# Smart Thermostats, Automation, and Time-Varying Prices

## Online Appendix

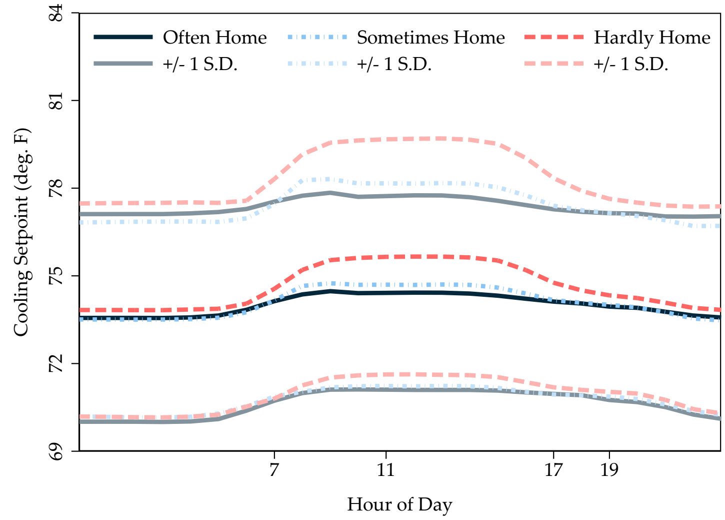
Joshua Blonz    Karen Palmer    Casey J. Wichman    Derek C. Wietelman

January 26, 2024

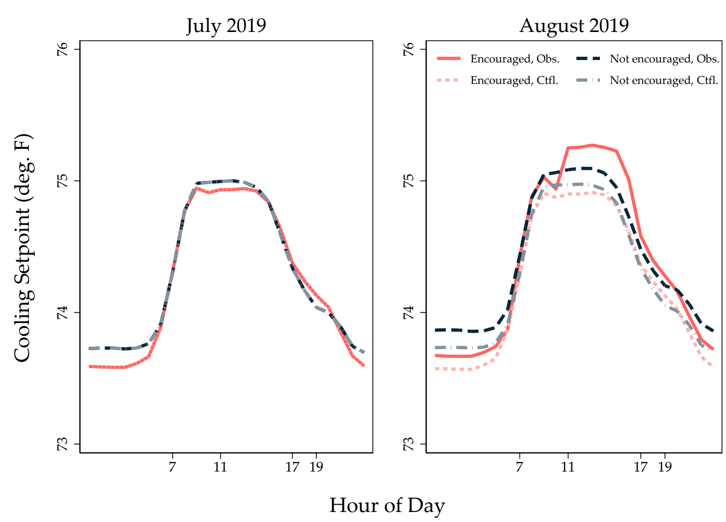
## Table of Contents

# A Additional Results

## A.1 Additional Figures



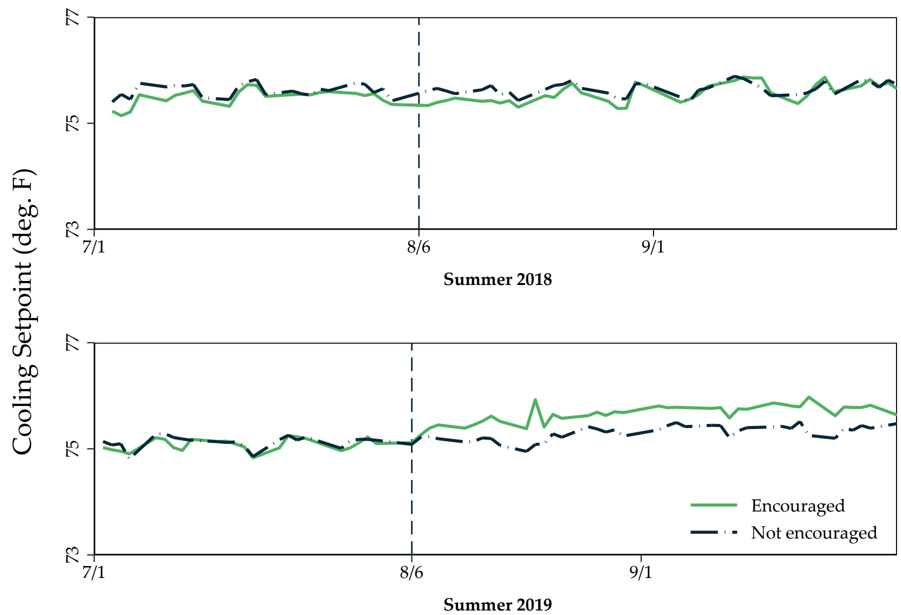

(a) **Counterfactual Setpoints by Occupancy Group**



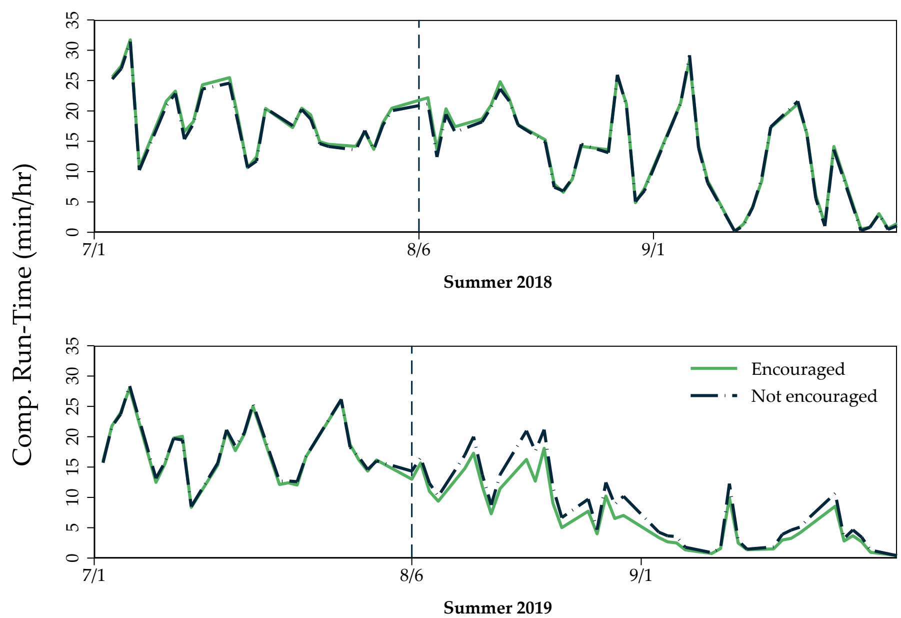

(b) **Effect of Setpoint Counterfactual Substitution**

Figure A.1. Counterfactual Setpoints

*Notes:* The figure illustrates the counterfactual cooling setpoints used to create the temperature wedge variables. Panel a shows the counterfactual setpoints by occupancy group. Panel b shows the effect of replacing the household's observed setpoints with the counterfactual setpoints in August 2019, when Eco+ has begun to directly affect setpoints for encouraged households. The observed and counterfactual lines lie over each other in July 2019 because no substitution takes place: Eco+ is not changing setpoints at that time, so no substitution is necessary.



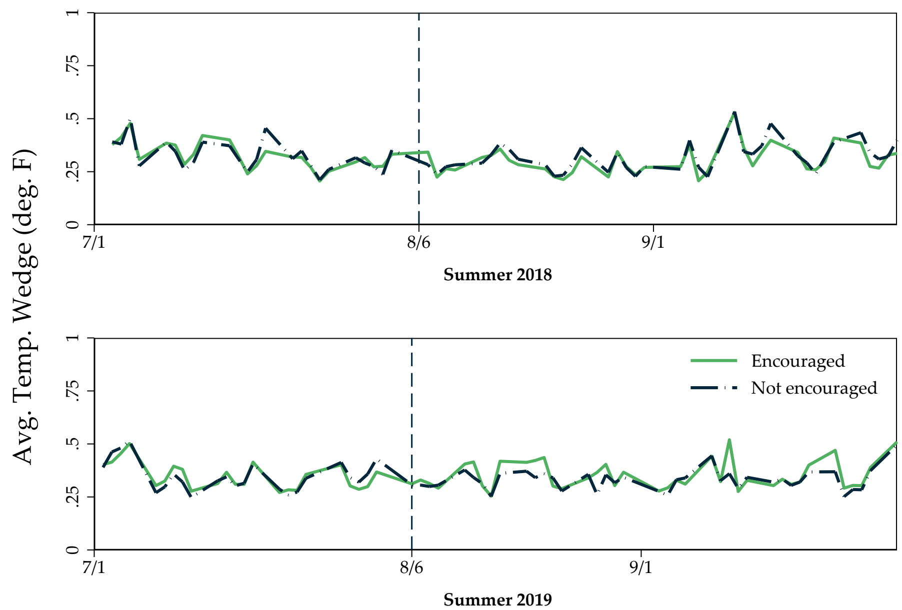

(a) **Cooling Setpoint**



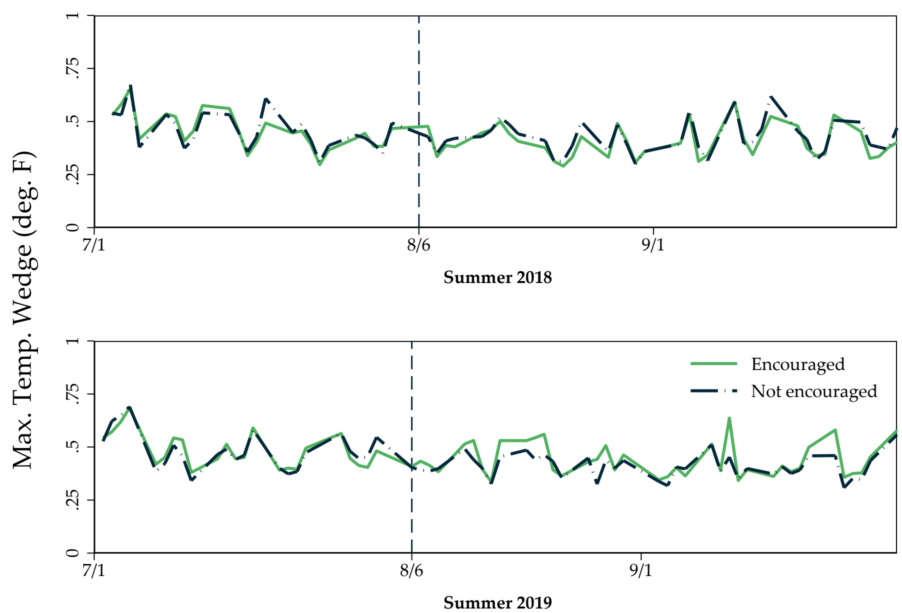

(b) **Compressor Run-Time**

Figure A.2. Daily Peak Period Averages by Encouragement Status

*Notes:* The figure presents daily averages during the peak period for setpoints and compressor run-time. These daily averages are characterized by three differences: before and after the encouragement date (8/6), by encouragement status, and by year (2018 vs. 2019).



