American Economic Journal: Applied Economics 2025, 17(1): 90–125
https://doi.org/10.1257/app.20210618

Smart Thermostats, Automation, and Time-Varying Prices[†]

By Joshua Blonz, Karen Palmer,
Casey J. Wichman, and Derek C. Wietelman*

Can automation complement economic incentives? We explore this question by randomly encouraging households to activate a feature on their existing smart thermostat that automates responsiveness to time-of-use electricity pricing. The feature reduces air conditioning use during the highest-priced afternoon period, raising indoor temperatures above a household's preferred temperature, primarily for customers who are typically home during the day. Customers infrequently override the feature when they experience discomfort, suggesting that they are willing to trade off monetary savings for small increases in discomfort. Automation thus enables low-cost changes in household energy use, with potentially large electricity supply-cost reductions at scale. (*JEL* D12, D91, L94, L98, Q48)

Electricity demand and supply fluctuate greatly throughout the day. Economists tend to advocate for time-varying prices that incentivize load shifting (Borenstein and Holland 2005; Fowlie et al. 2021; Blonz 2022; Borenstein and Bushnell 2022). The widespread rollout of smart meters has made time-varying pricing a viable and attractive option to efficiently balance electricity supply and demand by shifting consumption from high-cost hours to lower-cost hours. The efficiency gains from time-varying pricing are likely to grow over time with increased deployment of variable renewable resources, such as solar and wind (Borenstein 2019; Bushnell and Novan 2021). However, several obstacles inhibit its widescale adoption, including political feasibility constraints (Borenstein 2007; Joskow and Wolfram 2012) and

*Blonz: Federal Reserve Board (email: joshua.blonz@frb.gov); Palmer: Resources for the Future (email: palmer@rff.org); Wichman: Georgia Institute of Technology and Resources for the Future (email: wichman@gatech.edu); Wietelman: University of Maryland (email: dcwietel@umd.edu). Ilyana Kuziemko was coeditor for this article. The views in this paper are those of the authors and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff, Resources for the Future, Georgia Institute of Technology, University of Maryland, or Ecobee. We thank the Alfred P. Sloan Foundation for financial support under Grant No: G-2019-12443. We thank Ecobee for providing data access. We especially thank Kathryn Checkley, Kevin Trinh, Brent Huchuk, Valentin Randon, and others at Ecobee for their help in interpreting the data. We thank Severin Borenstein, Christopher Clapp, as well as seminar participants at the University of Chicago Energy Policy Institute, the 2020 AERE Virtual Summer Conference, the Online Agricultural and Resource Economics Seminar Series, the 2021 ASSA Meetings, the 2021 NBER Summer Institute, and the University of Chicago Experiments Group Lunch for helpful comments. Seema Jayachandran and two anonymous referees provided exceptional feedback. Any remaining errors are our own. This experiment is registered in the AEA RCT Registry under AEARCTR-0008019. Human subjects research, but IRB approval was not obtained for the following reason(s): All authors were employed by Resources for the Future, which does not have an IRB, when this research began. All data was anonymized and shared voluntarily for research purposes, there was minimal risk to human subjects, and randomization occurred as part of Ecobee's rollout of the Eco+ program.
[†]Go to https://doi.org/10.1257/app.20210618 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

the relatively small financial benefits from responding to time-varying prices, which rational consumers may ignore if attention costs are high (Jessoe and Rapson 2014; Sallee 2014; Harding and Lamarche 2016).

One potential solution to these roadblocks to widescale adoption of time-varying pricing is automation (Harding and Sexton 2017). Smart thermostats are being widely adopted by households interested in energy savings and smart devices.[1] Although evidence suggests that smart thermostats alone may not generate energy savings (Brandon et al. 2022), recent research has found that automated thermostats can increase consumers' responsiveness to time-varying prices (Gillan 2018; Bollinger and Hartmann 2019). However, automated smart thermostats may introduce costs to users, including increases in household discomfort and changes in consumer behavior that avoid or undo the effects of automation.

In this paper, we evaluate a randomized experiment implemented in partnership with Ecobee—a leading smart-thermostat company. Our experiment leverages the North American rollout of a suite of thermostat features, called "Eco+," which includes an algorithm that can automate the household's heating and cooling schedule if they pay time-of-use (TOU) electricity prices, which we refer to as the "Eco+ TOU feature" or the "TOU feature." Thermostat-level data allows us to observe in five-minute increments the average indoor temperature, how much the air conditioner runs (a proxy for energy usage), if the thermostat's motion sensor detects that an individual is in the home, thermostat setpoint, and a variety of other outcomes. These detailed data allow us to characterize both the potential energy saving benefits and the costs of the automated TOU feature, which include discomfort (measured as deviations from a customer's preferred temperature), and consumer responses to the feature's operation.

We demonstrate the effects of the Eco+ TOU feature on setpoints, cooling, and household discomfort. The TOU feature lowers energy usage over the course of a day primarily by moving thermostat setpoints upwards, reducing compressor usage by around 44 minutes (88 percent) during the peak period (11 AM to 5 PM). However, the change in energy use comes at the cost of increasing discomfort, which we measure in two ways: (1) the hourly average "temperature wedge" between experienced and preferred temperatures (*ATW*), and (2) the maximum hourly temperature wedge (*MTW*) intended to capture potential nonlinearities in discomfort. We find that the TOU feature causes an average increase in discomfort of roughly 0.20–0.25 degrees during the peak period in the posttreatment period. Results for *MTW* are similar to *ATW*, which suggests that experienced discomfort is relatively small even when accounting for nonlinearities. Notably, we find energy savings for all encouraged households, but the increases in discomfort are concentrated in homes where occupants are typically home throughout the day. These results imply that most customers (two-thirds of our sample) experience a win-win: reductions in energy use with no corresponding increase in discomfort.

[1] Some utilities subsidize the installation of smart thermostats. Estimates of smart-thermostat penetration suggest that approximately 18.3 percent of US households with broadband internet will have smart thermostats by 2023, with projections that about 38 million homes in the United States will have smart thermostats by 2026 (https://www.utilitydive.com/news/smart-thermostats-us-slow-adoption-misses-energy-savings/630901/, last accessed September 19, 2024).

These overall findings capture the net effect of the automated Eco+ TOU feature, but they do not differentiate between the algorithmic effect of the TOU feature and the behavioral changes that households may make in response to treatment. Households might respond to the higher indoor temperatures caused by the Eco+ TOU feature in three main ways: First, occupants could leave the home and go to another air-conditioned space (e.g., the movies or the store), however, we find strong evidence that encouragement does not change thermostat motion sensor activation. Second, occupants could adjust their Eco+ settings or turn off the Eco+ TOU features, which would permanently change the way their thermostat responds to higher temperatures. However, we find limited evidence of such adjustment behavior.

Finally, customers could temporarily override their thermostat by manually changing the setpoint to a level below their preprogrammed scheduled temperature. We identify these "thermostat overrides" (TOs) empirically to explore how our discomfort measures relate to override behavior. Both *ATW* and *MTW* climb quickly in the time leading up to an individual overriding their thermostat, suggesting that these measures capture some element of discomfort that is correlated with an individual taking action to override their thermostat's setting. On the relatively few days when TOs occur, customers offset the algorithmic energy savings and resulting discomfort. On days without TOs, the TOU feature reduces energy use by a similar amount to when we do not separately account for behavior. The tendency of customers to override their thermostats on occasion suggests that (a) the transaction costs to override a thermostat are not so high that it never happens and (b) the overall effect of TOs on the net benefits of the automated feature is small. These findings suggest that our results do not just reflect a default bias, but instead reflect a mix of the engineering specifications of the TOU feature and customer behavior. In the longer run, we find that people do not disable the automation, which suggests that Eco+ will provide savings in subsequent summers.

We elucidate the trade-off between energy savings and discomfort in context by calculating the implied energy cost savings for customers. The reduction in compressor usage from the TOU feature generates household-level savings of C$0.22–$0.29 (or 1.6–2.2 kWh) on each day in the posttreatment period, where the range captures different assumptions about air conditioning system efficiency. These energy savings, while modest at the household level, can translate into large savings when aggregated across a larger population of smart-thermostat owners. By applying the noncompliance rate and energy savings from our experiment to a stated governmental policy of installing 100,000 smart thermostats at no cost to Ontario households, we would expect a roughly 8–11 MW reduction in electricity demand from 4 PM to 5 PM on an average summer day in Ontario, which is about one quarter of the capacity of a small powerplant designed to meet peak demand. When extrapolated to a broader population, thermostat automation could offset meaningful amounts of peak-load generation, creating valuable energy savings for the grid (Blonz 2022).

Our findings make several contributions to the literature on automation and energy policy. Previous studies have estimated the energy savings from replacing a simple nonscheduling thermostat with a smart thermostat (List, Metcalfe, and Price 2018; Brandon et al. 2022); others have compared information with automation treatments in conjunction with the provision of new smart thermostats to study participants (Harding and Lamarche 2016; Bollinger and Hartmann 2019). By contrast, households in our sample have had a smart thermostat in operation for at least 12 months and have faced TOU prices for an even longer period. Customers' familiarity with the prices and hardware in our experiment, coupled with the widespread rollout of a real technology where consumers did not know they were being observed for an experiment, suggest a higher likelihood of the external validity of our results.

Second, we estimate energy-saving benefits, changes in discomfort, and consumer reactions to smart features on a widely available smart thermostat and TOU pricing. Many studies have found that households reduce their energy use in response to time-of-use pricing, but they cannot compare the energy savings to changes in indoor comfort (e.g., Jessoe and Rapson 2014; Burkhardt, Gillingham, and Kopalle 2019; Prest 2020). Our results suggest that most households in our experiment receive a win-win: a reduction in energy costs without a corresponding increase in discomfort. Even for the small subset of households that do experience increases in discomfort, the magnitudes are small.

Third, our setting and detailed data allow us to examine how individuals respond to an automated energy intervention. It is an open question to what extent consumers are willing to sacrifice some control over consumption decisions to produce private and social benefits. Our high-frequency data allow us to disambiguate the algorithmic and behavioral effects to show how individuals engage with smart technologies. We show that behavior does not meaningfully erode the energy savings from automation. In the short run, overrides are relatively infrequent, and in the longer run, we find no change in the user settings governing the automation. Both behavioral margins are critical to consider when designing automation technology, and our findings suggest that the TOU feature provides energy savings with a behavioral response that does not inhibit the long-run viability of thermostat automation.[2]

Fourth, our results have important applications because algorithms like the one we study can be provided at scale at a low marginal cost to the millions of households that have already installed smart technologies. These devices could be easily programmed to respond effectively to TOU or other time-varying prices with low attention or economic costs. The continued rollout of smart thermostats and algorithmic features, like those studied here, suggests that time-varying pricing might produce larger responses than previously estimated in the literature (e.g., Wolak 2011; Jessoe and Rapson 2015; Harding and Lamarche 2016; Gillan 2018; Fowlie et al. 2021; Blonz 2022).

I. Experimental Design

A. *Data Partner and Geographic Setting*

To understand how smart devices interact with time-varying prices, we study Ontario households equipped with smart thermostats. This province is an attractive region to study in this context for two primary reasons: many households in Ontario already have Ecobee thermostats, and more than 90 percent of households already

[2] Relatedly, Brandon et al. (2022) find that behavioral overrides of first-generation smart thermostats dampen their effectiveness of delivering energy savings when customers pay time-invariant prices. Our paper studies a different context, where the presence of TOU pricing combined with automation shows that the current generation of thermostats that are being widely adopted can provide economically relevant energy savings even after accounting for both short- and long-run behavioral responses.

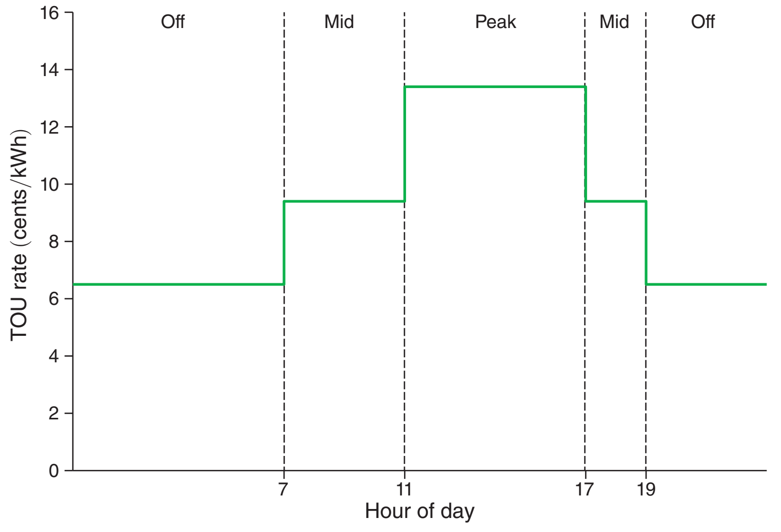

Figure 1. Ontario TOU Rates by Hour in Summer 2019

Notes: TOU electricity rates (in Canadian cents per kWh) by hour of day are presented for summer 2019. TOU hours and rates presented here are for weekdays during May through October. The Mid and Peak TOU rate periods invert during the winter months. Marginal TOU rates stayed largely constant between 2018 and 2019.

consume residential electricity on a TOU rate structure. Ontario households have been on TOU rates by default since 2012, although they continue to have the opportunity to opt out (Faruqui et al. 2015; Lessem et al. 2017). The TOU rates faced by customers in summer of 2019 are shown in Figure 1. There are three distinct prices based on the time of day when the electricity is consumed—peak, mid-peak, and off-peak.

