric tons of carbon dioxide into the atmosphere each year (EIA, 2019a, 2019b).

[2] In addition to the ability to schedule permanent temperature setpoints and interact with the thermostat remotely, the smart thermostat given to households in our experiment provided households with a social norm framing of their setpoint choices. Framing of setpoints is an increasingly common feature of more modern smart thermostats, and there is an extensive literature documenting the responsiveness of household energy consumption to social norm framing (e.g., Allcott, 2011; Ferraro and Price, 2013; Ayres et al., 2012; Costa and Kahn, 2013; Allcott and Rogers, 2014; Dolan and Metcalfe, 2015). Given this finding and the Peffer et al. (2013) result that individuals do not use the programmable features of their thermostats as intended, this feature should provide the best chance for the smart thermostats used in our experiment to cause a reduction in energy consumption.

from the various specifications is -0.001 (se 0.022) for electricity and 0.023 (se 0.026) for natural gas. We can rule out conservation effects larger than 4.3% for electricity and 2.9% for natural gas with 95% confidence. Hence, even our most optimistic estimates suggest savings that fall short of those predicted by engineering models.

To investigate whether this aggregate result masks significant, but offsetting, heterogeneous effects, we estimate the model across different subsamples such as day of the week, hour of the day, and ambient temperature/humidity quintiles. We find almost no evidence of heterogeneous treatment effects.

In order to explore potential mechanisms that would explain this null result, we use almost four million observations of treatment group heating, ventilation, and air conditioning (HVAC) system activity and user interactions with their smart thermostat in the form of scheduled temperature setpoints, temporary overrides, and HVAC system events. First, we provide descriptive evidence that users take advantage of the smart features of their devices. We do so by showing that they frequently schedule permanent setpoints, the pattern of those setpoints across hours of the day is intuitive, and the temperatures that they set are in line with Environmental Protection Administration (EPA) energy-efficiency guidelines. We then establish that users frequently override scheduled temperature setpoints, and when they do, override settings are less energy efficient than their previously scheduled counterparts.

Next, to more formally test the hypothesis that user behavior explains the discrepancy between the decrease in energy use purported by the engineering studies and our experimental estimates, we categorize smart thermostat households into flexible user type categories based on how intensively they use the energy-saving features of their thermostat.[3] We match these categories to our experimental, energy use data, interact user type with an indicator for treatment, and estimate difference-in-differences intention-to-treat (DDITT) models. Estimates indicate that some high-efficiency type users do realize significant savings by installing a smart thermostat, but that human behavior explains the discrepancy between engineering estimates and the null results in the experimental samples. Our findings suggest that engineering models fail to adequately incorporate how people actually use smart technologies and this limits the usefulness of their estimates in real-world settings. In doing so, we demonstrate that human behavior is a peril to scaling such technologies based engineering models.[4]

Moreover, our sample is comprised of those who expressed interest in a smart thermostat. While the margins that individuals in our experiments are operating on are very natural (List, 2020), and our samples are externally valid with respect to their base energy use, our null result is likely to be even stronger if smart thermostats were mandated to homes where people do not desire

[3] Specifically, we define types based on each household's relative position in the distributions of permanent setpoints and temporary overrides (e.g., high (low) types are those above (below) the median number of programmed setpoints). For both measures, we vary the cutoff between high and low types.

[4] Our results support the existing literature of large differences between engineering estimates and the the effects from economic studies with a focus on causality in the field (e.g., Fowlie et al. (2018); Davis et al. (2020)). A difference between their work and ours is that we observe people's actual behavior in the home in terms of how they use their smart thermostat.

such a conservation technology.[5]

Overall, while null results have usually pose a challenge for economists in terms of their informativeness (Abadie, 2020), we show that people's interaction with the smart technology is the reason why we observe the null effect (and the departure from engineering model predictions). While our findings are "statistical" nulls, they are not "policy" nulls because they are counter to widely-held prior beliefs based on the engineering literature, and this literature is driving public policy and expenditure.

We make several contributions to the literature. First, while there has been considerable research on smart grid investments (Joskow, 2012), much less work has been done exploring the impact of smart technologies on residential energy use. Initial assessments of these technologies have focused on changes in average energy use induced by in-home displays of real-time energy price or quantity information (see, e.g., Faruqui and Sergici, 2010; Jessoe and Rapson, 2014; Alberini et al., 2013).[6]

Second, while ours is not the first study to evaluate smart thermostats themselves, the existing literature is comprised of studies based on either engineering or econometric methodologies, and both approaches have known issues.[7] Regarding the former methodology, in their most naive form, engineering studies compare the energy use of an HVAC system simulated under two different scenarios: a smart thermostat optimally programmed for energy savings and a traditional thermostat set to maintain a fixed temperature (Urban et al., 2012; Urban and Gomez, 2013; Daken et al., 2016). Both the experimental and baseline scenarios are unrealistic because of the absence of human behavior.[8] As such, these studies estimate the upper bound on true energy savings. Thus, it is not surprising that device producers often justify their energy-saving claims based on the results of engineering studies. In contrast, our study is based on a field experiment that captures how individuals actually use both smart and traditional thermostats and allows us to estimate real-world savings as opposed to a hypothetical upper-bound.

The existing econometric literature primarily consists of white papers that thermostat producers use to claim 10-23% energy savings based on a combination of observational and experimental data (Apex Analytics, LLC, 2014, 2016; Aarish et al., 2015; Ho, 2014; Kelsven et al., 2016; Nest Labs, 2014, 2015; Schellenberg et al., 2017; Stewart and Jackson, 2015; Robinson et al., 2016; Ward et al., 2014). None have been subjected to peer review and, to varying degrees, all are unclear about salient features of the study, have methodological flaws (primarily related to selection), and/or draw incorrect conclusions from their estimates. The selection issues are likely to lead to upwardly biased estimates of savings, so again, it is not surprising that device manufacturers are eager to advertise the results of these studies.[9] These biased studies have likely influenced federal policy

[5]For the year 2009, Californians eligible for our thermostats experiment used 1.2 kWh per hour (EIA, 2010). This number is extremely similar to our sample, which ranges between 1.0 and 1.3.

[6]Harding and Lamarche (2016) is a notable exception. The authors consider the effect of technologies that automate temperature setpoint changes to dynamic pricing.

[7]See Allcott and Greenstone (2012) for a general discussion.

[8]This methodology is akin to implicitly assuming ideal behavior in the treatment group and no optimizing behavior in the control group.

[9]The Nest (2019) website advertises a 10 to 12% savings on heating and a 15% savings on cooling costs based

on smart thermostats.

Exceptions in terms of both clarity and quality are the white papers by Broaddus et al. (2016, 2018) and Park et al. (2017).[10] However, in both cases, the observed outcome is based on aggregate energy consumption data: Broaddus et al. (2016, 2018) observe monthly energy billing data and Park et al. (2017) observe weekly smart meter data. Following Agnew and Goldberg (2013), both studies include coarse measures to control for ambient weather conditions: counts of heating and/or cooling degree days. In contrast to the existing econometric literature, we use high-frequency energy use data to estimate a DDIV model. Ghanem and Smith (2021) formalize the benefits of using high-frequency hourly data over a more aggregate analog. They show that fixed effects estimators based on aggregate data are inconsistent when there is high-frequency temporal heterogeneity in the effects and/or confounders. Accounting for this variation is particularly important in our context because smart thermostats are designed to allow individuals to vary energy use in response to within-day changes in temperatures. Thus, in addition to our model specification addressing the selection issues that bias much of the existing literature towards findings of significant savings, our high-frequency data allows us to better control for differences in ambient weather conditions and more accurately estimate our coefficient of interest than existing studies.[11]

Third, our results have important policy implications as there are both government and industry funded subsidies of smart technologies. Between 2009 and 2014, the Department of Energy (DOE) invested $7.9 billion in smart technologies under the Smart Grid Investment Grant (SGIG) program (DOE, 2016).[12] Additionally, the joint EPA and DOE ENERGY STAR program certifies the energy efficacy of smart thermostats (and other household appliances). Based on these certifications, 170 utility companies subsidized the purchase of a smart thermostat in 2019 (EPA 2019). In 20 states, over half of all households are eligible for a smart thermostat rebate, and in the most generous case, all of the residents in Nevada are eligible to receive a smart thermostat for free (Bloomberg Finance L.P., 2019). Given the current information available, energy producers and policymakers alike are subsidizing these devices based on misleading information with public

on an internal study. Nest Labs (2015) reports estimates from a difference-in-differences (DD) regression model that compares the monthly energy use of a self-selected group of households that were early adopters of the Nest smart thermostat and enrolled in an energy-monitoring program to those who only enrolled in the monitoring program. The study's authors acknowledge potential sources of bias in their estimates, but fail to provide evidence that the behavior of their comparison group is a reasonable counterfactual for those who decide to install a Nest.

[10]These studies acknowledge self-selection in the treatment group and estimate ITT models on all those encouraged to install a smart thermostat in their experiment. The latter uses four different methodologies to estimate the effect of a smart thermostat on energy use, including a small-scale field experiment that uses a matched-pair randomization design to address selection after randomization. While significant, we note that estimated savings effects in these studies are generally smaller than in the previously cited studies, on the order of 1% to 6%.

[11]Novan et al. (Forthcoming) use similar high-frequency smart meter data to reexamine the effect of building codes on energy use. In contrast to the existing literature (Levinson, 2016; Kotchen, 2017), which analyzes lower-frequency data, the authors find that residential energy efficiency standards reduce electricity consumption.

[12]While more than two-thirds of these investments went towards outfitting households with smart meters and communication systems that allow utilities to integrate real-time market conditions into household consumption decisions via dynamic pricing plans or demand response messaging, a complementary set of investments targeted the development and dissemination of technologies such as smart thermostats that allow individuals to remotely communicate with their appliance and HVAC system.

funds that would be better spent on more effective policy interventions. Our findings highlight the benefits to policymakers of complementing engineering based approaches to evaluating energy efficiency programs with carefully designed field experiments to fully understand the impact of the programs they intend to subsidize.

An additional recommendation for policymakers is that they should attempt to match smart tech with smart economics. The households in our experiment were part of a standard tiered energy pricing tariff that millions of Americans currently face. Such tiered pricing has been shown to be sub-optimal from a welfare point of view (Borenstein, 2012), and so getting energy prices right will lead to large gains, especially so if consumers can easily optimize their energy use in the home. Evidence from Blonz et al. (2021) and Jessoe and Rapson (2014) suggests that there are benefits from combining time-varying prices (time-of-use or dynamic pricing) and smart tech. Our bottom line is that smart technology will not deliver savings alone, it needs smart pricing as a complement.

The remainder of this study is organized as follows: In Section 2 we describe the details of the field experiment, the sample of households in the study, and our data. The following section formalizes our empirical specification. Section 4 presents our model estimates, and Section 5 explores the mechanisms that drive our findings. The final section concludes.

2 Experimental Design

2.1 Smart Thermostat

The intervention in our field experiments occurs when a given household's existing thermostat is replaced by a smart device.[13] Smart thermostats are designed to increase consumer utility by improving the efficiency of the home's HVAC system and reducing adjustment costs. To these ends, the device in our experiment has two primary features common to most smart thermostats. First, the thermostat allows the user to program an extensive schedule of permanent temperature setpoints for each day of the week. Second, the user can either interact with the device directly or remotely via a web portal or smartphone app. Both lower the cost of adjusting temperature settings.[14]

While the effect of these features on energy usage is theoretically ambiguous depending the schedule the user sets and how she interacts with the device, there are several additional features of the thermostat used in our experiment that are designed to reduce energy consumption compared to a traditional thermostat. First, our smart thermostat is able to learn about how HVAC system operations affect indoor temperatures, then optimize the transition between programmed temperature setpoints. Second, when choosing setpoints, users receive messages that compare their settings to

[13] Specifically, surrogates of Opower/Honeywell installed a Honeywell Z-Wave Touchscreen Thermostat that communicates with a website portal and smartphone app designed and hosted by Opower. We do not observe anything about the pre-existing thermostat.

[14] Appendix Section A provides a more detailed description of the device. Panel (a) of Appendix Figure 12 displays the thermostat and associated applications. Panel (b) shows a screen-shot of scheduling using the smartphone app.

those of similar households. Analogous to the social comparison module studied in Allcott (2011), the thermostat interface presents: (i) descriptive norms with information on peer setpoint choices and (ii) injunctive norms with efficiency ratings of setpoints. Third, the thermostat app interface is designed to facilitate toggling to a less energy intensive setting when the user leaves home and toggling it back to the previous setting when the user returns. Finally, when a user overrides a permanent setpoint to make a temporary change that is more energy efficient than the scheduled one, she is prompted by a query asking if she wants to make this more energy efficient setting permanent.[15]

Newer smart thermostats may have additional energy saving features.[16] While we cannot say whether our experiment tests the efficacy of all smart thermostats or the combination of features in the experimental thermostat, we note that the thermostat in our experiment has all the core features of a current smart thermostat. Additionally, the analysis in Section 5 indicates that individuals make use of these features and do so largely as intended. Taken together, this suggests that our results are unlikely to be specific to the particular device installed as part of the experiment.[17]

2.2 The Field Experiments

Subjects were recruited in public places (e.g., malls, markets, and festivals) in two waves (or experiments). Recruitment for the first experiment took place across four counties in Northern California from July through October of 2012. Subjects in the second experiment were recruited from December of 2012 to February of 2013 in three Central California counties.[18] Appendix Figure 15 depicts the locations of homes in the experiments and provides visual evidence that treatment and control groups are spatially balanced across locations.[19]

Figure 1 illustrates the execution of the field experiments. It describes the assignment of households to treatment and control groups, as well as the subsequent installation decisions of treatment households. A total of 1,379 eligible households agreed to participate in our experiments: 815 as part of the Northern California experiment and 564 in the Central California experiment.[20] They

[15] Appendix Figure 13 highlights features of the smart thermostat. Panel (a) illustrates the social norm framing displayed when households choose setpoints. Panel (b) shows how households can remotely toggle the thermostat in response to leaving and returning home via a smartphone or personal computer.

[16] Daken et al. (2016) explain that some smart thermostats use the location of user cell phones to automatically adjust settings when users are away from home and/or optimize HVAC system settings in response to local weather conditions.

[17] To be eligible, an individual had to own her residence and have central air conditioning, a smart phone, and high-speed Internet. See Appendix Section I.1 for a summary of the eligibility requirements. For more information on canvassing, see Appendix Section I.2 for the original recruitment and enrollment guide.

[18] Subjects for the experiment in Northern California were recruited from the greater San Francisco/Sacramento area (Contra Costa, San Joaquin, Solano, and Yolo counties). Households in the Central California experiment are located in and around Fresno and Bakersfield (Fresno, Kern, and Madera counties).

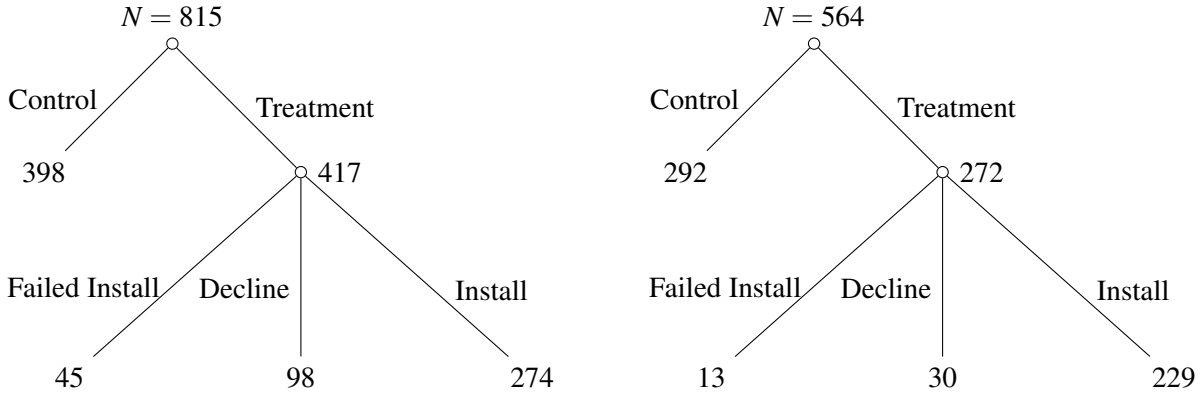
[19] We formally test balance in Section 2.6 and fail to reject the null of spatial balance in the counties where households are located.

[20] All household counts in this section are based on the households for which we observe electricity consumption. Aggregating across both experiments, there are a total of 1,379 unique households in the electricity sample, a total of 1,369 unique households in the natural gas samples, and a total of 1,385 unique households across both energy-

were randomized into either a treatment or control group. After group assignment, the experimenter had no further contact with the total of 690 control households across both experiments. The 689 total households assigned to the treatment groups were offered the smart thermostat described in the previous section and installation at no cost.

Professional installation of the smart technology is an important feature of our experiments over the encouragement or self-installation designs common to other experiments. Peffer et al. (2011, 2013) provide evidence that programmable thermostats are often installed incorrectly and list flawed installation as a reason they are not more effective. Additionally, Apex Analytics, LLC (2016) find that although cheaper, their self-installation design "led to substantial attrition among interested and qualified customers." In contrast to the 35% take-up rate in their experiment, on average across our experiments, the smart thermostat was successfully installed in 73% of treatment group homes. Of the remaining treatment homes, 19% percent declined, and 8% had complications that prevented installation (e.g., compatibility issues).[21]

Figure 1: Sample Randomization



(a) Northern CA Experiment

(b) Central CA Experiment

2.3 Energy Data

All households in the study were equipped with smart meters that enabled PG&E to record household-level data on hourly electricity use and daily natural gas consumption. The quantity of electricity consumed is measured in kilowatt hours (kWh), and the unit of measurement for natural gas is a

type samples. Stated another way, we observe 16 households with electricity consumption data, but not natural gas information and another six households that consume natural gas, but for which we have no electricity consumption information.

[21] Appendix Figure 14 plots the cumulative density function (CDF) of the difference in time between assignment and installation dates that illustrates how long it takes households in the treated groups to install the smart thermostat (conditional on eventual installation of the smart thermostat). Most households had the smart thermostat installed shortly after being assigned to the treatment group: 50% of households had their thermostat installed within 5 days, and 95% had it installed within 30 days.

therm (thm).[22] As we cannot observe temperature setpoints directly for control households with a traditional thermostat, and energy is the policy-relevant good, these measures are the main outcome variables in our analyses. In total, we observe an average of 11,908 hourly electricity use decisions for the 1,379 households in electricity sample and 495 natural gas use decisions for the 1,369 households in the natural gas sample over an 18 month period from July 2012 through December 2013.

2.4 Timing

Figure 2 presents two visual depictions of important timing issues associated with the experiments and data. Panel (a) of Figure 2 plots the flow of households into treatment and control groups over time. The horizontal axis spans the period of time over which we observe energy data. The grey shaded areas illustrate the periods of subject recruitment in each of the two experiments. The subfigure shows that treatment and control households are temporally balanced, as they were assigned at similar rates over time, and that there is very little attrition over the year and a half study period.[23]

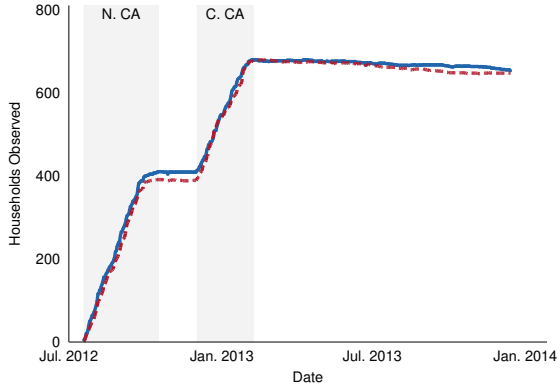
Unfortunately, we only observe energy readings starting on the first day of recruitment in Northern California experiment. Panel (b) illustrates the effect of this issue by plotting the number of electricity readings per day for each experiment relative to event time (where assignment to the treatment or control group occurs at time zero).[24] The figure shows that we do not observe a substantial pre-period for all households in the Northern California experiment, but we do for the Central California experiment. We report estimates in Section 4 both separately by experiment and based on a sample that combines data from both experiments to account for this issue. Estimates are not qualitatively different across specifications.

[22] A therm is a unit of heat energy equivalent to 100,000 BTUs.

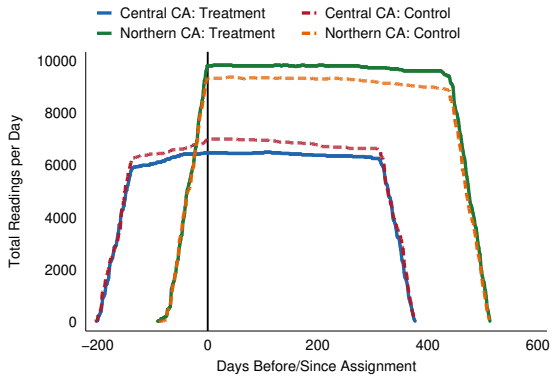
[23] We formally test balance in Section 2.6 and fail to reject the null of temporal balance in the month of assignment to experimental group.

[24] Plotting an analogous graph for natural gas readings changes the scale of the vertical axis but produces the same overall pattern.

Figure 2: Timing



(a) Number of Households Observed by Experimental Status and Date



(b) Total Electricity Readings by Experiment, Experimental Status, and Event Time

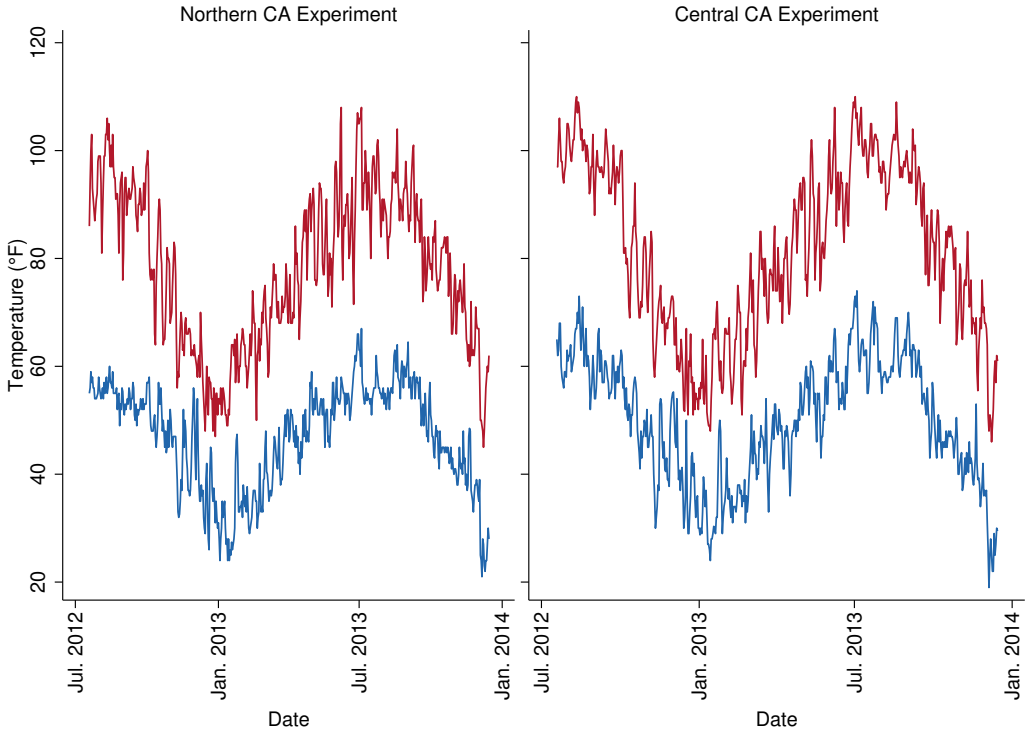
2.5 Additional Data

2.5.1 External Data

We supplement the main experimental dataset with information from several external sources and additional data collected as part of the experiment. First, we compile hourly temperature, humidity, and heat index readings for each county in the study from the National Oceanic and Atmospheric Administration (NOAA).[25] Appendix Table 3 summarizes the weather data. Temperatures in the combined sample (Panel C) average 63.7 degrees Fahrenheit (F), but range from below freezing to well over 100 degrees F. The various *Daily Measure* statistics for each of the three weather measures indicate that there is both spatial (between-county) and seasonal (within-county) variation in the data. The *Minimum* and *Maximum* statistics indicate that there is also daily variation in all three weather variables. Figure 3 visualizes this variation by plotting experiment-specific time series of the minimum and maximum daily outdoor temperatures over the sample period. The table and figures indicate that despite our sample being drawn from a temperate part of the country, there is substantial variation in the weather data. Summers are hot, humid and likely to require the use of air conditioning to ensure comfortable indoor temperatures. While the rest of the year is more moderate, there are many days cold enough to necessitate home heating.

[25] We are missing values for 0.09% of the temperature and 0.5% of the humidity observations in the sample. We interpolate these missing values using the predicted values from separate regressions of the given weather variable on location, day, and hour fixed effects. We calculate the heat index from the temperature and humidity readings (see: https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml for the formula).

Figure 3: Minimum and Maximum Daily Outdoor Temperatures $(^{\circ}F)$ by Date



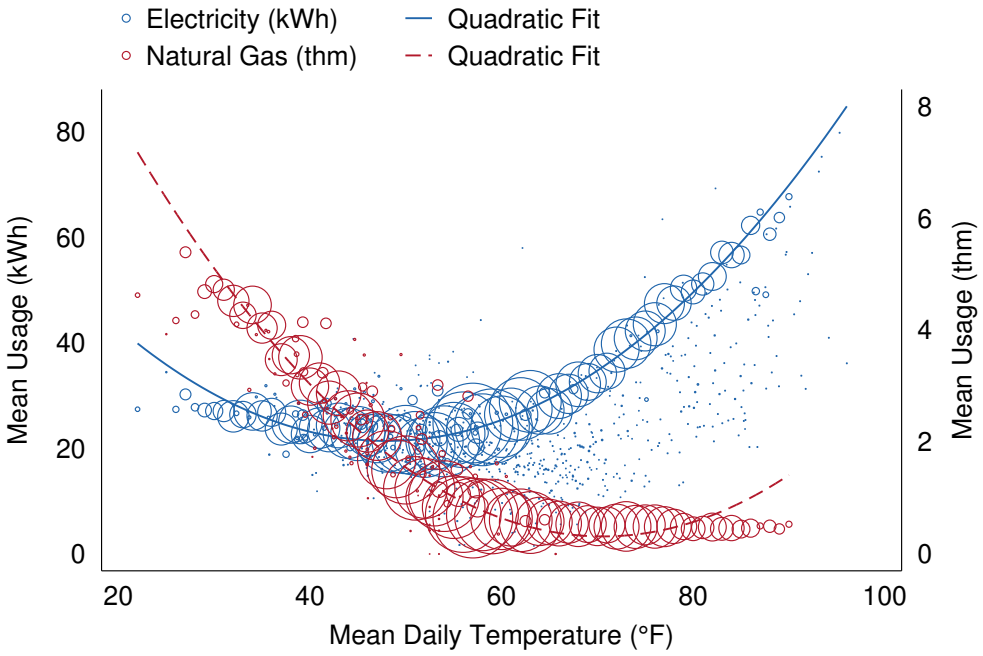
To confirm that this is the case and that we are able to identify the effects of HVAC system use in our smart meter data, Figure 4 plots the relationship between mean daily energy consumption and mean daily temperature for homes in the control group.[26] The blue markers represent electricity use (the energy source used for cooling; denoted on the left-hand vertical axis), and the red markers represent natural gas consumption (the predominate energy source for heating; denoted on the right vertical axis).[27] As one would expect, electricity use increases, and natural gas use decreases, with the temperature. Both relationships are non-linear, and the fitted-value lines indicate that quadratic models predict the data well.

These descriptive analyses indicate that there is sufficient variation in weather conditions in our sample and energy use responds to that variation, so our experimental setting meets the necessary conditions for assessing the efficacy of smart thermostats. They also inform our model specification. We estimate separate models of the effects of smart thermostats on electricity and natural gas use. For robustness, we include outdoor temperature and humidity measures, as well as location and time effects, as controls to mitigate the effects of residual variation on our estimates.

