Who Programs Their Thermostat and How? Five Behavioral Archetypes from Nine Years of California Smart Thermostat Telemetry

Juana Isabel Méndez, UC Berkeley, Tecnologico de Monterrey
Alan Meier, UC Davis
Kopal Nihar, Stanford University
Hideki Shimada, UC Davis, National Institute of Advanced Industrial Science and Technology
Therese Peffer, UC Berkeley

ABSTRACT

Thermostats play a critical role in residential demand flexibility; their performance depends not only on algorithms but on how occupants schedule, override, and adapt their systems. Current grid-interactive models employ idealized control patterns, overlooking the behavioral variability that drives actual HVAC operation. We studied thermostat scheduling behavior using nine years (2017–2025) of Ecobee Donate-Your-Data telemetry from 1,263 California households. Across the study period, households used 517 unique schedule names, including the three Ecobee defaults, Home, Sleep, and Away, and 514 user-defined labels. We classified these names into a rule-based taxonomy with 8 main categories and 19 subcategories, then grouped households by schedule configuration complexity, naming patterns, and setpoint calibration. These patterns yielded five behavioral archetypes, four statistically robust: the *Passive Default* (63.4%, factory defaults with minimal thermal differentiation), the *Active Customizer* (29.5%, progressive grid-interactive schedule expansion), the *Disciplined Minimalist* (3.2%, deliberate wide thermal deadband), the *Cold-Climate Customizer* (2.2%, extreme nocturnal setback), and the *Seasonal Switcher* (0.4%, two-program seasonal alternation). Schedule semantic content independently predicts heating and cooling setpoints and HVAC runtime after controlling for archetype, year, and outdoor temperature, establishing that schedule naming encodes behavioral intention. These archetypes provide an empirically grounded basis for behavior-aware Demand Response program design and establish schedule name analysis as a scalable behavioral segmentation tool. The vocabulary households assign to their schedules is itself a behavioral record, one that predicts HVAC outcomes independently of archetype and weather and is already embedded in existing utility datasets without surveys or additional instrumentation.

Introduction

The increasing penetration of variable renewable energy sources, rising electrification demands, and the growing frequency of extreme weather events have placed unprecedented stress on modern electric grids. In summer 2022, California experienced a heat wave that caused excessive peak demand and amplified grid reliability concerns, accelerating the search for scalable, non-generating demand-side solutions (Beltran et al., 2024). Smart thermostats have emerged as the primary residential platform for this flexibility, with over 3.5 million devices deployed statewide (DNV, 2024); however, actual savings from current load-management activities remain approximately 50% below the estimated 10.2 GW peak demand reduction potential, and the residential sector remains substantially underutilized relative to commercial and industrial customers (Murphy et al., 2024; Nambiar et al., 2024).

Nevertheless, the behavioral assumptions underlying Demand Response (DR) programs remain undifferentiated, as program designs typically treat enrolled households as a homogeneous population that responds uniformly to price signals, curtailment events, or automated setpoint adjustments (Li et al., 2025). Analysis of a large-scale thermostat dataset indicates that approximately 60% of users rely on scheduling. At the same time, nearly half of potential load reductions during direct load control programs are lost to overrides, with the magnitude of imposed setpoint adjustments identified as a key predictor of override timing (Huchuk et al., 2021; Kane & Sharma, 2019). These findings motivated efforts to classify users as compliant or non-compliant and to identify behavioral typologies for more targeted program design (Blonz et al., 2025; Huchuk et al., 2021); however, these typologies have remained focused on override behavior during DR events rather than on the scheduling configuration households maintain between events.

The Ecobee Donate Your Data (DYD) program provides device-level schedule names, setpoint time series, and disaggregated HVAC runtime at hourly resolution (Luo & Hong, 2022; Meier et al., 2019). Studies using this dataset have documented that households lower their thermostats approximately 1°C during sleep and residential AC runtime tracks grid-wide peak demand with high fidelity (Meier et al., 2019), and data-driven occupancy schedules derived from over 90,000 thermostat records have been developed to replace fixed-schedule modeling assumptions with representative behavioral profiles (Jung et al., 2023). Despite this data availability, a critical gap persists: existing studies treat thermostat scheduling as a static background element rather than a dynamic behavioral variable. The semantic content of user-defined schedule names has not been analyzed as a predictor of HVAC outcomes, and population-wide behavioral archetypes based on complete scheduling configurations have not been systematically constructed. Three research questions organize the analysis:

- RQ1: What distinct behavioral archetypes characterize thermostat scheduling among California smart thermostat users, and how are they distributed across the population?
- RQ2: Does the semantic content of user-defined schedule names independently predict HVAC setpoints and runtime after controlling for archetype, climate, and year?
- RQ3: What share of the smart thermostat population currently exhibits schedule labeling compatible with grid-interactive demand flexibility, and what does the behavioral profile of the remaining majority reveal about the limits of self-directed customization?

The findings have direct implications for how California utilities and program administrators design behavioral DR programs, segment enrolled populations, and set realistic expectations for demand flexibility from the residential smart thermostat fleet. Section 2 describes the data, sampling, clustering, and validation methodology. Section 3 presents the five behavioral archetypes. Section 4 discusses implications for program design. Section 5 concludes.

Methodology

This study draws on the Ecobee DYD program (Ecobee, 2025), which provides five-minute HVAC operational records per thermostat, including runtimes, setpoints, indoor temperature, and humidity, spanning January 2017–December 2025. Because households were not prompted to change their behavior and no intervention was applied, the corpus reflects ecologically valid in situ behavioral traces. Figure 1 shows the pre-processing flow. Starting from 22,282 raw identifiers obtained from Google BigQuery, sequential filters retained only devices with invariant building metadata, removed 197 devices with heating system configuration drift or inconsistent structural characteristics, resolved locations through the 2025 U.S. Census Gazetteer (Bureau, 2026), Photon/OSM, and Google lookup, and applied a final quality check requiring at least one valid 24-hour record per device. The resulting study population is 15,521 thermostats.