We partnered with Ecobee, an Ontario-based smart-thermostat company with a strong North American presence, to implement a randomized encouragement design with existing Ecobee thermostat owners in Ontario. Eligible households had previously volunteered to share their anonymized data as part of Ecobee's Donate Your Data (DYD) program. The DYD program encourages Ecobee customers across the United States and Canada to allow Ecobee to share their high-frequency (five-minute-interval) thermostat data, as well as anonymized information about their homes and thermostats, with researchers. As of 2019, there were 106,253 DYD customers with a country code listed as United States, Canada, or Mexico. Of these, 14,045 were in Canada, and more than half of these customers (7,702) were in Ontario. When a customer agrees to participate in the DYD program, all her data from the time the thermostat was initially installed are provided. DYD households in Ontario form the population from which we draw our experimental sample.

It is possible that the DYD population differs systematically from the general population of smart-thermostat customers.[3] These selection concerns are worth

[3] Meier et al. (2019) compare the Ecobee DYD sample in the United States with the data collected by the Energy Information Administration in its 2015 Residential Energy Consumption Survey (RECS) and find that the Ecobee data tend to include more single-family homes and fewer one-person households than the RECs sample, but otherwise, in housing characteristics (such as age and size) of the US single-family housing stock, both datasets are roughly comparable.

noting, but they do not affect the internal validity of our conclusions due to our randomized design. The main concern about the generalizability of our results would be a correlation between signing up for DYD and either the occupancy rate of the home or the choice of temperature settings.

In 2019, Ecobee introduced Eco+, a suite of algorithmic features designed to deliver energy savings without sacrificing in-home comfort. A main component of the Eco+ software upgrade in areas with time-varying electricity prices is the TOU feature, which aims to precool the home when electricity prices are lower and allows the indoor temperature to move higher when electricity prices are higher, thereby reducing the time that the compressor will run.[4] For the TOU feature to be active, households must both enroll in Eco+ and select information about their electric utility rate structure. Mechanically, the TOU feature reduces energy consumption for cooling by adjusting the thermostat's preprogrammed setpoint. For example, if a household has programmed their home to be 72 degrees between 11 AM and 5 PM, Eco+ may adjust that setpoint to be anywhere between 1 and 4 degrees higher based on the customer's settings.

The extent of the thermostat setpoint manipulation by Eco+ for enrolled customers is a function of how they set their "slider" scale, a numeric setting from 1 to 5 indicating how aggressively they want the algorithm to implement the TOU feature (with 1 being the weakest, and 5 the strongest). The extent to which the precooling and setbacks (adjustments to the temperature setting during high-price periods) are implemented is also a direct function of the chosen comfort setting. During the Eco+ enrollment process, customers are prompted to set a slider level. The default setting is 4, but customers can change their setting at any point through the app or at the thermostat. Those who set the slider to 1, the least aggressive setting, experience no precooling or setbacks, essentially turning off the TOU component of Eco+. Customers who choose intermediate slider settings see progressively greater precooling and setbacks.[5]

B. *Randomized Encouragement Design*

We partnered with Ecobee to randomly encourage eligible customers to enroll in the Eco+ feature update, while control households received no encouragement. Our experiment was built into Ecobee's North American rollout of the Eco+ algorithm in summer 2019. Prior to the rollout, in spring 2019, Ecobee provided us a list of eligible thermostats that participated in DYD and associated metadata to perform the randomization. We applied our own set of filtering criteria (removing accounts with multiple thermostats, fewer than 12 months of data, and multistage cooling systems), randomly assigned households into an "encouraged" group of 2,445 thermostats and held out the other 1,500 thermostats as a control group. After randomizing, we learned that 237 households had older-model thermostats that did not include motion sensing and that another 1,667 had a model for which motion sensors had to be installed separately, meaning that some households inevitably would not have motion sensor data, while

[4] The Eco+ software upgrade consists of five components, of which the TOU feature is one. Further details on the full suite of features can be found in online Appendix B.1.

[5] Further details on the latest iteration of the TOU feature of Eco+ can be found on Ecobee's website, https://support.ecobee.com/hc/en-us/articles/360035246672-eco-Frequently-Asked-Questions (last accessed April 20, 2023).

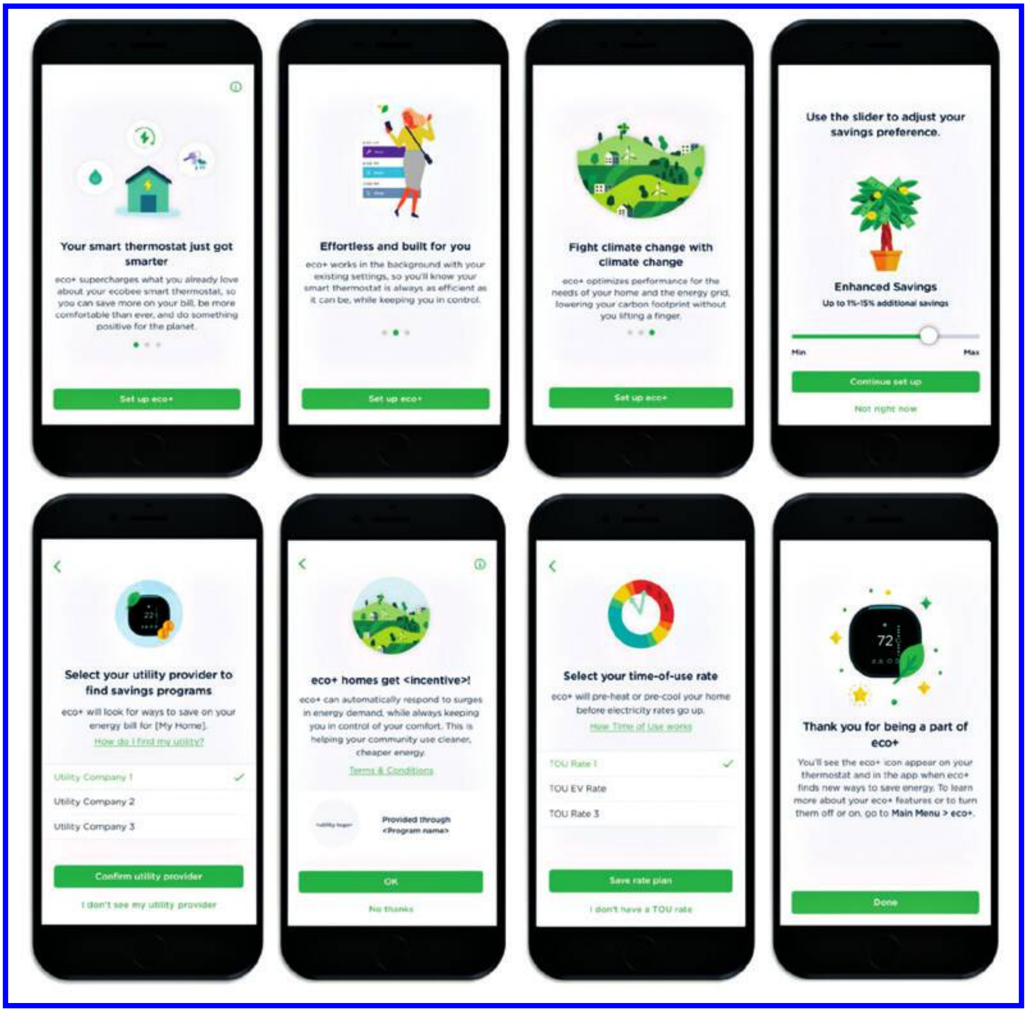

FIGURE 2. ECO+ ENCOURAGEMENT PROMPT THROUGH PHONE APPLICATION

Notes: The figure presents a series of screenshots that customers in the encouraged group received upon beginning the process to set up Eco+ after receiving the randomized encouragement. Prompts include such screens as the selection of the comfort or slider setting and associating the thermostat with the proper utility and time-of-use rate.

others would have such data if they chose to install motion sensors. Because these data were critical for measuring one of our outcomes, discomfort, we removed from our analysis thermostats with no motion sensor data.[6] Of the remaining thermostats, we were ultimately able to retrieve data on 2,133 thermostats, 1,319 of which were in our encouragement group and 814 in the not-encouraged group.

On August 6, 2019, Ecobee sent out the randomized encouragement. Figure 2 shows the series of prompts that customers saw on the Ecobee smartphone application

[6] Randomization into encouraged and not-encouraged groups was balanced among these thermostats, and all results for our analysis of compressor run time are consistent among these thermostats as well. Online Appendix Table A.1 shows that observables are largely balanced for the full universe of thermostats for which we were able to recover data. See online Appendix B.2 for more details.

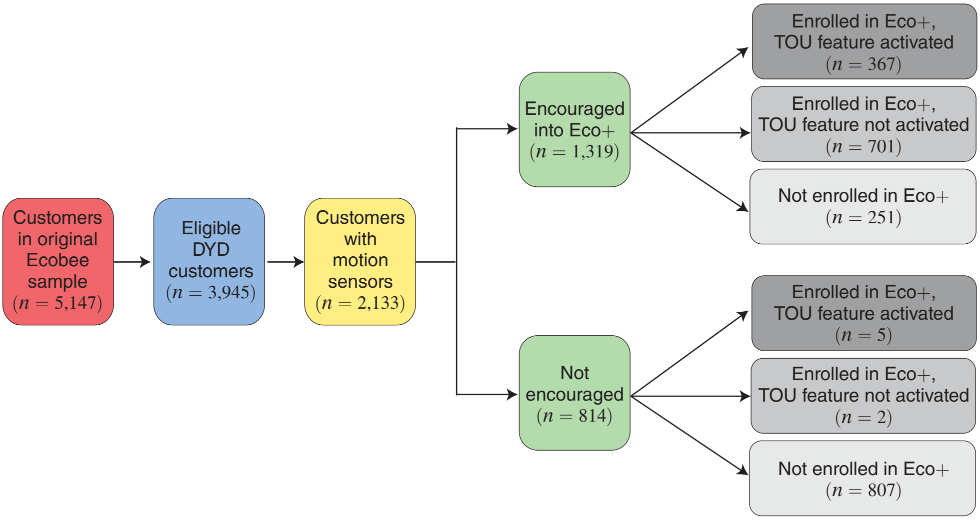

FIGURE 3. EXPERIMENTAL DESIGN AND SAMPLE RANDOMIZATION

Notes: The figure presents a flow chart that documents the experimental design and ultimate compliance with the Eco+ encouragement as well as TOU feature activation.

at the time of the experiment. Customers in the encouraged group were encouraged to enroll in Eco+ via email, and customers who used the Ecobee smartphone app received notifications about signing up for Eco+ through their phones as well. After completing the series of prompts and agreeing to the terms and conditions, the household was enrolled in Eco+ and its features began to activate.[7]

The encouragement into Eco+ had a high compliance rate: around 80 percent of encouraged households enrolled in Eco+ (and seven households discovered and enrolled in Eco+ despite not receiving any encouragement). Figure 3 shows the breakdown of encouragement acceptance and TOU feature activation for the thermostats for which we were able to collect data. Of the 80 percent of encouraged customers who enrolled, only 62 percent provided information on their electricity provider. This step can be seen as the lower left image in Figure 2. Of those individuals that provided their utility, only 61 percent indicated that they were on the TOU rate, which was the seventh image of Figure 2. Both steps are needed to activate the TOU feature. The large amount of attrition is likely due to individuals not knowing this information during setup and ending the setup or skipping this step. This results in approximately 367 people who activated the TOU features, who we call compliers, from the encouraged group (consumers who had TOU turn on at any point in the post-encouragement period) as shown in the rightmost grouping of Figure 3. Further details on our randomization, sample construction and sample attrition are presented in online Appendix B.2.

[7] The design can be considered a natural field experiment in that customers already had smart thermostats in their homes and did not actively know that they were part of an experiment (List 2007). This fact bolsters claims of external validity, as we can rule out customer knowledge of the experiment as a channel that could influence behavior. This is in contrast to much of the smart device literature, in which households must be recruited to join the experiment and install devices in their homes.

II. Data

A. *Smart-Thermostat and Eco+ Data*

In our randomized encouragement design, we have two primary sources of data. The first is Ecobee's DYD database (Ecobee 2019), which provides smart-thermostat data in five-minute intervals for all DYD customers. We observe compressor run time (our proxy for energy use), indoor and outdoor temperature and humidity levels, consumer heating and cooling setpoints, and motion detected in the home at the thermostat and at remote sensors that may be placed in other parts of the house. We focus our analysis on the warm months (July–September) of 2018 and 2019. We remove from the sample all observations on weekends and holidays when TOU rates are not in effect. We collapse our five-minute-interval data to the hourly level, with a thermostat-by-hour observation as the unit of analysis. This aggregation in time involves averaging setpoints and summing compressor run time and the amount of time per hour that smart-thermostat features like TOU are activated. Additionally, we calculate average hourly daytime and nighttime indoor and outdoor temperatures (Table 1).

We also observe data on when consumers override their preprogrammed (or algorithmic) thermostat settings to affect cooling system operation. These data come from two sources. The first source is the larger DYD dataset with observations for all households, both encouraged and control, included in the experiment. This information can be used to identify the time periods when customers manually adjust their thermostat setpoint down to override their preprogrammed settings during peak hours, thereby increasing air conditioning use in the interest of improving comfort. We discuss this measure of behavior in more detail below in Section IID. The second data source documents consumers' interactions with the Eco+ functions. For all households enrolled in Eco+, we observe daily data from the start of the experiment through January 2020, including when the user accepted the Eco+ terms and conditions, the daily slider scale setting for savings preferences, and daily dummy variables for which Eco+ features are enabled on that day. The daily slider setting allows us to observe the household's interaction with the Eco+ feature although we do not observe this setting for non-Eco+ households.