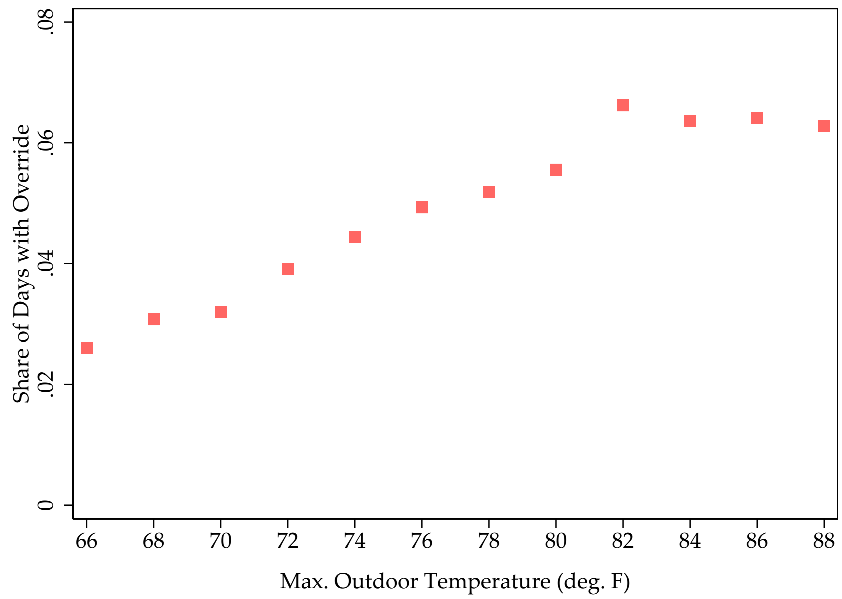

(a) **Average Temperature Wedge**



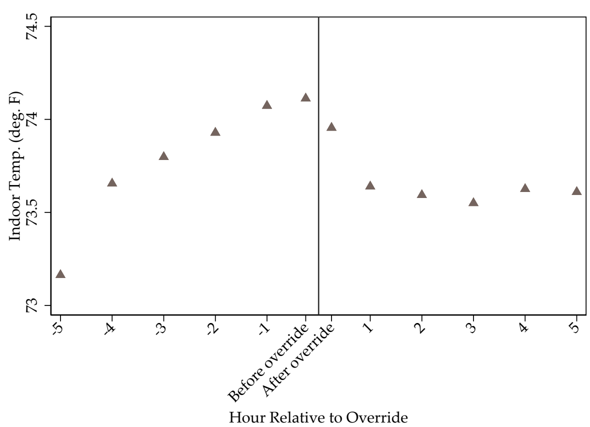

(b) **Maximum Temperature Wedge**

Figure A.3. Daily Peak Period Averages by Encouragement Status

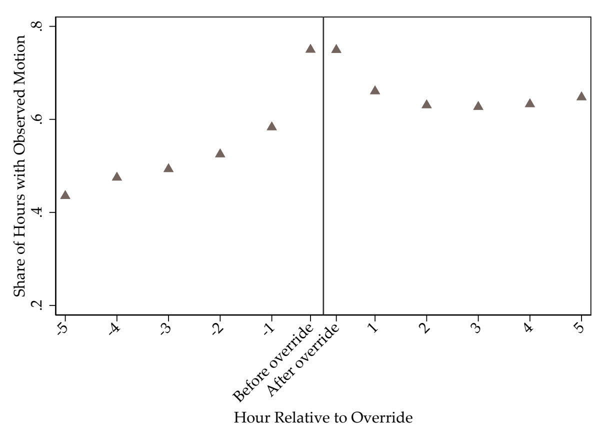
*Notes:* The figure presents daily averages during the peak period for average and maximum temperature wedges. These daily averages are characterized by three differences: before and after the encouragement date (8/6), by encouragement status, and by year (2018 vs. 2019).




(a) **Correlation between Overrides and Outdoor Temperature**



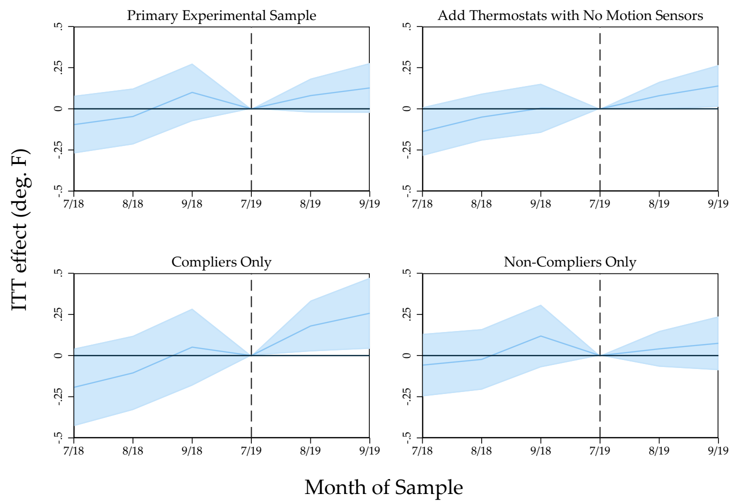

(b) **Indoor Temperature Relative to Override**



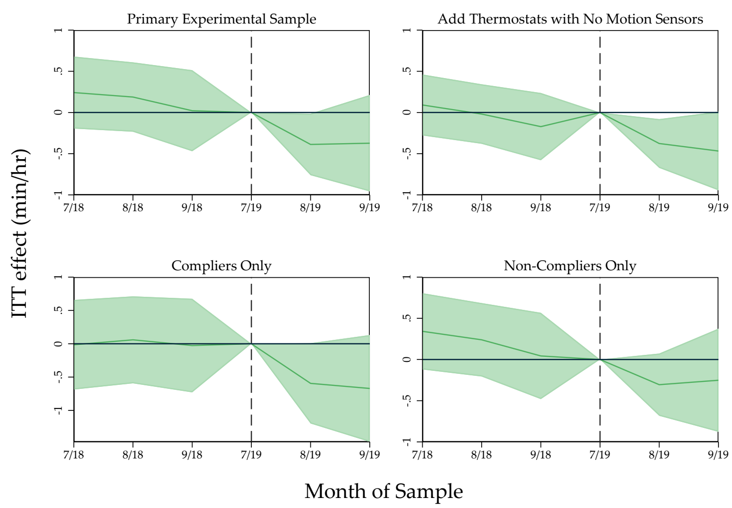

(c) **Observed Motion Relative to Override**

Figure A.4. Dynamics of Thermostat Overrides, Temperature, and Motion

*Notes:* The figure examines relationships between thermostat overrides and indoor and outdoor temperatures, as well as motion. Panel a presents the share of days in summer 2018 with an override event plotted in a series of 2 deg. F max outdoor temperature bins. Panels b and c present an equivalent plot to Figure 6 in the main text, where the profiles of indoor temperature and any observed motion in the hour are plotted in the hours leading up to and after an override event.




(a) **Cooling Setpoint**




(b) **Compressor Run-Time**

Figure A.5. Check for Parallel Pre-Trends in Setpoints and Compressor Run-Time

*Notes:* The figure presents a series of monthly intent-to-treat coefficient plots for setpoints (in panel a) and compressor run-time (in panel b) investigating parallel pre-trends in these outcomes. Trends are assessed separately for the base experimental sample, all thermostats including those dropped for not having motion sensor data, only those who experience TOU events ("Compliers"), and only those that do not experience TOU events ("Non-Compliers"). Regressions underlying these coefficient plots estimate monthly treatment effects relative to July 2019, the month before Eco+ was introduced. All specifications include household and hour-of-sample fixed effects. Standard errors are two-way clustered at the household and hour-of-sample level. 95% confidence bands are represented by the shaded areas surrounding the point estimates.