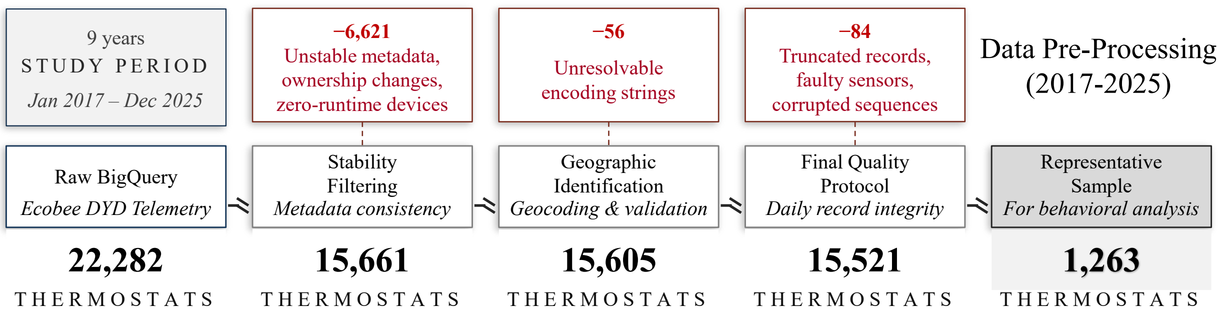

Figure 1. Data pre-processing pipeline from raw Ecobee Donate-Your-Data telemetry to final study population and analytic sample. *Source*: Own elaboration.

Sample Size Justification and Stratified Representative Sampling

The cleaned dataset held 15,521 thermostats, far more than needed to describe behavior reliably, so we narrowed it to a representative subset of 1,263 devices. That number is what keeps the smaller set accurate to within a ±3% margin at 95% confidence (Islam, 2018; Wadsworth, 1998), meaning the 1,263 devices behave statistically like the full 15,521. To keep the subset faithful to the whole, we drew it across Ecobee hardware models and ASHRAE climate zones, because scheduling interfaces differ by device generation and heating and cooling demand differs across California's regions, retaining the devices with the longest and most complete records. Figure 2 confirms that the geographic spread of the 1,263-device sample tracks the full population.

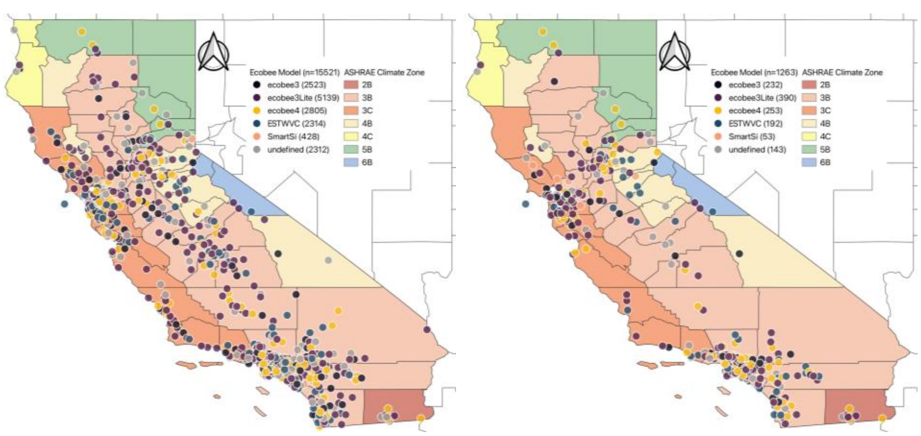

Figure 2. Geographic distribution of Ecobee thermostats in California by ASHRAE climate zone and device model. Left: full DYD population (n = 15,521). Right: stratified analytic sample (n = 1,263). Point colors indicate thermostat model; background shading indicates ASHRAE climate zone. *Source*: Own elaboration.

Schedule Name Taxonomy and Longitudinal Distribution

All subsequent analyses rest on five variables per device, namely the schedule name (the active program label), heating setpoint, cooling setpoint, indoor temperature, and disaggregated HVAC runtime. Throughout the nine-year period, households used 517 unique schedule names. The three Ecobee default modes, Home, Sleep, and Away, accounted for 94.3% of all hourly records; the remaining 514 names were user-defined.

To interpret these labels, we applied a rule-based taxonomy with 8 main categories and 19 subcategories. Schedule names were lightly cleaned before classification by converting text to lowercase, trimming extra spaces, and removing punctuation or separators. The examples in Table 1 are shown after this basic text cleaning. The taxonomy separates the default modes from user-defined variants of core routines, such as customized home, sleep, and away schedules. Other user-defined labels were grouped by apparent purpose, including activity, time-of-day, seasonal or comfort-related, occupancy-related, and grid- or system-oriented schedules. When multiple schedule names appeared in the same hourly record, the hour was split equally across active names to avoid double-counting.

Table 1. Schedule-name taxonomy used to classify thermostat operating modes.

Main category	Subcategory	n schedule names	Schedule names
Core	HOME	1	home
	SLEEP	1	sleep
	AWAY	1	away
Custom Core Variants	HOME CUSTOM	55	home weekend, home morning, homa, etc.
	SLEEP CUSTOM	60	bedtime, good night, sleep fan
	AWAY CUSTOM	11	afternoon aw, away again, nothere, etc.
Activity	ACTIVITY	43	work, yoga, shower, movie time
Grid / System-Oriented Settings	UTILITY RATE	98	save power, tou, flex alert, peak cost, etc.
	SETPOINT AS NAME	14	66f, 67f, 69ish, 73 heat, 75, 76, 81f, etc.
	SOLAR/GENERATION AWARE	10	beforesolar, eco solar, full solar, etc.
	EQUIPMENT/SYSTEM	7	aquastat1, stepfunction, circulating, etc.
Occupancy & Family Status	SPACE / LOCATION	17	back rooms, back warm, office, etc
	NAMED PERSON	13	aj, carleton, haxhere, kendrick, etc
	PET	2	puppy, turtle
Routine & Time-Based	TIME-OF-DAY	59	11 am, 3p cool, 4 pm, 4 pm heat, etc,
	WAKE UP CUSTOM	42	early mornin, early wake, first thing, etc.
Seasonal & Weather Specific	COMFORT	54	chillaxin, cold, comfort sett, comfy, etc.
	SEASON / CLIMATE	24	winter, cool sum day, spring sleep, etc.
Other	AMBIGUOUS	5	butter, sop, test2, tiny temp, v