B. *Details on Randomization and Balance*

Our ability to estimate causal effects rests in part on the integrity of the randomization. We explore the differences in pretreatment data between the encouraged and not-encouraged households in our sample for various home and thermostat features, as well as temperature readings and settings, in Table 1. The comparison suggests that the randomization performed well in terms of selecting households for encouragement into the experiment. The table shows no significant differences between the two groups for most of the house and thermostat characteristics included in the DYD database. There are, however, some minor differences in pre-encouragement nighttime compressor run times. In our empirical analysis, we use a difference-in-differences regression controlling flexibly for household-by-hour-of-week fixed effects to control for these small differences.

Table 1—Balance on Observables in Preperiod

Variable name	Not encouraged	Encouraged	Difference
Panel A. Household characteristics			
Floor area (100 sqft)	21.57	21.71	0.14
	(9.26)	(9.61)	(0.42)
Number of floors	2.61	2.56	−0.05
	(0.73)	(0.78)	(0.03)
Home age (years)	32.56	31.30	−1.27
	(31.86)	(29.73)	(1.39)
Household size	3.07	3.02	−0.05
	(1.18)	(1.23)	(0.07)
Model: Ecobee 3	0.67	0.67	−0.00
	(0.47)	(0.47)	(0.02)
Model: Ecobee 3 Lite	0.20	0.21	0.01
	(0.40)	(0.41)	(0.02)
Model: Ecobee 4	0.13	0.13	−0.01
	(0.34)	(0.34)	(0.01)
Panel B. Preperiod daytime characteristics			
Cooling setpoint (deg. F)	75.29	75.23	−0.06
	(4.04)	(3.95)	(0.14)
Compressor run-time (mins/hr)	13.89	14.07	0.19
	(20.40)	(20.45)	(0.39)
Avg. temp. wedge (deg. F)	0.35	0.35	−0.00
	(1.16)	(1.13)	(0.02)
Max. temp. wedge (deg. F)	0.47	0.47	−0.00
	(1.29)	(1.27)	(0.03)
Indoor temp (deg. F)	73.58	73.50	−0.08
	(3.01)	(3.05)	(0.09)
Outdoor temp (deg. F)	73.63	73.61	−0.02
	(8.07)	(7.99)	(0.06)
Hold (mins/hr)	19.98	20.58	0.60
	(28.05)	(28.27)	(0.95)
Thermostat override	0.05	0.06	0.00
	(0.23)	(0.23)	(0.00)
Panel C. Preperiod nighttime characteristics			
Cooling setpoint (deg. F)	74.10	74.00	−0.10
	(3.65)	(3.63)	(0.13)
Compressor run-time (mins/hr)	13.17	13.84	0.67
	(19.71)	(20.13)	(0.37)
Avg. temp. wedge (deg. F)	0.27	0.27	0.00
	(1.02)	(1.00)	(0.01)
Max. temp. wedge (deg. F)	0.37	0.37	0.00
	(1.15)	(1.14)	(0.02)
Indoor temp (deg. F)	73.40	73.28	−0.12
	(3.06)	(3.08)	(0.10)
Outdoor temp (deg. F)	67.11	67.16	0.05
	(7.29)	(7.25)	(0.10)
Hold (mins/hr)	20.60	21.00	0.40
	(28.23)	(28.38)	(0.95)
Thermostat override	0.05	0.05	0.00
	(0.23)	(0.23)	(0.00)

Notes: Sample means and standard deviations (in parentheses) are presented for characteristics of the randomized encouragement and control groups for the $n = 2{,}133$ thermostats used in the primary analysis. The third column presents the coefficient from regressions of each characteristic on the treatment indicator with the standard error below (in parentheses). Standard errors are clustered at the household level.

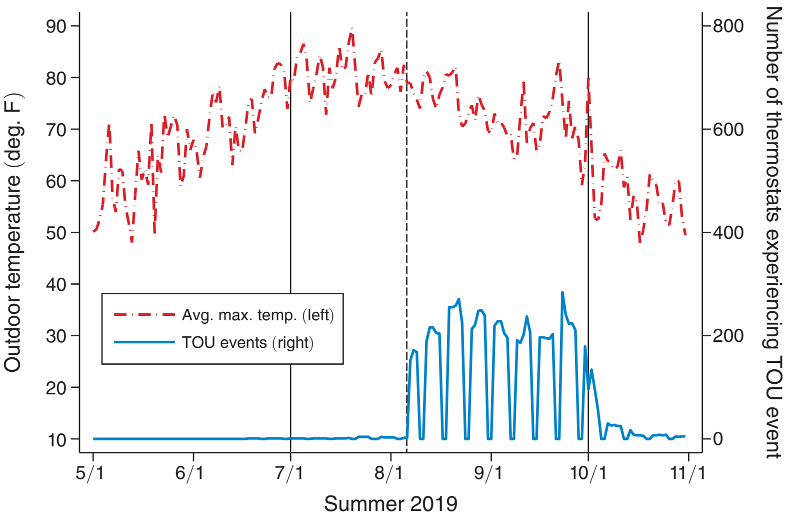

Figure 4. Temperature and TOU Activation

Notes: Average daily maximum temperatures by thermostat are indicated by the dashed pink time series and the left vertical axis. Daily counts of the number of thermostats experiencing a TOU event in 2019 (i.e., the TOU feature is observed at least once throughout the course of the day) are indicated by the solid blue time series and the right vertical axis. The vertical dashed line represent the date of encouragement (August 6, 2019). The vertical bold solid lines bound the sample used in the primary analysis for summer 2019 (July 1–September 30, 2019).

The automated responses to intraday variations in electricity prices and the associated monetary savings to consumers depend on how much the household is using its air conditioning system, which is a function of outdoor temperature.[8] In Figure 4, we present average daily maximum outdoor temperatures from May through October 2019 in the pink dashed line. The graph shows that on many days, the maximum temperature exceeds 80 degrees Fahrenheit, and the maximum temperature often exceeds 70 degrees. The average cooling setpoint in our sample for the months of July and August was 74–75 degrees, depending on the time of day. In September, both mean and average maximum temperatures fall, although we continue to observe days exceeding 80 degrees.

The solid blue line in Figure 4 shows the number of thermostats for which the TOU feature activates each day. The TOU activation is a function of both the outdoor temperature and the TOU rate structure. The weekly dips in the blue line reflect no TOU prices on the weekends. As outdoor temperatures decline in October, the TOU feature does not activate because households are no longer air conditioning their homes. The timing of our experimental treatment period, indicated by the vertical dotted lines, extends from August 6, when the encouragement was sent, to September 30, when temperatures begin to decline; we focus on weekdays only during that window.

[8] Despite Ontario's relatively cool climate, all households in our sample have air conditioning systems.

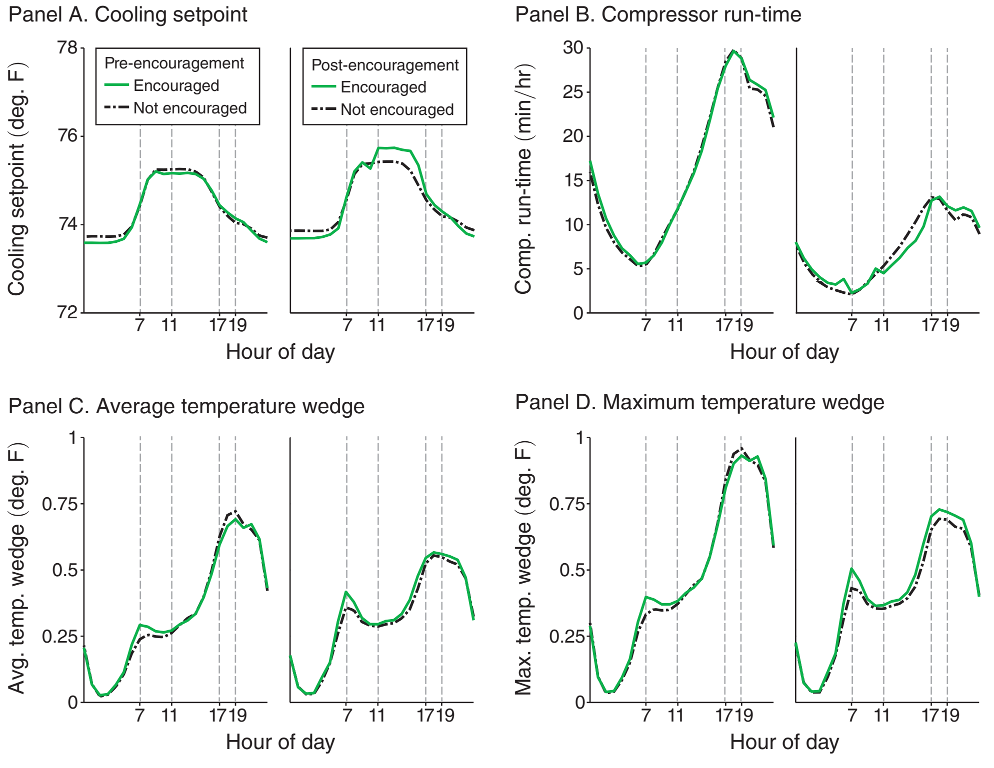

Figure 5. Hourly Profile of Outcome Variables by Encouragement Status

Notes: The figure presents average hourly outcomes by encouragement status. In each panel, average outcomes are presented separately for the "pre-encouragement" and "post-encouragement" periods on the left and right, respectively. "Pre-encouragement" refers to the period of July 1, 2019 up to August 6, 2019, the date of encouragement. "Post-encouragement" refers to the period from encouragement until the end of the experimental sample on September 30, 2019. The dashed lines indicate the TOU rate periods. 7–11 is the AM mid-peak rate period, 11–17 is the peak period, 17–19 is the PM mid-peak period, and 19–7 is the evening off-peak period.

The operation of the air conditioning system responds to variations in outdoor temperature over the course of the day and to the level of the thermostat setpoint. In panels A and B of Figure 5, we present thermostat setpoints and average compressor run time in minutes per hour over the course of a weekday in pre-encouragement (left column) and post-encouragement (right column) in 2019. The dashed vertical lines define the off-peak, mid-peak, and peak pricing periods. For the periods pre- and post-encouragement, we present setpoints and compressor run time separately for households in the encouraged group and the not-encouraged group. The pre-encouragement period includes data from July 1 up to the day of encouragement (August 6).

As shown in the figure, average cooling setpoints exhibit a similar pattern over the course of the day across the three time periods and the setpoints for the encouraged and not-encouraged groups are aligned prior to treatment. After the launch of the experiment average cooling setpoint is higher for encouraged households then for the not-encouraged group during the peak pricing hours and lower for the encouraged group than for the not-encouraged group in the off-peak hours, providing evidence of precooling in the off-peak hours for the encouraged group. Compressor run time results mirror the findings for setpoints. Minutes per hour of compressor operation were highest in July, the warmest month in our sample. Pretreatment compressor operation is very similar between encouraged and not-encouraged households, particularly during the peak pricing hours. After the experiment was launched, compressor run time for the encouraged group exceeds that of the not-encouraged group in the off-peak and mid-peak hours and falls below the not-encouraged group in the peak hours. This separation throughout the post-encouragement period, despite lower overall compressor operation due to cooler temperatures in September.

C. *Discomfort*

We develop two related metrics that capture aspects of thermal discomfort that individuals experience in their homes. To do so, we rely on the deviation between an energy user's preferred indoor temperature based on her past thermostat setpoints and her realized temperature (in degrees Fahrenheit) multiplied by an indicator for whether we observed motion within the home in a given hour. Formally, we construct an hourly measure of experienced discomfort—average temperature wedge (*ATW*)—from the five-minute interval data for household i during hour-of-sample t as follows:

$$ATW_{it} = \left(\frac{1}{12}\sum_{q=1}^{12}\left|T_{iqt} - T_{it}^{*}\right|\right) \times \mathbf{1}\{Motion_{it}\}, \tag{1}$$

where T_{iqt} is the average observed indoor temperature for household i in five-minute interval q during hour-of-sample t; T_{it}^{*} is the household's average preferred indoor temperature for each hour; and $\mathbf{1}\{Motion_{it}\}$ is an indicator for whether motion was observed in the home at any point within the hour. We sum the absolute difference between actual and preferred temperatures at each five-minute increment and divide that quantity by 12 (the number of five-minute intervals within the hour) to provide an hourly average that is zero if no motion is detected that hour. As an example, for a customer who prefers an indoor temperature of 72 degrees Fahrenheit but experiences a constant 74 degrees for 60 minutes, we would assign an ATW_{it} value of 2 degrees of discomfort for that hour.

Average deviations in temperature provide an easily interpretable proxy for discomfort but may not capture potential nonlinearities in discomfort. For example, a 6-degree temperature deviation for 10 minutes, with no deviation in the other 50 minutes of an hour, results in $ATW_{it} = 1$ degree. The experienced costs of this temperature pattern might be very different than experiencing a constant 1-degree temperature wedge for the entire hour. As a result, we also define a maximum temperature wedge (MTW_{it}) as the maximum across all 5-minute intervals in the hour:

$$MTW_{it} = \max\left\{\left|T_{iqt} - T_{it}^{*}\right|\right\} \times \mathbf{1}\{Motion_{it}\} \quad \forall q = 1,\ldots,12. \tag{2}$$

This maximum temperature wedge captures the extremes in discomfort for any hour when motion is detected within the home.