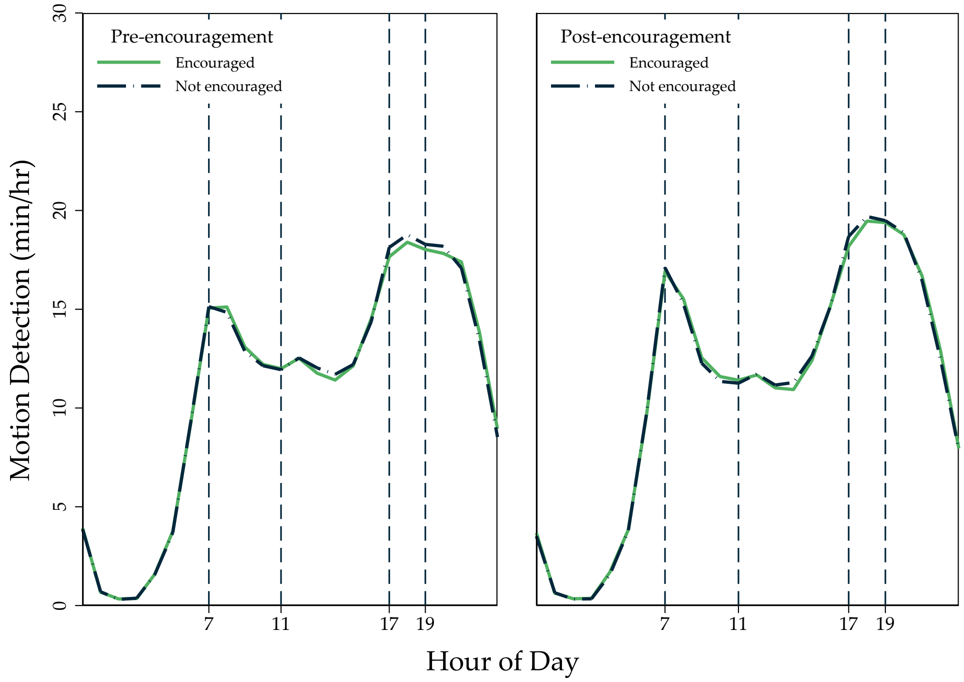



Figure A.6. Hourly Profile of Motion by Encouragement Status

*Notes:* The figure presents average hourly motion by encouragement group in a similar fashion to Figure 5. "Pre-encouragement" refers to the period of July 1, 2019 up to August 6th, 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 a.m. mid-peak rate period, 11-17 is the peak period, 17-19 is the p.m. mid-peak period, and 19-7 is the evening off-peak period.

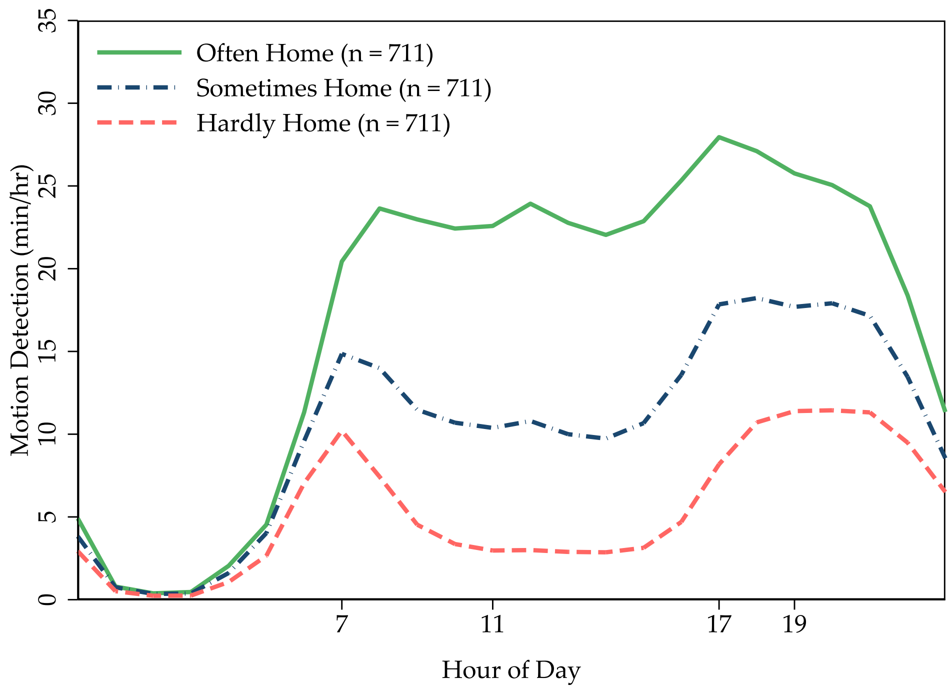



Figure A.7. Hourly Average Motion by Occupancy Group in July 2019

*Notes:* The figure presents average motion profiles for each of three defined occupancy groups (Often, Sometimes, and Hardly home). Occupancy groups are defined by average motion detection during the peak period in July 2019. Customers were sorted into classes by cutting each group at the 33rd and 66th percentile of motion detection, respectively.

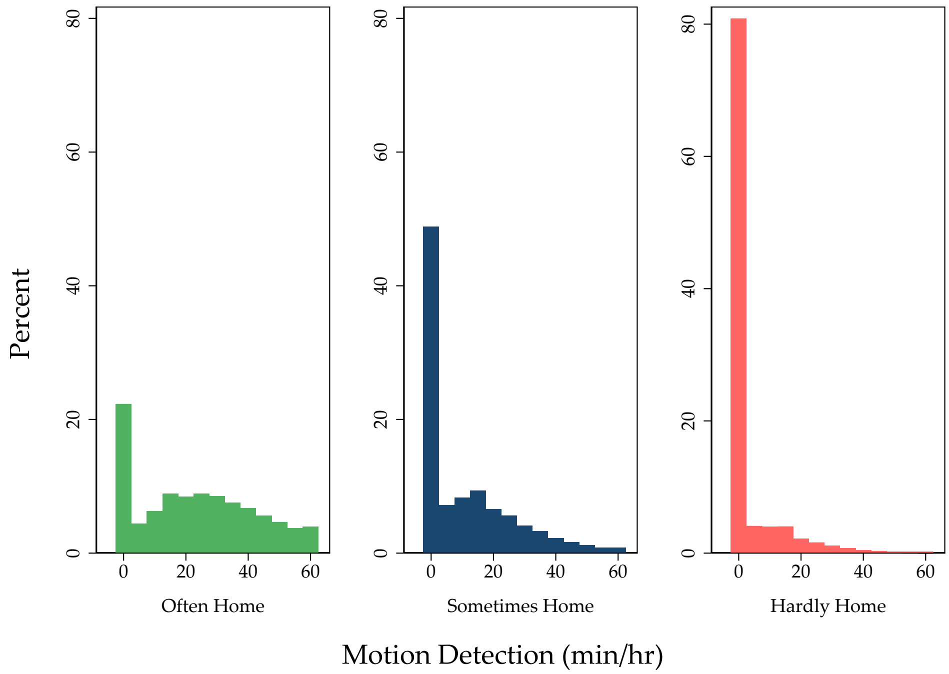
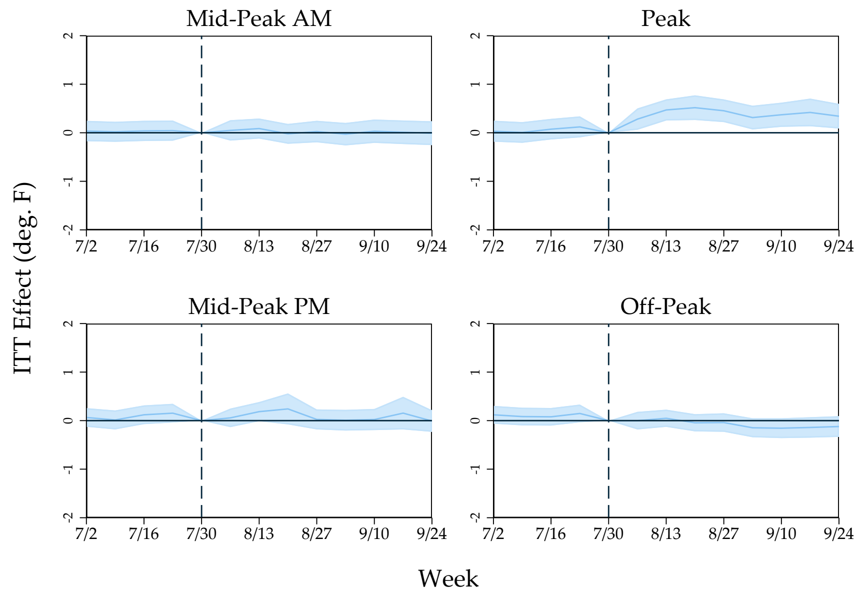
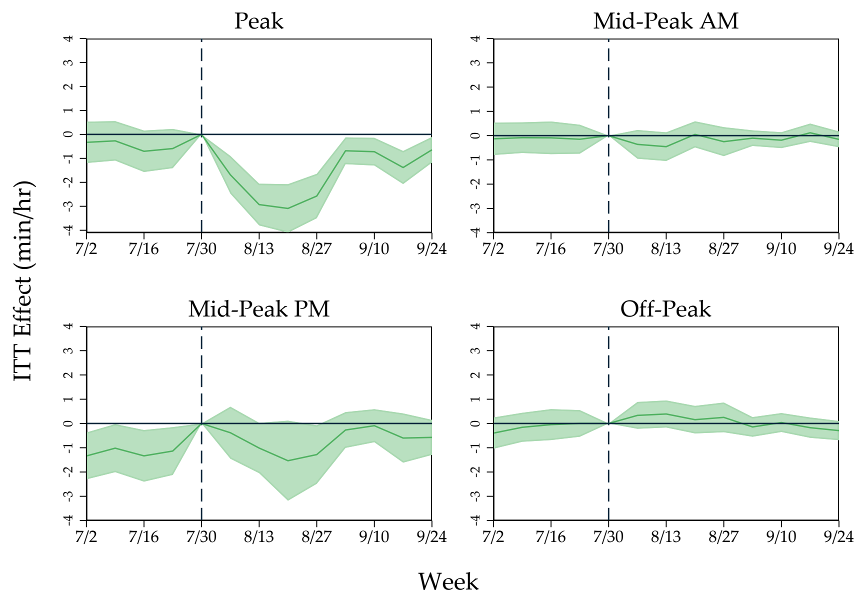



Figure A.8. Motion Histograms by Occupancy Group during Peak Period in July 2019

*Notes:* The figure displays the distributions of average hourly motion detection during the peak period for each of the three occupancy groups in July 2019.