[26] Analogous scatter plots based on treated households produce the same patterns. Additionally, Ge and Ho (2019) analyze high frequency, smart thermostat event log data (similar to the data we analyze in Section 5) and find that the home heating and cooling decisions of smart thermostat users are affected by weather conditions.

[27] The area of both markers are weighted by the number of observations in the given cell.

Figure 4: Average Daily Energy Use by Outdoor Temperature ($°F$)



Note: Marker areas and quadratic fit lines weighted by number of observations.

Second, we supplement the household electricity use measure with data from two sources that allow us to test whether smart thermostats have a differential effect on usage when there is critical demand load. To do so, we collect data on the average hourly real-time price of electricity from the California Independent System Operator (CAISO).[28] Electricity is produced from many sources with different production and external costs.[29] Appendix Figure 16 is a box and whisker plot of hourly spot prices by quintile that illustrates the variation in these costs in our data. Spot prices are relatively consistent over the first four quintiles, but increase substantially from the fourth to the fifth quintile. This is consistent with what we would expect during peak-load times, but the long whisker in the fifth quintile suggests that peak-demand times may comprise only a small fraction of the observations in our dataset.

To further identify times when the system is most strained, we also collect data on system-wide peak-alert messages from CAISO and utility-wide alerts from PG&E. The latter alerts (referred to as "SmartDays" by the utility) are issued at a finer spatial scale, but a more granular temporal level (daily) than the latter. In contrast, the CAISO alerts are issued hourly, but apply to a broader area. Since system-wide alerts may occur on days when it is less obvious that there is a need to reduce demand to avoid brownouts based on local conditions, we also identify CAISO alerts that were

[28]The real-time market for electricity in California clears every five minutes. We use this data to calculate the average spot price each hour.

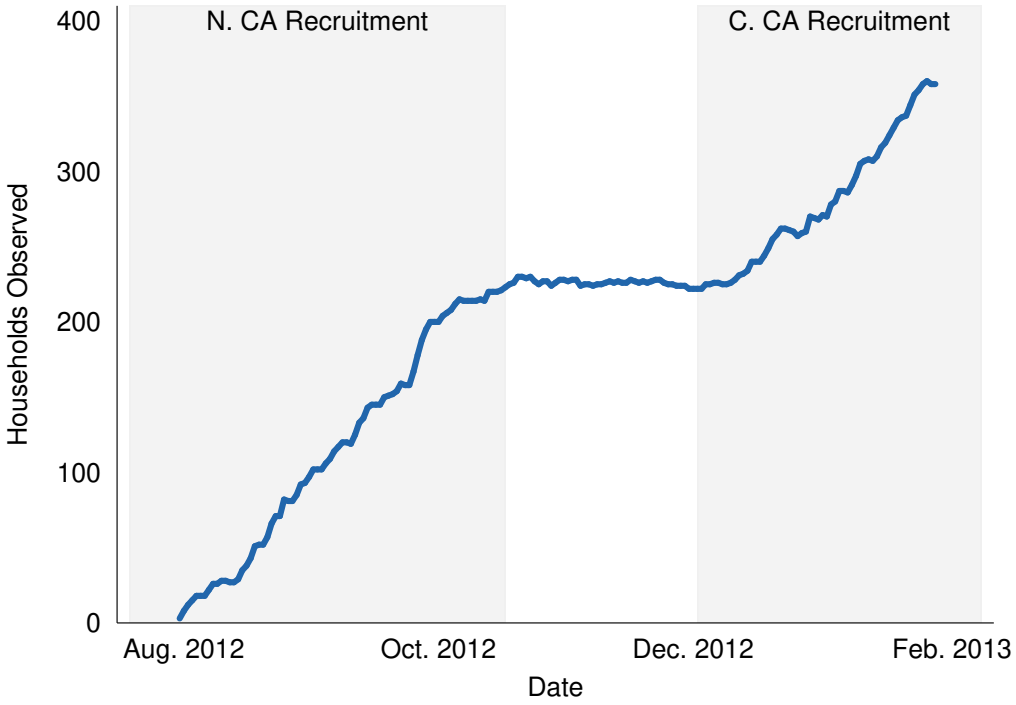
[29]California instituted a cap-and-trade carbon emissions program in 2012 (Shobe et al., 2014), so the price of electricity on the state's wholesale market reflects both the marginal cost of production and the prevailing market price for emissions as reflected in the price of carbon permits.

broadcast by local media outlets (e.g., the *Fresno Bee* or the *Bakersfield Californian*) to ensure that they reach a reasonable level of publicity to be salient to the households in our experiments. Conditioning on the additional information from these sources allows us to test whether smart thermostats reduce demand when the cost of electricity production to society is the greatest.

2.5.2 Internal Data

In addition to the external data we collect, we also observe a high-frequency, exact-time log of 3,967,558 HVAC system events, including user interactions with their smart thermostat, from 372 households. The unbalanced panel dataset spans from July 2012 to January 2013, and Figure 5 illustrates the number of households observed by calendar date. Recruitment and installation of smart thermostats first began in Northern California in July of 2012, whereas those in Central California began in December of 2012. Since this dataset is truncated in January of 2013, the majority of the observations in this dataset are generated by homes from Northern California, while only about 5% of the observations are from Central California homes.

Figure 5: Number of Households Observed in Events Data by Date



The system events and user interactions we observe include ambient temperature, HVAC state, and heating/cooling setpoints (which we classify into permanent setpoints and temporary overrides).[30] Permanent setpoints are thermostat temperature settings previously scheduled to occur

[30]Unfortunately, we do not observe who or how many people in the household have access to the app and/or interact with the thermostat.

automatically at specific times on a periodic basis. Temporary overrides are changes to the current setpoint which result from a concurrent interaction with the thermostat.[31] We aggregate these measures to hour-level observations. Appendix Table 4 summarizes the data. The table shows that while there are more observations from the Northern California experiment, settings in the two locations are remarkably similar.[32]

Finally, Opower and Honeywell conducted an online survey to collect baseline information on both treatment and control households in the experiments. We do not use these time-invariant household characteristics in our main analysis because they are redundant to household fixed effects, but we use them to test the validity of Opower and Honeywell's randomization process.

2.6 Balance

To test for balance, we estimate a linear probability model with an indicator for assignment to treatment as the dependent variable. Appendix Table 5 reports estimates from that model that summarize the results of our balance tests. Column (1) reports estimates based on a sample comprised of households from both experiments, and the estimates in Columns (2) and (3) are from models estimated on subsamples by experiment. The significance of each coefficient estimate represents the results of a single hypothesis test against a null of balance, and the reported F-statistics test the null hypothesis that all parameters in the given model are jointly equal to zero. We fail to reject the null for all single and multiple hypothesis tests across all three models. This indicates that control and treatment households are statistically balanced across observable, pre-experiment measures and is consistent with an appropriate randomization process.

We note that that households in the treatment group in the Northern California experiment used 5.5% less electricity per hour in the pre-period on average than those in the control. Accounting for means that are based on less than two weeks of data in Appendix Table 5 indicates that this difference is driven by the subset of households for which we observe only a limited number of pre-period electricity observations (see Section 2.4). Regardless, out of an abundance of caution, we estimate double-difference models to control for any potential pre-period imbalance.

2.7 Time-Trend and Event-Study Analyses

To illustrate basic temporal patterns in the data and the effect of experimental assignment on energy use Appendix Figure 20 plots the mean of residual energy consumption against event time (days before/after assignment to the treatment or control group) for each of the two experiments.[33]

[31] We do not directly observe whether system temperature changes are due to permanent setpoints or temporary overrides, but we are able to infer event types based on the precise timing of when the changes occur. See Appendix Section D for details.

[32] Average ambient temperatures are higher in Northern than Central California because of seasonal variation. The Northern California panel spans July through January, whereas the Central California panel runs from December through January.

[33] We project out household fixed effects prior to taking daily averages to adjust for pre-period differences in electricity use in a subset of Northern California homes for which we observe only a limited number of pre-period energy

Panel (a) displays electricity use, and Panel (b) illustrates the patterns in natural gas consumption. The figure shows that being assigned to receive free installation of a smart thermostat has no discernible impact on subsequent patterns of use. However, the raw data is too noisy to be visually conclusive.[34]

To provide further evidence of the validity of the experimental randomization and additional evidence of parallel pre-trends, Figure 2.7 plots the coefficient estimates and 95% confidence intervals from event studies of the effect of assignment to treatment.[35] Panels (a) and (b) plot electricity and natural gas estimates, respectively, based on data from the Northern California experiment. Panels (c) and (d) plot the Central California experiment analogs. Consistent with Appendix Figure 20, the event study plots show evidence of parallel pre-trends, but do not indicate large, persistent effects of being assigned to treatment on energy use.

We note that these figures do not account for incomplete take-up of the treatment, and they are based on temporally aggregated, day-level data. For these reasons, in the next section, we outline empirical models that allow us to instrument for smart thermostat installation and take advantage of the high-frequency nature of the electricity consumption data.

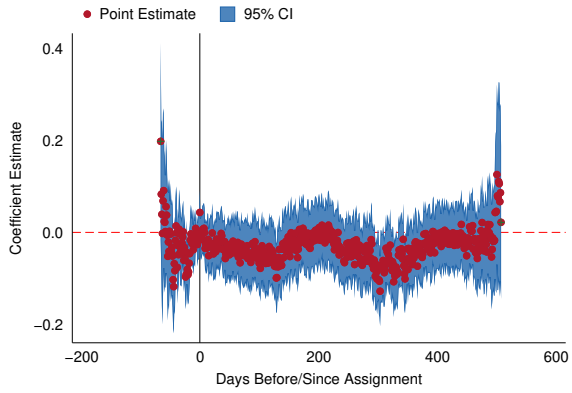
---

values.

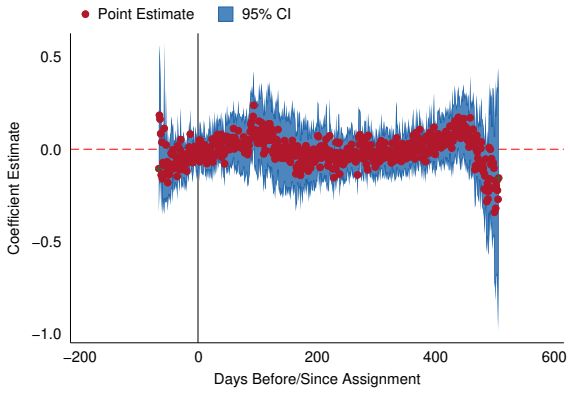
[34]For instance, the seasonal effects of summer for electricity use and winter for natural gas can be seen in the patterns in the data.

[35]The models being estimated include household fixed effects and daily time effects. Confidence intervals are based on standard errors clustered by household. To mitigate the visual effects of noisy coefficient estimates resulting from unbalanced lags and leads at the endpoints of the time window, we bin all lags and leads that are based on fewer than 30 observations (Schmidheiny and Siegloch, 2019; Clarke and Schythe, 2020).

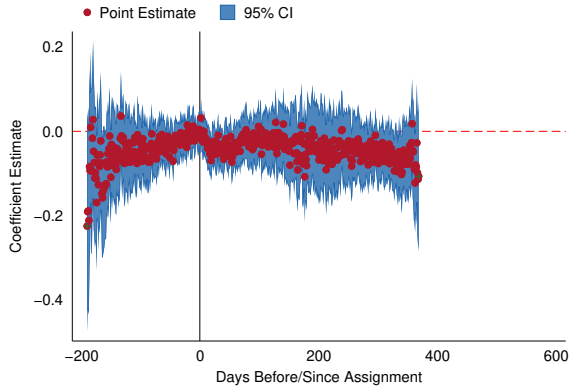
Figure 6: Event Study Estimates of Energy Use by Experiment



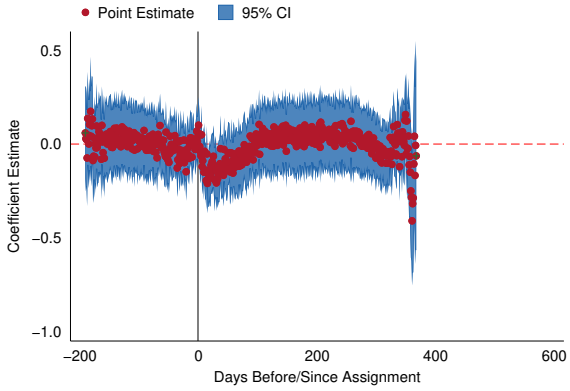
(a) N. CA Experiment, Electricity



(b) N. CA Experiment, Natural Gas



(c) C. CA Experiment, Electricity



(d) C. CA Experiment, Natural Gas

3 Empirical Model

Our field experiment randomizes receipt of a smart thermostat among eligible applicants. We observe a long time series of household-level energy use for treatment and control groups before and after experimental assignment. Both motivate our empirical strategy. Given the potential pre-period imbalance in electricity use discussed in Section 2.6, we estimate difference-in-differences (DD) models. To address noncompliance with experimental randomization, we augment our DD model with instrumental variables (IV) modeling techniques. We begin by formalizing our model specification, then discuss identification issues.

3.1 Model Specification

We model the effect of a smart thermostat on household i's consumption of energy type $j \in \{\text{kWh}, \text{thm}\}$ (electricity, natural gas) in time period t (e_{it}^{j}) using a DD model:

$$e_{it}^{j} = \alpha_i^j + \beta_t^j + \gamma^j S_i P_t + X_{it}\delta^j + u_{it}^j, \tag{1}$$

where S_i is an indicator equal to one if household i installs a smart thermostat, P_t is an indicator for post-assignment status in time period t, X_{it} is a vector of controls, α_i^j is a household fixed effect, β_t^j is a vector of time effects, and u_{it}^j is a household/time varying unobservable.[36] We cluster standard errors at the household level to account for serial correlation (Bertrand et al., 2004) and estimate the model separately for each energy type. When j denotes electricity, energy is measured in kWh and the time period is an hour. If j denotes natural gas, the energy unit is a therm and observations are recorded daily.

Our parameter of interest is γ^j, which measures the differential change in energy use across pre- and post-intervention periods for smart relative to traditional thermostat households. This specification implicitly assumes that smart thermostats have a constant effect for all households. Given that individuals in our treatment sample are each optimizing over their household's expected energy savings and installation costs when deciding whether or not to follow through with installation of the smart thermostat, our treatment is likely to result in heterogeneous effects and Roy (1951) selection on gains. Consistent with this underlying model of behavior, there is incomplete installation compliance among the treated households in our experiment (see Figure 1). To address concerns of bias from noncompliance, we estimate a DDIV model that uses the experimental randomization as an instrument for the installation of a smart thermostat. Formally, we estimate γ^j using two-stage least squares (2SLS) methods with $E\left[Z_{it}^j u_{it}^j\right] = 0$, where $Z_{it}^j = \left(\alpha_i^j, \beta_t^j, T_i P_t, X_{it}\right)'$, and T_i is an indicator for household i's treatment status in our experiment.[37]

3.2 Identification

If the assumption of parallel trends holds in our DD setting, our instrument is relevant and valid, monotonicity holds, and there is one-sided noncompliance in our experiment, our DDIV coefficient of interest, γ^j, identifies the ATT of a smart thermostat (Cornelissen et al., 2016). This is the average impact of a smart thermostat on the energy use of households that install one. We discuss our identifying assumptions in more detail in Appendix Section F.[38]

[36] We obtain similar results when estimating the model on the natural log of energy consumption ($\ln(e_{it}^j)$). If the randomization in our experiment is valid, our coefficient of interest is identified regardless of whether or not we include household fixed effects (α_i^j), time effects (β_t^j), or additional controls (X_{it}). Thus, we begin by estimating a basic specification of the model without any additional covariates that replaces α_i^j with $\alpha^j S_i$ and β_t^j with $\beta^j P_t$. Subsequent specifications add controls for the weather (which cannot be randomized a priori), household fixed effects, and various time effects to demonstrate robustness and to improve the efficiency of our estimates. Results are qualitatively similar across all specifications.

[37] Equation 1 is the second-stage equation, and the first stage is modeled as

$$S_i P_t = \theta_i^j + \kappa_t^j + \lambda^j T_i P_t + X_{it}\pi^j + w_{it}^j. \tag{2}$$

[38] First-stage results in Appendix Section G provide strong support for instrument relevance. Appendix Table 5 and Figure 2.7 provide evidence in favor of instrument validity and parallel trends. Monotonicity is a standard assump-

4 Results

We begin by reporting estimates of the parameters in Equation 1 for electricity and natural gas in the next section. We then re-estimate the model on restricted subsamples of the data to investigate whether our main results mask significant, but offsetting, heterogeneous treatment effects. In the subsequent section, we estimate the model separately by quintile of ambient weather conditions, day of the week, hour of the day, hour of the day by weekday/weekend, quintile of the price of electricity, and during peak-use alerts.

4.1 Main Estimates

Table 1 summarizes multiple estimates of the effect of a smart thermostat on energy use based on each of the two experiments and the combined sample of all households recruited during both experiments. Panel (A) reports estimated effects on hourly electricity usage, and Panel (B) reports analogous estimates based on daily consumption of natural gas. Each ATT estimate reported in Columns (1) through (6) is based on a separate DDIV regression corresponding to the experimental sample indicated in the given row and the controls indicated at the bottom of the table.[39] Column (1) reports estimates of a basic version of the DDIV model without any fixed effects, time effects or other additional controls.[40] Column (2) reports estimates from a model that adds an indicator for experimental wave to control for differences in recruitment conditions, as well as linear and quadratic county temperature and humidity readings to control for ambient weather conditions. Column (3) reports estimates from a model that adds household fixed effects to control for all of the time-invariant, unobserved characteristics of the home and household (e.g., age and square footage of the home, number of family members).[41] Column (4) reports estimates from a model that adds month-of-year (MOY) effects to control for aggregate, time-varying effects such as seasonal variation in weather patterns. Column (5) adds day-of-week effects to control for variation in daily usage patterns due to occupant work and schooling schedules. Finally, Column (6) replaces the time effects with day and hour-of-day effects.[42]

tion in IV settings that rules out irrational behavior. Finally, our experimental environment suggests that one-sided non-compliance is a reasonable assumption. In our context, this means that while some households randomized into treatment do not install a smart thermostat, no households in the control group install one. At the time of our experiment, smart thermostats were a nascent technology. According to the EIA's Residential Energy Consumption Survey (RECS), two to three years after our experiment, only 4.09% of all households in the survey and 10.58% of observationally similar households owned a smart thermostat. Regardless, note that if we relax the one-sided non-compliance assumption to one of monotonicity, our DDIV specification instead recovers the Local Average Treatment Effect (LATE) estimate of γ^j (Imbens and Angrist, 1994).

[39] See Appendix Tables 6, 7, and 8 for identical estimates with full regression diagnostics. The rk *LM* and Wald *F* statistics reported in those tables are first-stage diagnostic tests of under and weak identification, respectively, in models with non-i.i.d. errors. In all specifications, we reject the nulls of an under or weakly identified model. See Kleibergen and Paap (2006) for details.

[40] Relative to Equation 1, the model in Column (1) replaces α_i^j with $\alpha^j S_i$, β_t^j with $\beta^j P_t$, and restricts $\delta^j = 0$.

[41] Since the experimental wave indicator is perfectly collinear with recruitment wave, we drop the wave indicator from this and subsequent specifications.

[42] The estimates in Panel B are based on daily natural gas meter readings (thm), so the hour-of-day effects noted in Column (6) are included in the electricity model (Panel A) only. Estimates based on models that instead include week-

The coefficient estimate of -0.055 reported in Column (1) of the first row in Panel (A) indicates that a smart thermostat causes a 0.055 kWh decrease in electricity usage per hour. The cluster-robust estimate of the standard error of 0.058 reported in parentheses indicates that this estimate is statistically insignificant.[43] To put the magnitudes of these effects in context, Column (7) reports mean energy use in the control group in the corresponding sample. The estimated effect is equivalent to 4.986 percent of control group energy use of 1.103 kWhs per hour. The corresponding natural gas coefficient estimate in Panel (B) is equivalent to only 0.633 percent of control energy use. Both estimates are well short of the savings estimates from engineering studies touted by thermostat manufacturers. Across all specifications in both panels, the lack of economic or statistical significance indicates that smart thermostats do not reduce energy usage. In fact, both the electricity and natural gas estimates reported in Column (6) that are based on the combined sample from both experiments are positive, and the natural gas estimate is statistically significant.

4.2 Heterogeneity in Treatment Effects

In order to investigate the possibility of significant, heterogeneous effects that are not apparent in the aggregate, we estimate the model conditional on various sub-sample selection criteria. Given that the results in Table 1 do not indicate any substantial differences between experiments and given that all subsequent specifications include household fixed effects, all results presented in this section start from a sample that pools the observations from both experiments. This should also give our models the best chance of recovering a significant heterogeneous treatment effect. We treat the model reported in Column (5) of Table 1 as our preferred specification and use it as the basis for our subsequent analyses because it is applicable to both samples with hourly- and daily-level variation.

First, since smart thermostats will only have an effect on energy usage when there is a need for the HVAC system to heat or cool the house, moderate ambient temperature observations may attenuate a significant effect. To address this concern, Appendix Table 9 reports estimates by ambient temperature quintile. If the effect of a smart thermostat is only apparent when the HVAC system is in use, we would expect to find significant effects in the upper quintiles of temperature for electricity use and in the lower quintiles for natural gas. This is not the case. Only one of the 10 estimates is statistically significant, and the significant effect occurs in the second quintile of temperature for electricity consumption. Given the overall pattern of results, this finding is likely spurious.

---

of-year, month-by-year, week-by-year, and day-by-hour effects result in qualitatively similar results. Furthermore, estimates based on models that include weather controls, day-of-week effects, and household-by-MOY (or household-by-WOY) effects do not affect our findings. The specification identifies off of hourly (electricity) or daily (gas) variation in usage within a household at a particular time of year. Intuitively, identification comes from the change in consumption in a given month of a the year for a treated home before and after treatment, relative to that same change for a control home. We also estimate models that include ZIP Code-by-MOY and ZIP Code-by-WOY effects that similarly identify off of variation within a neighborhood at a particular time of year. Again, results are qualitatively similar.

[43] Standard errors are clustered at the household level.

Table 1: ATT Estimates of the Effect of a Smart Thermostat on Energy Use

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Power Use (kWh or thm)						Mean
Panel A: Electricity (kWh)							
N. CA Experiment: ATT ($\hat{\gamma}^{kWh}$)	-0.055	-0.061	-0.016	-0.016	-0.016	-0.003	1.103
	(0.058)	(0.058)	(0.046)	(0.046)	(0.046)	(0.041)	(1.196)
C. CA Experiment: ATT ($\hat{\gamma}^{kWh}$)	0.009	0.006	0.002	0.002	0.002	-0.001	1.191
	(0.029)	(0.028)	(0.025)	(0.025)	(0.025)	(0.023)	(1.273)
Both Experiments: ATT ($\hat{\gamma}^{kWh}$)	-0.031	-0.031	-0.003	-0.001	-0.001	0.026	1.140
	(0.036)	(0.035)	(0.022)	(0.022)	(0.022)	(0.017)	(1.230)
Panel B: Natural Gas (thm)							
N. CA Experiment: ATT ($\hat{\gamma}^{thm}$)	-0.009	0.009	0.085	0.075	0.075	0.069	1.422
	(0.061)	(0.063)	(0.068)	(0.066)	(0.066)	(0.055)	(1.761)
C. CA Experiment: ATT ($\hat{\gamma}^{thm}$)	-0.003	0.007	0.001	0.001	0.001	-0.021	1.129
	(0.044)	(0.031)	(0.027)	(0.026)	(0.026)	(0.026)	(1.332)
Both Experiments: ATT ($\hat{\gamma}^{thm}$)	0.062	0.065	0.028	0.023	0.023	0.055**	1.298
	(0.060)	(0.049)	(0.028)	(0.026)	(0.026)	(0.022)	(1.599)
Weather Controls		x	x	x	x	x	
HH Fixed Effects			x	x	x	x	
Month-of-Year Effects				x	x		
Day-of-Week Effects					x		
Day Effects						x	
Hour-of-Day Effects						x	

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. Columns (1) through (6) report ATT estimates of the effect of a smart thermostat on energy use ($\hat{\gamma}^{j}$) based on separate DDIV regressions corresponding to the experimental sample indicated in the given row and the controls indicated at the bottom of the table. To put the magnitudes of these effects in context, Column (7) reports mean energy use in the control group in the corresponding sample. The samples used to produce the estimates in Panel A are based on *hourly* electricity meter readings in kWh, while the samples underlying the estimates in Panels B are based on *daily* natural gas meter readings (thm). Thus, the hour-of-day effects noted in Column (6) are included in the electricity models (Panel A) only. Note that the estimates reported in Column (2) for samples that combine data from both experiments include an indicator equal to one for observations in the Northern California experiment. This indicator is perfectly co-linear with household fixed effects, so it is dropped from subsequent models. See Appendix Tables 6, 7, and 8 for full regression diagnostics. Based on the values of the rk *LM* and Wald *F* statistics reported in those tables, we reject the nulls of an under or weakly identified model across all specifications.

Similarly, Appendix Table 10 reports estimates by ambient humidity quintile. In contrast to the results by temperature quintile, the estimates in Columns (4) and (5) of Panel A indicate that smart thermostats have a significant, negative effect on electricity use when the humidity is high (but not necessarily the temperature). We would expect to find this pattern of results if smart thermostats are successful at reducing the level of humidity in treated homes without deviating from a pre-programmed schedule, but individuals in the control group are prone to over-adjusting their traditional thermostats to less energy-efficient setpoints in order to mitigate the discomfort caused by high humidity. Consistent with this explanation, we do not find similar, significant effects on the consumption of natural gas (in Panel B). Alternatively, as it takes more energy to cool humid air than dry air, the pattern temperature and humidity results is consistent with smart thermostats providing small energy-efficiency gains that are only evident when the HVAC system has to work hardest.[44]

Next, since smart thermostats may only have an effect on energy use during the weekdays when individuals have predictable schedules, Appendix Table 12 reports estimates by day of the week and by weekday/weekend. Across all days of the week and when we aggregate to the weekday/weekend level, we find no evidence that smart thermostats reduce energy consumption.

Similarly, smart thermostats may only have an effect during the times of day that individuals typically schedule permanent temperature changes (e.g., before leaving for work/school or after returning home). Appendix Table 13 reports estimates by hour of the day. We are only able to calculate estimates conditional on the hour of the day for the effects of a smart thermostat on electricity usage, as we observe natural gas use at the daily level. Again, there is scant evidence that smart thermostats have a significant effect on energy use. To further test whether smart thermostats have effects only during certain hours of the day on certain days of the week (e.g., weekdays), Appendix Table 14 reports estimates by hour of the day separately for weekdays (Panel A) and weekends (Panel B). Consistent with our previous findings, we do not find evidence of significant effects on electricity use by hour of the day and day of the week.

Finally, since a potential benefit of smart technologies is that they enable consumers to better respond to spikes in demand and network congestion that lead to increased wholesale prices and brownouts (Joskow, 2012), we estimate two sets of models that condition on times when the social benefits of reduced electricity consumption are the greatest. Appendix Table 15 reports results by quintiles of electricity spot prices. If smart thermostats save energy at the most beneficial times, we would expect to see a statistically significant, negative effect in the fifth quintile of prices when production and external costs are the greatest, but none of the reported effects are statistically significant. To further isolate periods of high demand, Appendix Table 16 conditions on times when the system operator or utility issued a peak-usage alert. Column (1) reports estimates based on a sample of hours when CAISO issued hourly alerts. Holladay et al. (2015) find that media coverage impacts consumer responses to price alerts, so the estimates in Column (2) further condition on a sample of hours when there was both a CAISO alert and local media coverage of that alert. Neither

[44] Appendix Table 11 reports estimates from analogous models that condition on quintiles of the heat index (the perceived temperature) to rule out effects by the combined effects of temperature and humidity on comfort. We do not find significant results.

of the resulting coefficient estimates is statistically significant.