For both ATW_{it} and MTW_{it}, our measures of T_{iqt} and $\mathbf{1}\{Motion_{it}\}$ are observed directly in our thermostat data. For preferred measures of indoor temperature T_{it}^{*}, we assign a counterfactual setpoint schedule using household-by-day-of-week-by-hour-of-day mean temperature setpoints in July 2019 (prior to our experiment). That is, we assume that a customer's scheduled setpoint on a Tuesday at 3 PM in July 2019 reflects her preferences at the same day-of-week and hour-of-day in our posttreatment period, August and September 2019.[9]

The resulting values of *ATW* and *MTW* over the course of the day in each month are displayed in panels C and D of Figure 5, separately for the encouraged and not-encouraged groups. Note that discomfort is, by definition, a function of occupancy and thus tends to increase in the evening when people are more likely to be at home. We see both the *ATW* and *MTW* measures of discomfort tend to be higher for the encouraged group than for the not-encouraged group in August during peak hours and also during the transition from off-peak to mid-peak periods in the early morning as precooling contributes to the wedge between desired and experienced temperatures.[10]

D. *Thermostat Overrides*

Our high-frequency thermostat data allows us to observe how and when people interact with their thermostats, including when they override their thermostat's programs and features. We are particularly interested in when consumers override their thermostat to respond to discomfort experienced during peak pricing periods of the day, which also tend to be the hours when outdoor temperatures are the highest.

We define a thermostat override (TO) as occurring when an individual overrides their thermostat's programming during the peak demand period between 11 AM and 5 PM To isolate behavior that happens in response to the Eco+ TOU feature, we limit TOs to times when hold was not engaged at 10 AM on the same day.[11] In most cases, a TO happens when an individual adjusts their thermostat down a degree or two to turn on their air conditioner. A TO can either override their prescheduled setpoints or the Eco+ TOU feature, which allows us to define this variable for all households (encouraged and control) during the peak hours both in the pretreatment

[9] We assume that these counterfactual set-points (T_{it}^{*}) are constant throughout the hour to minimize the effect that idiosyncratic changes in the five-minute data might have on our discomfort measures. Online Appendix B.3 describes the construction of our discomfort measure in further detail. Online Appendix Figure A.1 illustrates our counterfactual setpoints in addition to showing how the counterfactual substitution process affects setpoints in post-encouragement period.

[10] Online Appendix Figures A.2 and A.3 illustrate changes in the four primary outcome variables due to encouragement discussed here (setpoints, compressor run time, and average and maximum temperature wedges), where we graph daily peak period averages of each variable by encouragement status separately for Summer 2018 and 2019 and pre/post encouragement.

[11] Holds are generally temporary and expire when the next scheduled temperature starts, which is the default value for all Ecobee thermostats. Users can manually adjust their hold duration to be indefinite and to span time periods. Our approach of defining a TO as not having hold engaged at 10 AM allows us to focus on active engagement with the hold function during the peak time periods, which are most relevant for analyzing responses to Eco+ TOU control of the thermostat.

and treatment periods. About 1,649 of the 2,133 households in our sample override their thermostat at some point. As shown in panel B of online Appendix Figure A.4, these overrides happen more frequently when the weather is hot outside, but they are infrequent, in general. We see such overrides in about 5 percent of the summer days in our sample window (see Table 1). That statistic exceeds 6 percent on days when the outdoor temperature exceeds 80 degrees and is only 2.8 percent on days when the temperature does not exceed 70 degrees. Thermostat overrides are somewhat equally distributed throughout the hours of the peak period, with the most (20 percent) happening between 4 and 5 PM

Figure 6 relates TOs to our measures of discomfort using an event study that displays the time before and after an individual engages a TO event in our 2018 pre-encouragement data. We split the hour for when each TO event takes place into the minutes before (shown as "Before override" in Figure 6) and the minutes after (shown as "After override"), which allows us to see what indoor conditions lead to a TO.[12] The figure shows that individuals experience relatively low levels of discomfort in the hours before a TO. However, in the minutes proceeding an override, discomfort quickly climbs to high levels before the TO is initiated.[13] Once the thermostat has been overridden, the air conditioning turns on and the indoor temperature and discomfort slowly begins to fall.

In panels B and C of online Appendix Figure A.4, we unpack these changes in discomfort by showing similar figures for indoor temperatures and motion. Indoor temperature (panel B) climbs monotonically up to the point of the override, decreases abruptly, then flattens out. Panel C shows that motion sensor activation also increases gradually, with a jump right before an override. Taken together, these figures suggest that overrides can partially be explained by people coming home to relatively warm homes and overriding their thermostat settings to reduce indoor temperatures (thus producing more comfort). The increase in motion is not large enough to explain all overrides, suggesting that overrides are also driven by individuals that are home experiencing gradual increases in indoor temperatures.

Figure 6 (and online Appendix Figure A.4) provides supporting evidence that (a) TOs are related to indoor conditions and (b) our measures of discomfort capture a disamenity that people take actions to avoid using their air conditioning. Adjusting a thermostat on hot days is a relatively low-cost action that one can take to improve their comfort. Figure 6 shows that when *ATW* and *MTW* increase, regardless of the cause, individuals are observed taking measurable actions to improve their indoor environment.

III. Empirical Framework

In this section, we describe our empirical strategy to estimate the effect of the Eco+ TOU feature on our primary outcomes of interest. First, we estimate the

[12] On average, the Before and After override categories are approximately 30 minutes each, which suggests that the timing of a thermostat override is unrelated to what minute it is within the hour.

[13] We find a similar pattern of increasing discomfort before a TO event when looking at the five-minute raw thermostat data.

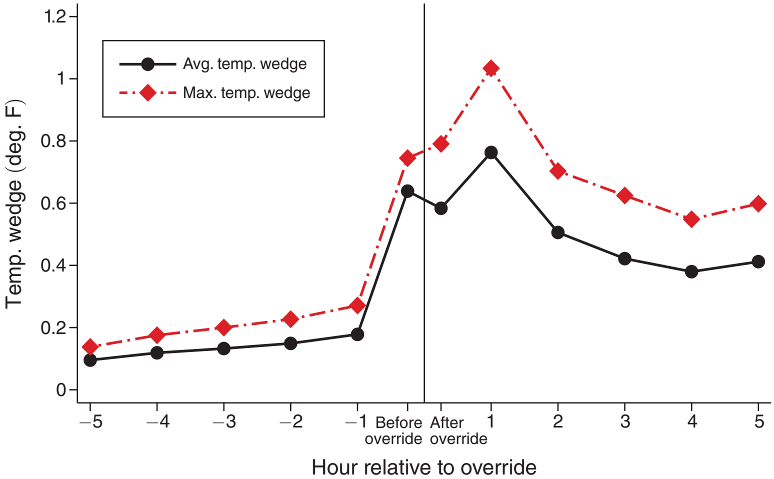

Figure 6. Dynamics of Thermostat Overrides

Notes: The figure plots the hourly average profiles of average and maximum temperature wedges in the hours up to and after a thermostat override event in peak periods. The hour for when each TO event takes place is split into the minutes before (shown as "Before override" in Figure 6) and the minutes after (shown as "After override"). On average, the before and after override categories are approximately 30 minutes each. Data represented is for summer 2018 only.

intent-to-treat (ITT) effects from our experiment, focusing on setpoint temperatures, compressor run time, and our two measures of discomfort as outcome variables. Next, we describe how we estimate local average treatment effects in the context of our experiment for both outcome variables during peak pricing periods. We then estimate the degree to which consumer behavior moderates the effectiveness of our algorithmic treatment.

A. *Estimating Intent-to-Treat Effects*

We estimate the effectiveness of the Eco+ TOU feature in a generalized difference-in-difference framework. We limit our sample to weekdays—when TOU pricing is in effect—from July through September in 2018 and 2019. We specify the following estimating equation:

$$(3) \qquad Y_{imdht} = \sum_{k=1}^{4} \left(\beta_k \times Encouraged_i \times Post_t \times \mathbf{1}\{Period_h = k\}\right) + \mathbf{W}'\gamma + \alpha_{imdh} + \lambda_t + \varepsilon_{imdht},$$

where Y_{idmht} is our outcome of interest for household i, in month m, on day of week d, during hour of day h, and hour of sample t. In our primary specifications, our outcome variables vary at the household (i)-by-hour (t) level. $Encouraged_i$ is a dummy variable that takes the value of one for customers in the encouragement group, and $Post_t$ is a dummy variable for the posttreatment period (treatment began on August 6, 2019). To measure the extent to which our encouragement affects our outcome variables over the day, we interact our encouragement indicator in the posttreatment period with indicators for each of our four pricing periods, $\mathbf{1}\{Period_h = k\}$, capturing separate treatment effects for off-peak, mid-peak AM, peak, and mid-peak PM periods. $\mathbf{W}$ is a vector of hourly outdoor weather controls, including outdoor temperature and relative outdoor humidity. α_{imdh} is a household-by-month-of-year-by-day-of-week-by-hour-of-day fixed effect, which is a set of fixed effects that absorbs time-variant characteristics at a fine-grained temporal level within the household. It allows us to control for occupant behaviors, preferences, and characteristics such as what hour an individual arrives home after work on Tuesdays in August. λ_t is an hour-of-sample fixed effect that picks up weather and other common shocks in our sample.[14] We two-way cluster all standard errors at the household and hour-of-sample level.

With our rich set of fixed effects, our ITT estimates are identified using both within-household variation before and after encouragement and by the difference in outcomes between encouraged and not-encouraged households in a given hour of the day for a specific day of the week in the same month of two consecutive years. The hour-of-sample fixed effect absorbs common shocks (e.g., a sports game or weather event) within each hour of the sample. As an example, this setup compares the difference between an encouraged and not-encouraged household's outcome on Tuesdays from 5 PM to 6 PM in August in 2018 (before encouragement) with the same Tuesday time window in 2019 (after encouragement), conditional on weather and household-specific time effects. This framework allows us to isolate the treatment effects while controlling for an individual household's daily and weekly schedule.

We focus on several outcomes of interest: thermostat setpoint (in degrees Fahrenheit), compressor run time (in minutes per hour), and our two discomfort variables: average and maximum temperature wedges (in degrees Fahrenheit multiplied by motion). This suite of outcome variables captures the full sequence of events that occur because of changes in thermostat settings, whether those changes were initiated by the Eco+ algorithm or household behavior. The thermostat setpoint is the target indoor temperature, which can be changed instantaneously. This variable jointly captures the degree to which the Eco+ algorithm adjusts the setpoint to reduce energy use, the customer's schedule of preprogrammed thermostat settings, as well as any manual changes to the setpoint that the customer makes. When the setpoint increases, the AC compressor will turn off until the house naturally warms up and the new setpoint temperature is reached. If the setpoint is decreased, the AC compressor may run to cool the house to that target temperature, which results in changes in discomfort (as measured by deviations from a counterfactual setpoint schedule).

[14] We explore alternative specifications to our econometric design in Tables A.2-–A.4, which shows how our results change when incorporating less comprehensive sets of fixed effects or limiting the sample to the post-encouragement period while controlling for mean outcomes in the pre-encouragement period (Hahn and Metcalfe 2021). Those results are largely consistent with our primary specifications but with larger standard errors, which suggests that our α_{imdh} term helps with the precision of our estimates. In addition, online Appendix Figure A.5 displays a series of monthly ITT plots that show that parallel pre-trends hold for setpoints and compressor run time.

B. *Estimating Local Treatment Effects*

Our experiment tests the role of the automated TOU feature on setpoints, compressor activity, and in-home comfort. The feature is activated primarily during the peak period of the day, when electricity is most expensive (see Figure 5), although we explore its effects during all pricing periods. In this section, we describe our strategy to estimate local average treatment effects (LATEs) that measure the overall magnitude of our treatment effects for households that complied with the treatment. To do so, we estimate the following two equations via 2SLS:

$$Y_{imdht} = \sum_{k=1}^{4} \left(\delta_k \times \widehat{TOU}_{it} \times \mathbf{1}\{Period_h = k\} \right) + \mathbf{W}'\gamma + \alpha_{imdh} + \lambda_t + \varepsilon_{imdht} \tag{4a}$$

$$\widehat{TOU}_{it} \times \mathbf{1}\{Period_h = k\} = \eta_k \times Encouraged_i \times Post_t \times \mathbf{1}\{Period_h = k\} + \mathbf{W}'\gamma + \alpha_{imdh} + \lambda_t + \varepsilon_{imdht}, \text{ for } k = 1, 2, 3, 4. \tag{4b}$$

Equation (4b) is the first stage and equation (4a) is the second stage. $\widehat{TOU}_{it}$ is a dummy variable equal to one for any weekday in the posttreatment period after the TOU feature activates for the first time. $\mathbf{1}\{Period_h = k\}$ is a dummy variable equal to one for each of the four pricing periods: off peak (7 PM–7 AM), mid-peak AM (7 AM–11 AM), peak (11 AM–5 PM), and mid-peak PM (5 PM–7 PM). $\widehat{TOU}_{it} \times \mathbf{1}\{Period_h = k\}$ is predicted in the set of four first-stage regressions (equation (4b)) using randomized encouragement interacted with the posttreatment period and the pricing period indicators as the instruments. All other variables are the same as defined in equation (3). In additional specifications, we estimate equations (4a) and (4b) with hourly indicators, rather than pricing-period indicators, but we estimate a separate regression for each hour of the day to reduce computational run time.