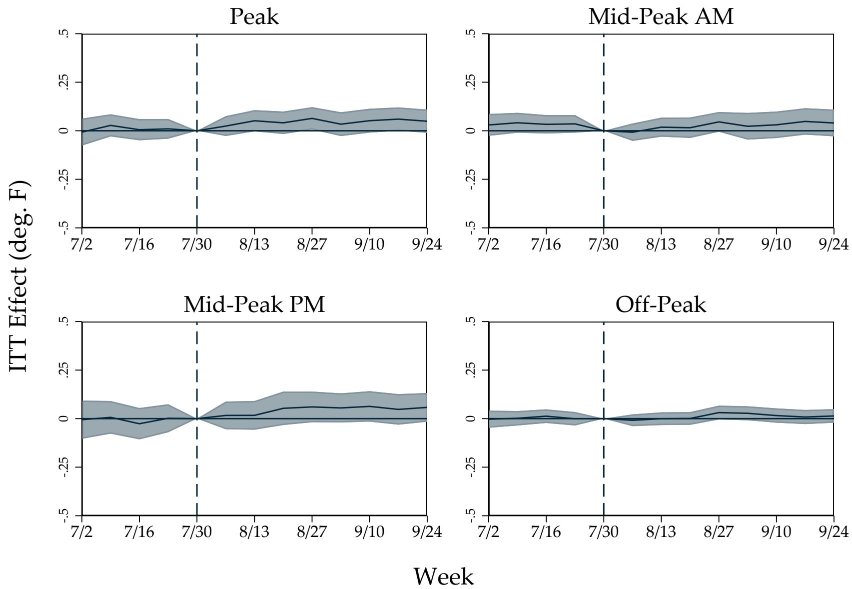

(a) **Cooling Setpoint**



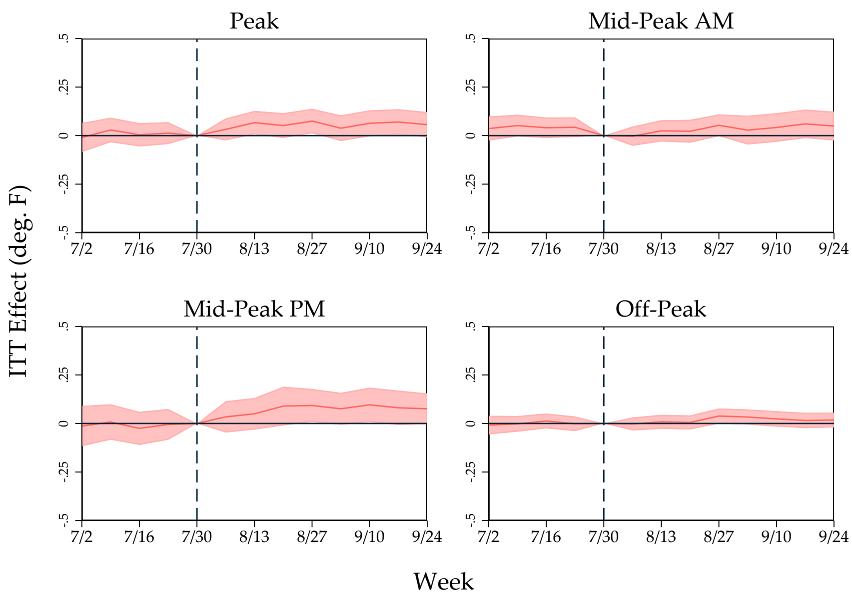

(b) **Compressor Run-Time**

Figure A.9. Event Study ITT Effects by Week of Sample

*Notes:* Coefficients from an event study regression (modified from Equation 2) of the effect of encouragement on cooling setpoints and compressor run-time are presented relative to the week before encouragement. Effects for each rate period are presented separately here. Data for both summer 2018 and summer 2019 are included, and the estimated specification includes household-by-month-by-day-of-week-by-hour-of-day and hour-of-sample FEs. Standard errors are clustered at the household and hour-of-sample level. 95% confidence bands are represented by the shaded areas surrounding the point estimates.



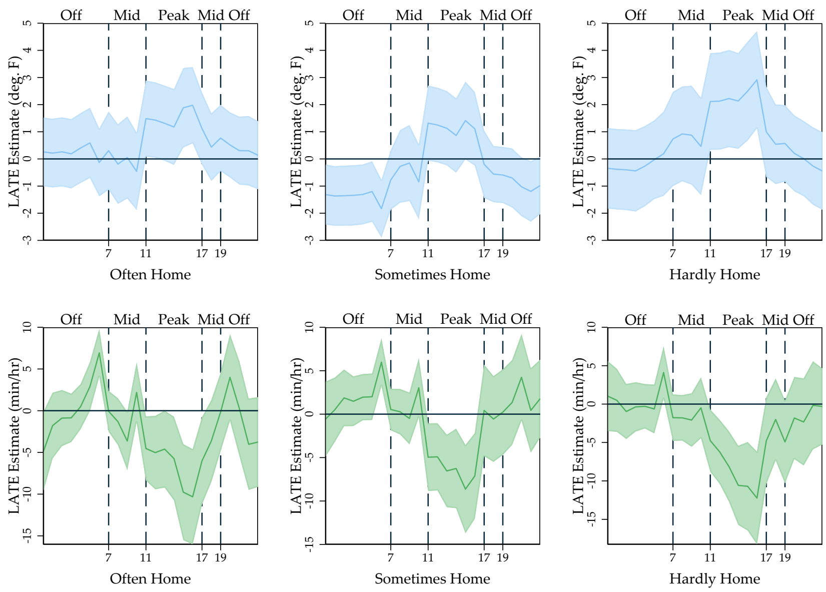

(a) **Average Temperature Wedge**



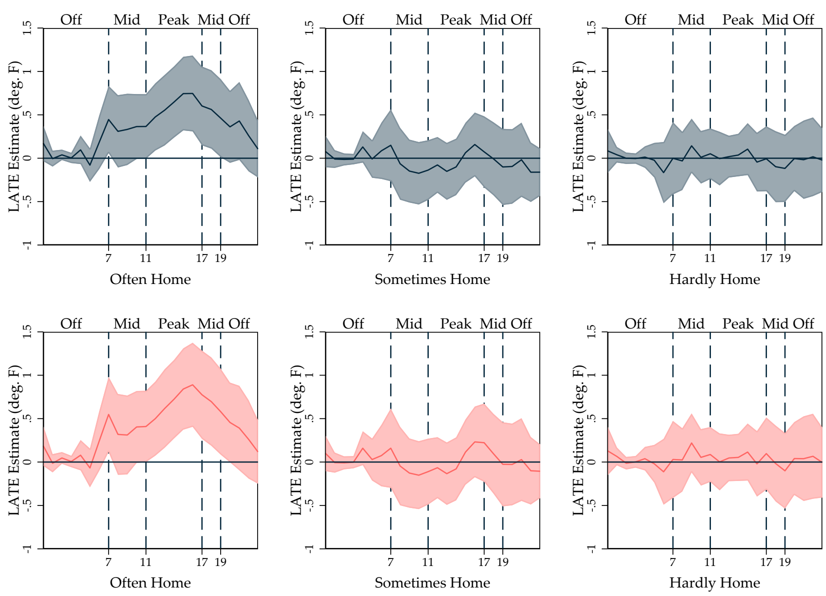

(b) **Maximum Temperature Wedge**

Figure A.10. Event Study ITT Effects by Week of Sample

*Notes:* Coefficients from an event study regression (modified from Equation 2) of the effect of encouragement on average and maximum temperature wedges are presented relative to the week before encouragement. Effects for each rate period are presented separately here. Data for both summer 2018 and summer 2019 are included, and the estimated specification includes household-by-month-by-day-of-week-by-hour-of-day and hour-of-sample FEs. Standard errors are clustered at the household and hour-of-sample level. 95% confidence bands are represented by the shaded areas surrounding the point estimates.




(a) **Setpoints and Compressor Run-Time**




(b) **Average and Maximum Temperature Wedges**

Figure A.11. LATE Effects by Occupancy Group

*Notes:* The figure presents local average treatment effects of TOU feature activation on each of the four primary outcomes of interest, also broken out by the occupancy heterogeneity groups. 24 hour-specific variations of Equations 4a and 4b are run for each outcome variable in each occupancy group separately, where $\widehat{TOU}$ is instrumented for with encouragement status in each hour: $Encouraged \times Post \times Hour$. 95% 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.