In contrast, we find significant, negative effects when PG&E issues a daily, utility-wide alert. Column (3) is based on a sample of all hours on alert days. PG&E advises their consumers to conserve electricity between 2:00pm and 7:00pm on these days, so the results in Columns (4) and (5) disaggregate this effect into off-peak and peak hours, respectively. The coefficient estimate in Column (3) (Column (4), Column (5)) indicates that smart thermostats reduce electricity consumption by 0.103 kWh (-0.073 kWh, -0.116 kWh) relative to mean electricity use of 1.775 kWh (1.402 kWh, 2.893 kWh) in the control group. These estimates imply that smart thermostats result in electricity conservation of 5.8 percent (5.2 percent, 4.0 percent). Although statistically and practically significant, these estimates fall short of advertised impacts even at times when smart thermostats should have their biggest impact on energy use. Appendix Figure 21 further disaggregates these effects by hour of the day. While the estimates are rarely statistically different from zero, the hourly estimates suggest that smart thermostats reduce electricity use before the peak alert period begins, but that smart thermostat households undo some of this benefit by ramping up their energy use before the peak period ends. The overall pattern of results on alert days is consistent with our previous findings that smart thermostats under-deliver on their promised energy efficiency claims.

5 Potential Mechanisms

In this section, we supplement our experimental analysis by analyzing user interactions with the smart thermostat. As shown in Appendix Table 4, this interactions data is heavily drawn from the Northern California wave of the experiment, so we focus our analysis on those households. We consider five questions with the interactions data of this sample. First, do users program their smart thermostat? Second, are programmed setpoints energy-efficient setpoints? Third, do users deviate from their programmed schedules? Fourth, do users deviate from their programmed schedule towards energy-efficient setpoints? Fifth, do households save energy when they use the smart thermostats scheduled setpoints and overrides as intended by engineers?

By answering these questions we aim to sharpen our understanding of the failure of the smart thermostat to deliver energy savings. While engineering models assume households will utilize the functionality of a smart thermostat and do so to conserve energy, economic models are ultimately agnostic and emphasize the potential for preferences to interfere with the response desired by an engineer. The first four questions consider whether households interact with the smart thermostat as engineers predict. The fifth question considers whether the subsamples of households that interact with their smart thermostat as an engineer assumes obtain the predicted energy savings.

We find that households schedule setpoints and that these setpoints are broadly in line with energy-efficient suggestions. However, the setpoint overrides made easy by the smart thermostat are common and these overrides are biased towards warmer setpoints in the winter and cooler setpoints in the summer. Finally, we find that households using the setpoint functionality as an engineering model assumes save as much as 10 to 20 percent on their consumption of natural gas. However, these savings are not found for electricity consumption, nor for households using the

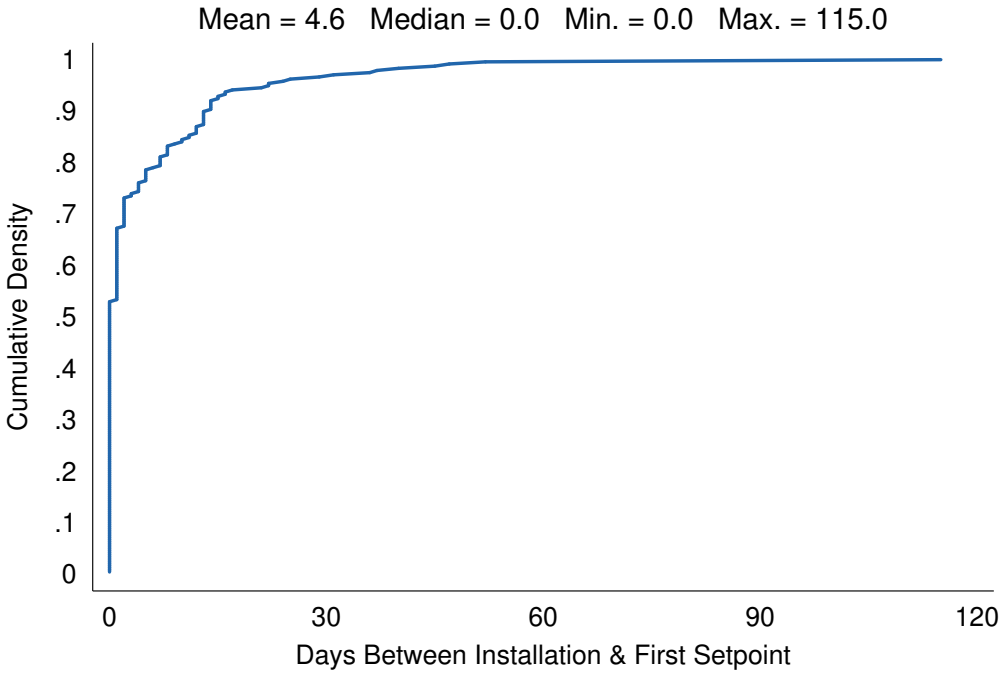
override functionality to obtain more energy-efficient setpoints.

5.1 Do Users Program Their Smart Thermostats?

Peffer et al. (2013) find that programmable thermostats fail to achieve their advertised savings due, in part, to poor usability.[45] If users do not program schedules for their smart thermostats to follow because the interfaces are too complicated or they do not understand how thermostats and/or their HVAC systems work, we would not expect the installation of a smart thermostat to affect energy consumption.

To determine the fraction of households who install the smart thermostat use the programmable features of the device and how long it takes them to begin doing so, Figure 7 plots the CDF of the time between the installation date and the first scheduled setpoint. The figure shows that almost all users who install a smart thermostat program at least one permanent setpoint, and most households do so almost immediately. The median time from installation to the first permanent setpoint is one day.

Figure 7: Distribution of Time from Installation to First Scheduled Setpoint

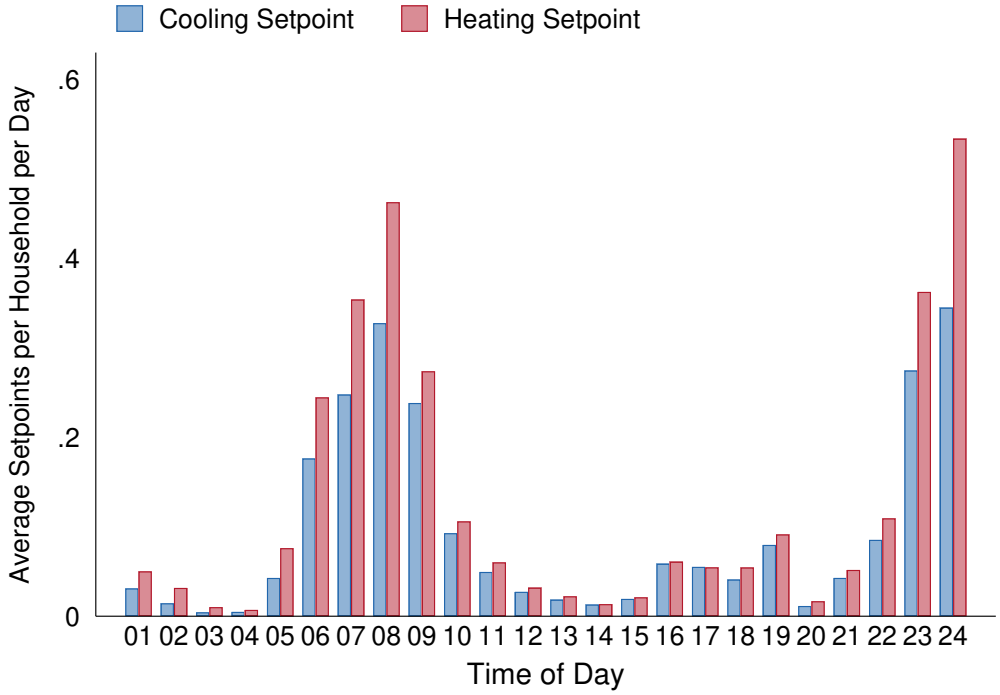


Note: Cumulative density conditional on observing the household in the HVAC events data.

[45] Programmable thermostats are a precursor technology to smart thermostats. The two types of thermostats share the ability to schedule permanent temperature setpoints in advance, but users cannot interact with programmable thermostats remotely, nor do they offer built-in setpoint framing. Peffer et al. (2013) report that they were so difficult to program that most users disabled their defining feature, and the ENERGY STAR program stopped certifying them in December 2009.

Additionally, users do not just quickly schedule a permanent setpoint, then fail to continue to use the smart features of the device. Individuals who have a smart thermostat installed as part of our experiment set an average of 3.749 (heating and cooling) setpoints per day. Figure 8 plots a measure of the frequency of permanent setpoints by hour of the day (denoted in military time) for both heating (red bars) and cooling (blue bars) setpoints. The figure provides visual evidence that setpoints occur frequently and when we would expect them: in the morning from about 5:00 AM until 10:00 AM when most users wake and leave for work and/or school. Similarly, there is a small increase in frequency of setpoints during the afternoon from 4:00 PM until 7:00 PM when users return home at the end of their days. Consistent with scheduling setpoints when most users go to sleep, we also observe frequent setpoints in the evening from about 10:00 PM until 12:00 AM. Thus, our analysis suggests that users do program their smart thermostats both quickly and frequently.

Figure 8: Average Permanent Setpoints per Household per Day by Time of Day

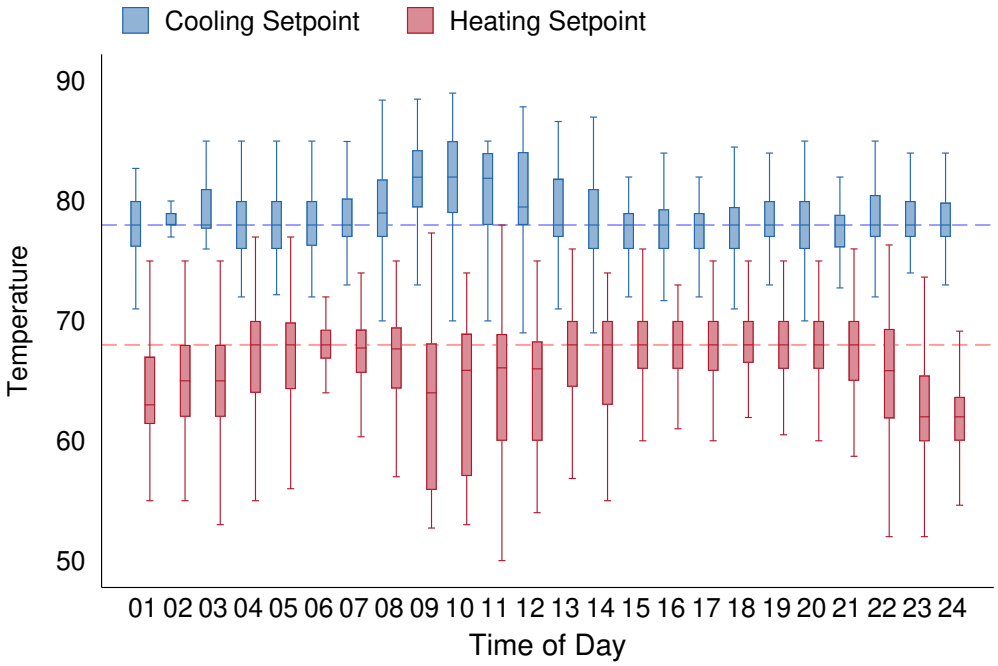


5.2 Are Programmed Setpoints Energy-Efficient Setpoints?

The previous analysis is consistent with users taking advantage of their device's scheduling feature, but is inconclusive as to whether or not they are programming setpoints to achieve energy savings. To inform the latter, Figure 9 is a box and whisker plot of heating and cooling setpoints by hour of the day. The dashed lines represent the cooling and heating temperature settings the DOE recommends for energy savings of 78 degrees F for cooling and 68 degrees F for heating

(DOE 2020). The figure illustrates that median (as well as the 25th and 75th percentiles of) temperatures are in line with the DOE's recommendations.[46] According to Appendix Table 4, cooling setpoints average 78.80 degrees F and are higher than heating setpoints, which average 63.95 degrees F. Additionally, the figure illustrates that there is temporal variation in setpoints over the course of the day consistent with individuals adjusting settings when they leave the house: cooling setpoints increase slightly starting at around 9:00 AM and drop back to baseline around 3:00 PM. Heating setpoints follow a similar, but opposite pattern with a more pronounced discrepancy between evening and daytime temperature setpoints. Overall, while the figure illustrates variation in setpoints across households, our analysis suggests that users program their smart thermostats to save energy.[47]

Figure 9: Box and Whisker Plots of Permanent Setpoints by Time of Day



Note: Dashed lines represent DOE recommended cooling/heating settings.

5.3 Do Users Deviate from Their Programmed Schedules?

Given that users seem to be programming their smart thermostats and doing so with energy savings in mind, we turn to an alternative explanation for our null findings. The remote features of the thermostat reduce the costs associated with both permanent and temporary setpoint changes. If users program their thermostats to reduce energy usage, but the ability to more easily adjust

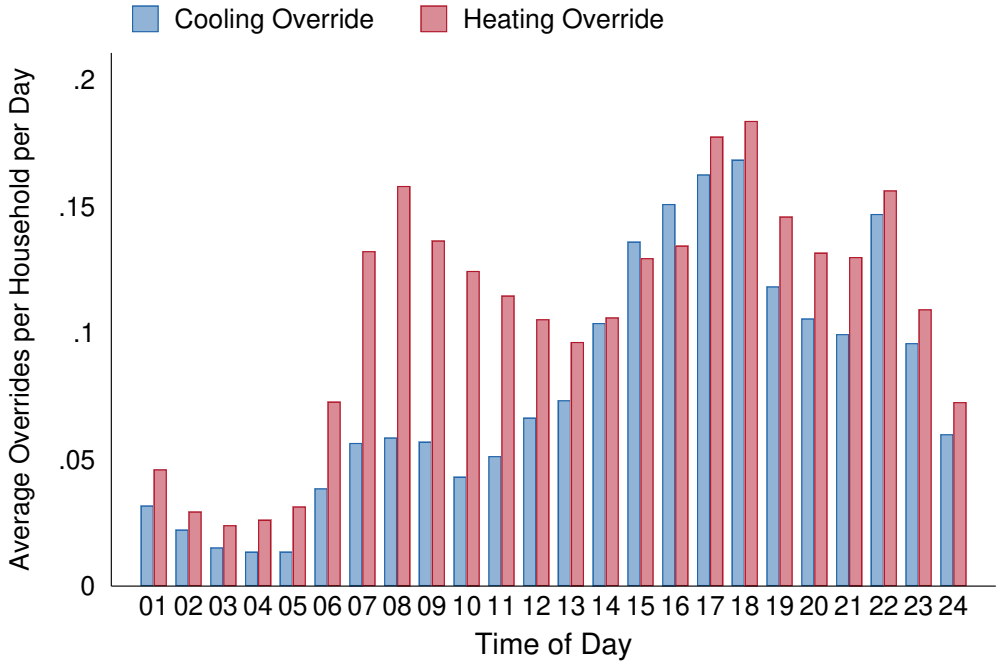
[46]The horizontal lines in the shaded boxes represent the median temperature setting, the ends of the boxes indicate the first and third quartiles, and the ends of the whiskers denote the upper/lower adjacent values.

[47]Regarding the variation in setpoints, Appendix Table 4 reports standard deviations of 4.12 degrees for cooling and 5.58 degrees for heating setpoints.

temperature settings via a computer or smart phone makes individuals more likely to deviate from their schedules, individuals may undo the benefits of their smart thermostat. If so, the effects of the scheduling and override features of smart thermostats have opposing effects on energy use and could result in a net null effect.

To explore this possibility, Figure 10 plots a measure of the frequency of setpoint overrides by time of the day.[48] As we would expect, overrides are more frequent when most individuals are likely to be awake, from about 6:00 AM to 11:00 PM. Heating overrides peak in the morning and early evening, while cooling overrides rise throughout the day until about 6:00 PM. More importantly given our focus, the figure illustrates that users often override their permanent schedule both when heating and cooling their homes. Compared to the previously noted 3.749 setpoints per day, users in our data temporarily override their permanent setpoints an average of 1.699 times per day. The hourly measures are also substantial relative to the number of permanent setpoints reported in Figure 8.

Figure 10: Average Temporary Overrides per Household per Day by Time of Day



Note: Days defined as those when the HVAC system cooled/heated the home.

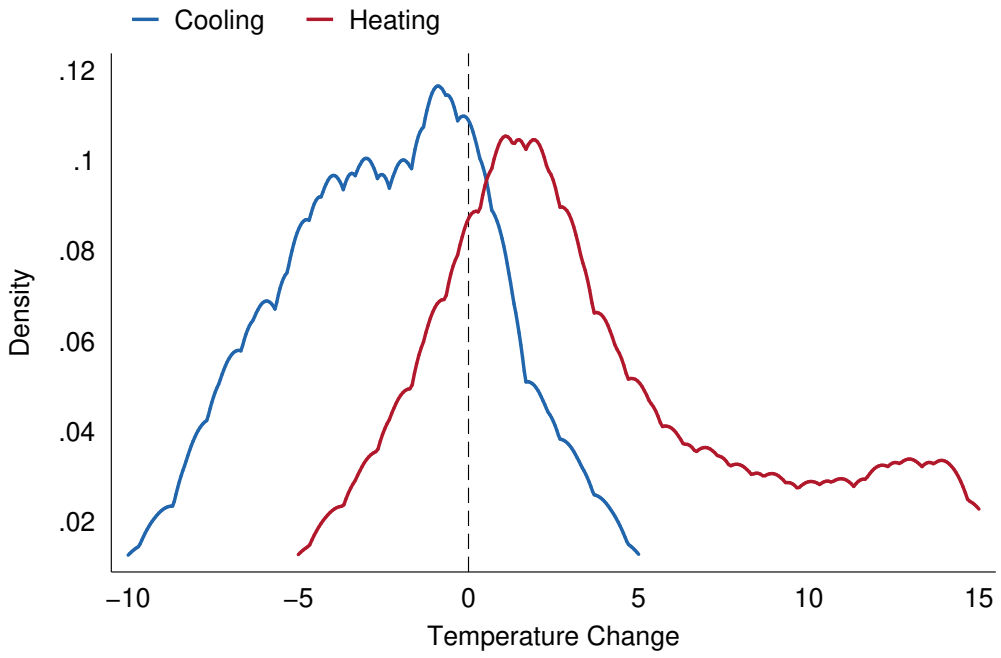
[48]The figure is the analog to Figure 8 for temporary overrides, save for our definition of "per day." While users program both heating and cooling setpoints every day, we typically only observe heating (cooling) overrides on heating (cooling) degree days. Given that we predominantly observe the HVAC system events data during the fall and winter, failure to address this issue results in heating and cooling override measures that are of different magnitudes. To account for this artifact in the data, we adjust the numerator of our measure to days on which the HVAC system heated or cooled the home to standardize the scales of the heating and cooling override measures.

5.4 Do Users Deviate From Their Programmed Schedule Towards Energy-Efficient Setpoints?

Evidence that smart thermostat users frequently override their setpoints offers a potential explanation for our null findings. The features of the smart thermostat that lower adjustment costs both make it easier to override in ways that increase energy use (e.g., users no longer have to get off the couch or out of bed and walk to the thermostat when they are uncomfortable) and to override to decrease energy use (e.g., by toggling the HVAC system off when leaving home). To determine which effect dominates, Figure 11 plots kernel densities of the difference between the override temperature a user sets and the permanent setpoint, conditional on a temperature override, by temperature setting (cooling or heating). The figure illustrates that when users override their permanently scheduled setpoints, they generally do so in ways that use more energy: when cooling, they set temperatures colder and when heating, they set it warmer.[49] Taken together with the previous figure, our analysis suggests that individuals undo the benefits of their preset smart thermostat schedule when they are uncomfortable in the moment. This suggests a potential explanation for our null experimental findings, and is consistent with the observational studies (Sachs et al., 2012; Peffer et al., 2013; Pritoni et al., 2015; Huchuk et al., 2020).

[49]There is a non-trivial mass at large override-setpoint temperature differences (e.g., greater than 10 degrees F). This is primarily driven by a small number of households that program setpoints (~55 degrees F) that essentially turn off the HVAC system in the morning and override those setpoints at varying times in the afternoon/evening every day. This is consistent with using the programmable features of the smart thermostat based on a consistent daily departure time and a variable return time. Additionally, we note that the figure plots override-setpoint temperature differences, not override-ambient temperature differences. The ambient temperature may not actually be as low as the setpoint, so the actual temperature change caused by the override may not be so extreme.

Figure 11: Density of Difference between Temporary Override and Permanent Setpoint Temperatures by Heating/Cooling



Note: Densities truncated at the 5th and 95th percentiles.

5.5 Do Households Save Energy When the Smart Thermostat is Used as Intended?

There is an extensive literature in economics in that shows that engineering estimates fail to live up to their predictions (Davis et al., 2014; Levinson, 2016; Fowlie et al., 2018; Alpízar et al., 2019; Davis et al., 2020). In this subsection we consider whether the non-response to the smart thermostat in our experimental analysis is due to the smart thermostat not being used as intended. To accomplish this, we estimate the effect of the smart thermostat on energy consumption for different subsamples of households in the treatment group. These subsamples are selected to split households who utilize the setpoint scheduling and overrides for energy-efficient temperatures from those who do not.

We begin by using the interactions data to classify households who installed a smart thermostat based on how diligently they use their device to achieve energy savings. We do so by defining three energy-efficiency types: high (H), low (L), and unknown types (?). Appendix Figure 18 illustrates how this classification builds on our existing experimental design. The unknown type is necessary because we do not observe all households who install a smart thermostat in the HVAC events data. The high and low types are based on the distributions of two measures of energy-efficiency: the average number of permanent setpoints and temporary overrides observed per hour. For both metrics, we specify models based on various cutpoints between high and low types. Appendix

Figure 19 plots the CDFs of both measures of behavior based on all households for which we observe interactions data. As an example, we define high-type households based on the permanent setpoint measure as those above the median and low types as those below the median. In contrast, for the other metric, we define high types as those below the median number of average overrides per hour and low types as those above the median.

Given these classifications, we interact indicators for type with treatment and estimate a DDITT model.[50] Letting $k \in \{\text{H, L, ?}\}$ index the three types, we estimate

$$e_{it}^{j} = \alpha_i^{j} + \beta_t^{j} + \sum_k \gamma_k^{j} T_i R_i^k P_t + X_{it}\beta_X^{j} + u_{it}^{j}, \tag{3}$$

where R_i^k is an indicator for household i being of type k and all other indexes, variables, and parameters are defined as in Equations 1 and 2. The parameters of interest in this model are the γ_k^j which are the the ITT effects of a smart thermostat on the consumption of energy j for households of type k.

Table 2 reports estimates of the γ_k^{thm} parameters based on this subsample. Panel (A) reports estimated effects from a model based on the permanent setpoint type classification, and Panel (B) reports analogous estimates based on the temporary override type definition. Column (1) reports estimates from a baseline DDITT model that does not differentiate by type. Consistent with our DDIV model estimates, the effects are not statistically significant. Columns (2) through (6) report estimates based on varying definitions of the high- vs. low-type percentile cutpoint.[51] The estimates in Panel (A) in these columns indicate that households above the 10th percentile of average permanent setpoints per hour enjoy statistically significant savings, with those above the 90th percentile seeing the greatest reduction in their natural gas use. In contrast, low types who program relatively few setpoints never reduce their energy consumption after installing a smart thermostat.

Interestingly, the high types that are above the median of permanent setpoints experience energy savings that are broadly in line with the predictions of engineering estimates. For example, Column (4) of Panel A shows the high types save about 0.15 thm per day and, on average, the control group consumes 1.5 thm per day. This 10 percent energy savings and the subsequent estimates in Columns (5) and (6) line up strikingly well with the engineering estimates discussed in fn. 10 and 12. Moving to Panel B, however, we see less consistent evidence that temporary overrides explain the response to the smart thermostat. In all columns except Column (6), we see no energy savings regardless of type at traditional levels of significance.

These results illustrate that engineering estimates can go astray because households do not utilize energy efficient technology as engineers assume. For example, in Column (4) of Table 2, Panel A we see that approximately 50% of households save the energy predicted by engineering estimates and the other 50% enjoy no such savings at all and this substantially attenuates the

[50] We are not able to estimate an analogous DDIV model because we do not have valid instruments for types.

[51] For instance, the estimates reported in Column (4) of Panel (A) define high-types as those with more than the median number of setpoints per hour and low-types as those below the median.

effect of the technology. Furthermore, in Panel B, we find that most specifications fail to find any evidence of temporary overrides providing households with energy savings. As a result, consumer and policymaker decisions based on these estimates are destined to fall short of their expected effects.

Table 2: ITT Estimates of the Effect of a Smart Thermostat on Natural Gas Use by Setpoint and Override Type

	(1)	(2)	(3)	(4)	(5)	(6)
		High/Low-Type Percentile Cutpoint				
	Baseline	10	25	50	75	90
	Power Use (thm)					
Panel A: Permanent Setpoint Type Classification						
ITT ($\hat{\gamma}^{thm}$)	0.051					
	(0.045)					
High Type ITT ($\hat{\gamma}^{thm}_H$)		-0.076	-0.108**	-0.146***	-0.178***	-0.302***
		(0.047)	(0.047)	(0.050)	(0.061)	(0.074)
Low Type ITT ($\hat{\gamma}^{thm}_L$)		-0.071	0.088	0.014	-0.025	-0.046
		(0.146)	(0.105)	(0.069)	(0.055)	(0.049)
N	805	805	805	805	805	805
$N \times T$	398,243	398,243	398,243	398,243	398,243	398,243
R^2	0.618	0.618	0.618	0.618	0.618	0.618
F statistic	520.944	397.030	398.793	397.866	398.103	398.168
Panel B: Temporary Override Type Classification						
ITT ($\hat{\gamma}^{thm}$)	0.051					
	(0.045)					
High Type ITT ($\hat{\gamma}^{thm}_H$)		0.058	-0.070	-0.068	-0.082	-0.095**
		(0.146)	(0.081)	(0.061)	(0.050)	(0.046)
Low Type ITT ($\hat{\gamma}^{thm}_L$)		-0.091*	-0.077	-0.086	-0.040	0.414***
		(0.046)	(0.050)	(0.054)	(0.074)	(0.095)
N	805	805	805	805	805	805
$N \times T$	398,243	398,243	398,243	398,243	398,243	398,243
R^2	0.618	0.618	0.618	0.618	0.618	0.618
F statistic	520.944	398.476	398.424	402.226	401.014	409.991
Weather Controls	x	x	x	x	x	x
HH Fixed Effects	x	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x	x
Day-of-Week Effects	x	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of the Northern California wave of the experiment. The sample underlying the estimates in both panels is based on *daily* natural gas meter readings (thm). The coefficient estimate for the unknown type ITT parameter ($\hat{\gamma}^{thm}_?$) is 0.014 across all specifications, and it is not statistically significant.

6 Conclusion

Our work informs the efficacy of a popular technology designed to conserve energy by exploring how smart technologies affect energy use—both through actual measurement and by investigating the mechanisms that prevent the realization of advertised energy savings. We provide evidence from a field experiment wherein residential households are randomized into either a treatment group that receives a smart thermostat or a control group. The smart thermostats given to the treatment group allow households to set more advanced schedules and adjust temperature settings remotely via a smart phone app. In addition, the smart thermostats provide households in the treatment group with information designed to promote energy-efficient setpoints.

In contrast to the commonly held prior that smart thermostats are an effective way to reduce residential energy use, we find little to no evidence that the installation of a smart thermosat reduces household energy consumption on average. This null result is robust to numerous specifications. We believe that the discord between the results of our field experiment and the extant belief stems from the source of the latter: engineering studies that do not adequately account for how individuals use their smart devices. We augment our experimental analysis with data on user interactions with their smart thermostat and find evidence that supports this belief.

There are many ways to extend our research. One avenue would be to better understand how different smart technology features, that often have opposing theoretical energy impacts, affect actual usage. Another would be to understand why smart thermostats are so popular given their costs and trivial energy-efficiency benefits. This avenue speaks to the energy efficiency gap literature as outlined by Allcott and Greenstone (2012). A further avenue would be to explore the impact that such technologies have on the price elasticity of energy demand (some preliminary evidence from Herter (2007) suggests that they do). If technology can enable people to better optimize their energy consumption, then price might become even more salient and therefore make people more marginal.