The $\hat{\delta}_k$ coefficients are our LATE estimates of the Eco+ automated TOU feature on our outcomes of interest (setpoints, compressor run time, and discomfort) for compliers during each of the pricing periods. The instrument satisfies the exclusion restriction based on the random assignment of households to encouraged and not-encouraged groups. Our LATE estimates account for both (1) compliance with our randomized encouragement design and (2) customers who enabled Eco+ but did not provide the required electricity rate information for the TOU feature to activate. 80 percent of encouraged households complied with encouragement and approximately 38 percent of those households supplied the requisite information about their energy utility's rate structure (see Section IB for a discussion of the enrollment steps). As a result, we expect first-stage coefficients of approximately 0.3 ($= 80\% \times 38\%$). In practice, we find that this estimate is slightly lower because not all customers complied with encouragement immediately, resulting in some "untreated" TOU days in the posttreatment period for compliers.[15]

C. *Estimating the Role of Behavior*

The LATE estimates in the previous section reflect the combined effects of the Eco+ programming and the potentially mitigating behavior that individuals might adopt in response to the algorithmic treatment. The success of the automated TOU feature depends on the willingness of customers to let the program control their thermostat and the degree of in-home temperature changes. But customers may react to higher indoor temperatures in several ways. First, if customers are uncomfortable, they may choose to leave their home, perhaps to work in an air-conditioned coffee shop or go to the movies. Second, customers may reduce the stringency or disable the TOU feature by adjusting the slider setting of Eco+. Third, customers may override the thermostat's features by adjusting the temperature, which places a "hold" on the indoor temperature setpoint and overrides any predetermined schedule or features that operate via Eco+. These actions take different levels of effort, ranging from high transaction costs for leaving the house to low transaction costs of overriding the thermostat via smartphone or the thermostat itself. We find that the latter is the most common way by which customers interact with their thermostat. We construct empirical tests of each of these actions to estimate the degree to which they change the effectiveness of TOU automation.

We observe motion within the home through a motion sensor on the thermostat and any other sensors a user installs in their home. We use these motion readings to determine if the TOU feature causes individuals to leave their home, which would reduce measured motion. To do so, we estimate a version of equation (3) with motion as the dependent variable. Because we code motion as a dummy variable, we estimate this specification as a linear probability model. We expect discomfort to be highest during the peak period, and we also expect daytime hours to be the most flexible in terms of individuals being willing to leave their home. So, we focus on whether our randomized encouragement reduces motion during peak hours.[16]

Second, we explore whether customers weaken, or disable, the Eco+ program. Recall that when enrolling in the Eco+ program, customers must choose a slider setting ranging from 1 to 5 in unit increments, with a default value of 4 during the initial setup. Choosing a value of 1 functionally disables the Eco+ features. We only observe daily slider settings for our encouraged households, so we cannot explore how slider settings change experimentally. Instead, we provide descriptive evidence on how encouraged households adjust their slider settings and the extent to which default effects persist throughout our sample.

[15] TOU activation happens relatively quickly, with the feature activating for 40 percent of compliers on the first day post encouragement, and for 82 percent of complier households within the first two weeks post-encouragement. The first instance of TOU feature activation took an average of 8 days for customers who eventually turned on the Eco+ TOU feature.

[16] We explore the distribution of our motion variables further in online Appendix Figures A.6–A.8.

Lastly, we leverage our ability to measure thermostat overrides (TOs) to explore how low-cost adjustment of the temperature setpoint reduces the effectiveness of the Eco+ TOU feature on hours when customers override their thermostat settings. Building on equations (3) and (4), we split the TOU indicator into periods when a TO is in effect and periods when it is not using the thermostat override variable described in Section IID. Specifically, we estimate:

$$
\begin{aligned}
(5) \qquad Y_{imdht} = {} & \mu \widehat{TOU}_{it} \times \mathbf{0}\{TO\}_{it} + \pi \widehat{TOU}_{it} \times \mathbf{1}\{TO\}_{it} \\
& + \nu_1 \mathbf{1}\{TO\}_{it} + \nu_2 \mathbf{1}\{TO\}_{it} \times Post_t + \mathbf{W}'\gamma + \alpha_{imdh} + \lambda_t + \varepsilon_{imdht},
\end{aligned}
$$

where $\mathbf{0}\{TO\}$ is an indicator for the time in the peak period when there is not a thermostat override and $\mathbf{1}\{TO\}$ is an indicator for the time when a thermostat override is in effect. We leverage our five-minute data to split the hour when a TO is enacted into two observations that represent the minutes before and the minutes after the override. For example, if a TO occurs at 9:15 AM, then we create a new Y observation for the first 15 minutes of the hour (with $\mathbf{0}\{TO\} = 1$) and a new observation for the latter 45 minutes of the hour (with $\mathbf{1}\{TO\} = 1$). Our approach allows us to separate the discomfort experienced before a TO, which could contribute to the override happening, from the discomfort experienced after the override goes into effect. To account for these split-hour observations, we weight observations in the regression by the number of minutes each observation occupies, allowing us to interpret the regression coefficients as changes in Y per hour. In practice, this weighting has a minimal impact, because only approximately 1.5 percent of hourly observations are split.

We interpret the $\hat{\mu}$ coefficient as the hourly effect of being a TOU complier on outcome Y when there is not a TO, and the $\hat{\pi}$ coefficient as the hourly effect of being a TOU complier on outcome Y when a TO is in effect. We include interactions for thermostat overrides captured by the $\hat{\nu}$ coefficients. All other variables are the same as in equation (3), but the sample is restricted to the peak period. We estimate equation (5) in our LATE framework where we instrument for $\widehat{TOU}_{it}$ with interactions between an indicator for encouragement and indicator for the posttreatment period for hours with and without thermostat overrides. We are primarily interested in whether $\hat{\mu}$ and $\hat{\pi}$ differ, which captures how thermostat overrides may mitigate the net effects of the Eco+ TOU feature. We explore the role overrides play for each of our outcome variables introduced previously: setpoints, compressor run time, and our two measures of discomfort.

IV. Results

We now turn to our primary results on the net effects of automated responsiveness to TOU pricing on setpoints, energy use, and in-home discomfort. We then present results on the role that individual behavior in the form of thermostat overrides and slider settings plays in moderating the net effects.

Table 2—Intent-to-Treat Estimates by Pricing Period

	(1) Setpoint	(2) Comp. run time	(3) Avg. temp. wedge	(4) Max. temp. wedge
Encouraged × *post* × *mid-peak AM*	0.012	−0.123	0.028	0.035
	(0.099)	(0.175)	(0.024)	(0.026)
Encouraged × *post* × *peak*	0.398	−1.728	0.049	0.059
	(0.103)	(0.277)	(0.023)	(0.025)
Encouraged × *post* × *mid-peak PM*	0.090	−0.661	0.049	0.079
	(0.092)	(0.348)	(0.029)	(0.032)
Encouraged × *post* × *off-peak*	−0.075	0.079	0.011	0.018
	(0.081)	(0.188)	(0.014)	(0.015)
Observations	6,037,377	6,037,377	5,884,617	5,884,617
Households	2,133	2,133	2,133	2,133
Pre-period control mean	74.7	13.5	0.32	0.43

Notes: The table presents intent-to-treat effects from estimating equation (3) for four primary outcomes of interest: setpoints (in deg. F), compressor run time (in min/hr), and average and maximum temperature wedges (in deg. F). All specifications include household-by-month-by-day-of-week-by-hour-of-day and hour-of-sample fixed effects. Standard errors are two-way clustered at the household and hour-of-sample level.

A. *The Effect of Automation on Setpoints*

First, we document in Table 2 how the randomized encouragement affects setpoints via estimation of equation (3). In the first column, setpoints increase by approximately four-tenths of a degree on average across all encouraged households in the peak period. This ITT effect is a combination of the change in setpoint driven by the Eco+ algorithm inclusive of any manual changes or overrides the customer makes during this period. There are no significant changes in setpoints in any other pricing period.

Next, in column 1 of Table 3, we estimate LATE coefficients for setpoints obtained from estimating equations (4a) and (4b), which better capture the changes that occur for LATE compliers in our experiment arising solely from the Eco+ TOU feature. The first-stage estimate for the peak period is 0.24, which accounts primarily for noncompliance with encouragement. The instrument is strong, as evidenced by a first-stage F-stat larger than 100, and it satisfies the exclusion restriction by virtue of randomized encouragement status. These coefficients show that indoor temperature setpoints increase by 1.68 degrees on average during the peak period, which is the average adjustment that the Eco+ TOU algorithm makes compared to household's July 2019 counterfactual setpoint. A change of 1.68 degrees is a nontrivial adjustment to in-home temperature settings and suggests that the algorithm is operating on margins that are large enough to change both energy use and discomfort. In all other pricing periods, the LATE estimate is small and insignificant, confirming that the TOU feature operates primarily during the peak period.

In panel A of Figure 7, we summarize these effects graphically over the course of the day by presenting hourly LATE coefficients. There is a notable and stable increase in setpoints throughout the peak period, with no statistical differences from zero during other hours of the day. This result primarily captures the success of the algorithm in increasing setpoints, but also reveals that any potential behavioral

TABLE 3—LATE ESTIMATES BY PRICING PERIOD

	(1) *Setpoint*	(2) *Comp. run time*	(3) *Avg. temp. wedge*	(4) *Max. temp. wedge*
$\widehat{TOU}$ × *mid-peak AM*	0.053	−0.520	0.117	0.149
	(0.416)	(0.741)	(0.100)	(0.110)
$\widehat{TOU}$ × *peak*	1.678	−7.279	0.205	0.249
	(0.431)	(1.185)	(0.096)	(0.105)
$\widehat{TOU}$ × *mid-peak PM*	0.378	−2.785	0.208	0.333
	(0.390)	(1.478)	(0.122)	(0.134)
$\widehat{TOU}$ × *off-peak*	−0.317	0.333	0.048	0.075
	(0.343)	(0.792)	(0.057)	(0.063)
Observations	6,037,377	6,037,377	5,884,617	5,884,617
Households	2,133	2,133	2,133	2,133
Peak control mean	75.3	8.25	0.33	0.41
Peak first-stage coefficient	0.24	0.24	0.24	0.24
Peak first-stage *F*-stat	103.5	103.5	99.4	99.4

Notes: The table presents local average treatment effects from estimating equations (4a) and (4b) for four primary outcomes of interest: setpoints (in deg. F), compressor run time (in min/hr), and average and maximum temperature wedges (in deg. F). $\widehat{TOU}$ is a binary variable that takes a value of 1 for a given household in all hours and on all days after the household experiences its first Eco+ TOU event. All specifications include four endogenous variables: the interactions of $\widehat{TOU}$ with the four pricing periods: $\widehat{TOU}$ × *Period*. In the first stage, these variables are instrumented for with randomized encouragement status interacted with the four pricing periods: *Encouraged* × *Post* × *Period*. All specifications include household-by-month-by-day-of-week-by-hour-of-day and hour-of-sample fixed effects. Standard errors are two-way clustered at the household and hour-of-sample level.

reaction to the thermostat's algorithm does not offset the algorithm's effectiveness completely.

B. *The Effect of Automation on Compressor Run Time*

We present ITT estimates of the effect of the TOU feature on compressor in column 2 of Table 2. On average, encouraged households' compressors run 1.73 fewer minutes per hour during the peak period than those of not-encouraged households. We also estimate reductions in energy use in the mid-peak PM period, suggesting some lingering effects that persist beyond the peak TOU period. Coefficients in the off-peak and mid-peak AM period are statistically similar to zero.[17]

The ITT effects provide an estimate—averaged across the hours within each pricing period—of how compressor run time changes for households in our encouragement group, including those who did not activate Eco+. In column 2 of Table 3, we present LATE coefficients, along with estimates of the first-stage coefficient, from estimating equations (4a) and (4b) for all pricing periods. The LATE coefficient in the peak period is −7.28, showing that compliers in our experiment ran their air conditioners for 7.28 fewer minutes during the typical peak-period hour in the posttreatment period. Relative to the compressor run time of control

[17] Online Appendix Figures A.9 and A.10 show ITT results from a modified event study version of equation (2) for the four primary outcomes, where weekly effects are estimated relative to the week prior to encouragement. Effects are strongest in August and begin to dissipate later into September with the arrival of cooler weather.