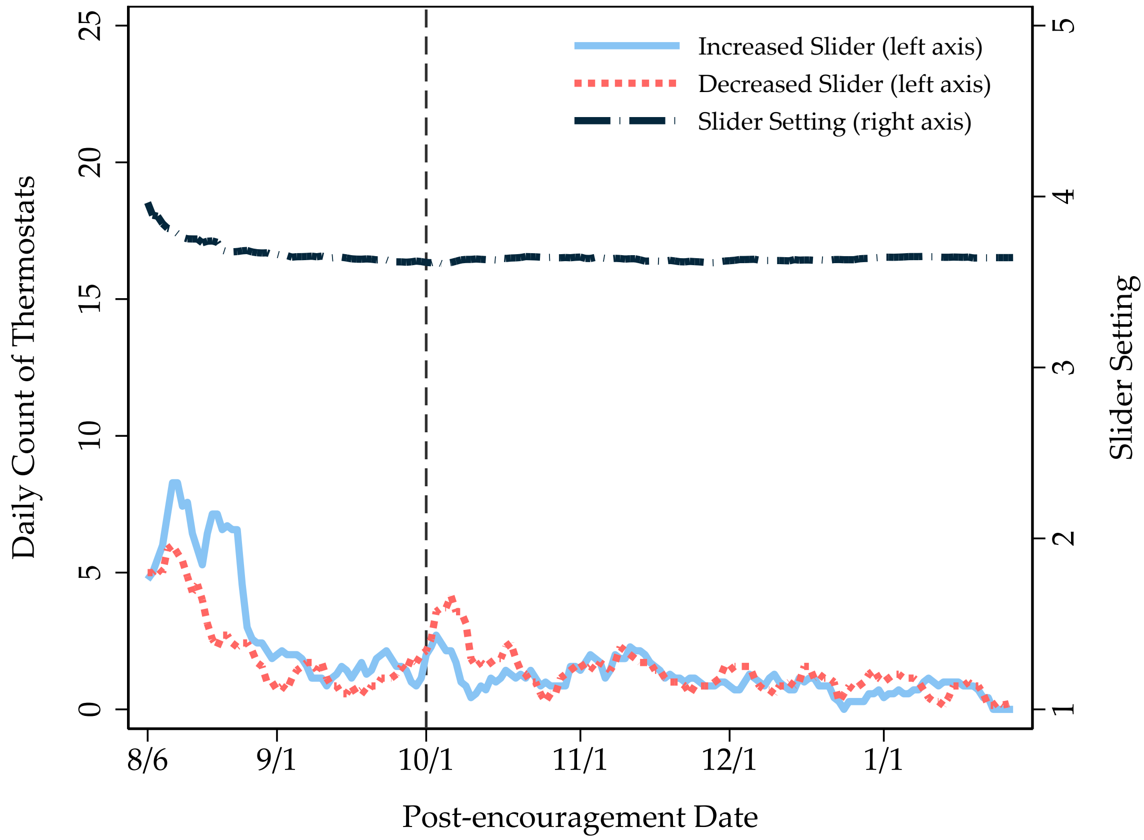



Figure A.12. Mean Daily Slider Values and Thermostat Changes

*Notes:* Averages of slider settings and daily counts of thermostats that changed slider settings from the previous day are presented here as seven-day moving averages. The vertical dashed line represents the end of the experimental analysis period on September 30, 2019. We continue to observe slider data until January 27th, 2020.

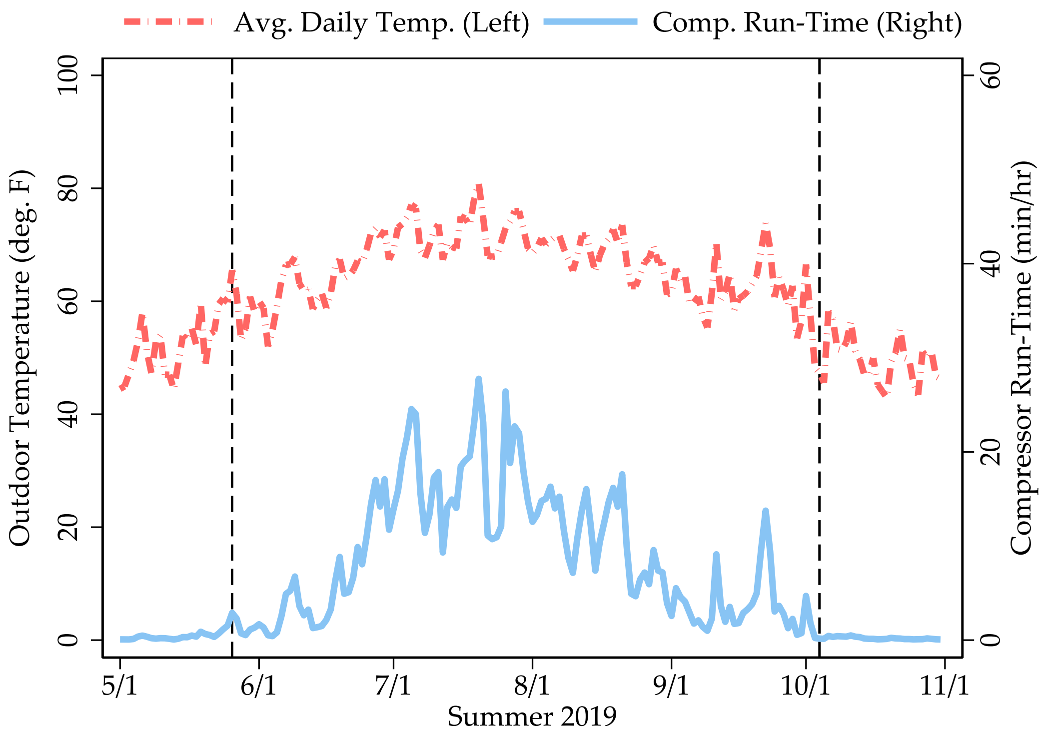



Figure A.13. Defining the Summer TOU Day Period

*Notes:* Average daily outdoor temperatures are indicated by the dashed pink time series and the left vertical axis. Average daily compressor run-times are indicated by the solid blue time series and the right vertical axis. The vertical dashed lines represent the start and end of the assumed "summer TOU day period" (May 26, 2019 to October 4, 2019). The start of the summer TOU day period is defined by the first instance when average hourly compressor run-times over the whole day went above one minute for two consecutive days. 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. Counting only the weekdays during this range gives 93 “summer TOU days.”

## A.2 Additional Tables

Table A.1. Balance on Observables in Preperiod Using All Thermostats

| Variable Name | Not encouraged | Encouraged | Difference |
|---|---|---|---|
| | | *Household Characteristics* | |
| Floor Area (100 Sqft) | 20.35 | 20.85 | 0.50 |
| | (10.24) | (9.96) | (0.36) |
| Number of Floors | 2.61 | 2.59 | -0.02 |
| | (0.73) | (0.75) | (0.03) |
| Home Age (Years) | 31.08 | 30.06 | -1.02 |
| | (30.95) | (29.42) | (1.07) |
| Household Size | 3.06 | 3.05 | -0.00 |
| | (1.21) | (1.25) | (0.06) |
| Model: Ecobee 3 | 0.43 | 0.42 | -0.00 |
| | (0.49) | (0.49) | (0.02) |
| Model: Ecobee 3 Lite | 0.44 | 0.43 | -0.01 |
| | (0.50) | (0.50) | (0.02) |
| Model: Ecobee 4 | 0.09 | 0.08 | -0.00 |
| | (0.28) | (0.28) | (0.01) |
| | | *Preperiod Daytime Characteristics* | |
| Cooling Setpoint (Deg. F) | 75.32 | 75.21 | -0.10 |
| | (4.09) | (4.06) | (0.12) |
| Compressor Run-Time (Mins/hr) | 13.90 | 14.33 | 0.42 |
| | (20.49) | (20.65) | (0.31) |
| Avg. Temp. Wedge (Deg. F) | 0.23 | 0.22 | -0.00 |
| | (0.95) | (0.93) | (0.02) |
| Max. Temp. Wedge (Deg. F) | 0.30 | 0.30 | -0.00 |
| | (1.06) | (1.05) | (0.02) |
| Indoor Temp (Deg. F) | 73.80 | 73.67 | -0.12 |
| | (3.10) | (3.10) | (0.07) |
| Outdoor Temp (Deg. F) | 73.63 | 73.63 | -0.00 |
| | (8.05) | (7.97) | (0.05) |
| Hold (Mins/hr) | 23.48 | 23.31 | -0.16 |
| | (29.07) | (29.05) | (0.79) |
| Thermostat Override | 0.05 | 0.05 | -0.00 |
| | (0.23) | (0.22) | (0.00) |
| | | *Preperiod Nighttime Characteristics* | |
| Cooling Setpoint (Deg. F) | 74.24 | 74.02 | -0.21 |
| | (3.76) | (3.75) | (0.11) |
| Compressor Run-Time (Mins/hr) | 12.43 | 13.23 | 0.81 |
| | (19.45) | (19.89) | (0.29) |
| Avg. Temp. Wedge (Deg. F) | 0.18 | 0.18 | 0.00 |
| | (0.83) | (0.82) | (0.01) |
| Max. Temp. Wedge (Deg. F) | 0.24 | 0.24 | 0.00 |
| | (0.94) | (0.94) | (0.01) |
| Indoor Temp (Deg. F) | 73.56 | 73.39 | -0.17 |
| | (3.12) | (3.12) | (0.08) |
| Outdoor Temp (Deg. F) | 67.15 | 67.21 | 0.06 |
| | (7.28) | (7.23) | (0.08) |
| Hold (Mins/hr) | 24.09 | 23.74 | -0.35 |
| | (29.19) | (29.12) | (0.78) |
| Thermostat Override | 0.05 | 0.05 | -0.00 |
| | (0.22) | (0.22) | (0.00) |

*Notes:* Sample means and standard deviations (in parentheses) are presented for characteristics of the randomized encouragement and control groups for all experimental thermostats, including those dropped from the primary sample (n=3,402). 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.