Grouping Households into Archetypes

An initial pass showed that the strongest divide among households was simply whether they used any custom schedule names, so each device was first assigned to one of three strata. These are Core (default names only, n=855), Core+Custom (at least one default and one custom name, n=403), and Custom (no defaults, n=5, too few to cluster and treated as a single descriptive group); sixteen devices with fewer than three years of coverage were described but not assigned. Within each stratum, households were then grouped by the similarity of their scheduling behavior using k-medoids clustering, with the number of groups selected by standard internal-validity criteria and stability confirmed across 300 resampled runs (Bouveyron et al., 2019). The resulting groups are stable, fewer than 3% of devices sit near a boundary, size gaps reflect genuine behavioral minorities, and membership is not reducible to climate zone. Sankey diagrams of year-to-year category transitions visualize how schedule composition evolves within each group.

Statistical Validation

We confirmed that the archetypes are statistically distinct, and that schedule names carry independent information, through a layered set of tests. First, the four main archetypes were compared across seven measures of scheduling behavior, HVAC runtime, and thermal preference using nonparametric Kruskal-Wallis tests (Wadsworth, 1998), with a Benjamini-Hochberg correction to guard against false positives from multiple comparisons; all seven measures differed significantly across archetypes. Second, linear regression models at the identifier-year level tested whether more elaborate scheduling predicted setpoints and runtime after controlling for archetype, calendar year, and mean outdoor temperature (Bouveyron et al., 2019). Third, weighted linear regression models at the identifier-year-schedule level tested whether schedule-name semantics predicted setpoints and HVAC runtime after controlling for archetype, calendar year, and mean outdoor temperature. All effects are reported with 95% confidence intervals and multiple-comparison-corrected p-values.

Results

How Households Name and Combine Schedules

The presence of custom names is itself a behavioral signal, indicating deliberate engagement beyond the device's out-of-the-box configuration. Therefore, each device-year was classified into one of three profile categories based on the schedule names active that year: Core only (defaults exclusively), Core + Custom (at least one default plus at least one user-defined name), or Custom only (no defaults present). Core-only consistently accounts for 75–80% of active identifiers; Core + Custom accounts for 20–25% and has been stable since 2020, following a gradual rise in 2017–2019; Custom-only remains below 2% in any year (Figure 3a).

Personalization is thus almost always additive rather than substitutive; users extend the default vocabulary rather than replacing it. At finer resolution, the Home–Sleep dyad is the modal program combination throughout the study period (~55–60% of devices through 2021, declining toward 50% by 2022–2025 as Home–Sleep + Custom grew to ~15% by 2024), while the Away–Home–Sleep triad and its custom variant account for a further 13–15% (Figure 3b).

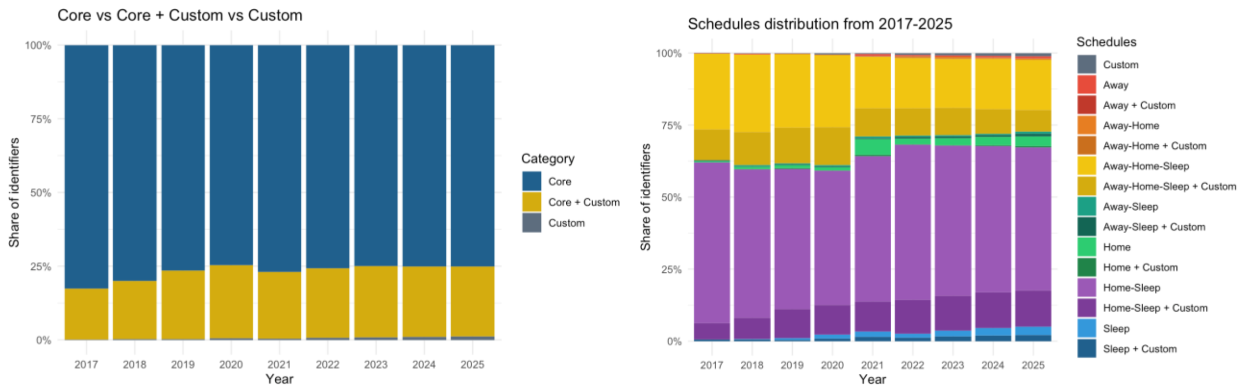

Figure 3. Annual distribution of schedule profile categories from 2017–2025: (a) (Core only vs. Core + Custom vs. Custom only); (b) Schedule program where each color represents a distinct named-program combination. *Source*: Own elaboration.

Archetype Separation

The archetypes differ most clearly in how households build and name schedules, with additional differences in runtime and comfort settings. In plain terms, the groups separate first by scheduling vocabulary, second by HVAC operation, and third by setpoint calibration. These differences are statistically significant across the main behavioral metrics.

More elaborate scheduling is linked to setpoint behavior but not directly to runtime. Each additional schedule a household maintains is associated with a lower heating setpoint, about 0.6°F, and a higher cooling setpoint, about 0.6°F, widening the comfort band. However, adding schedules does not by itself significantly change heating, cooling, or fan runtime after controlling for archetype, year, and outdoor temperature.

Schedule-name semantics add explanatory information beyond schedule count. At the subgroup level, SOLAR / GENERATION AWARE labels are associated with higher cooling runtime, consistent with precooling or solar-coordinated operation. PET and SETPOINT AS NAME labels are associated with higher heating runtime, while AWAY CUSTOM labels are associated with lower heating runtime. AWAY hours are also associated with lower cooling runtime. These associations indicate that schedule wording captures operational differences not explained by archetype, year, or outdoor temperature alone.