In summary, cooling and heating homes, powering transportation, and producing the wealth of goods and services enjoyed in modern economies are all heavily reliant on energy. Given that most of the world relies on non-renewable resources to produce energy and this reliance will not end any time soon (Covert et al., 2016), one of the greatest policy challenges of this century is how to address the negative externalities associated with energy production. Without efforts to promote energy conservation and associated reductions in greenhouse gas emissions, future generations will face a lower quality of life due to a degraded environment. Our hope is that this paper is a step towards ensuring that decision makers focus their energies on the smartest policies possible.

References

Aarish, Carlyn, Matei Perussi, Andrew Rietz, and Dave Korn, "Evaluation of the 2013-2014 Programmable and Smart Thermostat Program," Technical Report, The Cadmus Group 2015. 4

Abadie, Alberto, "Statistical Nonsignificance in Empirical Economics," *American Economic Review: Insights*, June 2020, *2* (2), 193–208. 4

Agnew, Ken and Mimi Goldberg, "Chapter 8: Whole Building Retrofit with Consumption Data Analysis Evaluation Protocol. The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures," Technical Report NREL/SR-7A30–53827, National Renewable Energy Laboratory April 2013. 5

Alberini, Anna, Silvia Banfi, and Celine Ramseier, "Energy Efficiency Investments in the Home: Swiss Homeowners and Expectations about Future Energy Prices," *The Energy Journal*, 2013, *34* (1), 49–86. 4

Allcott, Hunt, "Social norms and energy conservation," *Journal of Public Economics*, 2011, *95* (9), 1082 – 1095. Special Issue: The Role of Firms in Tax Systems. 2, 7

_ **and Michael Greenstone**, "Is There an Energy Efficiency Gap?," *Journal of Economic Perspectives*, February 2012, *26* (1), 3–28. 4, 31

_ **and Todd Rogers**, "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *American Economic Review*, October 2014, *104* (10), 3003–37. 2

Alpízar, Francisco, María Bernedo, and Paul J. Ferraro, "Input Efficiency as a Solution to Externalities: Engineers vs Behavioral Scientists in a Randomized Controlled Trial," Technical Report 2019. 28

Apex Analytics, LLC, "Energy Trust of Oregon Nest Learning Thermostat Heat Pump Control Pilot Evaluation," Technical Report 2014. 4

_ , "Energy Trust of Oregon Smart Thermostat Pilot Evaluation," Technical Report March 2016. 4, 8

Ayres, Ian, Sophie Raseman, and Alice Shih, "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage," *The Journal of Law, Economics, and Organization*, 08 2012, *29* (5), 992–1022. 2

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249. 17

Blonz, Joshua, Karen Palmer, Casey J. Wichman, and Wietelman Derek C., "Smart Thermostats, Automation, and Time-Varying Prices," Technical Report Working Paper 21-20, Resources for the Future (RFF) July 2021. 6

Bloomberg Finance L.P., "Sustainable Energy in America Factbook," Technical Report, The Business Council for Sustainable Energy 2019. Smart thermostat facts on page 112. 5

Borenstein, Severin, "The redistributional impact of nonlinear electricity pricing," *American Economic Journal: Economic Policy*, 2012, *4* (3), 56–90. 6

Broaddus, Phil, Barb Ryan, and Kelly Marrin, "PG&E Smart Thermostat Study: First Year Findings," Technical Report 2016. 5

_ , _ , and _ , "PG&E Smart Thermostat Study: Second Year Findings," Technical Report 2018. 5

Clarke, Damian and Kathya Tapia Schythe, "Implementing the Panel Event Study," IZA Discussion Papers 13524, Institute of Labor Economics (IZA), Bonn 2020. 15

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg, "From LATE to MTE: Alternative Methods for the Evaluation of Policy Interventions," *Labour Economics*, 2016, *41*, 47 – 60. SOLE/EALE Conference Issue 2015. 17

Costa, Dora L. and Matthew E. Kahn, "Energy Conservation "Nudges" and Environmentalist Ideology: Evidence From a Randomized Residential Electricity Field Experiment," *Journal of the European Economic Association*, 2013, *11* (3), 680–702. 2

Covert, Thomas, Michael Greenstone, and Christopher R. Knittel, "Will We Ever Stop Using Fossil Fuels?," *Journal of Economic Perspectives*, February 2016, *30* (1), 117–38. 31

Daken, Abigail A, Alan K Meier, and Douglas G Frazee, "Do Connected Thermostats Save Energy?," Technical Report, ACEEE Summer Study on Energy Efficiency in Buildings 2016. 4, 7

Davis, Lucas W., Alan Fuchs, and Paul Gertler, "Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico," *American Economic Journal: Economic Policy*, November 2014, *6* (4), 207–38. 28

_ , Sebastian Martinez, and Bibiana Taboada, "How effective is energy-efficient housing? Evidence from a field trial in Mexico," *Journal of Development Economics*, 2020, *143*, 102390. 3, 28

Department of Energy, "Smart Grid Investment Grant Program Final Report," Technical Report, U.S. Department of Energy, Office of Electricity Delivery & Energy Reliability December 2016. 5

_ , "Thermostats," https://www.energy.gov/energysaver/thermostats 2020. Accessed: July 21, 2020. 25

Dolan, Paul and Robert Metcalfe, "Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation," *Becker Friedman Institute for Research in Economics Working Paper*, 2015. 2

ecobee, "Savings from your ecobee," https://www.ecobee.com/savings/ 2013. Accessed: September 10, 2019. 2

Energy Information Administration, "Energy Consumption and Expenditures Tables," Technical Report, U.S. Department of Energy, 2015 Residential Energy Consumption Survey May 2018. 2

_ , "Annual Energy Outlook 2019 with Projections to 2050," Technical Report, U.S. Department of Energy, Office of Energy Analysis January 2019. 2

_ , "Monthly Energy Review," Technical Report, U.S. Department of Energy, Office of Energy Statistics August 2019. 2

Environmental Protection Agency, "ENERGY STAR Summary of HVAC & Smart Thermostat Programs," Technical Report, U.S. Environmental Protection Agency 2019. Accessed: July 24, 2020. 5

Faruqui, Ahmad and Sanem Sergici, "Household response to dynamic pricing of electricity: a survey of 15 experiments," *Journal of Regulatory Economics*, Oct 2010, *38* (2), 193–225. 4

Ferraro, Paul J. and Michael K. Price, "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment," *The Review of Economics and Statistics*, 2013, *95* (1), 64–73. 2

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram, "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program," *The Quarterly Journal of Economics*, 01 2018, *133* (3), 1597–1644. 3, 28

Ge, Qi and Benjamin Ho, "Energy Use and Temperature Habituation: Evidence from High Frequency Thermostat Usage Data," *Economic Inquiry*, 2019, *57* (2), 1196–1214. 11

Ghanem, Dalia and Aaron Smith, "What Are the Benefits of High-Frequency Data for Fixed Effects Panel Models?," *Journal of the Association of Environmental and Resource Economists*, 2021, *8* (2), 199–234. 5

Harding, Matthew and Carlos Lamarche, "Empowering Consumers Through Data and Smart Technology: Experimental Evidence on the Consequences of Time-of-Use Electricity Pricing Policies," *Journal of Policy Analysis and Management*, 2016, *35* (4), 906–931. 4

Heckman, James J. and Jeffrey A. Smith, "Assessing the Case for Social Experiments," *Journal of Economic Perspectives*, June 1995, *9* (2), 85–110. 52

Herter, Karen, "Residential implementation of critical-peak pricing of electricity," *Energy Policy*, 2007, *35* (4), 2121 – 2130. 31

Ho, Benjamin, "Understanding Energy Efficiency Benefits from Smart Thermostats in Southern California," Technical Report, EnergyHub 2014. 4

Holladay, J. Scott, Michael K. Price, and Marianne Wanamaker, "The perverse impact of calling for energy conservation," *Journal of Economic Behavior and Organization*, 2015, *110*, 1–18. 21

Honeywell, "Energy Savings Calculator," https://www.honeywellhome.com/en/energy-calculator 2013. Accessed: September 10, 2019. 2

Huchuk, Brent, William O'brien, and Scott Sanner, "Exploring smart thermostat users' schedule override behaviors and the energy consequences," *Science and Technology for the Built Environment*, 2020, *27* (2), 195–210. 27

Hudson, Sally, Peter Hull, and Jack Liebersohn, "Interpreting Instrumented Difference-In-Differences," Technical Report 2017. 52

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475. 18, 52

Ito, Seiro, "A Practical Guide to the Program Evaluation Methods," *Health Service and Poverty: Making Health Services More Accessible to the Poor*, 2007, pp. 143–203. 53

Jessoe, Katrina and David Rapson, "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use," *American Economic Review*, April 2014, *104* (4), 1417–38. 4, 6

Joskow, Paul L., "Creating a Smarter U.S. Electricity Grid," *Journal of Economic Perspectives*, February 2012, *26* (1), 29–48. 4, 21

Kelsven, Phillip, Robert Weber, and Eva Urbatsch, "Nest Learning Thermostat Pilot Program Savings Assessment," Technical Report, Bonneville Power Administration & Franklin Public Utility District 2016. 4

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 2006, *133* (1), 97 – 126. 18

Kotchen, Matthew J., "Longer-Run Evidence on Whether Building Energy Codes Reduce Residential Energy Consumption," *Journal of the Association of Environmental and Resource Economists*, 2017, *4* (1), 135–153. 5

Levinson, Arik, "How Much Energy Do Building Energy Codes Save? Evidence from California Houses," *American Economic Review*, October 2016, *106* (10), 2867–94. 5, 28

List, John A, "Non est Disputandum de Generalizability? A Glimpse into The External Validity Trial," Working Paper 27535 July 2020. 3

Nest, "It saves energy. And here's the proof.," https://www.nest.com/thermostats/real-savings/ 2019. Accessed: September 10, 2019. 2, 4

Nest Labs, "Energy Savings from the Nest Learning Thermostat: The Impact of Nest Learning Thermostat," Technical Report 2014. 4

_ , "Energy Savings from the Nest Learning Thermostat: Energy Bill Analysis Results," Technical Report 2015. 4, 5

Novan, Kevin, Aaron Smith, and Tianxia Zhou, "Residential Building Codes Do Save Energy: Evidence From Hourly Smart-Meter Data," *The Review of Economics and Statistics*, Forthcoming. 5

Park, Toby, Pantelis Solomon, Michael Sanders, Marius Dietsch, Julie Faller, Daniel Gibbons, Elisabeth Costa, and Alex Tupper, "Evaluating the Nest Learning Thermostat: Four Field Experiments Evaluating the Energy Saving Potential of Nest's Smart Heating Control," Technical Report, The Behavioural Insights Team 2017. 5

Peffer, Therese, Daniel Perry, Marco Pritoni, Cecilia Aragon, and Alan Meier, "Facilitating energy savings with programmable thermostats: evaluation and guidelines for the thermostat user interface," *Ergonomics*, 2013, *56* (3), 463–479. PMID: 23005033. 2, 8, 23, 27

_ **, Marco Pritoni, Alan Meier, Cecilia Aragon, and Daniel Perry**, "How people use thermostats in homes: A review," *Building and Environment*, 2011, *46* (12), 2529 – 2541. 8

Pritoni, Marco, Alan K. Meier, Cecilia Aragon, Daniel Perry, and Therese Peffer, "Energy efficiency and the misuse of programmable thermostats: The effectiveness of crowdsourcing for understanding household behavior," *Energy Research & Social Science*, 2015, *8*, 190 – 197. 27

Robinson, Jen, Ram Narayanamurthy, Bienvenido Clarin, Christine Lee, and Pranshu Bansal, "National Study of Potential of Smart Thermostats for Energy Efficiency and Demand Response," Technical Report, ACEEE Summer Study on Energy Efficiency in Buildings 2016. 4

Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146. 17

Sachs, Olga, Verena Tiefenbeck, Caroline Duvier, Angela Qin, Kate Cheney, Craig Akers, and Kurt Roth, "Field evaluation of programmable Thermostats," Technical Report, National Renewable Energy Lab.(NREL), Golden, CO (United States) 2012. 27

Schellenberg, Josh, Alana Lemarchand, and Alexandra Wein, "Xcel Energy Colorado Smart Thermostat Pilot - Evaluation Report," Technical Report, Nexant, Inc. 2017. 4

Schmidheiny, Kurt and Sebastian Siegloch, "On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications," CESifo Working Paper 7481, Institute of Labor Economics (IZA), Munich 2019. 15

Shobe, William, Charles Holt, and Thaddeus Huetteman, "Elements of emission market design: An experimental analysis of California's market for greenhouse gas allowances," *Journal of Economic Behavior and Organization*, 2014, *107*, 402–420. 12

Stewart, James and Jeremy Jackson, "Energy Savings from Honeywell Total Connect Comfort Thermostats-Estimates from 2nd National Impact Study," Technical Report, The Cadmus Group 2015. 4

Urban, Bryan and Carla Gomez, "A Case for Thermostat User Models," Technical Report, International Building Performance Simulation Association 2013. 4

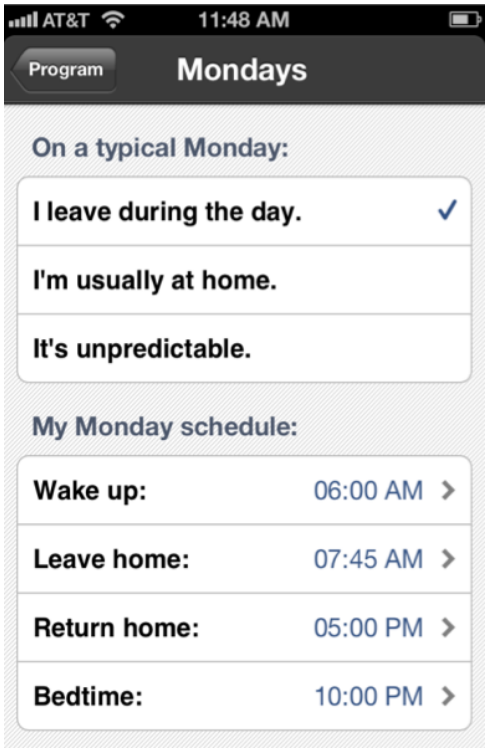
_ **, Diana Elliott, and Olga Sachs**, "Towards Better Modeling of Residential Thermostats," Technical Report 1, Proceedings of SimBuild 2012. 4

Ward, Bryan, James Stewart, and Jeremy Jackson, "Energy Savings from Honeywell Total Connect Comfort Thermostats," Technical Report, The Cadmus Group 2014. 4

A Smart Thermostat

Figure 12: Smart Thermostat Overview

(a) Interfaces: The left panel shows the web portal, the middle panel shows the smartphone app, and the right panel shows the thermostat.



(b) Permanent Setpoint Scheduling: Screenshot of the smartphone app scheduling interface.

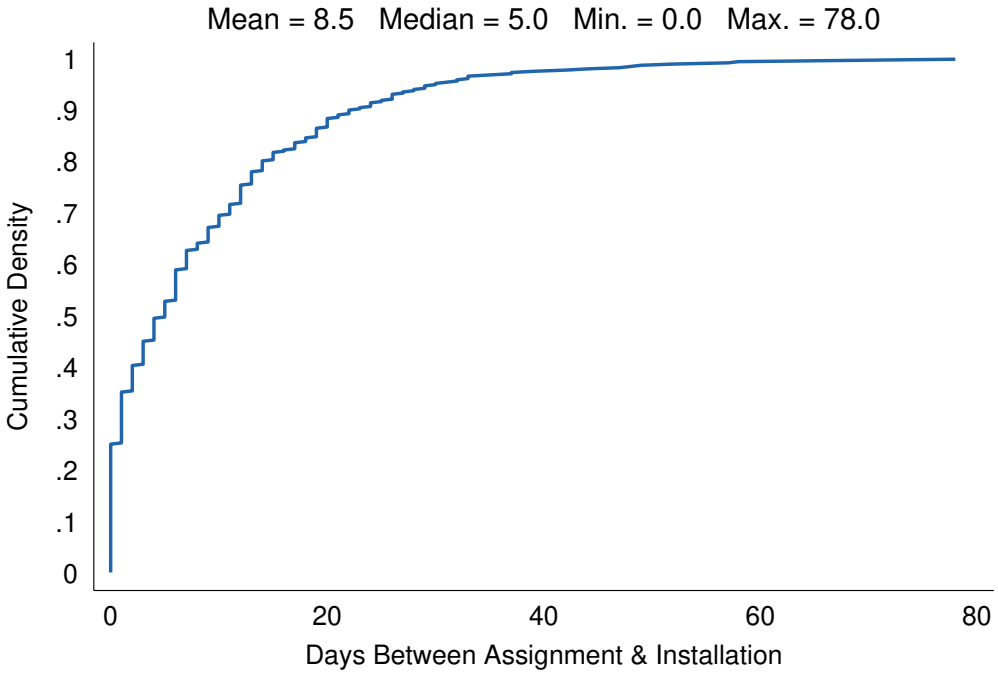
Figure 13: Smart Thermostat Features

(a) Setpoint Choice Messaging: Screenshots of smartphone app that shows the messaging associated with different temperature set points.

(b) Temporary Overrides: Screenshots of the smartphone app that facilitates changes to the temperature setpoint. The left panel shows the interface after the user indicates she is not home. The right panel shows the same interface when the user indicates she is at home.

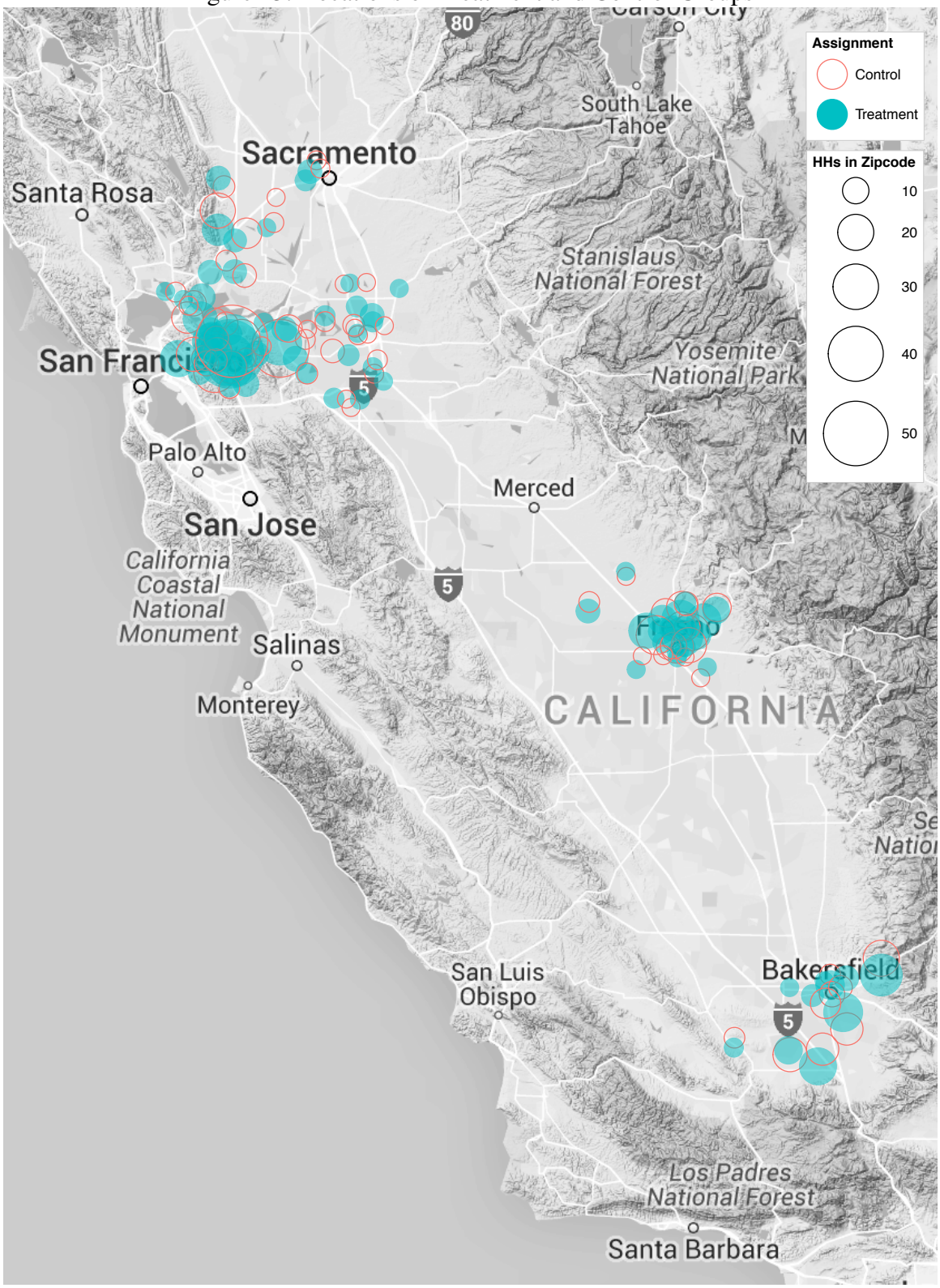
B Experimental Data

Figure 14: Conditional Distribution of Time from Assignment to Installation



Note: Cumulative density conditional on eventual installation.

Figure 15: Locations of Treatment and Control Groups

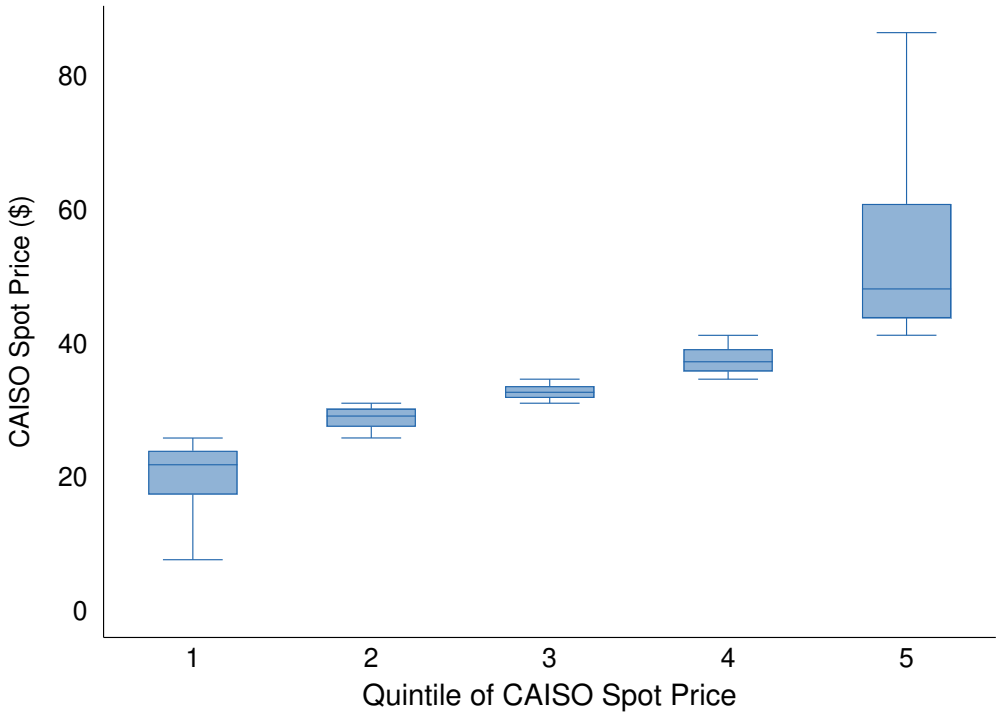


C External Data

Table 3: Daily Outdoor Weather Summary Statistics

Variable	Daily Measure	Mean	Std. Dev.	Between County Std. Dev.	Within County Std. Dev.	Min.	Max
Panel A: N. CA Experiment							
Temperature ($^\circ F$)	Mean	61.54	11.53	1.23	11.48	33.29	91.25
	Minimum	49.19	9.88	0.44	9.88	21.00	76.00
	Maximum	75.78	14.35	1.70	14.27	43.00	108.00
Relative Humidity ($\%$)	Mean	60.02	15.46	2.90	15.25	10.54	97.53
	Minimum	33.77	16.66	1.99	16.57	3.00	93.00
	Maximum	84.46	11.85	2.28	11.69	14.00	100.00
Heat Index ($^\circ F$)	Mean	60.61	11.57	1.16	11.52	31.62	90.30
	Minimum	48.42	10.25	0.45	10.24	18.97	76.01
	Maximum	74.19	13.63	1.54	13.57	40.82	108.61
N				4			
$N \times T$				2,060			
Panel B: C. CA Experiment							
Temperature ($^\circ F$)	Mean	66.58	14.36	2.56	14.20	32.63	96.04
	Minimum	54.20	12.91	3.66	12.56	19.00	85.00
	Maximum	79.85	16.05	1.17	16.02	45.00	110.00
Relative Humidity ($\%$)	Mean	51.11	17.08	6.55	16.22	13.33	96.78
	Minimum	28.86	15.74	1.78	15.67	2.00	90.00
	Maximum	73.15	16.87	10.15	14.70	22.00	100.00
Heat Index ($^\circ F$)	Mean	65.42	14.18	2.43	14.04	30.62	95.68
	Minimum	53.41	13.28	3.62	12.94	16.87	83.66
	Maximum	77.81	15.07	1.09	15.04	43.20	109.53
N				3			
$N \times T$				1,545			
Panel C: Both Experiments							
Temperature ($^\circ F$)	Mean	63.70	13.06	3.20	12.71	32.63	96.04
	Minimum	51.34	11.55	3.43	11.10	19.00	85.00
	Maximum	77.52	15.23	2.58	15.05	43.00	110.00
Relative Humidity ($\%$)	Mean	56.20	16.76	6.42	15.67	10.54	97.53
	Minimum	31.66	16.45	3.15	16.19	2.00	93.00
	Maximum	79.62	15.28	8.57	13.06	14.00	100.00
Heat Index ($^\circ F$)	Mean	62.67	12.97	3.04	12.66	30.62	95.68
	Minimum	50.56	11.90	3.41	11.48	16.87	83.66
	Maximum	75.74	14.38	2.31	14.22	40.82	109.53
N				7			
$N \times T$				3,605			

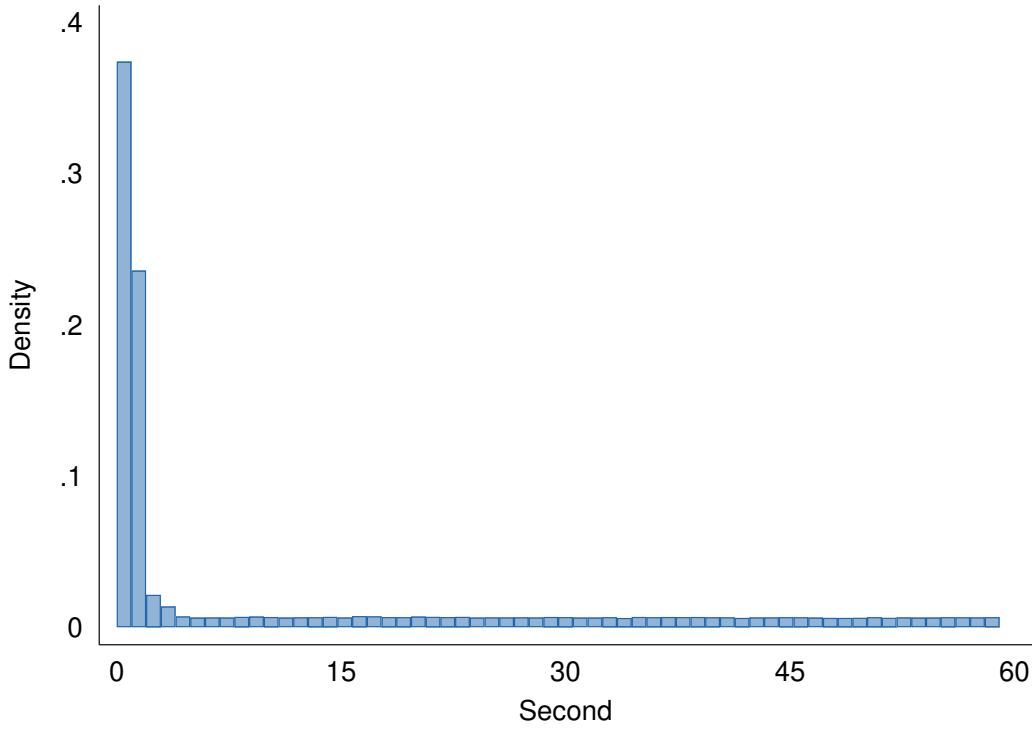
Figure 16: Box and Whisker Plots of CAISO Spot Price by Quintile of Price



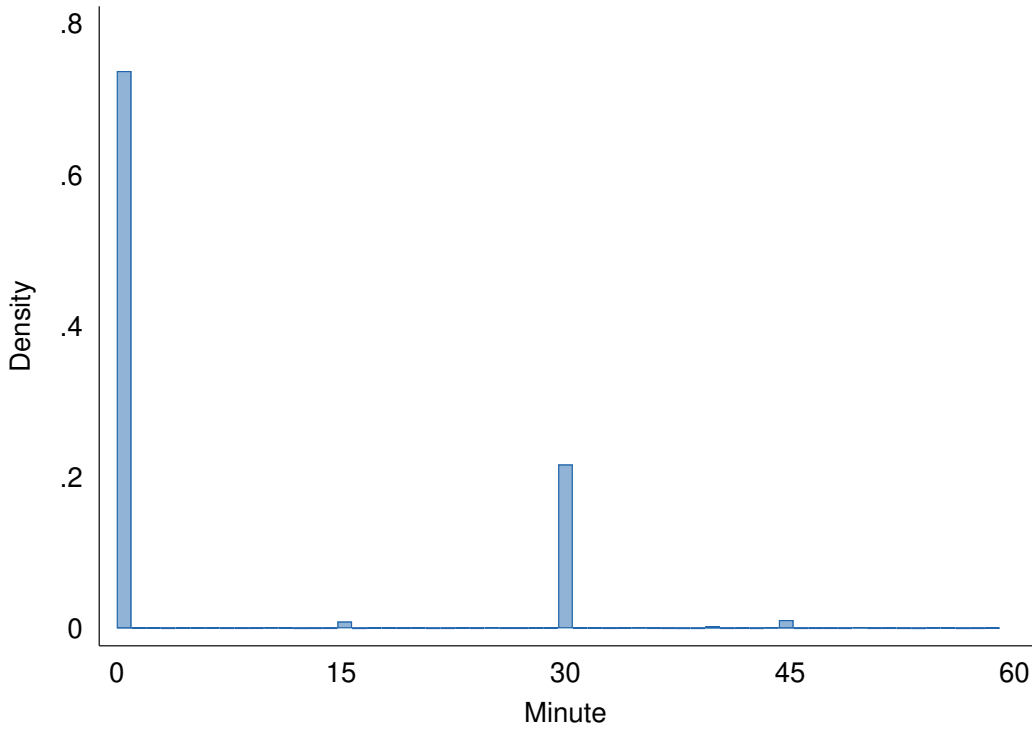
D Internal Data

The HVAC system events data does not label temperature changes as being the result of a permanent setpoint or temporary override. We infer this information based on the precise timing of when the change occurs. Appendix Figure 17 informs our approach to this classification. Panel (a) plots the density of the second of the minute at which temperature changes take place. The density is roughly uniform with a probability of about 0.70 across all seconds, save for a large increase in the probability of changes occurring at :00 through :02 (and to a lesser extent :03) seconds of the minute. Since we would expect temporary overrides to occur uniformly across seconds of the minute, we code temperature changes occurring at less than :03 seconds of the minute as permanent setpoints and all other temperature changes as temporary overrides. Panel (b) plots the density of permanent setpoints (as determined by our classification rule) by minute of the hour. Consistent with our priors, users schedule most setpoints on the hour or half hour (and to a lesser extent, at :15 and :45 minutes past the hour). This is both a finding and a confirmation of the validity of our approach to classifying setpoints and overrides.