Table A.2. ITT Fixed Effects Robustness Checks for Setpoints and Compressor Run-Time

| | Setpoint | | | Comp. Run-Time | | |
|---|---|---|---|---|---|---|
| | (1) From Paper | (2) HH FE | (3) No $\alpha$ | (4) From Paper | (5) HH FE | (6) No $\alpha$ |
| Encouraged × Post × Mid-Peak AM | 0.012 | 0.081 | 0.011 | -0.123 | -0.341 | 0.138 |
| | (0.099) | (0.091) | (0.157) | (0.175) | (0.273) | (0.205) |
| Encouraged × Post × Peak | 0.398 | 0.435 | 0.370 | -1.728 | -2.005 | -1.525 |
| | (0.103) | (0.099) | (0.166) | (0.277) | (0.286) | (0.360) |
| Encouraged × Post × Mid-Peak PM | 0.090 | 0.195 | 0.128 | -0.661 | -0.564 | -0.083 |
| | (0.092) | (0.090) | (0.149) | (0.348) | (0.345) | (0.478) |
| Encouraged × Post × Off-Peak | -0.075 | -0.039 | -0.111 | 0.079 | 0.178 | 0.658 |
| | (0.081) | (0.081) | (0.144) | (0.188) | (0.228) | (0.276) |
| Observations | 6,037,377 | 6,037,906 | 6,037,906 | 6,037,377 | 6,037,906 | 6,037,906 |
| Households | 2,133 | 2,133 | 2,133 | 2,133 | 2,133 | 2,133 |
| Pre-period Control Mean | 74.7 | 74.7 | 74.7 | 13.5 | 13.5 | 13.5 |

*Notes:* The table presents robustness checks for intent-to-treat effects estimated from Equation 2 to the inclusion of different levels of $\alpha$, the household fixed effect term, for setpoints and compressor run-time. Columns 1 and 4 reproduce the estimates using the household-by-month-by-day-of-week-by-hour-of-day fixed effect in Table 2. Columns 2 and 5 relax $\alpha$ to an overall household fixed effect. Columns 3 and 6 present results omitting the $\alpha$ term entirely. All specfications include the common hour-of-sample fixed effect. Standard errors are two-way clustered at the household and hour-of-sample level.

Table A.3. ITT Fixed Effects Robustness Checks for Temperature Wedges

| | Avg. Temp. Wedge | | | Max. Temp. Wedge | | |
|---|---|---|---|---|---|---|
| | (1) From Paper | (2) HH FE | (3) No $\alpha$ | (4) From Paper | (5) HH FE | (6) No $\alpha$ |
| Encouraged × Post × Mid-Peak AM | 0.028 | 0.033 | 0.028 | 0.035 | 0.039 | 0.036 |
| | (0.024) | (0.023) | (0.027) | (0.026) | (0.025) | (0.030) |
| Encouraged × Post × Peak | 0.049 | 0.027 | 0.023 | 0.059 | 0.033 | 0.031 |
| | (0.023) | (0.020) | (0.026) | (0.025) | (0.022) | (0.029) |
| Encouraged × Post × Mid-Peak PM | 0.049 | 0.021 | 0.017 | 0.079 | 0.045 | 0.043 |
| | (0.029) | (0.025) | (0.033) | (0.032) | (0.028) | (0.037) |
| Encouraged × Post × Off-Peak | 0.011 | 0.013 | 0.009 | 0.018 | 0.020 | 0.017 |
| | (0.014) | (0.015) | (0.015) | (0.015) | (0.016) | (0.017) |
| Observations | 5,884,617 | 5,885,435 | 5,885,435 | 5,884,617 | 5,885,435 | 5,885,435 |
| Households | 2,133 | 2,133 | 2,133 | 2,133 | 2,133 | 2,133 |
| Pre-period Control Mean | 0.32 | 0.32 | 0.32 | 0.43 | 0.43 | 0.43 |

*Notes:* The table presents robustness checks for intent-to-treat effects estimated from Equation 2 to the inclusion of different levels of $\alpha$, the household fixed effect term, for average and maximum temperature wedges. Columns 1 and 4 reproduce the estimates using the household-by-month-by-day-of-week-by-hour-of-day fixed effect in Table 2. Columns 2 and 5 relax $\alpha$ to an overall household fixed effect. Columns 3 and 6 present results omitting the $\alpha$ term entirely. All specifications include the common hour-of-sample fixed effect. Standard errors are two-way clustered at the household and hour-of-sample level.

Table A.4. ITT Robustness Checks Controlling for Pre-Period Hourly Averages

| | (1) Setpoint | (2) Comp. Run-Time | (3) Avg. Temp. Wedge | (4) Max. Temp. Wedge |
|---|---|---|---|---|
| Encouraged × Post × Mid-Peak AM | 0.046 | 0.073 | 0.019 | 0.026 |
| | (0.081) | (0.131) | (0.020) | (0.022) |
| Encouraged × Post × Peak | 0.428 | -1.641 | 0.030 | 0.038 |
| | (0.086) | (0.228) | (0.019) | (0.021) |
| Encouraged × Post × Mid-Peak PM | 0.111 | -0.401 | 0.038 | 0.064 |
| | (0.078) | (0.302) | (0.024) | (0.027) |
| Encouraged × Post × Off-Peak | -0.023 | 0.253 | 0.009 | 0.015 |
| | (0.069) | (0.154) | (0.011) | (0.012) |
| Observations | 1,867,602 | 1,867,602 | 1,867,059 | 1,867,059 |
| Households | 2,119 | 2,119 | 2,119 | 2,119 |

*Notes:* The table presents robustness checks for intent-to-treat effects estimated from Equation 2. For each outcome variable, pre-period hourly averages of that outcome are included as a series of 24 covariates (omitted from the table), and the specifications are estimated only on observations post-encouragement. All specifications include the common hour-of-sample fixed effect. Standard errors are two-way clustered at the household and hour-of-sample level.

Table A.5. LATE Estimates by Pricing Period, Keeping NA Flags

| | (1) Setpoint | (2) Comp. Run-Time | (3) Avg. Temp. Wedge | (4) Max. Temp. Wedge |
|---|---|---|---|---|
| $\widehat{TOU}$ × Mid-Peak AM | 0.053 | -0.520 | 0.105 | 0.139 |
| | (0.416) | (0.741) | (0.098) | (0.108) |
| $\widehat{TOU}$ × Peak | 1.678 | -7.279 | 0.205 | 0.251 |
| | (0.431) | (1.185) | (0.093) | (0.102) |
| $\widehat{TOU}$ × Mid-Peak PM | 0.378 | -2.785 | 0.214 | 0.342 |
| | (0.390) | (1.478) | (0.117) | (0.130) |
| $\widehat{TOU}$ × Off-Peak | -0.317 | 0.333 | 0.052 | 0.082 |
| | (0.343) | (0.792) | (0.056) | (0.061) |
| Observations | 6,037,377 | 6,037,377 | 6,037,113 | 6,037,113 |
| 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 | 103.5 | 103.5 |

*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: $\widehat{TOU} \times Period$. In the first-stage, these variables are instrumented for with randomized encouragement status interacted with the four pricing periods: $Encouraged \times Post \times Period$. Observations that were removed for potentially faulty motion sensor data are added back to the temperature wedge specifications in Columns 3 and 4. 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.

Table A.6. Setpoint and Compressor Run-Time LATE Effects by Occupancy Group

| | Cooling Setpoint | | | Comp. Run-Time | | |
|---|---|---|---|---|---|---|
| | (1) Often | (2) Sometimes | (3) Hardly | (4) Often | (5) Sometimes | (6) Hardly |
| $\widehat{TOU}$ × Mid-Peak AM | -0.073 | -0.511 | 0.748 | -0.695 | 0.806 | -1.526 |
| | (0.696) | (0.615) | (0.848) | (1.254) | (1.176) | (1.396) |
| $\widehat{TOU}$ × Peak | 1.544 | 1.182 | 2.332 | -6.629 | -6.370 | -8.803 |
| | (0.683) | (0.666) | (0.877) | (2.010) | (1.852) | (2.113) |
| $\widehat{TOU}$ × Mid-Peak PM | 0.775 | -0.379 | 0.773 | -4.710 | 0.027 | -3.591 |
| | (0.635) | (0.548) | (0.770) | (2.393) | (2.366) | (2.567) |
| $\widehat{TOU}$ × Off-Peak | 0.317 | -1.185 | -0.132 | -0.169 | 1.759 | -0.493 |
| | (0.590) | (0.503) | (0.691) | (1.365) | (1.296) | (1.460) |
| Observations | 2,011,880 | 2,014,498 | 2,010,999 | 2,011,880 | 2,014,498 | 2,010,999 |
| Households | 711 | 711 | 711 | 711 | 711 | 711 |
| Peak Control Mean | 74.7 | 74.9 | 76.3 | 9.70 | 8.21 | 6.87 |
| Peak First Stage Coefficient | 0.24 | 0.25 | 0.22 | 0.24 | 0.25 | 0.22 |
| Peak First-Stage F Stat | 36.2 | 38.9 | 31.5 | 36.2 | 38.9 | 31.5 |

*Notes:* The table presents local average treatment effects from estimating Equations 4a and 4b for two primary outcomes: setpoints (in deg. F) and compressor run-time (in min/hr). Effects are estimated separately for each of the occupancy group categories (Often, Sometimes, and Hardly Home households). $\widehat{TOU}$ 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: $\widehat{TOU} \times Period$. In the first-stage, these variables are instrumented for with randomized encouragement status interacted with the four pricing periods: $Encouraged \times Post \times Period$. 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.