Thus, the archetypes are clearly distinct, more elaborate scheduling is associated with wider setpoint calibration, and schedule-name semantics independently predict HVAC operation. The following subsections describe each archetype in detail.

Passive Default Archetype (n = 801, 63.4%)

The *Passive Default Archetype* (cluster Core_C2) is the largest archetype, representing nearly two-thirds of the sample (Figure 4). These households use the three Ecobee default schedules exclusively. Indoor temperature medians differ minimally across categories (HOME ~72°F, SLEEP ~70°F, AWAY ~71°F), and despite HOME–SLEEP heating setpoints programmed at 67–68°F, realized indoor temperatures remain elevated above setpoint, indicating that programmed nocturnal setback is not achieved in practice, most plausibly due to building thermal mass buffering overnight temperature drops. Cooling setpoint distributions show greater dispersion in AWAY than in HOME or SLEEP, suggesting moderate heterogeneity in absence-period programming within an otherwise uniform group. Setpoint values show no inter-annual drift and no detectable response to the 2020 pandemic period, consistent with a population that configured the thermostat at installation and has not revisited settings.

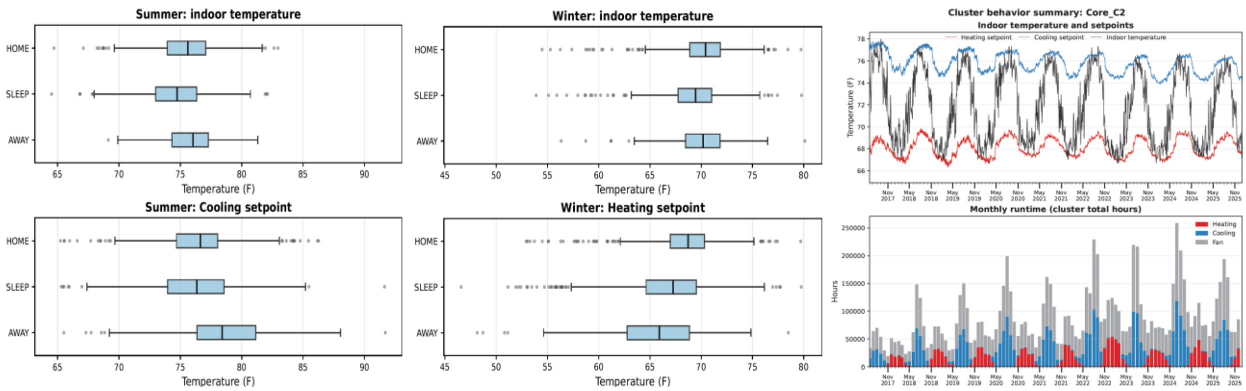

Figure 4. *Passive Default Archetype* (Core_C2 cluster) profile (n = 801). Seasonal boxplots of indoor temperature and setpoints by schedule category (summer and winter, left panels); daily average indoor temperature, cooling setpoint, and heating setpoint over 2017–2025 with monthly aggregate HVAC runtime by mode (right panels). *Source*: Own elaboration.

Disciplined Minimalist Archetype (n = 40, 3.2%)

The *Disciplined Minimalist Archetype* (Core_C1) uses the same three default schedules as the Passive Default Archetype but produces a qualitatively different thermal outcome (Figure 5). A clear HOME–SLEEP–AWAY setback cascade is evident in winter heating setpoints: HOME median ~65°F, SLEEP ~60°F, and AWAY ~58°F, representing a progressive 5–7°F stepdown across occupancy states. The resulting thermal deadband between heating and cooling setpoints is the widest in the sample, and indoor temperature floats freely within it, engaging conditioning only at the thermal extremes. Aggregate monthly runtime is an order of magnitude lower than the *Passive Default Archetype*, with cooling nearly absent even in summer and heating modest relative to cluster size. This archetype represents deliberate tolerance of indoor temperature variation as an energy-minimization strategy, achieved entirely within the factory-default schedule structure without any custom programming.

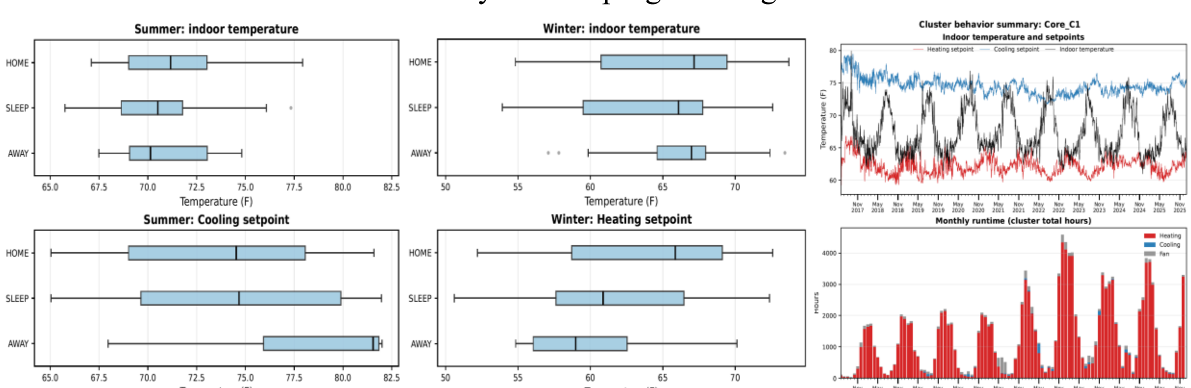

Figure 5. *Disciplined Minimalist Archetype* (Core_C1 cluster) profile (n = 40). *Source*: Own elaboration.

Active Customizer Archetype (n = 373, 29.5%)

The *Active Customizer Archetype* (Core+Custom_C1) is the most behaviorally complex group, the only one whose schedule portfolio spans all 17 taxonomy categories alongside the three core defaults. Schedule adoption is not a one-time reconfiguration but an incremental layering of behavioral specificity across enrollment years.