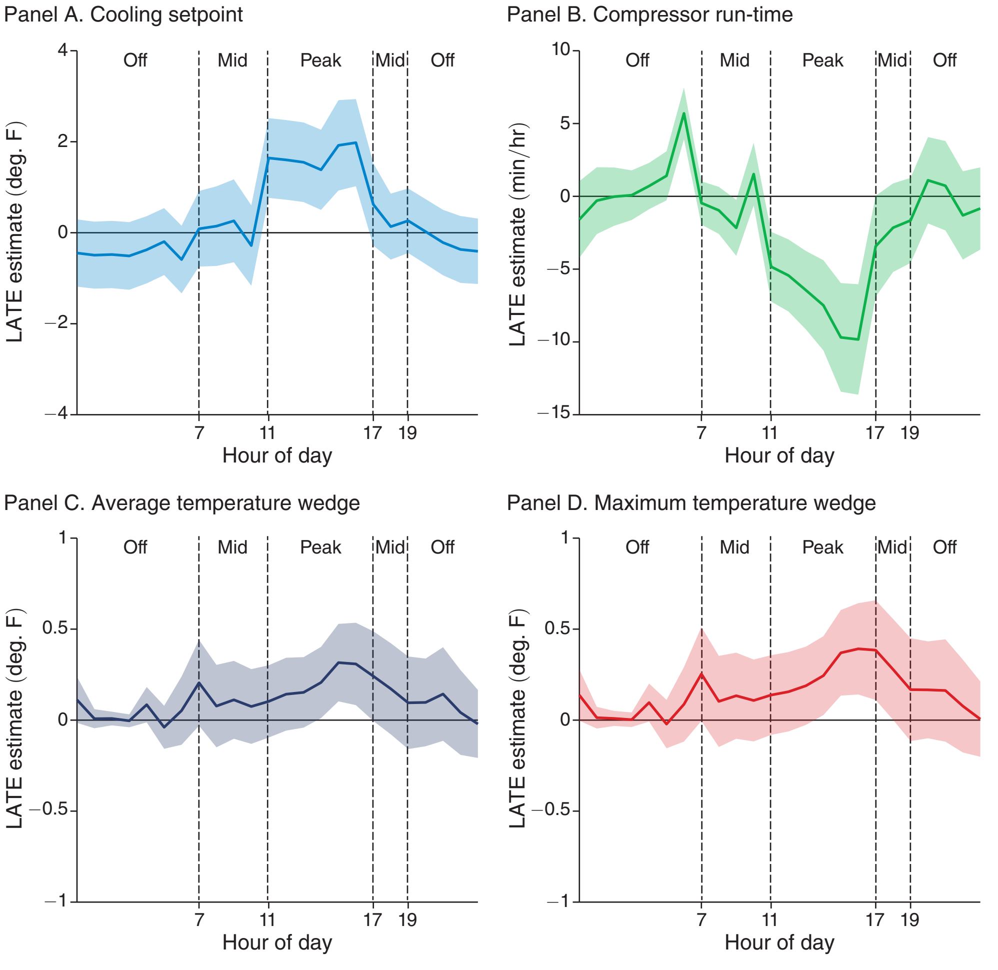

Figure 7. Hourly LATE Estimates of Automated TOU Feature on Outcomes of Interest

Notes: The figure presents local average treatment effects of TOU feature activation on each of the primary outcomes of interest. Twenty-four hour-specific variations of equation (4) are run for each outcome variable, where $\widehat{TOU}$ is instrumented for with encouragement status in each hour: *Encouraged* $\times$ *Post* $\times$ *Hour*. Ninety-five percent confidence intervals for the hourly coefficients are denoted by the shaded areas. TOU electricity rate periods are denoted by the dashed lines. All specifications include household-by-month-by-day-of-week-by-hour-of-day and hour-of-sample fixed effects. Standard errors are two-way clustered at the household and hour-of-sample level.

households, this effect translates to an 88 percent ($= -7.3/8.25$) reduction in run time. These energy savings spill over into the mid-peak PM period in which we see smaller, but significant (p-value of 0.06), reductions in compressor run time (2.79 minutes per hour). We do not find evidence that the TOU feature significantly changes compressor run time, on average, during other pricing periods.

To highlight how the treatment effects evolve over the course of the day, we plot hourly LATE coefficients from estimation of equations (4a) and (4b) in panel B of Figure 7 with compressor run time as our outcome variable. We observe two main results. First, compressor run time increases briefly, but sharply, at 6 AM, the hour before the mid-peak pricing period, which is a precooling effect. We estimate a similar spike in the hour immediately before the peak period as well, although this effect is not statistically different from zero. Both peaks are also mirrored by small but statistically insignificant decreases in thermostat setpoints in panel A of Figure 7. Aside from the precooling effect observed at 6 AM, all off-peak coefficients are close to zero and statistically insignificant. Second, we estimate large reductions in compressor run time during the peak period that become larger later in the day. Each of these hourly coefficients is statistically different from zero at the $p < 0.01$ level. This result is notable because our treatment effects are estimated relative to not-encouraged households, which have paid TOU prices for years and have the same type of smart and easily programmable thermostat. In other words, we anticipate that some not-encouraged households are already optimizing their schedule for the prevailing TOU rates. The automated TOU feature facilitates additional adjustments on top of the behavior of relatively savvy energy consumers. Overall, the time profile in panel B of Figure 7 shows that the automated TOU feature in the Eco+ rollout can substantially reduce, and to a lesser extent shift, compressor usage throughout the day in response to the TOU rates.

C. *The Effect of Automation on Discomfort*

In addition to compressor run time, we are interested in the extent to which changes in the time profile of energy use driven by automated TOU responsiveness affect in-home comfort, which aligns more closely with what consumers value than do measures of electricity use alone. We present ITT estimates for both measures of discomfort in Table 2. During the peak period, coefficients for *ATW* and *MTW* for all encouraged households are 0.049 and 0.059, respectively, suggesting increases in peak period discomfort. In off-peak and morning mid-peak periods, this effect is small and statistically similar to zero. In the evening mid-peak period, discomfort is significantly positive and the coefficient is 0.049 for *ATW* and 0.079 for *MTW*.

The LATE coefficients for our measures of discomfort (and the corresponding first-stage results) are presented in Table 3, alongside the previously discussed results for setpoints and compressor run time. The results for discomfort mirror that of compressor run time. During the peak period, we see increases of both *ATW* and *MTW* of approximately 0.21–0.25 degrees on average in the posttreatment period.[18] Both coefficients are statistically significant. Although these estimates are relatively small degree changes, these effects translate to an approximately 62 percent and 61 percent increase in discomfort for average and maximum temperature wedges, respectively, during the typical peak hour for compliers. Additionally, these discomfort results are different from the average LATE for setpoints in the peak period (1.68 degrees) for a few reasons. First, both measures of discomfort are scaled by whether motion is observed in the home, which moderates this coefficient. Second, as shown in Figure 7, setpoints change immediately after the peak period begins

[18] Online Appendix Table A.5 shows that these coefficients are robust to the inclusion of previously dropped data due to potentially faulty motion sensor data.

at 11 AM, but it takes about three hours on average for indoor temperatures to rise enough to produce discomfort.

We also estimate that discomfort spills over into the mid-peak PM period despite setpoints returning close to normal in these periods. The coefficient is similar across the peak and mid-peak PM periods for *ATW* and slightly larger in the mid-peak PM period for *MTW*. The *MTW* effects are larger in the mid-peak PM period than in the peak period because the mid-peak PM period is only two hours compared to a six-hour peak period, and temperatures are high at 5 PM after limited air conditioning usage during the peak period.

The increase in discomfort caused by the TOU feature can be seen with hourly detail in panels C and D of Figure 7. As shown, the time profile of average temperature wedges throughout the day largely mimics changes in energy use, although it moves in the opposite direction. During off-peak periods, the effect on discomfort is statistically similar to zero. During the mid-peak AM period, discomfort is positive, but in all cases, the 95 percent confidence intervals overlap with zero. During the peak period, discomfort starts to increase beginning at 11 AM and is statistically significant in the late-afternoon hours. These effects are similar, although slightly larger in absolute magnitude, for our maximum temperature wedge also shown in Figure 7. The discomfort in the mid-peak PM period starts to decline after 5 PM, and is statistically insignificant by 7 PM

Overall, the algorithmic effects on discomfort are relatively modest. On average, customers that enabled the Eco+ TOU feature experience less than half a degree of deviation from their preferred temperature per hour. This statistic is only slightly larger when focused on maximum deviations from preferred temperatures. This result implies that the energy savings that occur as a result of the program come with relatively small discomfort costs.

D. *Occupancy Heterogeneity*

We construct indicators of occupancy heterogeneity to explore what types of households see energy savings and/or experience discomfort. To do so, we separate households into one of the three groups—Often, Sometimes, and Hardly Home customers—based on the tercile of motion observed within the home during the peak period in July 2019 before encouragement. We then estimate our primary LATE regressions on compressor run time and our measures of discomfort and plot the resulting hourly coefficients in Figure A.11 in the online Appendix, with corresponding coefficients by rate period presented in online Appendix Tables A.6 and A.7. We find that setpoints increase during the peak period for all occupancy groups during the peak period, with the largest increases for the Hardly Home group. For all occupancy types, we estimate nearly identical results to the full sample for compressor run time: there is a brief increase before the beginning of the mid-peak AM and peak pricing period, with 5- to 10-minute-per-hour decreases in compressor run time during the peak pricing period. There are no meaningful differences in how the TOU feature operates across the occupancy groups.

For discomfort, however, we do observe differences across household types. First, Often Home households experience a statistically significant increase in discomfort of about 0.5 degrees per hour at 7 AM for *ATW* and *MTW* when prices increase. Second, we estimate increases in peak-period discomfort, which increases monotonically from 11 AM until 5 PM similar to the full sample, and we find that the magnitude of the ITT effect is substantially larger for the Often Home customers, averaging between 0.4 and 1.0 degrees per hour for both *ATW* and *MTW*. These increases in discomfort persist into the evening mid-peak period but decrease towards zero. In other words, Often Home households experience more discomfort during the peak periods than the full sample, in addition to some discomfort in mid-peak periods. For Sometimes Home and Hardly Home households, we observe no statistically significant changes in discomfort for any hour of the day. Because we continue to observe reductions in peak-period compressor usage for Sometimes Home and Hardly Home households, this result suggests that automation can save energy with little to no discomfort costs because these occupants are not home to experience them. Thus, although we estimate peak-period energy savings for all groups, the costs of discomfort are borne entirely by the customers who are typically home throughout the day. This result is somewhat mechanical because our measures of discomfort are scaled by observed motion. However, it provides insight that the Eco+ TOU algorithm can reduce energy consumption used to cool an empty house and that there are not significant discomfort costs when people return to a slightly warmer home.

E. *Does Behavior Erode the Effectiveness of Automation?*

The previous sections estimate the net effects of encouragement to enable the Eco+ TOU algorithm. The resulting effects on setpoints, compressor run time, and discomfort reflect both the automation of the Eco+ TOU feature and the behavior of the thermostat users. In this section, we disentangle the degree to which consumer behavior mitigates the effectiveness of the feature. We first focus on whether customers leave the home or disable the automated feature, providing evidence that this channel of adjustment is inconsequential. Next, we provide evidence that thermostat overrides do erode the impacts of the feature, but that these overrides are not frequent enough to mitigate the effectiveness of the algorithm in an appreciable way.

First, in Figure 8, we present hourly results from estimating equations (4a) and (4b) with motion as the outcome variable. Because our previous results provide evidence that the TOU feature increases in-home discomfort during the peak period, we might expect customers to leave their home as a result of this discomfort. If this story were true, we would see a decrease in motion for encouraged households during the peak periods (and perhaps the evening mid-peak period). As shown, we observe no statistically significant changes in motion for the encouraged group (relative to the control group) throughout the day. The coefficients are tightly estimated in the nighttime, when occupants are likely sleeping and although there is more variation throughout the day, the magnitudes are small (at most, we see an approximately two-minute relative increase in motion per hour). We present LATE coefficients for each pricing period in online Appendix Table A.8 for both the number of minutes motion is observed within the hour as well as an indicator for any motion observed within the hour. These results confirm the findings of Figure 8; we find

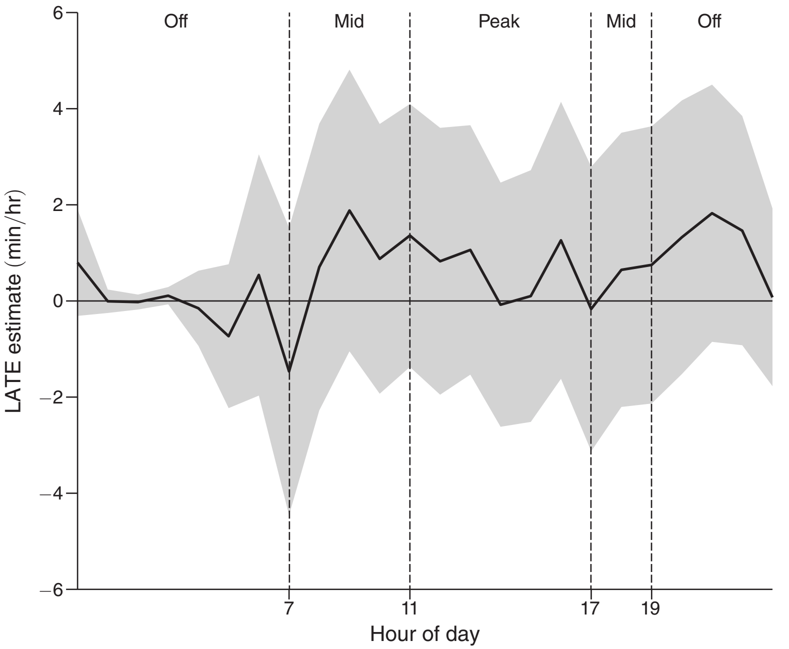

Figure 8. Hourly LATE Estimate of Automated TOU Feature on Motion

Notes: The figure presents local average treatment effects of TOU feature activation on motion detection (min/hr). Twenty-four hour-specific variations of equation (4) are run for motion detection, where $\widehat{TOU}$ is instrumented for with encouragement status in each hour: *Encouraged* × *Post* × *Hour*. Ninety-five percent confidence intervals for the hourly coefficients are denoted by the shaded areas. TOU electricity rate periods are denoted by the dashed lines. The specification includes household-by-month-by-day-of-week-by-hour-of-day and hour-of-sample fixed effects. Standard errors are two-way clustered at the household and hour-of-sample level.

no statistical changes in motion due to encouragement. This result is not surprising because the experienced temperature wedge is less than half of a degree (see LATE coefficients in Table 3) and leaving home is a relatively costly behavioral response. We suspect that experienced discomfort would have to be much greater to trigger customers to leave their home.

Next, we explore whether customers simply disable the TOU feature by adjusting the feature's settings. As mentioned previously, customers can adjust how strongly the feature operates, and they can effectively turn off the feature by choosing the lowest slider setting in the Eco+ settings as a result of discomfort. We explore changes in slider settings directly, but we only observe these values for customers in the encouraged group. In online Appendix Figure A.12, we show the average slider setting across all customers enrolled in Eco+ TOU over the entire six-month period for which we have slider data. We see virtually no change over time in the average slider setting (right axis), which is just below 4, the default setting. We also see that of the roughly 1,100 households enrolled in Eco+ over this time horizon, typically fewer than 10 adjust their slider values on the average day, and roughly the same number of households adjust the values *up* (indicating a willingness to sacrifice even more comfort for savings) as adjust their settings down from August 2019 (the start of the experimental period) through January 2020. Our experimental regression sample ends at the end of September, because temperatures become cool enough to obviate the need for air conditioning, but we study slider settings through January 2020 to explore longer-run behavioral responses to the Eco+ TOU feature.