Figure 17: Timing of HVAC System Events

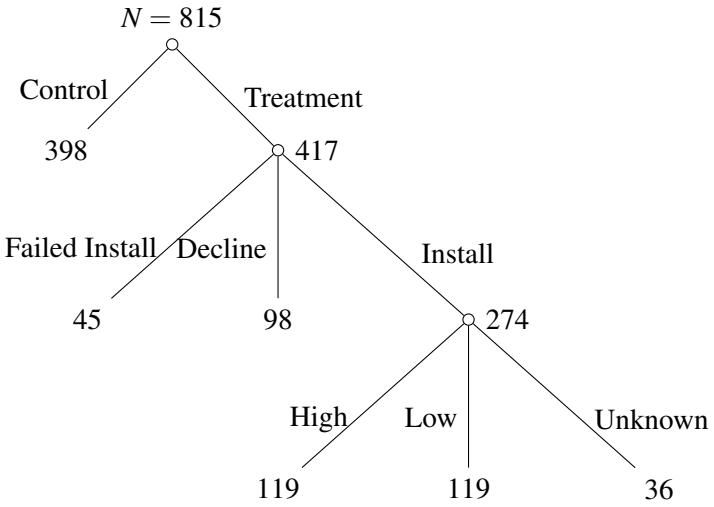


(a) Density of Temperature Changes by Second of the Minute



(b) Density of Permanent Setpoints by Minute of the Hour

Figure 18: Modified Sample Randomization with Energy-Efficiency Types (N. CA Experiment)



Note: Counts of high- and low-efficiency types are based on a definition that divides types at the median.

Figure 19: Distributions of Permanent Setpoints and Temporary Overrides

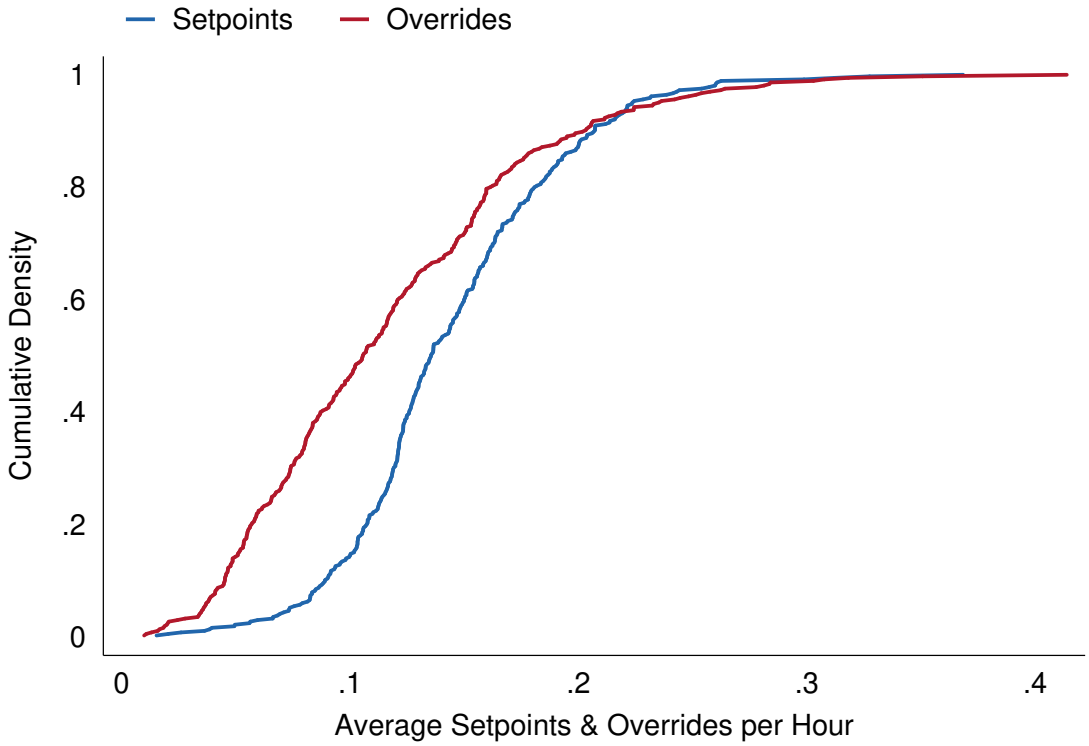


Table 4: User Interactions Summary Statistics by Experiment

	Both Experiments			N. CA Experiment			C. CA Experiment		
Variables	Mean	Std.Dev.	Obs.	Mean	Std.Dev.	Obs.	Mean	Std.Dev.	Obs
Ambient Temp.	68.92	5.18	346,199	69.08	5.20	320,601	66.88	4.36	25,598
Cooling Setpoints	78.87	4.12	57,197	78.91	4.10	54,164	78.22	4.52	3,033
Heating Setpoints	63.93	5.59	79,165	63.88	5.60	73,359	64.64	5.44	5,806
Cooling Overrides	77.51	3.97	15,622	77.51	3.87	14,414	77.56	4.99	1,208
Heating Overrides	67.48	4.21	48,035	67.37	4.14	41,497	68.16	4.58	6,538
N		372			238			134	
$N \times T$		386,134			357,392			28,742	

E Balance and Time-Trend Analysis

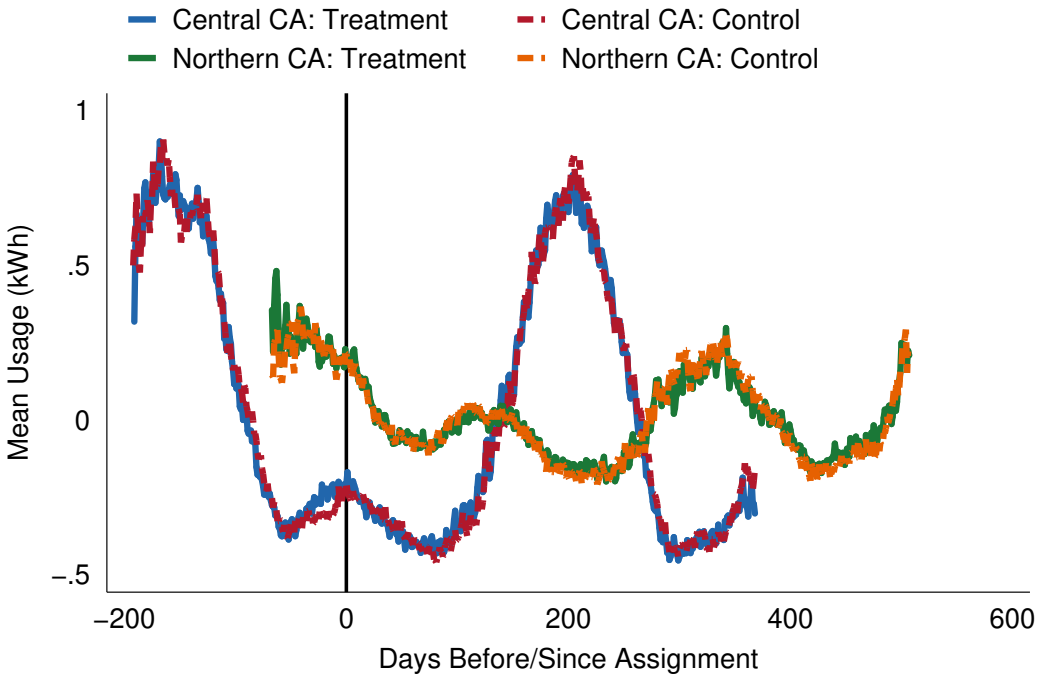
Table 5: Balance Table

Variable	(1) Both Treatment Indicator	(2) N. CA Treatment Indicator	(3) C. CA Treatment Indicator
Household Characteristics			
Family in the Household Indicator	0.025	-0.028	0.083
	(0.053)	(0.071)	(0.080)
Pets in the Household Indicator	0.013	0.019	0.005
	(0.029)	(0.038)	(0.045)
HER Experiment Indicator	-0.021	-0.004	-0.046
	(0.031)	(0.040)	(0.048)
HER Recipient Indicator	-0.006	0.027	-0.062
	(0.039)	(0.049)	(0.063)
Home Characteristics			
Multi-Family Home Indicator	-0.017	-0.022	0.039
	(0.080)	(0.091)	(0.166)
Year Home Built (Year / 1,000)	0.230	-0.589	1.363
	(0.800)	(1.110)	(1.170)
Size of Home (Sq. Ft. / 10,000)	0.286	0.377	-0.061
	(0.246)	(0.324)	(0.433)
Pool Indicator	-0.006	0.037	-0.082
	(0.033)	(0.044)	(0.052)
Electric Heat Indicator	0.014	-0.068	0.126
	(0.094)	(0.125)	(0.140)
Pre-Period Energy Use			
Mean (kWh)	-0.036	-0.054	0.010
	(0.028)	(0.034)	(0.048)
Mean (thm)	-0.029	0.002	-0.054
	(0.032)	(0.050)	(0.040)
N	1,385	821	564
R^2	0.011	0.015	0.021
F	0.664	0.689	0.799

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

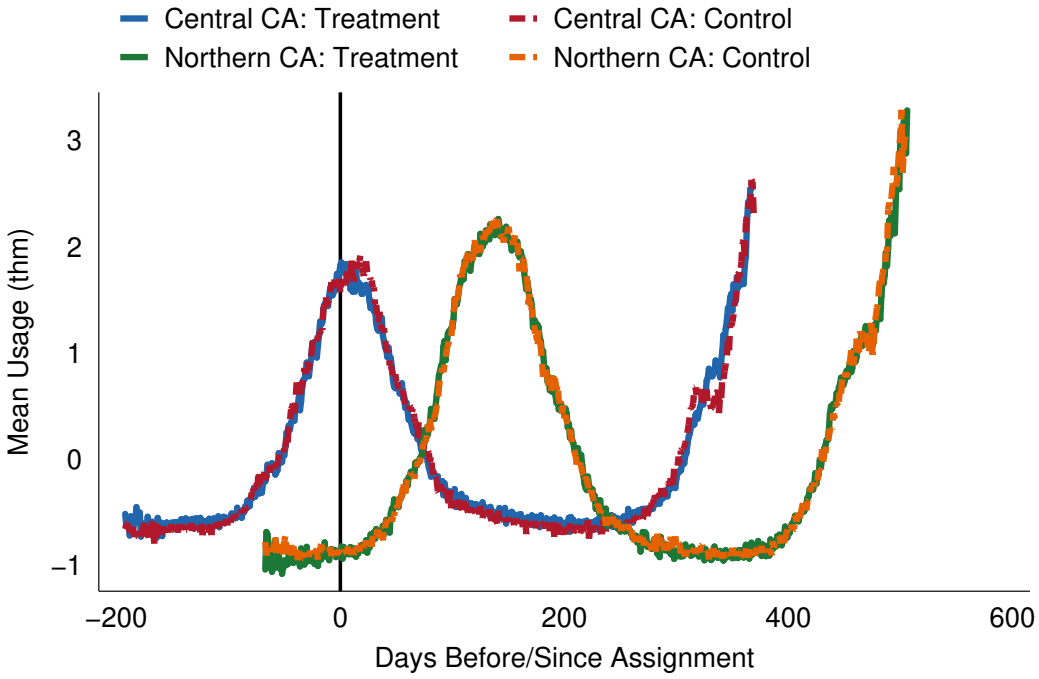
The table reports linear probability model estimates of the probability of assignment to treatment. The *HER Experiment Indicator* variable is equal to one for households that participated in the Home Energy Report experiment, and the *HER Recipient Indicator* variable is equal to one for households that were assigned to the treatment group in that experiment. We interpolate missing values of continuous variables (year built, home size, and pre-period energy use). We also code as zero and include an indicator for missing values of binary variables (heating type) and mismeasured values of pre-period electricity means in Northern California that are based on less than two weeks of data (see Section 2.4). Models include indicators for month and county of recruitment, as well as the aforementioned indicators for missing/mismeasured values. All omitted coefficient estimates are statistically insignificant. The F-statistic tests the null hypothesis that all parameters are jointly equal to zero. We fail to reject the null in all three models.

Figure 20: Average Residual Energy Use by Experimental Status and Wave



Note: Only means based on 30 or more homes per day are included in the figure.

(a) Electricity



Note: Only means based on 30 or more homes per day are included in the figure.

(b) Natural Gas

F Discussion of Identifying Assumptions

Our DDIV empirical specification identifies the ATT of a smart thermostat if our experimental instrument is relevant and valid, there are common or parallel trends, monotonicity holds, and there is one-sided experimental noncompliance.[52] We provide evidence that each of these assumptions is reasonable in this section. First, instrument relevance requires that assignment to treatment affects the probability that a household installs a smart thermostat. We report the first-stage F statistics with all of our results tables. As one would expect of a field experiment, we always easily reject the null of weak instruments.

Second, the instrument validity assumption in a DDIV model can be thought of as two separate conditions (Hudson et al., 2017). The first is the traditional IV assumption that the instrument is exogenous and the only way assignment to the treatment group affects energy use is through the installation of a smart thermostat. The second is the assumption implicit in all DD analyses that post-period randomization does not affect the pre-period values of outcomes (energy use) or treatment (smart thermostat installation). Both assumptions are satisfied by the nature of our experiment: households are randomly assigned to a treatment or control group. Assignment occurs both (shortly) after the household first interacts with the experimenter and after the household's pre-period energy use decisions have been made. The analyses in Section 2.6 and 2.7 are consistent with an appropriate randomization process.

Third, the common or parallel trends assumption requires that the unobserved, counterfactual trend in energy use that would have been experienced by the treated group is parallel to observable, untreated trend in the comparison group. In the context of our experiment, this means that the energy consumed by control group households is a good proxy for the energy homes who installed a smart thermostat would have used in a counterfactual world without a smart thermostat. While this assumption is fundamentally untestable because of the counterfactual outcomes problem, it is satisfied if there is appropriate randomization (Hudson et al., 2017). Nonetheless, we provide additional support for this assumption by showing evidence of parallel pre-trends via the event studies in Section 2.7.

Finally, if there is two-sided noncompliance in an experiment, the estimates are confounded by substitution bias (Heckman and Smith, 1995). The standard in the literature is to relax the noncompliance assumption to one of monotonicity (or uniformity). In our case, this means that the experimental treatment makes all households in more (or less) likely to get a smart thermostat than they would have been otherwise. Under this alternative assumption, the DDIV specification recovers the LATE estimate of γ^j (Imbens and Angrist, 1994). This is an estimate of the average impact of a smart thermostat on the energy consumption of households that were induced to install one by our experiment.

Our experimental environment allows us to make the stronger assumption that there is one-sided experimental noncompliance that allows us to identify the ATT of a smart thermostat. The

[52] Alternatively, we can recover the ITT estimate of γ^j by replacing the S_i in Equation 1 with T_i. This is an estimate of the average effect of being randomized into the treatment group in our experiment. We estimate DDITT models in Section 5.5 as we do not observe additional instruments for household energy-efficiency types.

assumption of one-sided noncompliance is tenuous to the extent that "the need for treatment under question is widely acknowledged and there is competition over implementation" (Ito, 2007). This is not the case in our context as smart thermostat technology was in its infancy at the time of our study. Using data from the EIA's Residential Energy Consumption Survey (RECS), we find that two to three years after our experiment, only 4.09% of all households in the survey and 10.58% of households observationally similar to those in our study own a smart thermostat.[53] Additionally, while we are unable to directly observe whether any households in the control group upgrade their thermostat, we never observe control households using a smart thermostat on Opower platform. Thus, the available evidence supports the validity of the assumption of one-sided noncompliance in our experimental context.

[53]The RECS is not conducted annually, so we use data from the 2015 survey as it is the closest possible survey iteration subsequent to the time period observed in our data. The previous iteration of the survey in 2009 did not ask questions about smart devices. We define "observationally similar" households by restricting the RECS sample to homes that would pass Opower's initial eligibility screening to join the trial (to the extent possible given the measures available). Specifically, we condition on owner-occupied, single-family homes located in the Pacific Division (state of residence is not observed) that have a functioning central furnace or heat pump, central air conditioning, and an electrical connection. We are not able to condition on whether or not the household has a high-speed Internet connection or whether the occupants plan to move in the next year, as those questions are not part of the RECS survey.

G Main Results with Full Regression Diagnostics

Table 6: N. CA Experiment-ATT Estimates of the Effect of a Smart Thermostat on Energy Use

	(1)	(2)	(3)	(4)	(5)	(6)
	Power Use (kWh or thm)					
Panel A: Electricity (kWh)						
ATT ($\hat{\gamma}^{kWh}$)	-0.055	-0.061	-0.016	-0.016	-0.016	-0.003
	(0.058)	(0.058)	(0.046)	(0.046)	(0.046)	(0.041)
Constant	1.294***					
	(0.035)					
N	815	815	815	815	815	815
$N \times T$	9,729,849	9,729,849	9,729,849	9,729,849	9,729,849	9,729,849
F statistic	44.591	270.070	343.596	353.483	350.375	171.299
rk LM statistic	391.219	391.264	313.225	313.190	313.190	269.656
rk Wald F statistic	379.956	380.003	670.871	670.765	670.766	639.637
Panel B: Natural Gas (thm)						
ATT ($\hat{\gamma}^{thm}$)	-0.009	0.009	0.085	0.075	0.075	0.069
	(0.061)	(0.063)	(0.068)	(0.066)	(0.066)	(0.055)
Constant	0.523***					
	(0.020)					
N	805	805	805	805	805	805
$N \times T$	398,243	398,243	398,243	398,243	398,243	398,243
F statistic	801.768	568.771	674.486	519.934	520.789	22.446
rk LM statistic	386.783	386.896	313.868	313.885	313.886	270.288
rk Wald F statistic	377.042	377.090	672.580	672.617	672.609	641.179
Weather Controls		x	x	x	x	x
HH Fixed Effects			x	x	x	x
Month-of-Year Effects				x	x	
Day-of-Week Effects					x	
Day Effects						x
Hour-of-Day Effects						x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of the Northern California wave of the experiment. The sample used to produce the estimates in Panel A is based on *hourly* electricity meter readings in kWh, while the sample underlying the estimates in Panel B is based on *daily* natural gas meter readings (thm). Thus, the hour-of-day effects noted in Column (6) are included in the electricity model (Panel A) only. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 7: C. CA Experiment-ATT Estimates of the Effect of a Smart Thermostat on Energy Use

	(1)	(2)	(3)	(4)	(5)	(6)
	Power Use (kWh or thm)					
Panel A: Electricity (kWh)						
ATT ($\hat{\gamma}^{kWh}$)	0.009	0.006	0.002	0.002	0.002	-0.001
	(0.029)	(0.028)	(0.025)	(0.025)	(0.025)	(0.023)
Constant	1.292***					
	(0.030)					
N	564	564	564	564	564	564
$N \times T$	6,691,885	6,691,885	6,691,885	6,691,885	6,691,885	6,691,885
F statistic	49.321	411.636	539.983	392.831	389.611	249.936
rk LM statistic	394.996	395.009	384.992	384.985	384.985	374.160
rk Wald F statistic	677.494	677.449	1,352.535	1,352.620	1,352.619	1,365.852
Panel B: Natural Gas (thm)						
ATT ($\hat{\gamma}^{thm}$)	-0.003	0.007	0.001	0.001	0.001	-0.021
	(0.044)	(0.031)	(0.027)	(0.026)	(0.026)	(0.026)
Constant	1.101***					
	(0.034)					
N	564	564	564	564	564	564
$N \times T$	279,061	279,061	279,061	279,061	279,061	279,061
F statistic	3.488	357.120	408.612	280.326	281.833	15.312
rk LM statistic	393.909	393.941	390.416	390.404	390.404	379.295
rk Wald F statistic	675.636	675.284	1,376.620	1,376.557	1,376.527	1,388.599
Weather Controls		x	x	x	x	x
HH Fixed Effects			x	x	x	x
Month-of-Year Effects				x	x	
Day-of-Week Effects					x	
Day Effects						x
Hour-of-Day Effects						x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of the Central California wave of the experiment. The sample used to produce the estimates in Panel A is based on *hourly* electricity meter readings in kWh, while the sample underlying the estimates in Panel B is based on *daily* natural gas meter readings (thm). Thus, the hour-of-day effects noted in Column (6) are included in the electricity model (Panel A) only. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 8: Both Experiments-ATT Estimates of the Effect of a Smart Thermostat on Energy Use

	(1)	(2)	(3)	(4)	(5)	(6)
	Power Use (kWh or thm)					
Panel A: Electricity (kWh)						
ATT ($\hat{\gamma}^{kWh}$)	-0.031	-0.031	-0.003	-0.001	-0.001	0.026
	(0.036)	(0.035)	(0.022)	(0.022)	(0.022)	(0.017)
Constant	1.293***					
	(0.024)					
N	1,379	1,379	1,379	1,379	1,379	1,379
$N \times T$	16,421,734	16,421,734	16,421,734	16,421,734	16,421,734	16,421,734
F statistic	67.704	538.083	818.852	749.195	743.903	547.479
rk LM statistic	738.263	749.372	611.958	612.274	612.275	488.960
rk Wald F statistic	790.294	819.435	1,948.381	1,951.624	1,951.629	1,931.185
Panel B: Natural Gas (thm)						
ATT ($\hat{\gamma}^{thm}$)	0.062	0.065	0.028	0.023	0.023	0.055**
	(0.060)	(0.049)	(0.028)	(0.026)	(0.026)	(0.022)
Constant	0.963***					
	(0.028)					
N	1,369	1,369	1,369	1,369	1,369	1,369
$N \times T$	677,304	677,304	677,304	677,304	677,304	677,304
F statistic	126.946	685.010	910.597	687.556	686.021	87.637
rk LM statistic	733.785	744.065	618.764	619.162	619.163	497.269
rk Wald F statistic	790.386	817.152	1,976.210	1,980.104	1,980.097	1,958.933
Wave Indicator		x				
Weather Controls		x	x	x	x	x
HH Fixed Effects			x	x	x	x
Month-of-Year Effects				x	x	
Day-of-Week Effects					x	
Day Effects						x
Hour-of-Day Effects						x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of both waves of the experiment. Note that the estimates reported in Column (2) are based on a model that includes an indicator for the first wave of experiment (N. CA). This indicator is perfectly co-linear with household fixed effects, so it is dropped from subsequent models. The sample used to produce the estimates in Panel A is based on *hourly* electricity meter readings in kWh, while the sample underlying the estimates in Panel B is based on *daily* natural gas meter readings (thm). Thus, the hour-of-day effects noted in Column (6) are included in the electricity model (Panel A) only. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

H Heterogeneous Treatment Effects Estimates

H.1 Ambient Weather Estimates

Table 9: ATT Estimates of the Effect of a Smart Thermostat on Energy Use by Ambient Temperature Quintile

	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
	Power Use (kWh or thm)				
Panel A: Electricity (kWh)					
ATT ($\hat{\gamma}^{kWh}$)	-0.036	-0.033*	-0.024	-0.008	0.009
	(0.022)	(0.019)	(0.019)	(0.024)	(0.044)
N	1,376	1,379	1,379	1,379	1,378
$N \times T$	3,345,085	3,541,064	3,239,489	3,102,224	3,193,872
F statistic	1.522	1.610	14.502	16.745	24.597
rk LM statistic	368.164	652.296	681.468	600.120	545.434
rk Wald F statistic	1,379.806	1,920.331	1,966.682	1,879.175	1,769.185
Panel B: Natural Gas (thm)					
ATT ($\hat{\gamma}^{thm}$)	-0.054	-0.013	0.005	-0.008	0.010
	(0.064)	(0.038)	(0.023)	(0.018)	(0.015)
N	1,364	1,366	1,369	1,368	1,365
$N \times T$	145,525	147,440	120,087	138,512	125,737
F statistic	22.958	0.550	6.339	6.145	0.431
rk LM statistic	360.657	435.244	563.227	699.424	403.356
rk Wald F statistic	1,375.353	1,587.271	1,323.568	1,802.507	1,377.126
HH Fixed Effects	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x
Day-of-Week Effects	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce the estimates in Panel A is based on *hourly* electricity meter readings in kWh, and temperature quintiles are calculated from the distribution of *hourly* average ambient temperature readings. The sample underlying the estimates in Panel B is based on *daily* natural gas meter readings (thm), and temperature quintiles are calculated using the distribution of *daily* average ambient temperature readings. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 10: ATT Estimates of the Effect of a Smart Thermostat on Energy Use by Ambient Humidity Quintile