The Sankey diagram in Figure 6 traces schedule evolution from 2017 through 2025. HOME and SLEEP dominate as structurally stable flows throughout. AWAY contracts progressively from 2021 onward while ROUTINE consolidates as a persistent fourth-tier category from 2019, indicating a gradual shift from generic absence setback toward routine-based occupancy programming. GRID OR SYSTEM and WEATHER SPECIFIC maintain thin but consistent flows across all years, and custom variants (HOME CUSTOM, SLEEP CUSTOM, AWAY CUSTOM, ACTIVITY, OCC & FAM STATUS) emerge from 2019–2020 onward.

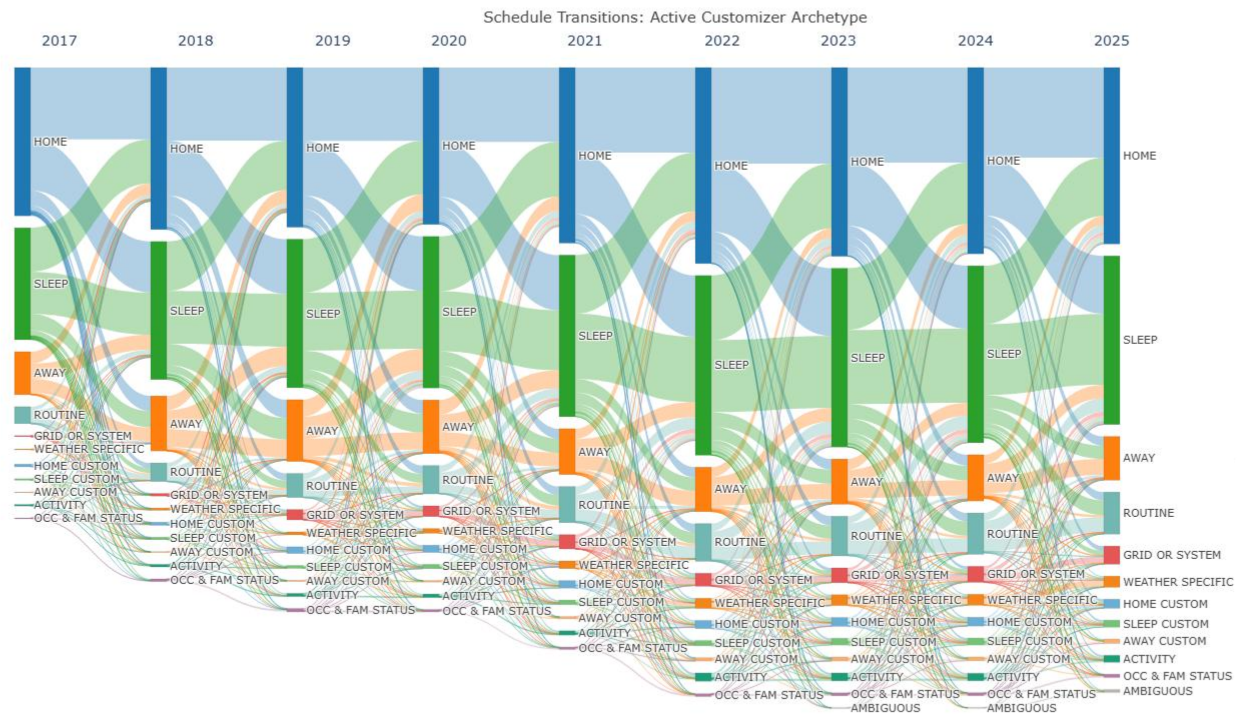

Figure 6. Sankey diagram of annual schedule category distributions for the Active Customizer Archetype (n = 373), 2017–2025. Band width is proportional to the share of devices using each taxonomy category per year. *Source*: Own elaboration.

Despite this categorical breadth, realized indoor temperatures remain compressed (70–72°F), with behavioral differentiation concentrated in setpoint programming (Figure 7). AWAY and AWAY CUSTOM schedules show elevated cooling setpoints (~79–80°F) consistent with absence-driven setback, while occupancy-referenced categories (HOME CUSTOM, ROUTINE, OCC & FAM STATUS) cluster closer to 75–77°F. In winter, AWAY-type schedules drop toward 61–65°F versus 67–69°F for HOME, ACTIVITY, and ROUTINE categories.

This archetype is distinguished by persistent fan-only operation across all months, indicating integration of ventilation or air-quality logic beyond pure thermal conditioning, a behavioral signature absent in all other archetypes.

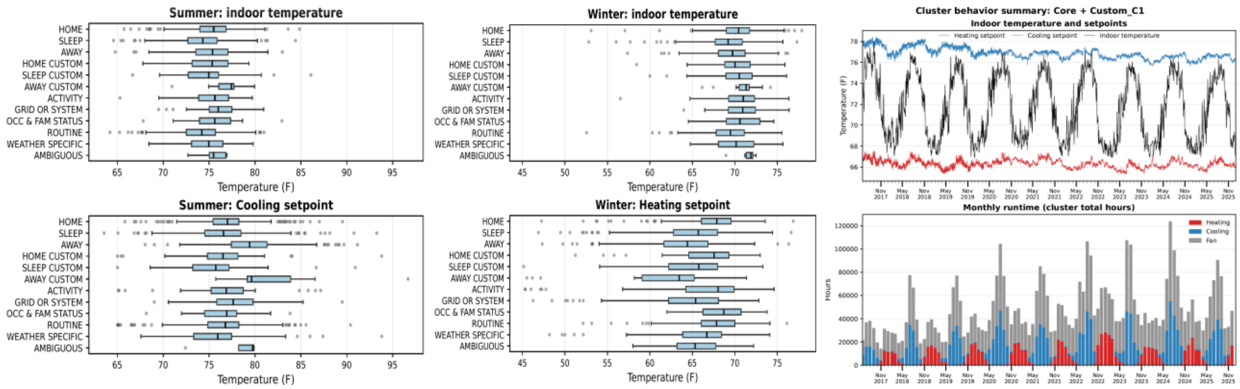

Figure 7. *Active Customizer Archetype* (Core+Custom_C1 cluster) profile (n = 373). *Source*: Own elaboration.