Table A.7. Temperature Wedge LATE Effects by Occupancy Group

| | Avg. Temp Wedge | | | Max. Temp Wedge | | |
|---|---|---|---|---|---|---|
| | (1) Often | (2) Sometimes | (3) Hardly | (4) Often | (5) Sometimes | (6) Hardly |
| $\widehat{TOU}$ × Mid-Peak AM | 0.364 | -0.061 | 0.031 | 0.396 | -0.043 | 0.082 |
| | (0.185) | (0.176) | (0.154) | (0.205) | (0.193) | (0.165) |
| $\widehat{TOU}$ × Peak | 0.589 | -0.040 | 0.028 | 0.664 | -0.007 | 0.048 |
| | (0.193) | (0.160) | (0.133) | (0.213) | (0.176) | (0.145) |
| $\widehat{TOU}$ × Mid-Peak PM | 0.587 | 0.037 | -0.053 | 0.741 | 0.164 | 0.038 |
| | (0.224) | (0.206) | (0.191) | (0.250) | (0.225) | (0.208) |
| $\widehat{TOU}$ × Off-Peak | 0.168 | -0.023 | -0.017 | 0.189 | 0.010 | 0.010 |
| | (0.098) | (0.095) | (0.103) | (0.108) | (0.104) | (0.112) |
| Observations | 1,995,852 | 1,975,716 | 1,913,049 | 1,995,852 | 1,975,716 | 1,913,049 |
| Households | 711 | 711 | 711 | 711 | 711 | 711 |
| Peak Control Mean | 0.47 | 0.35 | 0.18 | 0.59 | 0.44 | 0.21 |
| Peak First Stage Coefficient | 0.24 | 0.25 | 0.22 | 0.24 | 0.25 | 0.22 |
| Peak First-Stage F Stat | 35.8 | 37.1 | 29.3 | 35.8 | 37.1 | 29.3 |

*Notes:* The table presents local average treatment effects from estimating Equations 4a and 4b for two primary outcomes: average and maximum temperature wedges (in deg. F). Effects are estimated separately for each of the occupancy group categories (Often, Sometimes, and Hardly Home households). $\widehat{TOU}$ 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: $\widehat{TOU} \times Period$. In the first-stage, these variables are instrumented for with randomized encouragement status interacted with the four pricing periods: $Encouraged \times Post \times Period$. 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.

Table A.8. LATE Effects for Motion Detection

| | (1) Motion Minutes | (2) Motion Indicator (0/1) |
|---|---|---|
| $\widehat{TOU}$ × Mid-Peak AM | 0.499 | 0.026 |
| | (1.202) | (0.034) |
| $\widehat{TOU}$ × Peak | 0.754 | -0.002 |
| | (1.229) | (0.036) |
| $\widehat{TOU}$ × Mid-Peak PM | 0.245 | -0.006 |
| | (1.386) | (0.034) |
| $\widehat{TOU}$ × Off-Peak | 0.498 | 0.005 |
| | (0.532) | (0.018) |
| Observations | 5,884,617 | 5,884,617 |
| Households | 2,133 | 2,133 |
| Peak Control Mean | 12.2 | 0.49 |
| Peak First Stage Coefficient | 0.24 | 0.24 |
| Peak First-Stage F Stat | 99.4 | 99.4 |

*Notes:* The table presents local average treatment effects from estimating Equations 4a and 4b for two primary outcomes: motion detection (min/hr) and a binary motion indicator for any motion within the hour. $\widehat{TOU}$ 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: $\widehat{TOU} \times Period$. In the first-stage, these variables are instrumented for with randomized encouragement status interacted with the four pricing periods: $Encouraged \times Post \times Period$. 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.

Table A.9. Transition Matrix of Chosen Slider Setting Values over Time

| | End Choice | | | | | |
|---|---|---|---|---|---|---|
| Start Choice | 1 | 2 | 3 | 4 | 5 | Total |
| 1 | 100 | 5 | 5 | 13 | 19 | 142 |
| 2 | 5 | 13 | 1 | 1 | 1 | 21 |
| 3 | 10 | 7 | 46 | 4 | 4 | 71 |
| 4 | 47 | 29 | 40 | 459 | 89 | 664 |
| 5 | 14 | 2 | 7 | 15 | 215 | 253 |
| Total | 176 | 56 | 99 | 492 | 328 | 1151 |

*Notes:* This transition matrix shows user-chosen slider setting values by thermostats on the first day for which we observe data for each thermostat, as well as on the final day for which we observe slider data (January 27th, 2020). Movement toward the bottom-left corner represents users' "throttling down" settings of the algorithm (towards 1), whereas movement toward the top right corner represents users' adopting more aggressive settings (toward 5). 1,151 unique thermostats are represented in the matrix.

Table A.10. Benefit-Cost Analysis Parameters

| Parameter | Value | Calculation | Source |
|---|---|---|---|
| **Assumptions** | | | |
| (1) Unit Capacity Ratio (BTU/hr) | 33202 | - | NEUD Table 27 |
| (2) SEER, Low (BTU/W*hr) | 12 | - | NEUD Table 27 |
| (3) SEER, High (BTU/W*hr) | 16 | - | NEUD Table 27 |
| (4) Peak Electricity Price (C$) | .134 | - | OEB 2019 |
| (5) Number of Summer TOU Days | 93 | - | ecobee data |
| (6) Scaling: New thermostats | 100000 | - | Authors' assumptions |
| (7) Scaling: Observed TOU compliance rate | .24 | - | Authors' calculations |
| (8) Scaling: Hypothetical TOU compliance rate | .5 | - | Authors' assumptions |
| **Calculations** | | | |
| Run-time to kWh conversion, Low | 2.767 | (1) / (2) * .001 | Authors' calculations |
| Run-time to kWh conversion, High | 2.075 | (1) / (3) * .001 | Authors' calculations |
| Scaling: New compliers | 24000 | (6) * (7) | Authors' calculations |
| Scaling: Hypothetical new compliers | 50000 | (6) * (8) | Authors' calculations |

*Notes:* The table presents underlying assumptions and calculations that are used to construct Table 5 and Table A.11. The source link for (1)-(3) is found here: https://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/showTable.cfm?type=CP&sector=res&juris=on&rn=27&page=0. The source link for (4) is found here: https://www.oeb.ca/consumer-information-and-protection/electricity-rates/historical-electricity-rates (last accessed January 18, 2024).

Table A.11. Aggregate Benefits and Costs of Eco+ TOU Adoption

| **Panel A:** Agg. summer impacts | | *Change in kWh per hh* | *Change in $ per hh* |
|---|---|---|---|
| | Low efficiency | -202.59 | -$27.15 |
| | | (-312.75, -92.40) | (-$41.91, -$12.39) |
| | High efficiency | -151.92 | -$20.36 |
| | | (-234.54, -69.31) | (-$31.43, -$9.29) |
| **Panel B:** Scaling impacts | | *Peak demand reduction (MW)* *(Compliance rate = 0.24)* | *Peak demand reduction (MW)* *(Compliance rate = 0.5)* |
| | Low efficiency | 10.88 | 22.67 |
| | | (7.37, 14.39) | (15.35, 29.99) |
| | High efficiency | 8.16 | 17.00 |
| | | (5.52, 10.79) | (11.51, 22.49) |

*Notes:* The table presents estimates of aggregate benefits and costs of the Eco+ TOU feature at scale under varying assumptions. 90% confidence intervals are presented below coefficient estimates in parentheses. Further underlying parameter assumptions (including with respect to Low and High efficiency cases) are found in Appendix Table A.10.

# B Experimental Details and Data Construction

## B.1 Details on the Eco+ Software

Ecobee's proprietary "Eco+" software is a suite of algorithmic features designed to preserve comfort while increasing energy efficiency. Eco+ includes five distinct features: Feels Like, Time of Use (TOU), Community Energy Savings, Smart Home & Away, and Schedule Assistant. We discuss the TOU feature in depth in the paper. We describe each of the other four features in further detail here.[1]

The "Feels Like" feature attempts to account for the effect of humidity on indoor temperatures by averaging over the past 10 days of humidity data collected by the thermostat. On particularly humid days in the summer, the thermostat will cool longer than it normally would to ensure that the temperature experienced in the home is representative of the preferred temperature setting. If humidity is far lower than the rolling average, then the compressor will not run as long as it normally would at that temperature setting.

The Community Energy Savings feature attempts to link thermostat schedules to demand response events announced by local utilities. For this feature to activate properly, users must properly associate their thermostat with the utility that serves them when setting up Eco+, and must be served by a utility that offers such a program. This process is illustrated by the sixth screen in Figure 2. If these conditions are met, on days when demand response events are called, the feature will precool the home by adjusting setpoints and then implementing a setback to raise setpoints and prevent compressors from running during the high-price demand response period. These precool and setback schedules are dependent on slider settings and tailored around the announced times of the demand response event. Demand response events are rare in our data, occurring only on roughly 0.5% of thermostat-days in the sample.

Smart Home & Away are features that predate the development of the Eco+ algorithm but have been rolled into the Eco+ algorithm. The essence of these features is to rely on motion detection to help the thermostat adapt to times when presence in the home does not match scheduled home/away times. For example, if motion is not detected during a scheduled "home" period, then Smart Away will kick in. Set points are allowed to hover 1 to 4 degrees away from the customer's preferred set point during home periods so that the preferred temperature can be quickly recovered if motion is indeed detected. The amount of time it takes to trigger Smart Away depends on the customer's chosen slider setting when setting up Eco+. For the least aggressive slider setting (1), it takes two hours of no motion detection during a scheduled Home period for Smart Away to trigger, while at the most aggressive slider setting (5) it only takes one hour of no motion for Smart Away to trigger. Smart Home is unaffected by the slider setting and is turned on when motion is detected after the away setting has been active for an hour and Smart Home has not been previously activated in the past two hours.

[1] Ecobee periodically releases software updates that may make changes to the way these features operate in practice. The descriptions offered here are our best attempt to describe the features as they operated during our experiment in summer 2019. For information on Eco+ directly from Ecobee, see: `https://support.ecobee.com/hc/en-us/articles/360035246672-eco-Frequently-Asked-Questions`. (last accessed January 18, 2024).

The Schedule Assistant feature takes advantage of the motion sensor capabilities to assess whether a household's motion detection schedule matches up with the setpoint schedule that the household has selected. For example, if a household's existing schedule indicates that they are "home" during a certain portion of the day, but no motion is ever detected by the thermostat during those hours, an automated email is sent to the household with suggestions for new temperature schedules that they can adopt.

## B.2 Details on Randomization and Experiment

We worked in conjunction with Ecobee to develop and implement the randomized encouragement design for Ontario households in spring and summer 2019. Ecobee provided us with an initial list of 5,147 thermostats in their Donate Your Data database for Ontario. We then applied a series of filtering criteria to these thermostats in order to create the sample of thermostats included our initial randomized encouragement sample.