We present further evidence of changes in slider settings in online Appendix Table A.9. This table is a transition matrix that shows how consumers adjusted or did *not* adjust their slider settings from their initial choices (when they signed up for the Eco+ program) relative to the last day we observe data (January 27, 2020). Most households are clustered along the diagonal, implying that 72 percent of consumers do not deviate from their initial slider values. For the 318 customers who do adjust their settings, 55 percent adjust the settings downward, opting for less savings and more comfort. The other 45 percent of the adjusters increase their slider settings, suggesting an even greater willingness to surrender control of the thermostat after having experienced its effects than they had when they initially signed up. Only 7 percent (76/1151) of households deactivated the Eco+ program by turning their slider settings to 1 during the experiment, while another 100 households effectively deactivated Eco+ during their enrollment (i.e., by choosing the lowest setting during setup).

Finally, we explore the role of thermostat overrides, which is the behavioral adjustment with the lowest transaction costs. In Table 4, we present results from estimation of equation (5). These coefficients estimate what happens on hours with and without thermostat overrides during the peak pricing period.[19,20] The coefficients for $\widehat{TOU} \times \mathbf{1}\{TO\}$ captures the average hourly effect of being a complier when a thermostat override is in effect compared to the control group. The coefficients for $\widehat{TOU} \times \mathbf{0}\{TO\}$ capture the average hourly effect of being a complier when a thermostat override is not active.[21] First, we see that on non-override hours, the encouraged group have thermostat setpoints that are 1.71 degrees higher than the not-encouraged group, which is similar to our primary coefficient in Table 2. On hours with overrides, however, this coefficient is -0.40 and statistically similar to zero, suggesting that when encouraged household's override their settings in the peak period, they effectively offset the algorithm's change in their setpoint degree-for-degree. These coefficients are statistically different from each other (p-value of 0.04). The changes in setpoints on TO hours lead to similar differences in compressor run time. On non-override hours, complier households have 7.26 fewer minutes of compressor run time, on average, while on override hours this same effect is 0.82 fewer minutes of run time. Despite the large difference in magnitude, these coefficients are not statistically different from each other (p-value of 0.22) due to the large standard error on the $\widehat{TOU} \times \mathbf{1}\{TO\}$ coefficient.

[19] We focus solely on the peak-pricing period for this analysis because of the way in which we define thermostat overrides (see Section IID).

[20] We include, but do not interpret, the $\mathbf{1}\{TO\}_{it}$ By Minute and $\mathbf{1}\{TO\}_{it} \times Post_t$ control coefficients in the table for completeness.

[21] The first-stage estimate for thermostat override days is 0.20, which is lower than the 0.24 on non-override days. The lower first stage is mechanical, because not everyone who turns on the Eco+ TOU feature will have an override day.

Table 4—LATE Estimates of Thermostat Override Behavior

	(1) *Setpoint*	(2) *Comp. run time*	(3) *Avg. temp. wedge*	(4) *Max. temp. wedge*
$\widehat{TOU} \times \mathbf{1}\{TO\}$	−0.398	−0.816	0.115	0.252
	(1.042)	(5.372)	(0.182)	(0.219)
$\widehat{TOU} \times \mathbf{0}\{TO\}$	1.707	−7.261	0.204	0.246
	(0.434)	(1.177)	(0.096)	(0.106)
$\mathbf{1}\{TO\}$ by minute	−1.247	9.523	0.333	0.449
	(0.064)	(0.403)	(0.019)	(0.021)
$\mathbf{1}\{TO\} \times post$	0.086	0.724	−0.298	−0.390
	(0.148)	(0.857)	(0.031)	(0.036)
Observations	1,514,598	1,514,598	1,476,379	1,476,379
Households	2,133	2,133	2,133	2,133
Control mean	75.3	8.14	0.33	0.41
$\mathbf{0}\{TO\}$ first-stage coefficient	0.24	0.24	0.24	0.24
$\mathbf{0}\{TO\}$ first-stage F-stat	206.0	206.0	199.3	199.3
$\mathbf{1}\{TO\}$ first-stage coefficient	0.20	0.20	0.20	0.20
$\mathbf{1}\{TO\}$ first-stage F-stat	39.5	39.5	36.9	36.9

Notes: The table presents local average treatment effects from estimating equation (5) for four primary outcomes of interest: setpoints (in deg. F), compressor run-time (in min/hr), and average and maximum temperature wedges (in deg. F). $\widehat{TOU}$ is a binary variable that takes a value of 1 for a given household in all hours and on all days after the household experiences its first Eco+ TOU event. In the first-stage, $\widehat{TOU} \times \mathbf{1}\{TO\}$ and $\widehat{TOU} \times \mathbf{0}\{TO\}$ are instrumented for with randomized encouragement status interacted with override status: $Encouraged \times Post \times TO$. All specifications include household-by-month-by-day-of-week-by-hour-of-day and hour-of-sample fixed effects. Standard errors are two-way clustered at the household and hour-of-sample level.

Because the air conditioning system is running more frequently during override hours, we expect that discomfort will be mitigated. Of course, discomfort is likely a primary *cause* of thermostat overrides, but also an outcome of the override, as demonstrated graphically in Figure 6. In column 3 of Table 4, we estimate that during override hours our average temperature wedge is 0.12 degrees and not significant. On hours without overrides, the average temperature wedge is 0.20 degrees, which is almost identical to the results in Table 3, which do not take TOs into account. These coefficients are not significantly different from each other, however, due to the large confidence interval on the coefficient for override hours. Column 4 of Table 4 shows the results for maximum temperature wedge. The coefficients for thermostat override and non-override days are both about 0.25, although only the non-override day coefficient is significant. The TO coefficients are likely a similar magnitude because, on many override days, the maximum temperature typically climbs to a high level in the minutes before the thermostat is overridden.

Overall, we explore three natural ways customers could mitigate the effectiveness of the automated TOU-responsiveness feature, all of which are important to understand the benefits and costs of scaling up automated energy programs like these. First, we find that encouraged customers are no more likely to leave their home during the peak periods than not-encouraged customers. Second, among encouraged customers, we find no evidence that customers disable the feature during our study period. Lastly, we do find evidence that customers override their thermostat settings temporarily, and this margin of adjustment exactly offsets the effectiveness of automation in generating energy savings (and, as a result, increases the customers' level of comfort). These override events, however, are relatively uncommon, which suggests that consumer behavior will not substantively reduce the effectiveness of automated energy programs.

V. Discussion

In this section, we contextualize the results of our analysis and discuss some implications for energy policy. So far, we have focused on variables and units of measurement easily calculated using data collected by the smart thermostat. Setpoints, compressor run time, and our derived discomfort measures, although reflective of the energy services consumers experience, are not standard units of measurement in the literature. To improve comparability with other studies, we convert compressor run time to estimates of kWh usage and energy cost savings. We then conduct back-of-the-envelope calculations to characterize the energy savings–discomfort trade-off that households experience and we perform additional analysis to put those results into context.

A. *Energy and Monetary Savings of Eco+*

In many ways, compressor run time is a more intuitive unit of measurement than kWh usage for how much air conditioning a household is using. Most people are aware of when their air conditioner is running, but they may not know how that translates to electricity usage or their monthly utility bills. However, despite its salience, compressor run time does not lend itself to cost-benefit analysis or comparisons with other studies. Unfortunately, we do not have utility bills or the data on each household's air conditioning unit that might allow us to directly calculate changes in energy use. Instead, we rely on back-of-the-envelope conversions based on survey data of Ontario air conditioner characteristics. To convert compressor operation to energy consumption, we use data from Canada's National Energy Use Database (Natural Resources Canada 2019) to generate a scaling factor that converts compressor run time to kWhs based on different assumptions for the efficiency of air conditioning units. We assume that the average size of air conditioners in the Eco+ experiment matches the averages of units across Ontario.[22]

To perform this calculation, we estimate *daily* LATE coefficients for our preferred specification (equations (4a) and (4b)), which measures the average hourly effect of the automated TOU feature for compliers in the posttreatment period.[23] Daily LATE coefficients are hourly averages across all hours of the day, inclusive of increases in run time due to precooling and decreases during the peak period, that can easily be scaled up to an aggregate daily effect by multiplying by 24. In panel A of Table 5, the daily LATE coefficient shows that the Eco+ TOU feature is responsible for reducing compressor run time by 1.97 minutes per hour (or 47.3 minutes per day) in the posttreatment period. Based on the characteristics of typical low- and high-efficiency air conditioning systems in Ontario homes (corresponding

[22] See online Appendix B.4 for further details on the benefit-cost calculations. Additionally, online Appendix Table A.10 outlines parameter values and assumptions used in the benefit-cost calculations.

[23] Formally, these estimates follow our specification in equations (4a) and (4b) with hourly run time as the dependent variable, but we only estimate one coefficient for the day instead of separate coefficients for the four pricing periods. All controls and fixed effects remain the same.

TABLE 5—PRIVATE BENEFITS AND COSTS OF ECO+ TOU FEATURE

	Hourly run-time LATE (min/hr)		*Daily change* (kWh/day/hh)
Panel A. Private benefits (*kWhs*)			
Changes in energy use	−1.97	Low efficiency	−2.18
	(−3.04, −0.90)		(−3.36, −0.99)
		High efficiency	−1.63
			(−2.52, −0.75)
			Daily change (C$/day/hh)
Panel B. Private benefits (*C$*)			
Changes in energy costs		Low efficiency	−C$0.29
			(−C$0.45, −C$0.13)
		High efficiency	−C$0.22
			(−C$0.34, −C$0.10)
			Hourly discomfort (deg/hr)
Panel C. Private costs (*discomfort*)			
Avg. temp. wedge		Sometimes/hardly home	0.01
			(−0.17, 0.19)
		Often home	0.51
			(0.21, 0.81)
Max. temp. wedge		Sometimes/hardly home	0.05
			(−0.14, 0.24)
		Often home	0.59
			(0.26, 0.92)

Notes: The table presents estimates of the private benefits and costs of the Eco+ TOU feature. 90 percent confidence intervals are presented below coefficient estimates in parentheses. Discomfort LATE estimates in panel C are estimated on only the mid-peak and peak samples combined. Further assumptions underlying these calculations are detailed in online Appendix Table A.10.

to SEER ratings of 12 and 16, respectively), we calculate that the 1.97 minutes of compressor run time saved by the automated TOU feature per hour corresponds to an average reduction of 2.18 kWh per TOU day for low-efficiency households and 1.63 kWh for high-efficiency households. Multiplying those savings by the electricity rates that households pay during peak periods—where the majority of the energy conservation arises—provides savings of C$0.29 per day for households with low-efficiency air conditioners and C$0.22 for households with high-efficiency air conditioners (Table 5, panel B). Summer 2019 in Ontario had 93 days hot enough for TOU to turn on for the average Eco+ household, which would result in a total summer savings ranging from C$20.36 to C$27.15 (online Appendix Table A.11, panel A).[24,25] These savings are larger than those in Fowlie et al. (2021), who found

[24] See online Appendix Figure A.13 for a visual representation of summer TOU days. We define the start of the summer TOU day period by looking for the first instance when average hourly compressor run times over the whole day went above one minute for two consecutive days. We define the end of the summer TOU day period as the instance when average hourly compressor run times were under one minute for two consecutive days and never rose back above this threshold. Based on this definition, the summer TOU day period for 2019 was from May 26 to October 4. Counting only the weekdays during this range gives 93 days when TOU might turn on.

[25] The savings estimates are based on a private marginal cost of electricity to the thermostat owner. We use the private marginal cost in Table 5 because adoption of the Ecobee thermostat and Eco+ are primarily being driven by private cost-benefit analysis. An alternate formulation could use the social marginal cost (SMC) of electricity. We consider scenarios where the social marginal cost is C$0.05, C$0.10, and C$0.15. Using our estimated reductions in aggregate energy use, the automated TOU feature would reduce social costs by C$10.14, C$20.27, or C$30.41 for the average low-efficiency household for the entire summer (i.e., $2.18 \times 93 \times SMC$) or C$7.58, C$15.16, or C$22.74 for the average high-efficiency household for the entire summer (i.e., $1.63 \times 93 \times SMC$) for each of the assumed social marginal costs.

that households saved around US$2 per month on TOU pricing when they were not provided with smart thermostats.[26] We also note that Ontario has a relatively mild summer climate, which means the savings are likely a lower bound on the savings that could accrue in a warmer region.

B. *Energy Savings–Comfort Trade-offs*

The energy savings from Eco+ come at the potential cost of discomfort. This is not unique to our setting. Many other energy-saving programs and technologies come with a discomfort trade-off, but it is typically ignored in program evaluations because it is difficult to measure. For example, the Opower household social comparison program has been shown to reduce household energy consumption by 1 to 2 percent at relatively low cost to the utility (e.g., Allcott 2011; Allcott and Rogers 2014), but likely also causes discomfort, which has not been included in cost-benefit analyses. Studies of time-varying pricing (e.g., Jessoe and Rapson 2014; Harding and Lamarche 2016; Bollinger and Hartmann 2019) suffer from the same inability to measure the discomfort caused by reduced energy consumption, much of which likely comes from reduced air conditioning usage and higher indoor temperatures. Fowlie, Greenstone, and Wolfram (2018) come closest to measuring the indoor temperature effects of an energy-saving intervention: they use a one-time post-retrofit temperature measurement to explore changes in indoor temperature from weatherization. They find no changes after the weatherization but also acknowledge that they have a limited and selected subsample. Our experiment, in contrast, gives us the ability to measure discomfort caused by the automated thermostat feature and better understand the trade-off between energy savings and indoor comfort: we can calculate a dollar value that reflects the trade-offs households are currently making between energy savings and indoor comfort.