Cold-Climate Customizer Archetype (n = 28, 2.2%)

The *Cold-Climate Customizer Archetype* (Core+Custom_C2) extends the schedule portfolio to 10 taxonomy categories but is almost entirely heating-dominated: aggregate winter heating reaches 2,500–3,000 hours/month across 28 devices, with cooling minimal even in summer (Figure 8). The defining feature is SLEEP CUSTOM, which maintains indoor temperatures near 64°F with an IQR below 2°F, representing a 5–7°F nocturnal setback below all other categories in the cluster and the most precise deliberate setback observed across the full sample. Despite this heating-dominant runtime, cooling setpoint heterogeneity is paradoxically high: HOME CUSTOM spans a cooling setpoint IQR of approximately 20°F, suggesting that warm-season customization is highly individualized within an otherwise cold-focused group. A gradual indoor temperature decline from ~70°F in 2017 to ~67–68°F by 2023–2025 is consistent with progressive cold-season behavioral adaptation over the enrollment period, though direct causal attribution remains beyond the scope of this observational study.

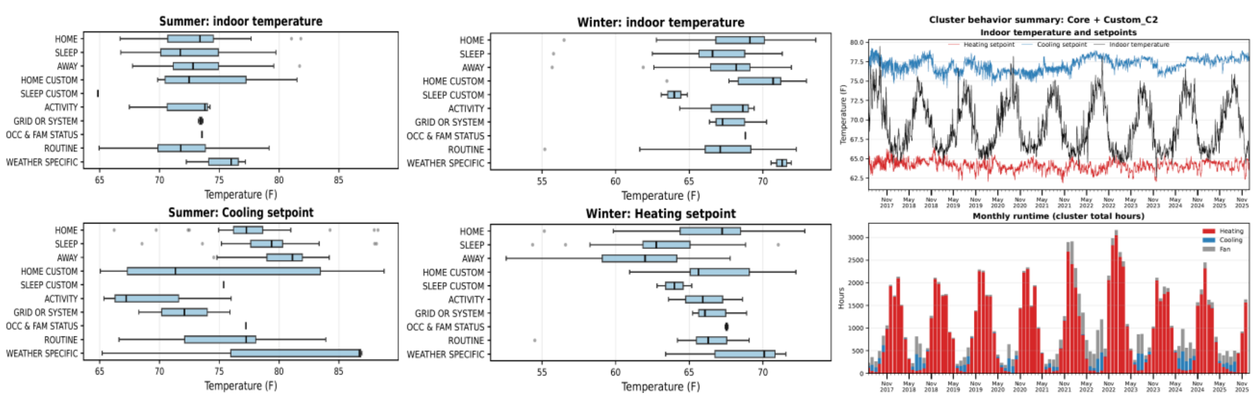

Figure 8. *Cold-Climate Customizer Archetype* (Core+Custom_C2 cluster) profile (n = 28). *Source*: Own elaboration.

Seasonal Switcher Archetype (n = 5, 0.4%)

The *Seasonal Switcher Archetype* (Custom_C1), depicted in Figure 9, replaced the Ecobee defaults with exactly one user-defined schedule category, WEATHER SPECIFIC, making it the only archetype with no HOME, SLEEP, or AWAY labels in its portfolio. Data span 2022–2025, the most recent enrollment window in the sample. Summer indoor temperature converges to ~70.5°F with an IQR below 0.5°F, and summer cooling setpoints cluster tightly at ~72–72.5°F. Winter heating setpoints span ~65–67°F with greater dispersion, while heating runtime dominates winter months and cooling dominates summer with the cleanest seasonal alternation observed across all archetypes and no shoulder overlap or interannual drift. The combination of a single taxonomy category, extremely low intra-cluster variance, synchronized activation from 2022, and geographic concentration is more consistent with centralized property management than independent household customization.

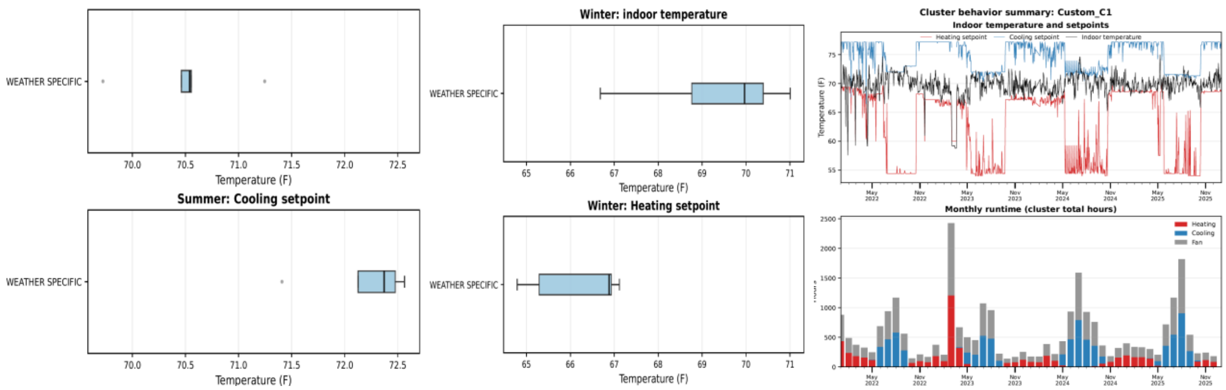

Figure 9. *Seasonal Switcher Archetype* (Custom_C1 cluster) profile (n = 5). *Source*: Own elaboration.

Discussion

A Behavioral Engagement Gap, not a Technology Gap

The *Passive Default Archetype* represents the most consequential finding for DR program design. Nearly two-thirds of the sample operate hardware-identical devices to those in the *Active Customizer Archetype* but produce none of the grid-interactive outcomes. The gap is behavioral: these households accepted the factory configuration at installation and have not revisited it, consistent with the well-documented failure of programmable thermostats to deliver projected energy savings in practice (Meier et al., 2019), and with the finding that fewer than half of monitored households correctly configure their thermostats into a mode that enables automated temperature control (Li et al., 2025). Nine years of setpoint time series, with no interannual drift and no detectable response to the 2020 occupancy shock, confirm this rigidity empirically.