	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
	Power Use (kWh or thm)				
Panel A: Electricity (kWh)					
ATT ($\hat{\gamma}^{kWh}$)	0.050	-0.010	-0.021	-0.041**	-0.066***
	(0.048)	(0.024)	(0.019)	(0.018)	(0.020)
N	1,379	1,379	1,379	1,379	1,379
$N \times T$	3,313,684	3,333,963	3,255,920	3,239,969	3,278,198
F statistic	45.607	3.514	8.612	4.219	7.804
rk LM statistic	521.960	564.647	595.843	638.333	623.192
rk Wald F statistic	1,763.238	1,860.182	1,910.165	1,944.091	1,612.296
Panel B: Natural Gas (thm)					
ATT ($\hat{\gamma}^{thm}$)	0.004	-0.010	-0.005	0.047	-0.022
	(0.017)	(0.025)	(0.036)	(0.044)	(0.067)
N	1,367	1,369	1,369	1,369	1,367
$N \times T$	141,016	133,650	132,648	153,013	116,975
F statistic	0.930	0.188	0.149	14.963	65.458
rk LM statistic	380.444	564.518	647.032	611.390	550.812
rk Wald F statistic	1,356.189	1,740.682	1,908.480	1,522.235	1,306.659
HH Fixed Effects	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x
Day-of-Week Effects	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce the estimates in Panel A is based on *hourly* electricity meter readings in kWh, and humidity quintiles are calculated from the distribution of *hourly* average ambient relative humidity readings. The sample underlying the estimates in Panel B is based on *daily* natural gas meter readings (thm), and humidity quintiles are calculated using the distribution of *daily* average ambient relative humidity readings. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 11: ATT Estimates of the Effect of a Smart Thermostat on Energy Use by Ambient Heat Index Quintile

	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
	Power Use (kWh or thm)				
Panel A: Electricity (kWh)					
ATT ($\hat{\gamma}^{kWh}$)	-0.036	-0.030	-0.026	-0.009	0.009
	(0.022)	(0.019)	(0.019)	(0.024)	(0.043)
N	1,376	1,379	1,379	1,379	1,378
$N \times T$	3,296,464	3,272,861	3,296,156	3,273,130	3,283,123
F statistic	1.491	1.632	13.865	17.538	24.681
rk LM statistic	367.624	636.517	691.267	604.526	546.840
rk Wald F statistic	1,381.488	1,927.034	1,955.091	1,883.345	1,770.575
Panel B: Natural Gas (thm)					
ATT ($\hat{\gamma}^{thm}$)	-0.060	-0.004	-0.004	-0.003	0.009
	(0.066)	(0.044)	(0.024)	(0.018)	(0.015)
N	1,364	1,366	1,369	1,367	1,365
$N \times T$	135,502	136,401	134,876	135,317	135,204
F statistic	18.708	6.519	10.808	12.692	0.289
rk LM statistic	351.296	404.357	586.160	702.841	413.818
rk Wald F statistic	1,364.503	1,468.623	1,403.564	1,797.169	1,406.956
HH Fixed Effects	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x
Day-of-Week Effects	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce the estimates in Panel A is based on *hourly* electricity meter readings in kWh, and heat index quintiles are calculated from the distribution of *hourly* average ambient heat index readings. The heat index is calculated using temperature and humidity readings. See https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml for the exact formula. The sample underlying the estimates in Panel B is based on *daily* natural gas meter readings (thm), and heat index quintiles are calculated using the distribution of *daily* average ambient heat index readings. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 12: ATT Estimates of the Effect of a Smart Thermostat on Energy Use by Day of the Week and Weekday/Weekend

	(1) Sunday	(2) Monday	(3) Tuesday	(4) Wednesday	(5) Thursday	(6) Friday	(7) Saturday	(8) Weekday	(9) Weekend
	Power Use (kWh or thm)								
Panel A: Electricity (kWh)									
ATT ($\hat{\gamma}^{kWh}$)	-0.006	-0.014	-0.007	-0.001	0.005	0.005	0.010	-0.002	0.002
	(0.024)	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)	(0.023)	(0.022)	(0.023)
N	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379
$N \times T$	2,338,599	2,331,710	2,331,777	2,331,619	2,362,409	2,362,700	2,362,920	11,720,215	4,701,519
F statistic	682.877	596.313	582.663	552.920	692.183	555.214	688.254	694.402	732.848
rk LM statistic	610.468	605.939	604.236	600.642	624.896	620.751	616.090	611.775	613.368
rk Wald F statistic	1,946.894	1,941.134	1,936.473	1,933.095	1,972.928	1,966.289	1,952.711	1,951.954	1,950.200
Panel B: Natural Gas (thm)									
ATT ($\hat{\gamma}^{hm}$)	0.024	0.015	0.026	0.024	0.018	0.017	0.033	0.022	0.028
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.030)	(0.031)	(0.026)	(0.027)
N	1,369	1,369	1,369	1,369	1,369	1,369	1,369	1,369	1,369
$N \times T$	94,575	96,480	96,480	96,474	97,760	97,764	97,771	484,958	192,346
F statistic	646.968	296.353	634.809	519.942	661.750	588.390	592.766	616.326	737.276
rk LM statistic	622.274	611.094	609.913	606.781	630.704	626.904	623.225	617.610	622.893
rk Wald F statistic	1,981.320	1,965.987	1,959.102	1,960.855	1,998.501	1,993.533	1,984.444	1,978.155	1,983.795
Weather Controls	x	x	x	x	x	x	x	x	x
HH Fixed Effects	x	x	x	x	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce the estimates in Panel A is based on *hourly* electricity meter readings in kWh, while the sample underlying the estimates in Panel B is based on *daily* natural gas meter readings (thm). Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 13: ATT Estimates of the Effect of a Smart Thermostat on Electricity Use by Hour of the Day

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	12:00	1:00	2:00	3:00	4:00	5:00	6:00	7:00	8:00	9:00	10:00	11:00
	Power Use (kWh)											
Panel A: AM												
ATT ($\hat{\gamma}^{kWh}$)	-0.022	-0.012	-0.020	-0.030	-0.015	0.009	0.003	-0.003	0.005	-0.029	-0.041	-0.042
	(0.027)	(0.023)	(0.021)	(0.021)	(0.021)	(0.023)	(0.024)	(0.027)	(0.030)	(0.036)	(0.039)	(0.042)
N	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379
$N \times T$	684,283	684,283	682,930	684,283	684,283	684,283	684,283	684,283	684,284	684,290	684,291	684,295
F statistic	265.776	246.431	217.654	202.070	184.084	174.976	155.358	175.907	179.396	206.676	254.733	305.851
rk LM statistic	614.284	614.247	614.352	614.225	614.218	614.175	614.184	614.190	614.150	614.089	613.938	613.624
rk Wald F statistic	1,956.322	1,956.072	1,956.121	1,956.061	1,956.057	1,955.827	1,955.958	1,955.923	1,955.732	1,955.709	1,955.444	1,955.012
Panel B: PM												
ATT ($\hat{\gamma}^{kWh}$)	-0.028	-0.005	0.015	0.021	0.046	0.076*	0.048	0.035	-0.004	-0.032	-0.022	-0.019
	(0.045)	(0.047)	(0.048)	(0.047)	(0.045)	(0.042)	(0.039)	(0.036)	(0.034)	(0.032)	(0.031)	(0.027)
N	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379
$N \times T$	684,296	684,299	684,301	684,304	684,307	684,308	684,308	684,308	684,308	684,308	684,308	684,308
F statistic	383.020	453.596	527.898	615.775	665.775	701.382	748.405	680.792	617.310	548.500	409.111	308.266
rk LM statistic	612.714	611.982	611.318	610.725	610.204	609.946	609.792	609.666	609.584	609.590	609.614	609.611
rk Wald F statistic	1,952.825	1,950.654	1,949.534	1,948.465	1,947.225	1,946.455	1,946.583	1,946.181	1,946.106	1,946.458	1,947.013	1,947.114
Weather Controls	x	x	x	x	x	x	x	x	x	x	x	x
HH Fixed Effects	x	x	x	x	x	x	x	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x	x	x	x	x	x	x	x
Day-of-Week Effects	x	x	x	x	x	x	x	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce the estimates is based on *hourly* electricity meter readings in kWh. Panel A reports estimates from models that are conditional on the given hour in the AM. Panel B reports analogous estimates based on the given hour in the PM. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 14: ATT Estimates of the Effect of a Smart Thermostat on Energy Use by Hour of the Day and Weekday/Weekend

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	12:00	1:00	2:00	3:00	4:00	5:00	6:00	7:00	8:00	9:00	10:00	11:00
						Power Use (kWh)						
Panel A: Weekdays Only												
AM: ATT ($\hat{\gamma}^{kWh}$)	-0.026	-0.015	-0.025	-0.036*	-0.017	0.011	0.000	-0.010	0.003	-0.032	-0.039	-0.042
	(0.028)	(0.024)	(0.022)	(0.021)	(0.021)	(0.023)	(0.025)	(0.027)	(0.030)	(0.036)	(0.040)	(0.042)
PM: ATT ($\hat{\gamma}^{kWh}$)	-0.027	-0.009	0.010	0.016	0.051	0.085*	0.053	0.036	-0.007	-0.035	-0.022	-0.027
	(0.045)	(0.048)	(0.049)	(0.050)	(0.047)	(0.044)	(0.041)	(0.038)	(0.035)	(0.033)	(0.032)	(0.028)
N	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379
$N \times T$	~488,000	~488,000	~488,000	~488,000	~488,000	~488,000	~488,000	~488,000	~488,000	~488,000	~488,000	~488,000
Panel B: Weekend Days Only												
AM: ATT ($\hat{\gamma}^{kWh}$)	-0.010	-0.006	-0.007	-0.017	-0.009	0.004	0.011	0.013	0.011	-0.022	-0.047	-0.041
	(0.028)	(0.024)	(0.022)	(0.022)	(0.021)	(0.023)	(0.024)	(0.027)	(0.032)	(0.038)	(0.042)	(0.044)
PM: ATT ($\hat{\gamma}^{kWh}$)	-0.032	0.005	0.027	0.031	0.034	0.053	0.037	0.032	0.002	-0.026	-0.023	0.000
	(0.048)	(0.050)	(0.050)	(0.048)	(0.045)	(0.043)	(0.040)	(0.037)	(0.036)	(0.034)	(0.031)	(0.029)
N	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379	1,379
$N \times T$	~196,000	~196,000	~196,000	~196,000	~196,000	~196,000	~196,000	~196,000	~196,000	~196,000	~196,000	~196,000
Weather Controls	x	x	x	x	x	x	x	x	x	x	x	x
HH Fixed Effects	x	x	x	x	x	x	x	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x	x	x	x	x	x	x	x
Day-of-Week Effects	x	x	x	x	x	x	x	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce all estimates is based on *hourly* electricity meter readings in kWh. All panels report estimates of the coefficient of interest from separately estimated models that are conditional on readings corresponding to the given hour in the AM or PM. The results in Panel A are based on a sample comprised of weekdays only. Panel B reports analogous estimates based on weekend days only. Across all specifications, the min(rk LM statistic)=609.463 and the min(rk Wald F statistic)=1,942.447. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications. Full regression diagnostics are available from the authors by request.

H.3 Price and Peak-Alert Estimates

Table 15: ATT Estimates of the Effect of a Smart Thermostat on Energy Use by Price Quintile

	(1)	(2)	(3)	(4)	(5)
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Power Use (kWh)				
ATT ($\hat{\gamma}^{kwh}$)	0.027	-0.002	0.003	-0.007	-0.012
	(0.027)	(0.020)	(0.021)	(0.024)	(0.026)
N	1,379	1,379	1,379	1,379	1,379
$N \times T$	3,274,085	3,272,108	3,273,626	3,272,248	3,272,504
R^2	0.133	0.114	0.096	0.088	0.081
F statistic	692.939	784.327	788.042	719.974	630.113
rk LM statistic	713.547	557.232	514.273	466.017	585.746
rk Wald F statistic	1,827.171	1,849.146	1,798.582	1,651.200	1,938.912
Weather Controls	x	x	x	x	x
HH Fixed Effects	x	x	x	x	x
Month-of-Year Effects	x	x	x	x	x
Day-of-Week Effects	x	x	x	x	x

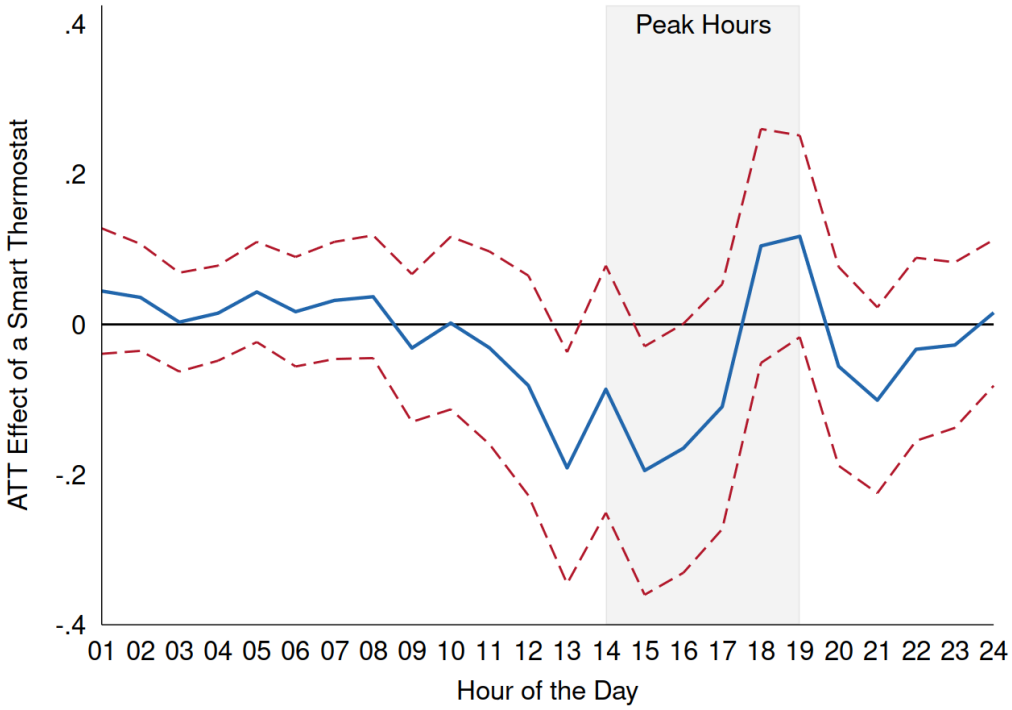
Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce the estimates is based on hourly electricity meter readings in kWh. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Table 16: ATT Estimates of the Effect of a Smart Thermostat on Energy Use on Peak Alert Days

	(1) CAISO Alert Hours	(2) CAISO Alert Hours w/ Local Media Coverage	(3) PG&E Alert Days	(4) PG&E Alert Days: Off-Peak Hours	(5) PG&E Alert Days: Peak Hours
			Power Use (kWh)		
ATT ($\hat{\gamma}^{kwh}$)	-0.046	-0.011	-0.103***	-0.073**	-0.116*
	(0.052)	(0.069)	(0.033)	(0.029)	(0.065)
N	1,378	1,376	1,378	1,378	1,378
$N \times T$	227,005	158,614	475,920	356,940	118,980
R^2	0.015	0.019	0.015	0.012	0.016
F statistic	126.639	117.259	236.227	204.697	135.732
rk LM statistic	595.465	630.654	525.434	529.324	531.045
rk Wald F statistic	1,580.882	1,585.741	1,777.016	1,775.353	1,785.294
Weather Controls	x	x	x	x	x
HH Fixed Effects	x	x	x	x	x
Day Effects	x	x	x	x	x
Hour-of-Day Effects	x	x	x	x	x

Note: Standard errors clustered at the household level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. All estimates are based on a sample comprised of both waves of the experiment. The sample used to produce the estimates is based on hourly electricity meter readings in kWh. Based on the values of the rk LM and Wald F statistics, we reject the nulls of an under or weakly identified model across all specifications.

Figure 21: ATT Estimates of the Effect of a Smart Thermostat on Energy Use on PG&E Peak Alert Days by Hour of the Day



I Recruitment and Enrollment

I.1 Subject Eligibility

Appendix Table 17 summarizes the eligibility requirements for participation in the experiment. Participants had to own their residence and have central air conditioning with a single thermostat. They also had to have a smart phone and high-speed Internet. Finally, individuals who were planning to move in the near future were excluded from the experiment.

Table 17: Subject Eligibility Summary

	Eligible	Not Eligible
Rent or own?	Own	Rent
Home Type	House or Condo	Apartment or Other
Phone	iPhone or Android	Blackberry or Other
# of Thermostats	1	≥ 2
A/C	Central Air	Box Unit, Fans, Other
Heating	Air Vents	Baseboard or Other
High-speed Internet?	Yes	No
Plan to move in next year?	No	Yes

I.2 Trial Recruitment and Enrollment Guide

FOR ONLINE PUBLICATION: ONLINE APPENDIX B

UTILITY Smart Thermostat Trial
Recruitment and Enrollment Guide

November 19, 2012

Introduction

Experimental Trial Information

UTILITY is running an experimental thermostat trial with Opower and Honeywell, offering eligible customers a free remote-controlled thermostat solution (a thermostat controlled by a smartphone and web application). The goal of the experiment is to test the energy savings and customer experience of the thermostat solution. Customers gain a thermostat and app that helps them save energy, by creating a customized, energy efficient schedule that fits their lifestyle.

For this trial, 1 in 2 qualifying customers will receive the thermostat solution. Customers who meet the eligibility qualifications must complete the online enrollment process to determine if they will receive a thermostat or not. At the end of the online enrollment process the system will randomly flip a coin to determine which customer will receive the remote-controlled thermostat and which will not. All customers who enroll for a chance to participate are benefiting the trial (even those who do not receive a thermostat), and it is important that all qualified customers complete the full enrollment process.

Customers should be encouraged to enroll for a chance to receive this exciting solution, which allows them to control their thermostat on-the-go. UTILITY, Opower, and Honeywell are grateful for the time each customer takes to enroll online for a chance to participate, and all customers should be thanked for their time regardless of the outcome.

Customers should be encouraged to answer all qualification and enrollment questions honestly. If a customer provides inaccurate information during enrollment it negatively impacts the trial and the customer will ultimately be turned down for the trial.

Talking Points for Recruitment Events

Initial Communication

Initial communication should be a call to action, provide quick benefits (FREE remote-controlled thermostat), provide a fun atmosphere and garner attention.

- Do you own an iPhone or an Android? If so, would you be interested in a free thermostat controlled by your smartphone?
- How would you like to gain better control of your energy use at home? You can control your thermostat at home from right here! Want to know how?
- Sign-up for a free remote-controlled thermostat, a $500 dollar value and take control of your energy consumption and improve the comfort of your home.
- I know you're in a hurry but this opportunity will allow you to take control of your energy use and you'll always come home to a house at the perfect temperature.
- Save energy while you're away and stay comfortable while you're at home, all by using your smartphone or the web.
- How would you like to control your heating/cooling by your iPhone or Android and

through the internet from anywhere in the world?

After Initial Communication

After initial communication, you should be focused on getting the customer more excited about the offering by providing key information and benefits unique to the opportunity.

- We are conducting a trial on behalf of UTILITY that allows you to interact with your heating & cooling system using your smartphone or the web. That means you can control your home's comfort at your fingertips from wherever you are. All you need is your smartphone of the web. Are you ready to take control?

- Did you know that a typical family spends almost half (49%) of its energy cost on heating and cooling? (*Source:* Energy Star)-- How would you like to have the opportunity to be selected for a special trial UTILITY is conducting to provide a limited number of customers a thermostat controlled by your smartphone? That's right you can control the comfort of your home at anytime or any place using your smartphone or the web.

- How would you like to be one of the lucky UTILITY customers who receives a free thermostat controlled on-the-go from your smartphone or the web? This is over a $500 value completely free with professional installation and a 1-year warrantee. UTILITY is conducting this trial to allow customers a unique way to reduce energy use and save money. The process for signing up only takes a few minutes of your time. Let's see if you qualify.

- Check out this free thermostat controlled by your smartphone. You'll have complete control over your comfort, and you can see how your temperature settings stack up against other participants in the trial.

Overcoming Initial Objections

Objection: "I don't have time"

- You'll never come home to a cold house again and sign-up only takes a few minutes.

Objection: "I still don't have time"

- Okay; here's how you can see if you qualify and sign-up from home (postcard)

Objection: "I don't want to give out my personal information"

- You're information is completely confidential and will be only used to determine if you qualify for the free thermostat.

Objection: "I'm not interested"

- Here is a free pen, compliments of UTILITY. Have a great day!

Initial Eligibility Screening

	Eligible	Not eligible

Do you rent or own your home?	Own	Rent
What kind of home do you own?	Single family, Townhome, Condo	- Apartment - Other
What kind of phone do you have?	- iPhone - Android	- Blackberry - Other
How many thermostats do you have in your home?	One (1)	Two (2) or more
How do you cool your home?	Central air	- Window box unit - Fans - Other
What is the main way you heat your home?	Air vents	- Baseboard - Other - None
Are your heating and air conditioning systems functional and have you used them the last 6 months?	Yes	No
Do you have high-speed internet access (Cable, DSL, satellite, Broadband)?	Yes	No
Do you have an available *ethernet* port on your internet router?	Yes	No
Do you plan to move to a new home in the next 12 months?	**No**	Yes
Will other adults in your household object to enrolling in this program?	**No**	Yes

Customer Does NOT Pass Initial Eligibility Screening

- Thank you for your interest, but unfortunately you don't meet the eligibility requirements for this trial. However, UTILITY is developing a number of residential energy efficiency programs that you may qualify for. Please fill out this post card in to enable them to contact you in the future for other offerings. Thank you and please accept this free pen, compliments of UTILITY. We appreciate your time!

- If you do know someone else who may be interested, please let them know about this free trial and they can sign-up right away. (Staffer hands the customer a post card.)

Customer Passes Initial Screening

- Great! You've pre-qualified to participate in the selection process, which only takes a few minutes. Would you like to learn how the thermostat and app works? (demo)

- Let's get you signed-up and see if you are selected to join the UTILITY Smart Thermostat Trial, with a free remote- controlled thermostat and professional installation. The sign-up process just takes a few minutes and we can help you complete it here.

- You'll need your UTILITY account number for enrollment. You can use my phone to retrieve your utility account number from UTILITY. You will also be asked to provide

the last four digits of the Social Security Number of the UTILITY account holder—this may be you or a housemate. Staffer provides customer phone & contact number (1-888-743-0011).

Customer is Selected to Join the Trial

Encourage customers to take the first available appointment. Explain that technicians are only in the area for a limited amount of time.

- Congratulations! You've been selected to participate in the UTILITY Smart Thermostat Trial. A customer service representative will contact you with further information about your free installation. You will receive an email reminder with the date and time of your installation appointment, but you may want to write it down now, so you don't forget.
- Tell your friends and family to see if they are eligible and sign-up online! (postcard)
- Here is a free lens cleaner or smartphone holder for your smartphone, compliments of UTILITY. We appreciate your time!
- You will be contacted within a few days to confirm your eligibility and appointment time. (Honeywell CSR will conduct a follow-up call to confirm appointment time & answer any additional questions)

Customer is NOT Selected for the Trial

Thank you for your interest in the Smart Thermostat Trial. Unfortunately, this is currently a trial so participation cannot be granted for everyone.

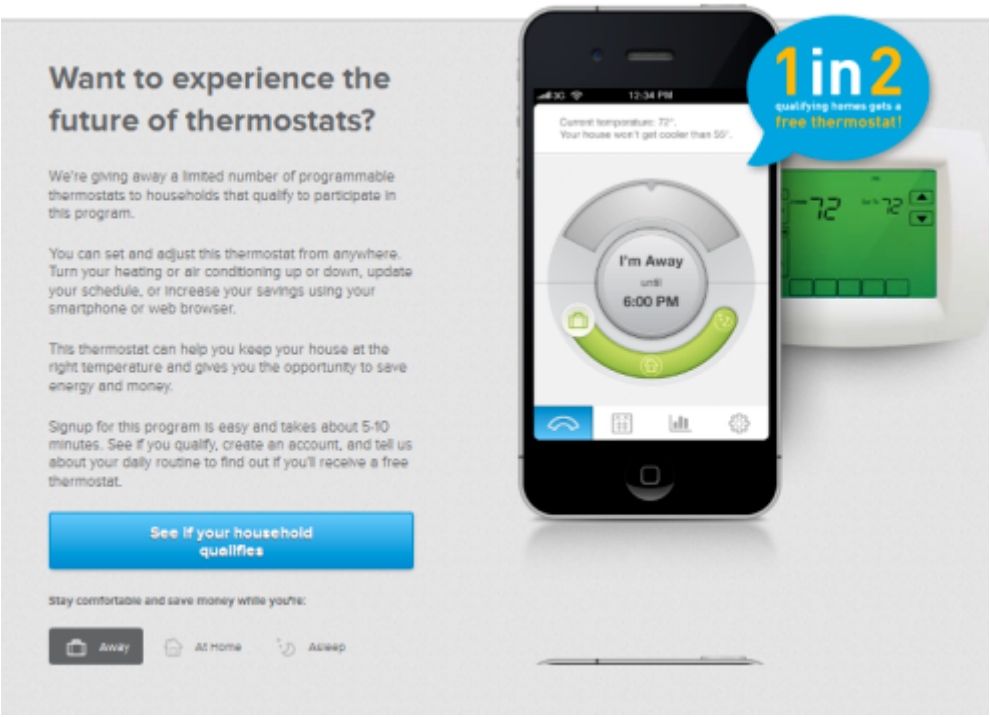
- In the event the trial is extended, would you like to leave your contact information, which will only be used to contact you regarding other opportunities to participate in UTILITY residential trials or programs?
- Please accept this free pen, compliments of UTILITY. Have a great day.
- Tell your friends and family to see if they are eligible and sign-up! (postcard).
- Here is a free lens cleaner or smartphone holder for your smartphone, compliments of UTILITY. We appreciate your time!

How Online Enrollment Works

If a customer passes the initial qualification screening, direct them to the Opower Web application to enroll online. Eligible customers have a 1 in 2 chance of being selected to receive a thermostat.

Enroll online at: https://thermostat.opower.com/

The customer begins by clicking "See if your household qualifies."

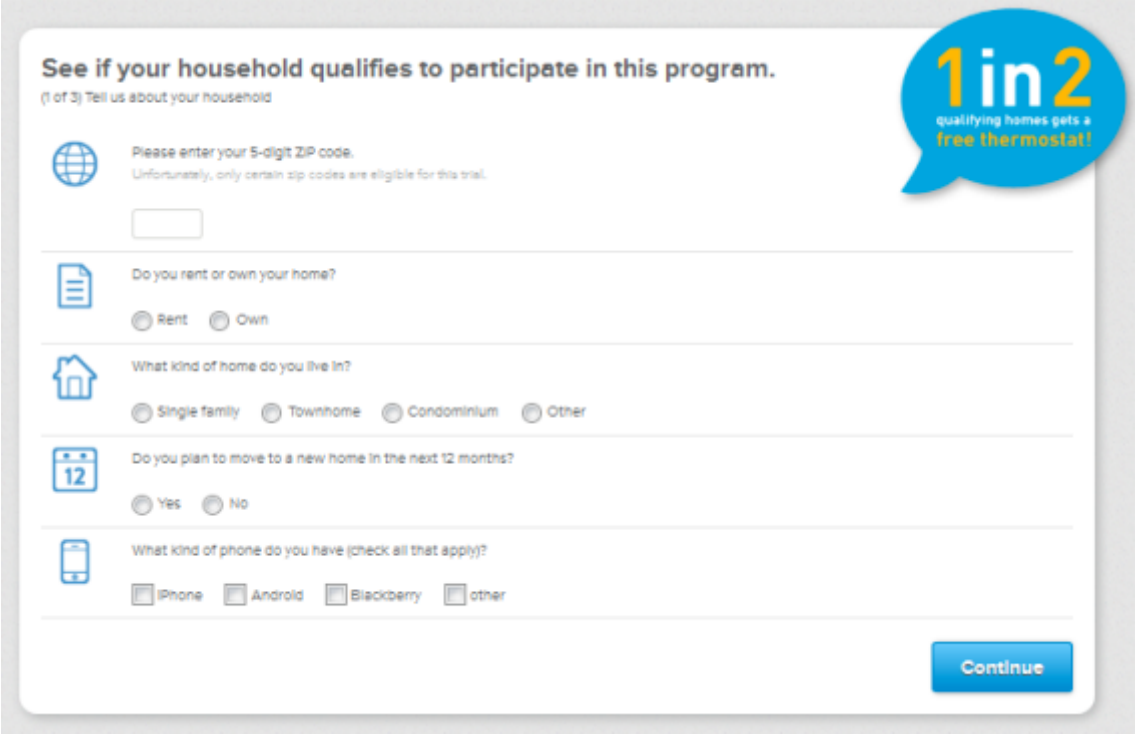


Verifying if the Household Qualifies

In order to verify that they can participate in the program, customers must answer a series of questions about their home.