First, we drop any users that have multiple thermostats, which drops 671 thermostats. We then drop households with multistage cooling systems, which drops 127 thermostats. We then drop 1 thermostat that was flagged with a non-Canadian country code. We then drop any households who had their Ecobee thermostat for less than a full year at the time of randomization, which drops another 345 thermostats. Finally, there were some cases of duplicate thermostat identifiers in the file. We address this by keeping the first record of any duplicate thermostat identifier, which drops 58 more thermostats from the file. This leaves us with 3,945 unique thermostat observations in our randomized encouragement sample. We then implemented the randomization. We created an encouragement group of 2,445 thermostats and held out a control group of 1,500 thermostats that were not encouraged to sign up for Eco+.

We then returned this finalized sample to Ecobee to implement the experiment. Ecobee sent out the randomized encouragements on August 6, 2019. This date coincided with Ecobee's larger rollout of Eco+ across North America. Thermostats in the encouraged group were presented with a series of prompts asking them to sign up for Eco+, configure their TOU rate settings with their local utility, set their comfort slider setting, and accept the Eco+ terms and conditions.

We received two primary experimental report files back from Ecobee. The first was a decoder file that contained the identifiers of the thermostats that we included in our experiment as they appeared in the DYD database, so that we could retrieve the exact thermostat-level files for only those thermostats included in the experiment. After cleaning this file according to the same procedure as the original file, we are left with 3,403 thermostats: 2,101 thermostats in the encouraged group and 1,302 in the control group. Of these 3,403, we are able to retrieve thermostat-level data files from the DYD database for 3,402 thermostats in 2019 and 3,396 thermostats in 2018.[2] We aggregate these thermostat-year files together to form the summer 2018/2019 panel that we use as our primary analysis dataset, the first data source described in Section 3.1.

[2] There are multiple potential reasons why we observe attrition from the original sample of 3,945 thermostats to the new sample of 3,402 thermostats. For one, thermostats may have decided to withdraw from DYD between the time we conducted the randomization in June and the time we received access to the thermostat-level data initially in November 2019. It is also possible that some non-DYD thermostats were allowed into the sample provided to us, so their data would never have appeared in DYD in the first place.

The second experimental file provided to us by Ecobee is a set of daily data for thermostats that accepted the Eco+ encouragement. These data include information like the date of acceptance of the Eco+ terms and conditions, daily data on slider settings, and indicators for whether each feature of Eco+ was turned on or off by day through the end of January 2020. These data are the second data source described in Section 3.1. We have data on 3,834 thermostats in this file, much closer to our original randomization file. 2,007 of these thermostats have the more detailed daily data, 2,001 of which are in our randomized encouragement group. This indicates that only a few thermostats in the control group discovered and signed up for Eco+ on their own. However, when testing for the presence of these thermostats in Ecobee's DYD database we are left again with 3,396 thermostats in 2018 and 3,402 in 2019, as indicated by the original decoder file.

After randomization and upon beginning our analysis, we discovered that older Ecobee thermostat models that did not have motion sensor capabilities were included in the randomization sample. Additionally, some Ecobee models that do have motion sensor capabilities but require users to purchase and install separate motion sensors were included in the randomization sample as well. Data reported from these thermostats are valid for outcomes such as setpoints, compressor run-time, etc, but if no motion sensors are installed then the data are lacking the key motion variable that we need to calculate temperature wedges. We further trim the 3,402 thermostats down to exclude thermostats missing the motion sensor data entirely. We also restrict our data to only keep thermostats with motion sensor data in July 2019, the month before Eco+ encouragement took place (used to define our occupancy groups). Imposing these sample restrictions trims the sample to the n=2,133 thermostats sample size used in the primary analysis.

Dropping these thermostats from our sample did not significantly impact the randomization component of our experiment. The percentage of thermostats that we had to drop was nearly identical in each of the encouraged and not encouraged groups. Of the 2,133 thermostats included in the primary analysis, 62 percent (1,319) of thermostats were encouraged to sign up for Eco+ and 38 percent were not encouraged. Of the thermostats that we drop due to a lack of motion sensor data entirely, 62 percent (757) were encouraged and 38 percent were not encouraged. In addition to the balance table for the primary sample (Table 1), we also present a balance table for the full sample of thermostats. Table A.1 illustrates that statistical balance continues to hold for key variables of interest in the sample that includes the previously dropped thermostats.

### B.3 Construction of Discomfort

Equations 1 and 2 in the paper define our measures of average and maximum temperature wedges that we use to characterize experienced household discomfort. There are two core components of the equation: how we define motion and how we define the motion-unadjusted temperature wedge itself.

The first primary component of the discomfort definition is the construction of the dummy variable for whether motion is observed within the hour or not: $1[Motion_{it}]$. Motion is provided to us in the DYD data as a series of dummy variables for whether motion is observed at the 5-minute level across a number of sensors as well as at the thermostat itself. There are data on motion sensors available for up to ten separate motion sensors plus the motion sensor itself. Given computational constraints, we keep data from the first five motion sensors as well as the thermostat itself, which covers essentially all motion sensors available to us in the raw data.

Our ability to construct our measure of temperature discomfort necessarily depends on the accuracy of the thermostat-level motion sensor data that we obtained from the DYD database. Although we could still calculate the unadjusted measure of temperature wedges, we cannot complete the calculation without accurate motion sensor data. When cleaning the five-minute-level readings to construct our panel used in the analysis, we did not assume that a missing motion sensor reading indicated no motion. As described in Appendix B.2, we remove thermostats from our sample that do not have motion sensor readings in the DYD database. In order to further screen for potentially unreliable motion sensor readings, we drop motion sensor readings for thermostat-days where motion is triggered for more than 70% of the day, or if motion is triggered for more than 7.5% of the 1–4 a.m. overnight time block. In this case, we treat the final value of the motion variable as missing, rather than imputing a 0/1 value. This is the primary reason for the small difference in sample sizes between compressor run-time and temperature wedge regressions presented in the paper, such as in Tables 2 and 3.

The second primary component of the discomfort definition is the relevant indoor temperature wedge between actual and preferred temperatures: $|T_{iqt} - T^*_{it}|$, referenced in Section 3.3 in the paper. We cannot simply use the customer's observed setpoint as their measure of "preferred" temperature, because Eco+ makes direct changes to the thermostat's setpoints when the TOU feature is operating. Therefore, to construct our measure of preferred temperature, we take a household's day-of-week-by-hour-of-day average setpoint in July 2019 (before Eco+ was pushed to our encouraged group). We then substitute in this average for the equivalent hours in August and September 2019. Expanding the example given in Section 3.3, say one observes a given household on four Tuesdays at 3 p.m. in July 2019. The household's setpoint during these four hours are 73, 74, 74, and 73 degrees F., respectively. We average these four observations (73.5), and use this value as our measure of preferred temperature. If we did not make this substitution, and the household accepts Eco+, then we might mistakenly conclude that the household's preferred temperature during this time of the day is, say, 76 degrees, if the TOU feature raises setpoints by 2-3 degrees during the peak period. Figure A.1 displays how this counterfactual substitution process works to smooth out the impact of Eco+ on setpoints. In panel B, one can see that there is no impact of the substitution process in July 2019, because Eco+ is not yet working to directly change setpoints. This is illustrated by the fact that the observed and counterfactual lines for both the encouraged and not encouraged groups lie on top of one another. In August 2019, the counterfactual setpoint for the encouraged group falls below the observed setpoints in the data, which shows clear signs of algorithmic manipulation (pre-cooling drop followed by spike up during the peak period). By using the counterfactual setpoints as our measure of "preferred" temperature, we ensure that the algorithmic manipulation of setpoints is not biasing our measure of “preferred" indoor temperature.

## B.4 Electricity Conversion Calculations

In Section 6, we consider the private benefits and costs of Eco+ adoption and calculate the energy savings/discomfort tradeoff that such households make. To do this, we must convert from compressor run-times that we observe in the DYD data to electricity consumption (in kWh). Table A.10 outlines the various assumptions we make when undergoing these conversions.[3]

First, we need a value for a unit capacity ratio (in BTU/hour) to convert compressor run-times into electricity consumption. We take the 2019 estimate from Canada's National Energy Use Database (NEUD) for central air-conditioning units: 33,202 BTU/hr (Natural Resources Canada 2019). Next, we need an assumption about the energy efficiency performance of the compressors in our sample. From NEUD, we observe average seasonal energy efficiency ratios (SEER) by year (expressed in terms of BTU/W*hr). For a lower-bound assumption on Ontario-specific SEERs, we select a value of 12, as this is below the average SEER value for central air-conditioning systems in Ontario in recent years (12.5 in 2018, 12.7 in 2019). For an upper-bound estimate, we select a SEER value of 16, which is above the required SEER of 13 for new AC systems in Canada and also higher than the 14.5 SEER required to be considered "efficient".[4]

Dividing the unit capacity ratio by the appropriate SEER and adjusting for units (W*hr to kWh) yields a conversion factor that we use to convert compressor run-times to kWh. These are provided in the first two lines under the "Calculations" panel of Table A.10. For sensitivity, we present results for both SEER assumptions as the "Low efficiency" (12 SEER) and "High efficiency" (16 SEER) cases in Table 5 in the main paper and for aggregate impacts presented in Table A.11.

[3]Table A.10 also presents a number of parameter values vital to our calculation of aggregate summer impacts, as well as hypothetical peak demand reduction impacts of scaling up Eco+, which we present in Table A.11 and discuss in Sections 6.1 and 6.3, respectively.

[4]Source: `https://www.hrai.ca/consumer-tip/what-exactly-is-seer-rating-`. (last accessed January 18, 2023).