Similar to compressor run time, we estimate daily *ATW* and *MTW* LATE coefficients. However, we limit our regression sample to the hours of 7 AM to 7 PM, which is when we see the changes in discomfort (see Figure 7).[27] In contrast to the peak period results in Table 3, these daily regressions capture the average hourly discomfort that occurs across this twelve-hour period. We present these results in panel C of Table 5. We split the results into two groups using the insight from Section IVD that changes in discomfort are only experienced by Often Home

[26] In addition to Fowlie et al. (2021), which is focused on TOU pricing and most comparable to our setting, there are a number of papers that examine critical peak pricing programs, which generally use day-ahead price alerts or nudges to encourage users to conserve electricity on the hottest days of the year. Burkhardt, Gillingham, and Kopalle (2019) find a welfare benefit of around US$32 per household across a summer when consumers face 27 peak event days and a peak price of US$0.64 per kWh. The similarity of the results in our paper shows that automation can provide a similar magnitude of benefits. Brandon et al. (2019) study the role that social nudges can play, finding that social nudges can cause a 4 to 7 percent reduction in peak demand.

[27] We estimate this regression following equations (4a) and (4b), although we only estimate one coefficient and we drop hours outside of 7 AM to 7 PM Note that we could estimate these LATE regressions for the full 24-hour period and scale them to the 7 AM to 7 PM period by dividing by 0.5 (i.e., 12/24). These approaches are functionally equivalent. Reporting LATE coefficients for the 7 AM to 7 PM period allows us to avoid artificially attenuating our estimates by averaging in zeroes from the nighttime period. Moreover, we continue to report coefficients as hourly averages for the 7 AM–7 PM period because aggregating degrees of discomfort over time results in unintuitive units.

households. Hardly and Sometimes Home households do not experience meaningful discomfort from the Eco+ TOU feature. In other words, two-thirds of the households in our sample see no significant increase in discomfort from the automated TOU feature, but all household occupancy types experience a C$0.22–0.29 savings from reduced compressor run times, which means there is no material trade-off for us to evaluate.[28] This is a win-win: for a substantial portion of our sample, customers see energy savings without a corresponding decrease in comfort.

Customers who are often home, however, experience an average hourly increase in indoor temperatures of 0.5–0.6 degrees above their preferred temperature during the 7 AM to 7 PM period on a typical day in the posttreatment period. These Often Home households trade this off for a savings of around C$0.22–0.29 per TOU day—a trade-off of roughly C$0.37–0.57 per degree of discomfort within our experiment. The trade-off includes household behavioral responses, which we showed in Section VIE consists of households manually adjusting their thermostat temperature settings. The relative infrequency of thermostat overrides suggests that the transaction cost to avoid discomfort is relatively small and is not substantially influencing our estimated trade-off between discomfort and monetary savings.

C. *Implications*

So far, we have shown that the automated TOU feature significantly reduces peak-period energy usage for treated customers in our sample. Some of those customers experience no discomfort costs, others are willing to bear small discomfort costs for energy savings, and consumer reactions to this discomfort do not appreciably erode the energy-savings benefits from the automated feature. The changes in electricity demand, although modest at the customer level, become more noteworthy when aggregated across larger numbers of thermostat users. To extrapolate these energy savings beyond our sample, we note that in 2017, Ontario's Environment and Climate Change minister initiated a C$40 million program to install 100,000 smart thermostats at no cost to customers across the province through the Green Ontario Fund.[29] Taking the 100,000-thermostat target as given and applying the first-stage compliance rate from our experiment (24 percent) results in 24,000 eligible thermostats activating the TOU feature. Applying our LATE coefficients to those thermostats would result in a reduction in electricity demand from 4 PM to 5 PM of 8–11 MW, depending on assumptions about air conditioning efficiency. This magnitude is comparable to offsetting a quarter of peak generation from a small natural gas peaker plant in Ontario (Blonz 2022). Policies that reduce frictions in enrolling in the program would generate larger savings. If compliance increased from 24 percent to 50 percent, then energy savings could increase to 17–23 MW per peak hour (online Appendix Table A.11, panel B).

[28] Online Appendix Figure A.11 shows that all three occupancy groups experience almost identical compressor run time savings. As a result, we did not break the energy savings in panel A of Table 5 into different occupancy groups because it would have produced the same result.

[29] Source: https://www.thestar.com/news/queenspark/2017/08/30/eligible-ontario-homeowners-to-get-smart-thermostats-under-new-program.html (last accessed September 19, 2024).

Notably, these potential savings come from a software update designed to improve on an existing smart thermostat designed to help people save money through programming and optimization of their cooling. Energy savings may have been realized when the smart thermostat was installed, and our estimated savings are additional.[30] This software update could be pushed to hundreds of thousands of customers with very low marginal cost, which would increase the potential benefits of widespread TOU pricing.

Because the economic incentives for customers to adjust behavior can be small, TOU prices alone may have limited effects on electricity demand. Automation may provide an additional benefit by automatically adjusting household electricity use during peak periods without asking customers to bear large reductions in comfort. Those small reductions in energy use per household can translate to meaningful supply-side cost reductions at scale.

VI. Conclusion

In this paper, we evaluate the large-scale rollout of a new smart-thermostat feature designed to automate consumers' responsiveness to TOU pricing. Our randomized experiment and unique data on motion and indoor temperatures allow us to estimate causal effects of the automated TOU pricing feature on air conditioner run time, households' thermal comfort, and consumer reactions to experienced discomfort. We find that the Eco+ TOU feature is effective at shifting load throughout the day by strategically adjusting thermostat setpoints and reducing compressor usage when electricity is most expensive. On average, households save C\$0.22 to C\$0.29 per summer day in the posttreatment period. The discomfort costs of this feature are borne primarily by people who are typically home during the peak pricing period. Those customers experienced average increases in discomfort less than a half-degree degree warmer than their revealed preference setpoint. Yet, for a large portion of the population, the automated feature delivers energy cost reductions with no increase in discomfort costs. Importantly, we show that behavioral change driven by these increases in discomfort does not inhibit the overall effectiveness of the algorithm. Electricity customers do not leave their house or disable the feature as a result of the encouragement (and resulting discomfort). We do see, however, that on relatively warmer days customers are more likely to override their thermostat settings, but these events are rare and do not undermine the effectiveness of the feature in delivering energy conservation in an appreciable way.

The potential for the use of software to optimize the operation of electrical devices in response to time-varying electricity prices is large. TOU pricing continues to roll out across North America, and other devices, such as hot water heaters and pool pumps, could be adapted to optimize their operation in response to time-varying prices. Our work suggests that such programs could deliver important energy savings for customers and could be designed in a way that customers will not be averse to surrendering control of smart devices to a price-responsive algorithm

[30] The energy savings from installing a smart thermostat depend on what thermostat it replaced. Our findings are not affected by those initial energy savings.

that makes relatively small changes. Such automation could help pave the way for more widespread acceptance of time-varying electricity prices.

Decarbonizing the electricity system to avoid the worst consequences of global climate change will require substantial reliance on variable and intermittent sources of renewable energy, such as wind and solar. In a renewables-abundant future, the efficiency gains from dynamic pricing for balancing demand and supply at each moment in time will be even greater than they are today. Smart devices that enable automated responses to time-varying prices can lower the cost of adjusting consumption and help overcome political opposition to more widespread implementation of time-varying and ultimately dynamic pricing of electricity. With federal and state energy policy targeting almost complete decarbonization of the electricity system in the next 15 years, smart technologies have the potential to be integral to achieving these goals at a reasonable cost.

REFERENCES

Allcott, Hunt. 2011. "Social Norms and Energy Conservation." *Journal of Public Economics* 95 (9–10): 1082–95.

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

Blonz, Joshua A. 2022. "Making the Best of the Second Best: Welfare Consequences of Time-Varying Electricity Prices." *Journal of the Association of Environmental and Resource Economists* 9 (6): 1087–1126.

Blonz, Joshua A., Karen Palmer, Casey J. Wichman, and Derek C. Wietelman. 2025. *Code for: "Smart Thermostats, Automation, and Time-Varying Prices."* Nashville, TN: American Economic Assocation; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E197966V1.

Bollinger, Bryan K., and Wesley R. Hartmann. 2019. "Information versus Automation and Implications for Dynamic Pricing." *Management Science* 66 (1): 290–314.

Borenstein, Severin. 2007. "Wealth Transfers among Large Customers from Implementing Real-Time Retail Electricity Pricing." *Energy Journal* 28 (2): 131–49.

Borenstein, Severin. 2019. "Charging with the Sun." Energy Institute Blog, University of California–Berkeley, March 11. https://energyathaas.wordpress.com/2019/03/11/charging-with-the-sun/ (accessed January 18, 2024).

Borenstein, Severin, and James B. Bushnell. 2022. "Do Two Electricity Pricing Wrongs Make a Right? Cost Recovery, Externalities, and Efficiency." *American Economic Journal: Economic Policy* 14 (4): 80–110.

Borenstein, Severin, and Stephen Holland. 2005. "On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices." *RAND Journal of Economics* 36 (3): 469–93.

Brandon, Alec, Christopher M. Clapp, John A. List, Robert Metcalfe, and Michael Price. 2022. "The Human Perils of Scaling Smart Technologies: Evidence from Field Experiments." NBER Working Paper 30482.

Brandon, Alec, John A. List, Robert D. Metcalfe, Michael K. Price, and Florian Rundhammer. 2019. "Testing for Crowd out in Social Nudges: Evidence from a Natural Field Experiment in the Market for Electricity." *PNAS* 116 (12): 5293–98.

Burkhardt, Jesse, Kenneth Gillingham, and Praveen K. Kopalle. 2019. "Experimental Evidence on the Effect of Information and Pricing on Residential Electricity Consumption." NBER Working Paper 25576.

Bushnell, James, and Kevin Novan. 2021. "Setting with the Sun: The Impacts of Renewable Energy on Conventional Generation." *Journal of the Association of Environmental and Resource Economists* 8 (4): 759–96.

Ecobee. 2019. *Confidential Thermostat Data from Randomized Encouragement Trial*. Toronto, Ontario: Ecobee Donate Your Data Program. https://www.ecobee.com/en-us/donate-your-data/ (accessed November 5, 2019).

Faruqui, Ahmad, Sanem Sergici, Neil Lessem, and Dean Mountain. 2015. "Impact Measurement of Tariff Changes When Experimentation Is Not an Option: A Case Study of Ontario, Canada." *Energy Economics* 52 (Part A): 39–48.

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

Fowlie, Meredith, Catherine Wolfram, Patrick Baylis, C. Anna Spurlock, Annika Todd, and Peter Cappers. 2021. "Default Effects and Follow-On Behavior: Evidence from an Electricity Pricing Program." *Review of Economic Studies* 88 (6): 2886–2934.

Gillan, James Michael. 2018. "Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption." PhD diss. University of California–Berkeley.

Hahn, Robert W., and Robert Metcalfe. 2021. "Efficiency and Equity Impacts of Energy Subsidies." *American Economic Review* 111 (5): 1658–88.

Harding, Matthew, and Carlos Lamarche. 2016. "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* 35 (4): 906–31.

Harding, Matthew, and Steven Sexton. 2017. "Household Response to Time-Varying Electricity Prices." *Annual Review of Resource Economics* 9: 337–59.

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

Jessoe, Katrina, and David Rapson. 2015. "Commercial and Industrial Demand Response under Mandatory Time-of-Use Electricity Pricing." *Journal of Industrial Economics* 63 (3): 397–421.

Joskow, Paul L., and Catherine D. Wolfram. 2012. "Dynamic Pricing of Electricity." *American Economic Review* 102 (3): 381–85.

Lessem, Neil, Ahmad Faruqui, Sanem Sergici, and Dean Mountain. 2017. "The Impact of Time-of-Use Rates in Ontario." *Public Utilities Fortnightly* 87: 56–61.

List, John A. 2007. "Field Experiments: A Bridge between Lab and Naturally Occurring Data." *BE Journal of Economic Analysis and Policy* 6 (2).

List, John, Robert Metcalfe, and Michael Price. 2018. "Smart Meters: Do Their Prices Matter to Their Adoption and Do They Save Energy?" Unpublished.

Meier, Alan, Tsuyoshi Ueno, Marco Pritoni, Leo Rainer, Abigail Daken, and Dan Baldewicz. 2019. "What Can Connected Thermostats Tell Us about American Heating and Cooling Habits?" In *ECEEE Summer Study Proceedings,* 649–57. Stockholm, Sweden: European Council for an Energy Efficient Economy.

Natural Resources Canada. 2019. "Table 27: Cooling System Stock by Type, New Unit Efficiencies, Stock Efficiencies, and Unit Capacity Ratio." Office of Energy Efficiency National Energy Use Database. https://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/showTable.cfm?type=CP&sector=res&juris=on&rn=27&page=0 (accessed January 18, 2024).

Prest, Brian C. 2020. "Peaking Interest: How Awareness Drives the Effectiveness of Time-of-Use Electricity Pricing." *Journal of the Association of Environmental and Resource Economists* 7 (1): 103–43.

Sallee, James M. 2014. "Rational Inattention and Energy Efficiency." *Journal of Law and Economics* 57 (3): 781–820.

Wolak, Frank A. 2011. "Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment." *American Economic Review: Papers and Proceedings* 101 (3): 83–87.