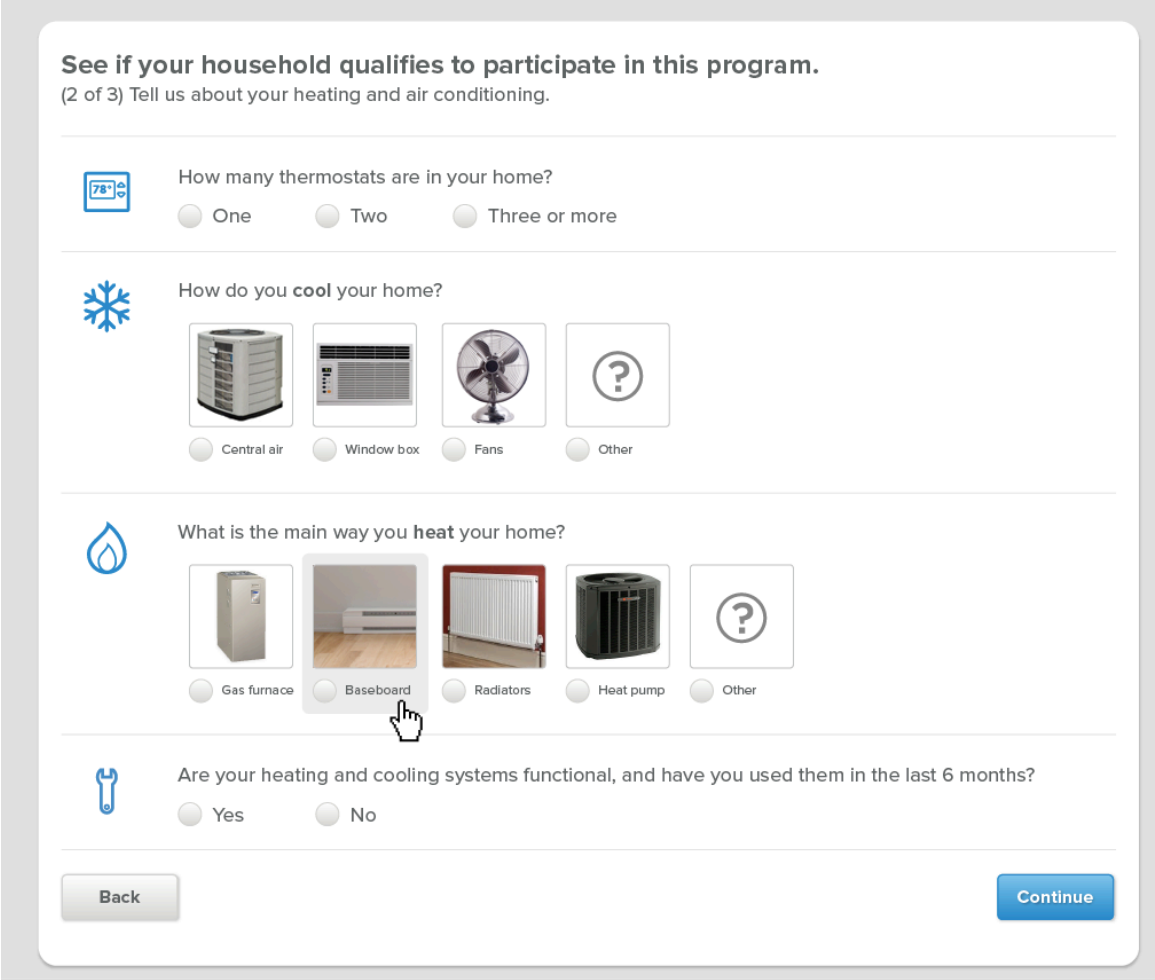
On the first verification screen, they are asked to provide the following information:

- Zip code: Qualified zip codes are those within the greater Fresno and Bakersfield areas, see list provided by Honeywell.
- Whether they rent or own: Customers must own their own home.
- What kind of home they live in: Customers can select any option except "other."
- Whether they plan on moving in the next year: Customers must plan on remaining in the same home.
- What kind of phone they have: Customers must have an iPhone or Android phone if the utility program requires a smartphone.



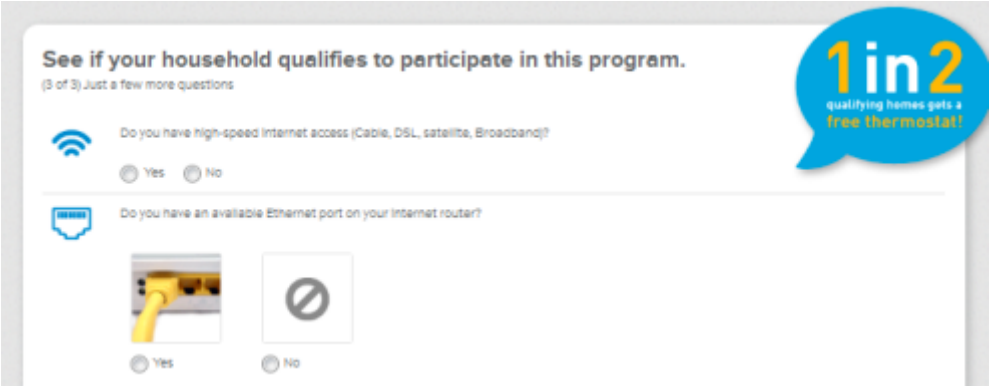
If a customer qualifies based on the answers to the questions above, they are asked to provide the following additional information:

- Number of thermostats: Customers can have only one thermostat.
- Primary cooling system: Customers must have central air.
- Main way they heat their home: Customers must have a gas furnace.
- If their air conditioning and heat are currently working: Customers must have an operational air conditioner and heater that they have used in the last 6 months.



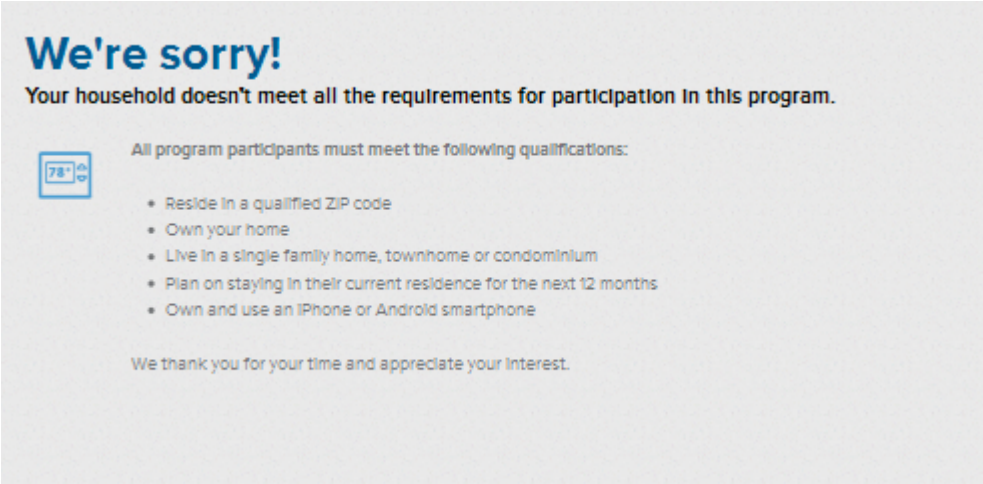
Finally the customer is asked, if they:

- Have high-speed Internet access: Customers must have high-speed access.
- Have an available Ethernet port on their router: Customers must have an available port.
- Are in agreement with the terms and conditions of the program: Customers must agree to the terms. Terms vary by utility.



When they complete the final verification screen, they are told if they are eligible to receive an account. They must meet all of the qualifications to be considered for the program.

If a customer answers any of the qualification questions with a response that makes them ineligible, they are excluded from the program.

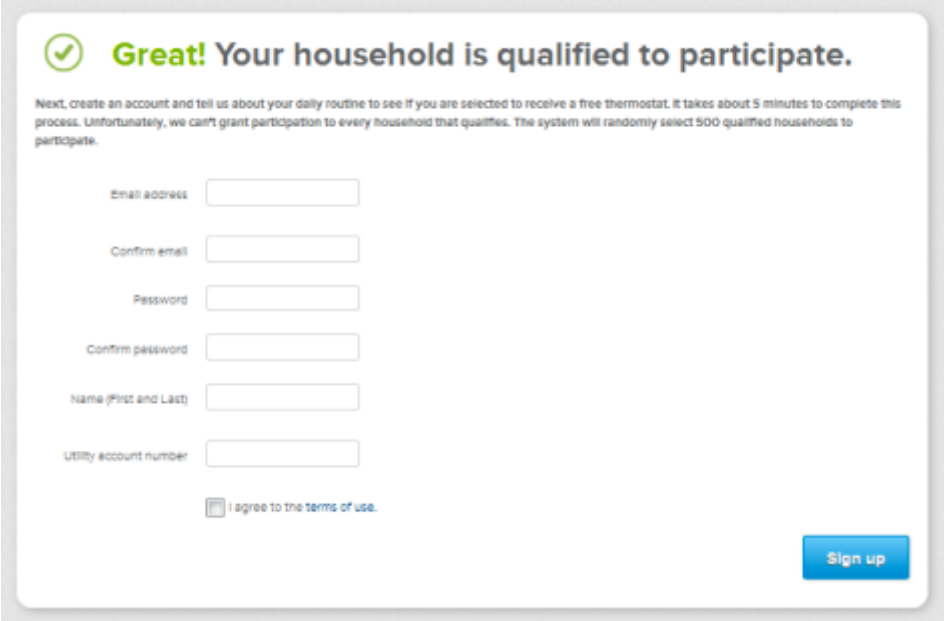


Creating an Account

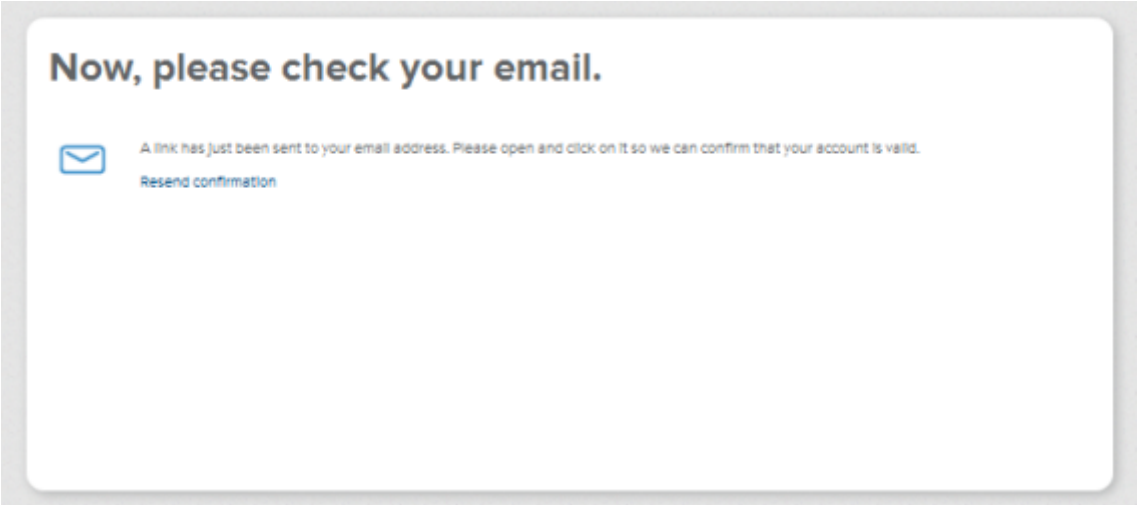
Customers who are eligible for the program are required to enter the following information to create an account:

- The email address they will use to access the Web application. Basic validation is performed to verify that the email address is well-formed.
- A unique password. The password must be at least eight characters long. Passwords must not be or contain the customer's name or email address.
- Customers enter the same password again and are prompted to correct the password if it is not identical in the two password fields.
- The full name of the utility account holder exactly as it appears on the utility bill. The customer enrolling in the program must enter the name of the utility account holder as it appears on the utility bill, even if they are not the account holder.
- The utility account number exactly as it appears on the utility bill. This includes spaces or any other characters included in the data.

Customers are prompted to agree to the Opower Terms of Use.



Customers submit their account information, and then a new page prompts the customer to check their email.

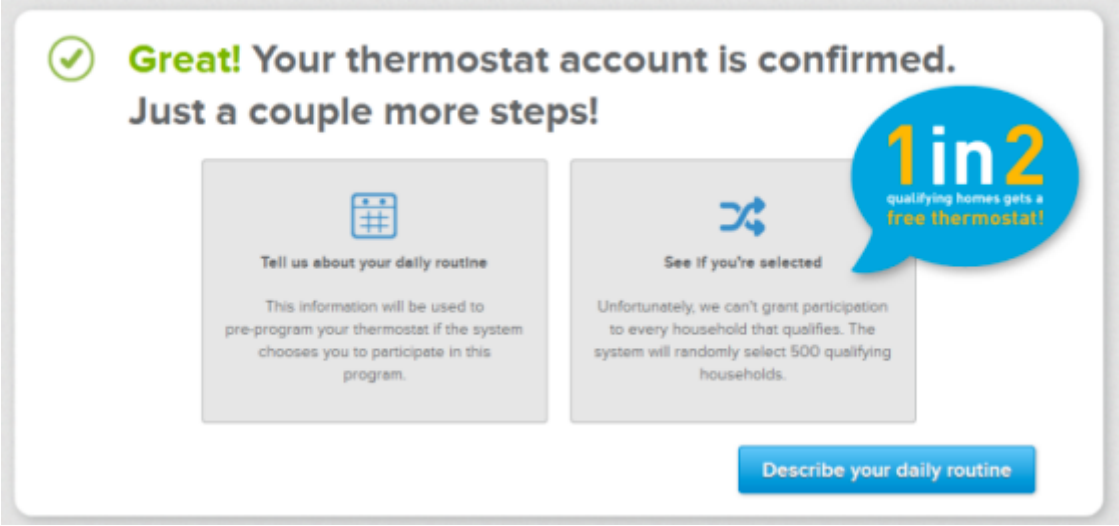


Customers should receive an email message at the address they specified. If the customer does not receive the email, they have the option to "Resend confirmation" in the Web application. The email is titled "Your Thermostat," and it will arrive from an @opower.com email address. The customer may need to check their junk/spam folder for the email.

The customer must click "Confirm my account" to complete their registration and verify their email address. If nothing happens when the button is clicked, the customer can copy and paste the customer-specific URL provided in the email to their Internet browser to confirm the account.

Thermostat Registration

Once the customer has confirmed their account, they are provided with more information about the program and asked to describe their daily routine.



Qualifying Questions

The customer begins to program their thermostat by providing the following information:

- Whether multiple people live in their home. Opower tailors the language in the application to the number of people in the household.

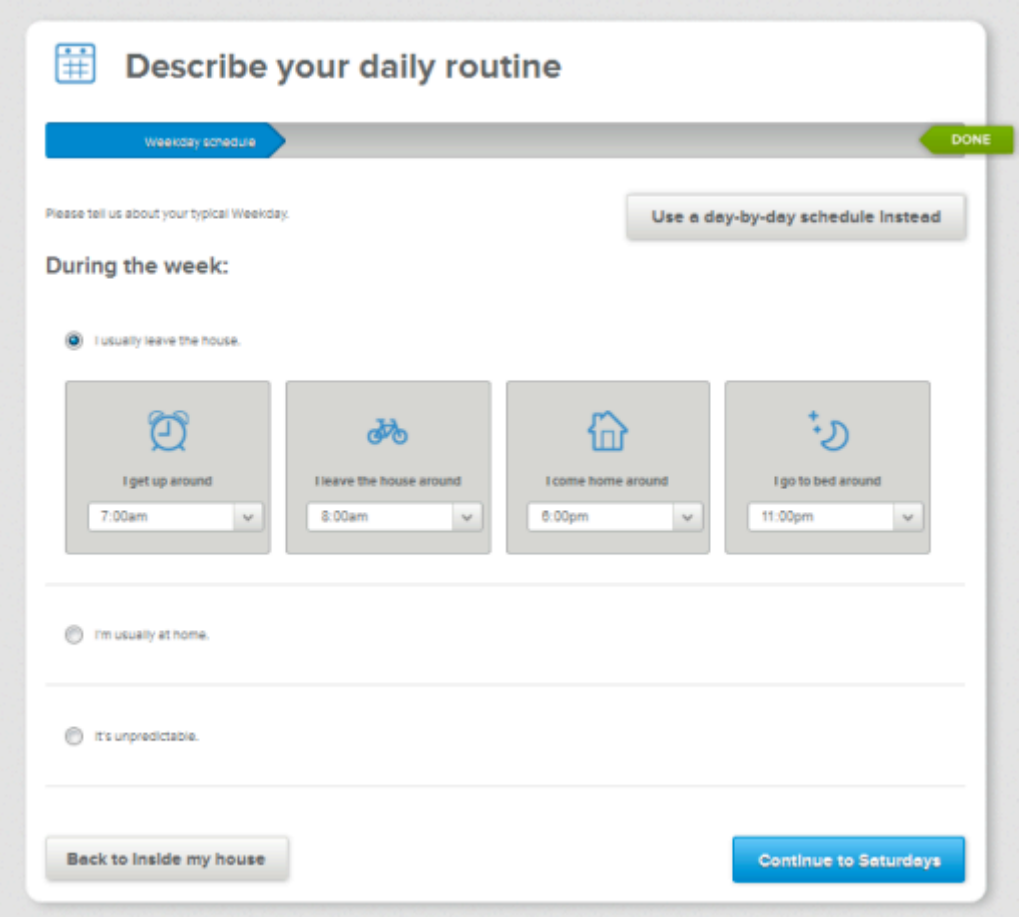
- Whether they have pets. If the customer has pets, the default away temperature of the home is adjusted to a safe temperature for household pets. For homes with pets, the default away temperature is 82 instead of 85 for cooling and 60 instead of 55 for heating.
- Their mobile phone number. Customers are sent a text message to this number with a link to the Opower mobile application..

Setting an Initial Schedule

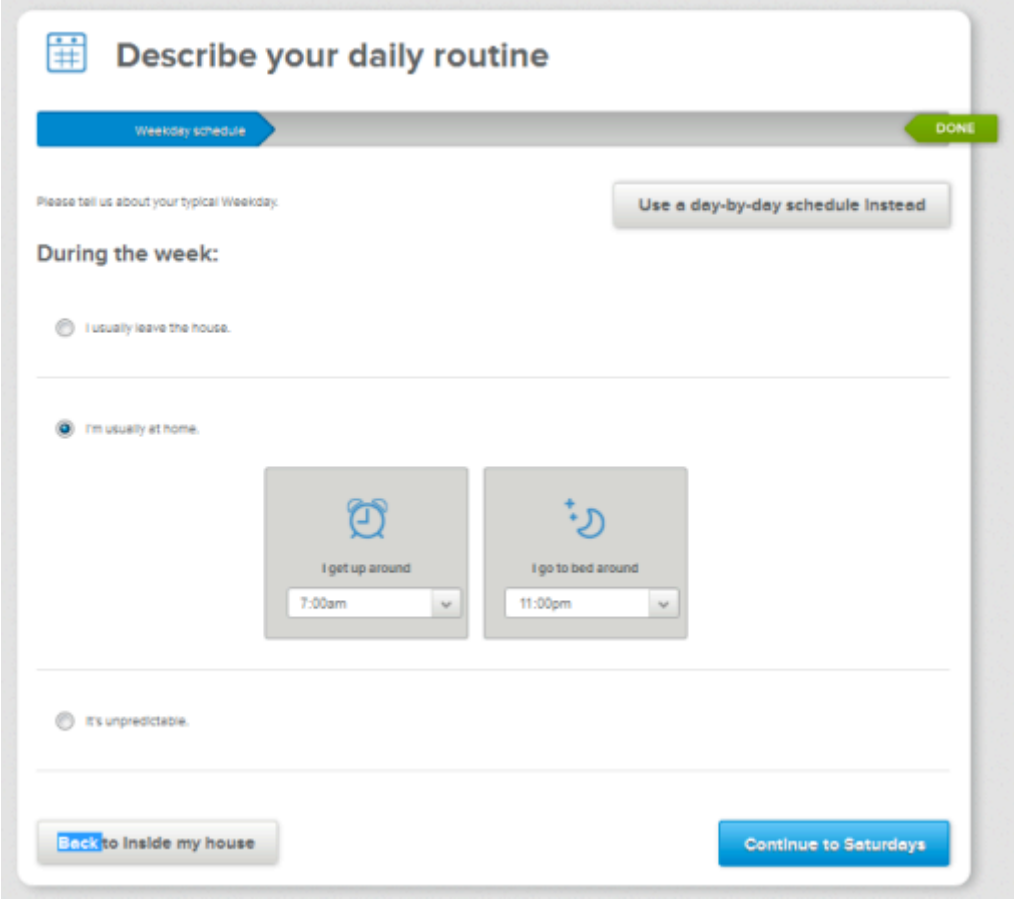
After completing the qualification questions, the customer is prompted to create a personalized schedule. By default, customers set a schedule for all weekdays and then Saturday and Sunday.

For all weekdays, Saturday, and Sunday, the customer has the following options:

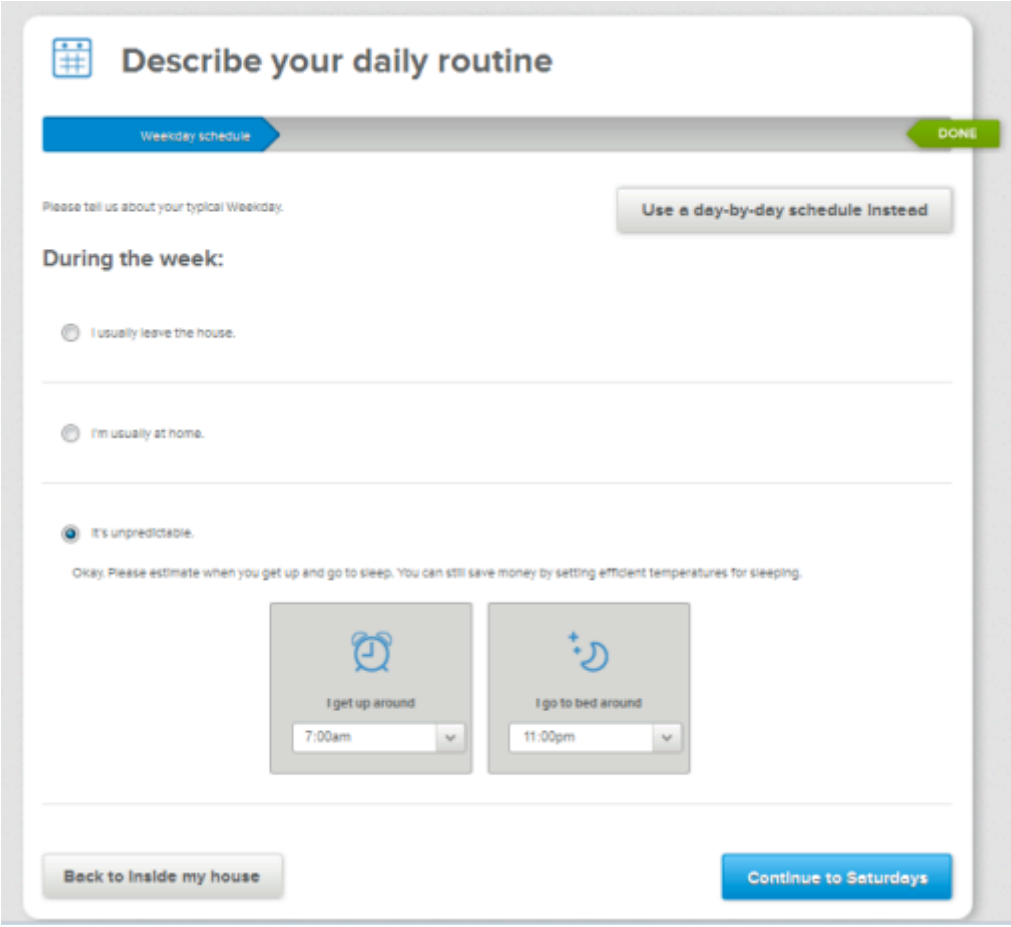
- They can set a schedule for when they typically wake, leave the home, return home, and go to sleep.



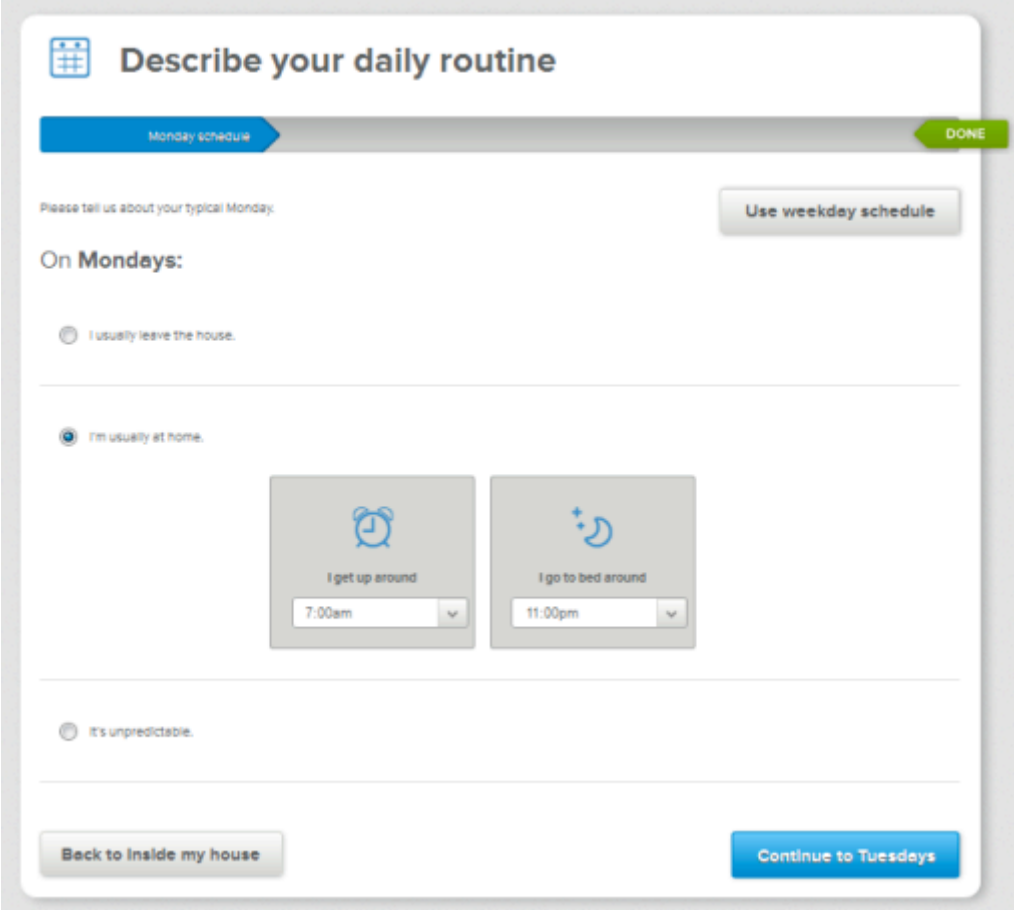
- They can indicate they are home all day and set the time for when they usually wake and go to sleep.



- They can indicate their schedule is unpredictable. In this case, they are still asked when they typically wake and go to sleep.