Several mechanisms plausibly explain this rigidity, and the observational design does not allow adjudication between them. Installation friction and interface complexity create a locally optimal stopping point that requires occupancy prediction, which many users do not sustain; absent feedback and satisfaction anchoring, there is no experiential signal that the configuration should change (thebigbry1971, 2024). Each corresponds to a distinct barrier identified in the behavioral DR literature, specifically technological, cognitive, feedback, and motivational barriers (Huchuk et al., 2021; Li et al., 2025).

This rigidity reframes the intervention problem as one of choice architecture rather than engagement. Whoever sets the factory default governs the grid behavior of most of the fleet. Against the roughly 3.5 million smart thermostats deployed in California (DNV, 2024), the *Passive Default Archetype* corresponds to on the order of 2.2 million households reachable through one design decision, changing the default schedule, rather than through behavior-change campaigns.

The program consequence follows directly. Passive nudges such as bill inserts, app notifications, or peer comparison achieve average event-hour reductions of approximately 2.6% (Murphy et al., 2024) and are unlikely to move these households toward schedule customization. The highest-return interventions instead reduce or eliminate the configuration burden, namely manufacturer-level grid-interactive default schedules negotiated with utilities, enrollment-time pre-population of utility-rate and precooling schedules, automated schedule suggestion based on inferred occupancy patterns, and direct load control offered on an opt-out basis, already the control strategy of 62% of U.S. Wi-Fi thermostat programs (Murphy et al., 2024).

Implementing these mechanisms at scale would nonetheless require ensuring user acceptance through transparent enrollment, complying with data privacy regulations for inferred occupancy patterns, and achieving integration with existing smart home ecosystems via standardized APIs. Table 2 translates each archetype into a recommended program action, so that behavioral segmentation maps directly onto intervention design.

Table 2. Recommended program action by archetype.

Archetype	**Share**	**Recommended program action**
Passive Default	63.4%	Redesign factory defaults toward grid-interactive presets; pre-populate utility-rate and precooling schedules at enrollment; offer direct load control on an opt-out basis
Active Customizer	29.5%	Identify and reinforce with rate signals and feedback; reuse this group's ROUTINE, GRID OR SYSTEM, and WEATHER SPECIFIC schedule vocabulary as a template library for onboarding less engaged households
Disciplined Minimalist	3.2%	Prime automated-DR candidates; wide deadband and low override risk make them easy to enroll
Cold-Climate Customizer	2.2%	Target heating-season setback and winter-peak programs; cooling engagement is heterogeneous within this group and may respond to individualized warm-season outreach
Seasonal Switcher	0.4%	Likely a centrally managed property; engage property managers rather than individual occupants

An Emerging Self-Organized Labeling for Grid-Interactive Scheduling

In contrast, the *Active Customizer Archetype* shows progressive schedule expansion, with SOLAR and UTILITY RATE categories growing from 2019 onward through spontaneous user-initiated behavior. The statistically confirmed association between SOLAR-labeled hours and cooling runtime is consistent with precooling strategies that shift thermal load ahead of peak demand periods, the load-shifting behavior that programs such as Sonoma Clean Power's GridSavvy Rewards seek to elicit through direct thermostat control (Beltran et al., 2024).

This reframes a common assumption in DR program design where grid-interactive behavior typically requires utility program support to emerge. The program implication is to identify these households, validate and reinforce their behavior with utility-provided rate signals and feedback, and use the *Active Customizer Archetype's* schedule vocabulary as a design template for pre-populated schedule libraries that could lower the engagement barrier for the *Passive Default Archetype*, targeting behaviorally engaged households first before transitioning to less-engaged groups (Huchuk et al., 2021).

Archetype segmentation bears directly on the Virtual Power Plant buildout now underway. The U.S. Department of Energy estimates that 80–160 GW of VPP capacity, enough to serve 10–20% of national peak load, must be deployed by 2030 to support load growth at lower grid cost, roughly tripling current scale (U.S. Department of Energy, 2025). California has adopted a statutory load-shift goal of 7,000 MW by 2030 under Senate Bill 846 (California Energy Commission, 2023). Both targets depend on enrolling residential devices at scale, and the archetypes indicate where enrollment effort pays off.

Active Customizers are already exhibiting the precooling and rate-responsive behavior VPPs aggregate and need only validation and rate signals; Disciplined Minimalists, with wide deadbands and low override risk, are prime automated-DR candidates; the Passive Default majority is unreachable through recruitment campaigns and is better served by grid-interactive factory defaults and opt-out enrollment. Schedule-name screening offers VPP aggregators a zero-cost first filter for this targeting, since the signal already resides in their telemetry.

Schedule Complexity, Setpoint Calibration, and Operational Efficiency

Figure 10 positions the five archetypes along two independent dimensions, the number of schedule categories a household uses and the width of its thermal deadband. Archetypes with identical vocabulary occupy very different positions on the thermal axis, and archetypes with the widest vocabulary do not necessarily achieve the widest deadband, which carries direct implications for how schedule complexity should be interpreted in DR program evaluation.

The clearest case is the contrast between the *Disciplined Minimalist* and the *Passive Default Archetypes*. Both use the same three default programs while producing divergent thermal outcomes, which directly parallels the field evidence that energy consumption can vary more than tenfold among identically constructed homes due to differences in thermostat use alone (Jung et al., 2023). The variance lies in calibration, not interface complexity; for instance, the *Seasonal Switcher Archetype* produces a stable 10–13°F seasonal differentiation with only two programs, demonstrating that minimal vocabulary achieves high thermal differentiation when setpoints are configured once with clear seasonal intent.

Two results from the analysis reinforce these distinctions. First, the number of schedules a household maintains does not predict how long the system runs, which parallels the finding that manual adjusters can save as much energy as schedule users when setpoint calibration is the operative mechanism (Huchuk et al., 2021). Second, the meaning of a schedule name is not arbitrary, since it carries behavioral intent that translates into measurable operational differences, a signal that event-based compliant or non-compliant typologies do not capture (Blonz et al., 2025; Huchuk et al., 2021).