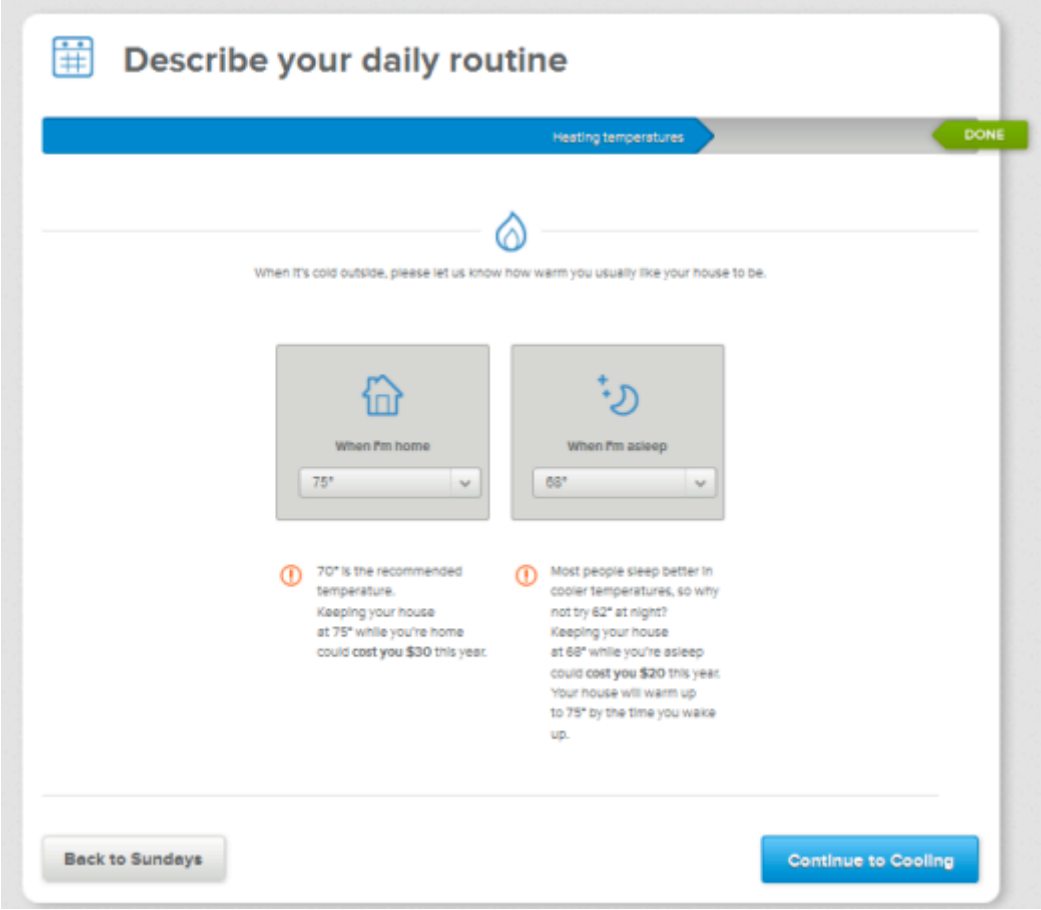
Instead of setting the same schedule for all weekdays, a customer can also create a day-by-day schedule for each weekday separately. The same schedule options are available on a daily basis.



Setting Initial Temperatures

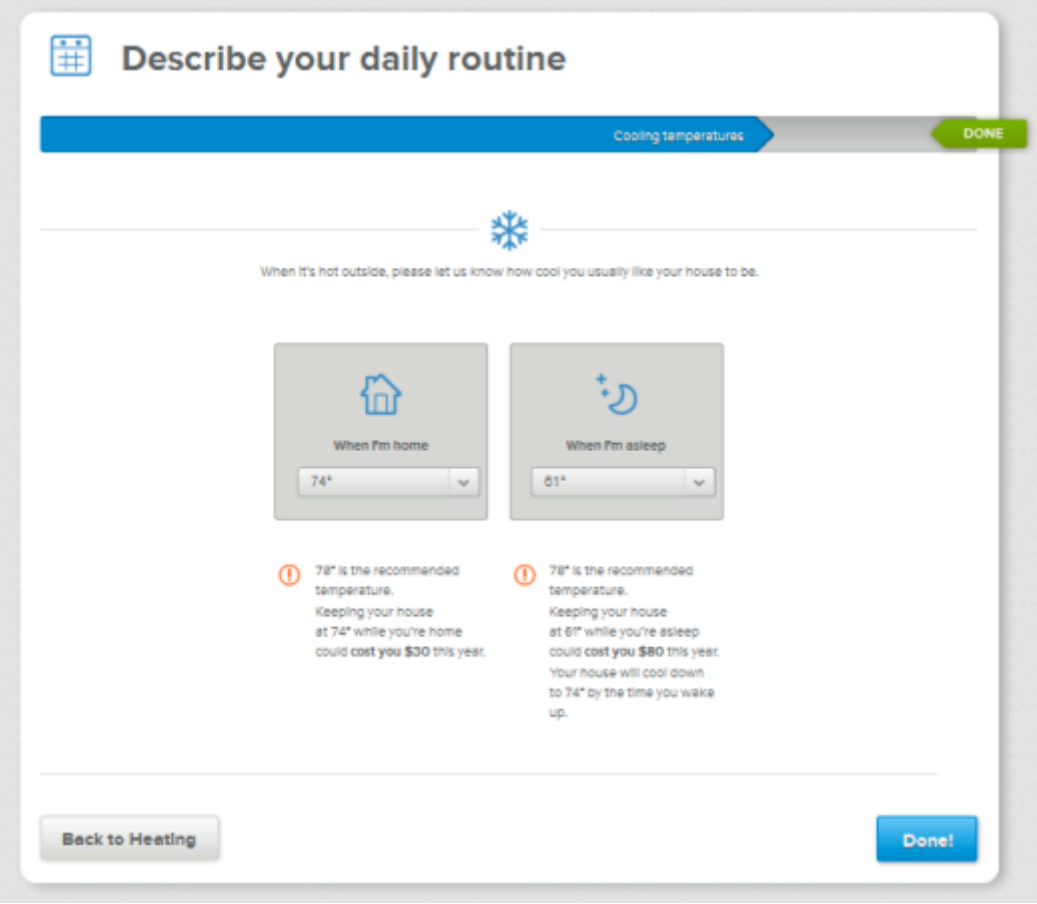
Customers are prompted to set their home and sleep temperatures for heating and cooling. The default temperatures for these settings are based on the suggested Energy Star settings (ENERGY STAR® Program Requirements for Residential Climate Controls, Version 1.0 Partner Commitments, DRAFT 2).

On the heating page, customers are asked how warm they would like their home to be when they are home and asleep.



If the home temperature is greater than the recommended setting (less efficient), an insight appears to tell them how much money they will spend during the winter keeping the home at this higher temperature. If the away temperature is higher than the recommended setting, they are prompted to try setting the temperature lower since the house will warm up to a comfortable setting before they wake up.

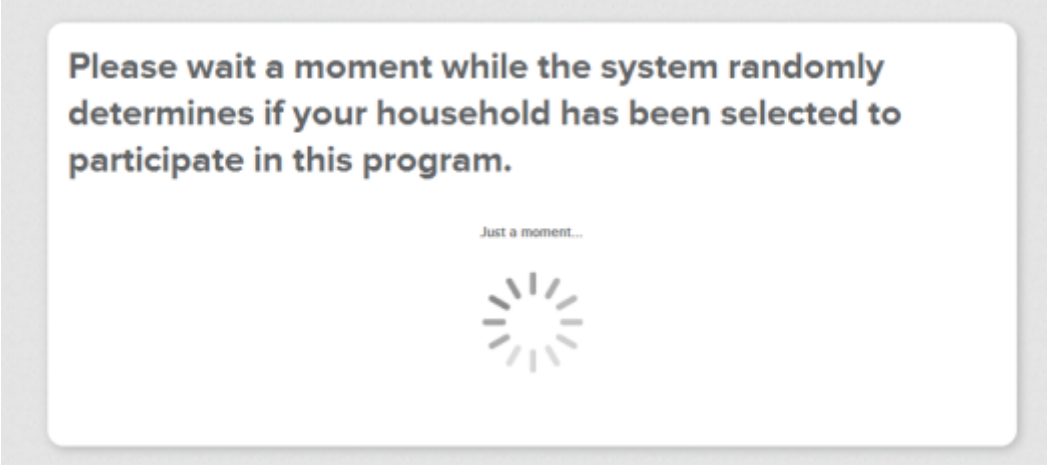
On the cooling page, customers are asked how cool they would like their home to be when they are home and asleep.



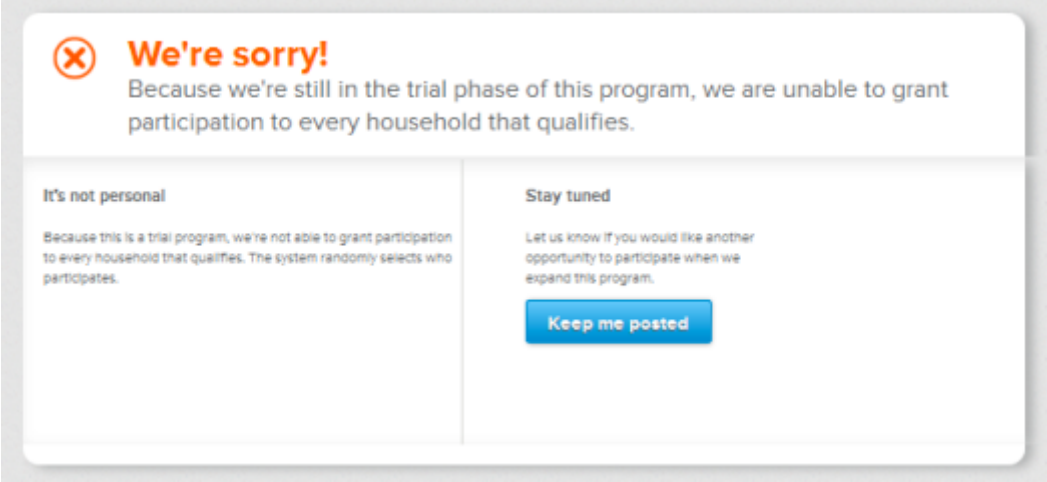
If the home temperature is less than the recommended setting (less efficient), an insight appears to tell them how much money they'll spend during the summer keeping the home at this lower temperature. If the away temperature is lower than the recommended setting, they are prompted to try setting the temperature higher since the house will cool down to a comfortable setting before they wake up.

Installation

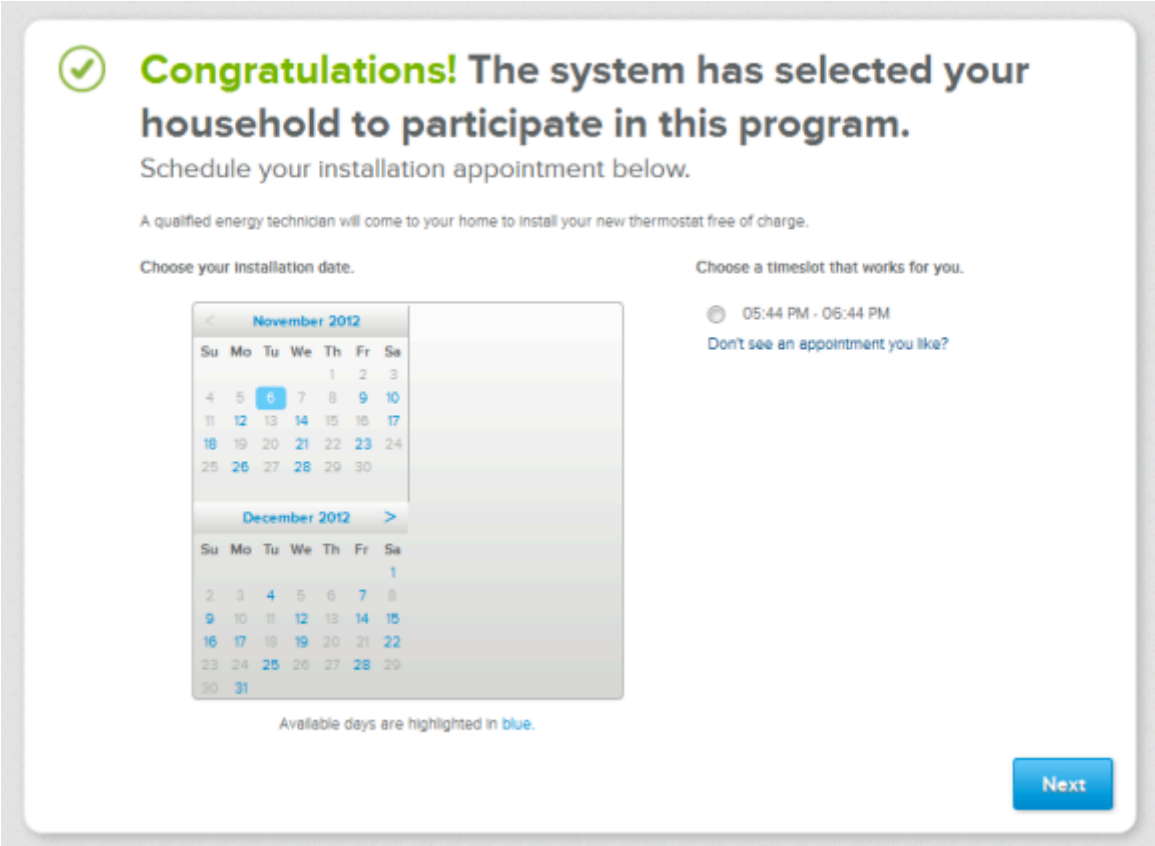
After submitting their temperature settings, the customer is randomly selected to be part of the test or control group.



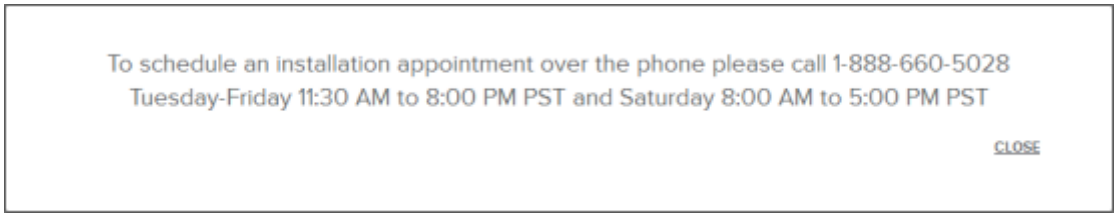
If they are part of the control group, they will not receive a thermostat. Customers in the control group may opt to sign up for a waiting list and may receive a thermostat if the program is expanded.



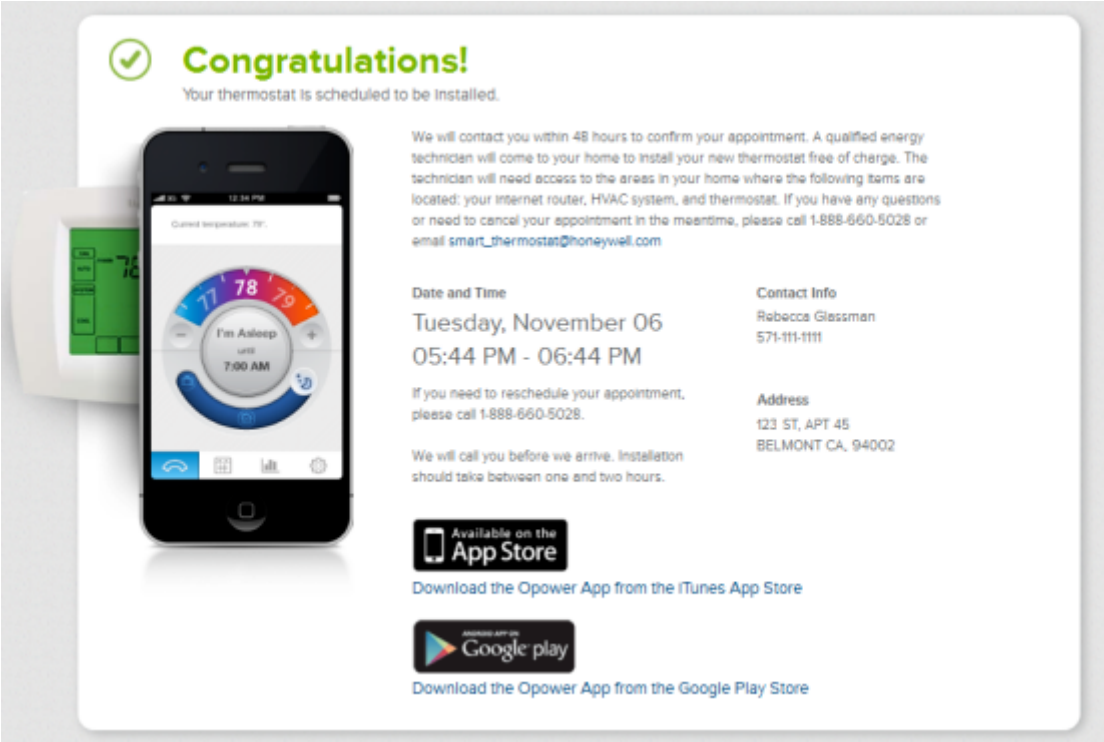
If they are randomly selected into the test group, they will receive a thermostat and become part of the program. Customers participating in the test group can schedule an appointment to have their thermostat installed.



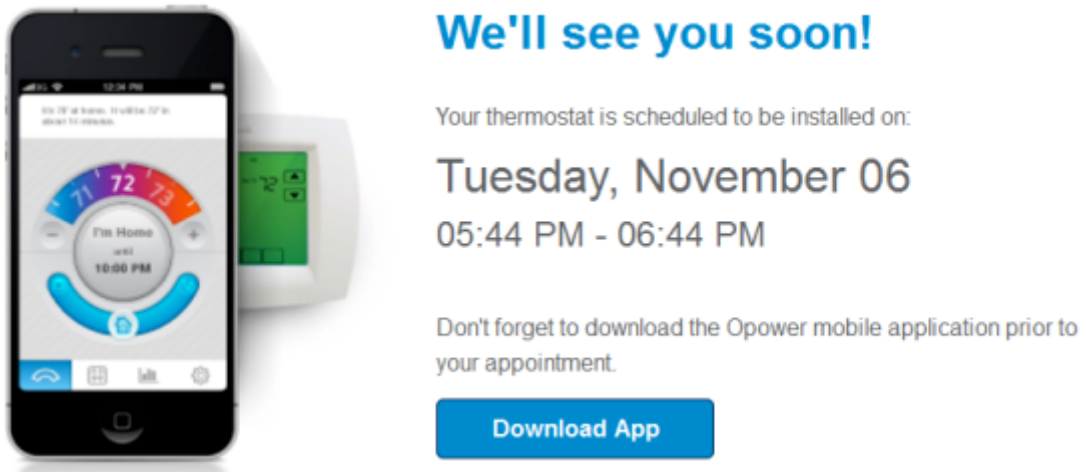
If none of the times available on the screen are convenient for the customer, they can click "Don't see an appointment you like?" to see a phone number they can call to schedule the appointment (1-888-660-5028).



Once they have selected the date and time for their appointment, they will see a confirmation screen. This includes information on how to reschedule the appointment and where to download the mobile application.



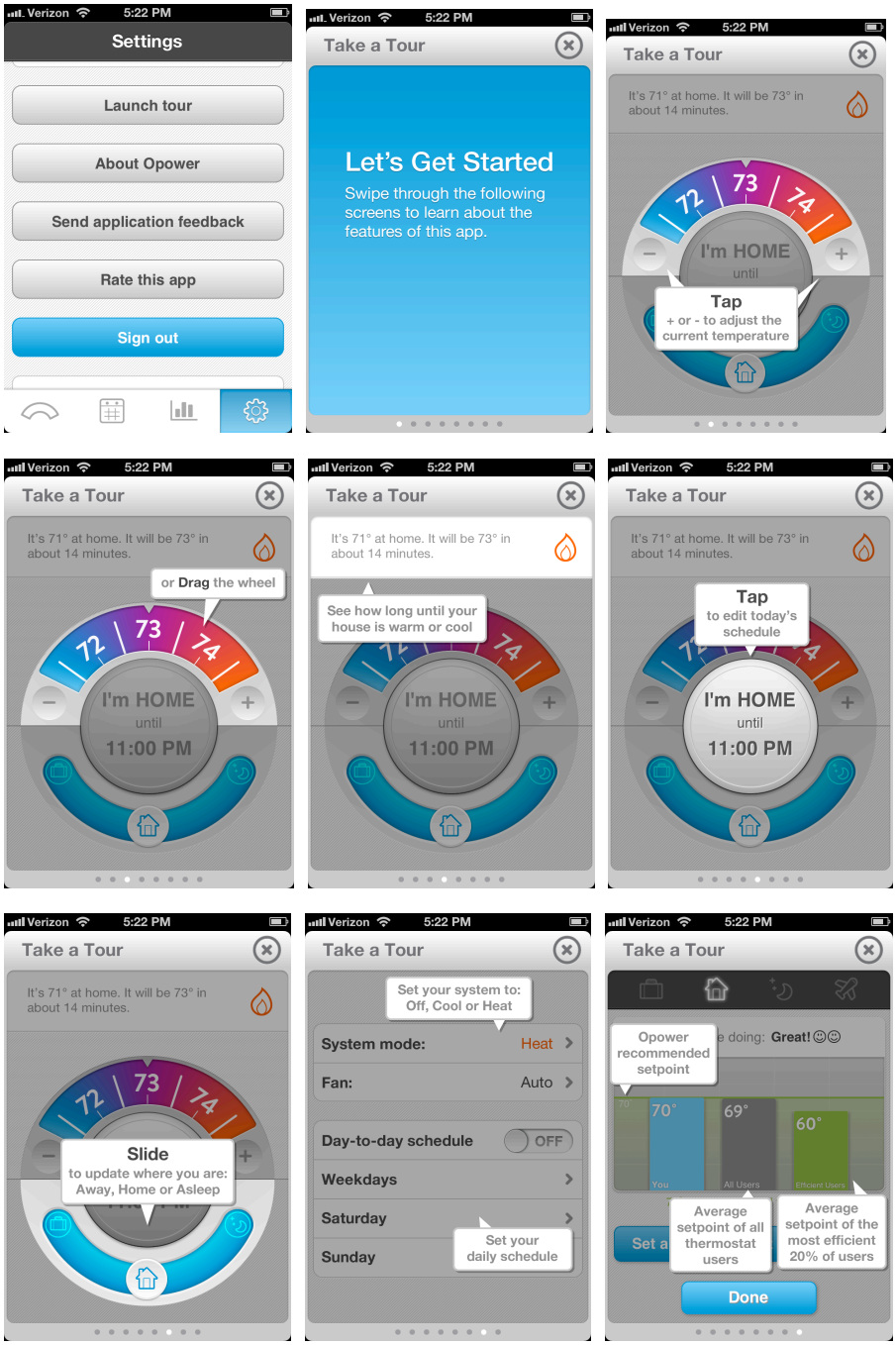
The customer will also receive an email confirmation for their appointment and a reminder to install the mobile application in advance of the appointment.



Mobile Application Tour

The mobile application tour can be launched at anytime, using the Opower mobile app on the iPod Touches, and later on the customer's smartphone. Click on the Settings tab, click "Launch

tour," slide through the tour pages, and click "Done" to exit. The tour provides an overview of some of the main application functions and customer messages.



Answering Customer FAQs

This section will help you answer customer questions about the program, mobile and Web applications, and thermostat. A full set of customer FAQs can be found at https://thermostat.opower.com/faq.

What is this thermostat program?

Opower and Honeywell have partnered to create a smart thermostat solution, which allows utility customers to program and monitor heating and cooling energy usage, not just from the thermostat itself, but also via Internet-connected devices like smartphones. This solution also gives you the ability to create optimal thermostat schedules that fit your lifestyle and provides customized recommendations to help you trim your energy bills.

How can I save?

A programmable thermostat can help reduce your heating and cooling costs. You can save all year long if you ensure your thermostat is set at the optimum program settings that match your lifestyle. You can manipulate your temperature setting and conserve energy, even while you are away, through the use of the Internet or your smartphone. Setting your programmable thermostat to the highest comfortable temperature in the summer and lowest comfortable temperature in the winter can help you reduce your energy bill.

What are the estimated savings based on?

The estimated costs and savings calculations are based on average heating and air conditioning usage and utility billing rates in your area. These are only estimations and are not a guarantee of savings from your utility company.

What other benefits does this program provide?

This thermostat program also benefits the community by helping to educate customers about energy use and energy efficiency goals. The energy customers save will not only help the environment, but also help reduce the need for new power plants and the occurrence of power outages.

Are there any safety or privacy concerns I should be aware of related to this thermostat program?

The Honeywell VisionPro thermostat used for this program was rigorously tested prior to being installed in customers' homes. These devices go through numerous quality control checks by multiple parties, to ensure they meet a high level of customer safety, reliability, and satisfaction.

It is also our top priority to protect our customers' information. We apply the same privacy protection standards to all data collected by the company from customers. We treat each customer's personal information and data as confidential, consistent with all regulatory requirements, including those established by the Public Utilities Commission. Therefore, be assured that your information is kept private.

Can I get this device for my other properties and/or business?

The smart thermostat program is only available for residential use at this time. Only a single thermostat is available for each program participant.

How many devices can I access the applications from?

Only a single wall-mounted thermostat is available for each program participant. You can install and access the mobile application from as many smartphones as you would like, but the application must be registered with the same username and password. Similarly, you can use the

Web application from any supported web browser on any computer. If more than one member of your household uses the application at the same time, the changes are preserved for the last person who saves their changes.

Can people see if I am home or not?

No. We apply the same privacy protection to this data as other all other data collected by the company for customers. The only way someone can see your status and schedule is if you give them your login credentials to the web or smartphone application.

If I work from home or have a severe illness for which I have special temperature needs, can I still benefit from this program?

You will always have control of your thermostat, so you can set safe and comfortable temperatures that are suitable for your lifestyle. An easy way to save energy is to lower your heating temperatures and raise your cooling temperatures when you are away. Depending on your personal needs, you may also be able to use more efficient temperatures while you are asleep.

How safe is the program? Can anyone hack into the system?

It is our top priority to protect our customers' information. Our system employs industry-standard defense mechanisms against brute-force attacks, code injection, and other malicious activity. We apply the same privacy protection standards to all data collected by the company from customers. We treat each customer's personal information and data as confidential, consistent with all regulatory requirements, including those established by the Public Utilities Commission. Therefore, be assured that your information is kept private.

What smartphones support the mobile application?

The mobile application is currently supported on the Apple iPhone 3GS or later, running IOS 4.3 or later, and Android phones running 2.2 or above. To locate your operating system on your iPhone, open the *Settings* app, click on "About," and see what "Version" your iPhone is running (needs to be 4.3 or above). To locate your operating system on your Android, open the *Settings* app, click on "About phone," and see what "Android version" your phone is running (needs to be 2.2 or above).

How do I make a one-time change to my schedule?

You can use the "Thermostat" page of the mobile application or the " My Thermostat" page of the Web application to manually change your temperature, change your current state (away, home, asleep), or set a new time to come home, wake, or go to sleep. On the thermostat on the wall, you can also manually change your temperature.

How can I change my email address and/or password?

Open the Web application, and then select "My account" to change your password or email address.

I now have three ways to change my thermostat. How are they different?

You can use your thermostat to manually change temperatures, turn on and off your heating and AC, and control your fan. The Web application has the same functionality as the thermostat and also allows you to register for an account, set a vacation schedule, and change your account

settings, primary schedule, default temperatures, state (home, away, asleep), and schedule for today. The mobile application has all of the functionality of the thermostat and Web application, plus it allows you to compare your temperature settings, set a passcode, and set and receive notifications.

Which browsers are supported for the Web application?

The current major release and previous major release of the four desktop browsers with the largest market share are supported. Currently, this means Internet Explorer, Safari, Mozilla Firefox, and Google Chrome are supported.

Will my house really be comfortable enough when I get home?

Yes. You just set the time you will return home and your thermostat does the rest. Your home will be heated or cooled for you before you return home after being away or on vacation. Your smart thermostat learns the amount of time it takes to heat or cool your house before you arrive, based on the actual temperature in your home and past usage.

Can I enroll in the program using my smartphone?

You can only enroll in the program using the Web application. If you are selected for the program, you will receive information about how to install the mobile application.