Overall, programs that push users to create more schedules without guiding setpoint values risk adding complexity without efficiency; whereas, households that have already programmed aggressive setback states may be better positioned to absorb automated DR adjustments without triggering override (Kane & Sharma, 2019).

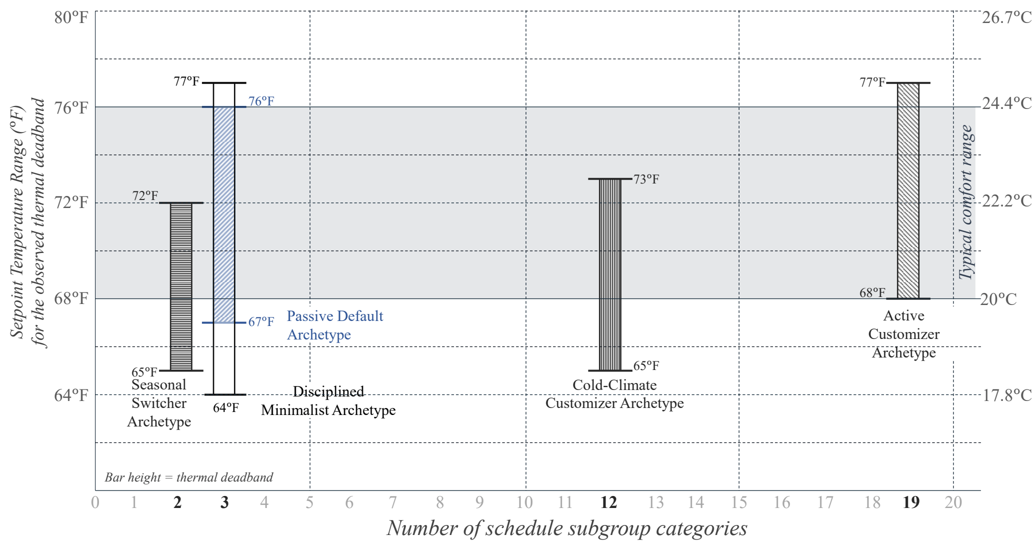

Figure 10. Cross-archetype comparison. *Source*: Own elaboration.

Limitations

Six limitations bound generalizability. The Ecobee DYD sample skews toward newer construction, larger homes, and higher-income households, so the five archetypes may not represent the full behavioral distribution of California's smart thermostat fleet. All associations are observational; causal attribution of HVAC outcomes to schedule categories is not warranted. The data come from a single manufacturer because Ecobee's Donate Your Data program is, to our knowledge, the only initiative through which a thermostat manufacturer shares device-level behavioral telemetry for research. Interface and vocabulary effects specific to Ecobee therefore cannot be separated from general behavior, and comparable data-sharing programs from other vendors, such as Google Nest or Resideo, would be a high-value contribution to understanding how households actually use their thermostats. The category taxonomy classifies only English-language schedule names, excluding non-English labels, numerical names, and idiosyncratic abbreviations. The rapid expansion of California TOU rate structures and Virtual Power Plant programs since 2023 may have accelerated adoption of grid-oriented schedule categories beyond what the 2017–2025 Sankey diagram captures, particularly UTILITY RATE and SOLAR/GENERATION AWARE schedules within the Grid/System-Oriented Settings group.

Finally, archetypes with small sample sizes, such as the *Cold-Climate Customizer* and the *Seasonal Switcher*, are subject to elevated estimation uncertainty and structural confounding. The *Seasonal Switcher*, consistent with a single managed property, is reported descriptively, and characterizations of both groups remain provisional pending replication in larger samples.

Conclusion

This study analyzed nine years of thermostat scheduling behavior from 1,263 California residential smart thermostat devices (2017–2025). We identified five statistically distinct behavioral archetypes and a rule-based semantic taxonomy with 8 main categories and 19 subcategories of user-defined schedule names. Both were linked to measured HVAC outcomes through layered statistical validation. Answers to each of our specific research questions are shared below.

Our research centered on three specific research questions. The first research question asked what distinct behavioral archetypes characterize thermostat scheduling among California smart thermostat users and how they are distributed across the population. Our findings showed five archetypes with a strongly skewed distribution. The *Passive Default Archetype* (63.4%) retains factory schedules with minimal thermal differentiation, the *Active Customizer Archetype* (29.5%) progressively expands a personalized scheduling vocabulary, the *Disciplined Minimalist Archetype* (3.2%) keeps default programs with a deliberately wide deadband, the *Cold-Climate Customizer Archetype* (2.2%) concentrates on extreme nocturnal setback, and the *Seasonal Switcher Archetype* (0.4%) alternates two seasonal configurations consistent with a centrally managed property. Nearly two thirds of devices therefore operate essentially as shipped.

Our second research question focused on whether the semantic content of user-defined schedule names independently predicts HVAC setpoints and runtime after controlling for archetype, climate, and year. It does. Schedule names carry operational information that archetype, year, and outdoor temperature do not absorb, with significant runtime associations for SOLAR/GENERATION AWARE (+307 sec/hr cooling, consistent with precooling), PET (+372 sec/hr heating), SETPOINT AS NAME (+198 sec/hr heating), and AWAY CUSTOM (-101 sec/hr heating). A schedule name encodes behavioral intent that translates into measurable operational differences.

Our third research question asked what share of the population exhibits schedule labeling compatible with grid-interactive demand flexibility and what the remaining majority reveals about the limits of self-directed customization. At most the 29.5% in the *Active Customizer Archetype* qualifies, having developed solar- and rate-oriented vocabulary since 2019 without utility prompting. The *Passive Default Archetype* (63.4%) shows no grid-interactive vocabulary across nine years that included a major occupancy shock, indicating that self-directed customization has reached its ceiling and defining the gap program design must address.

These answers translate directly into program mechanisms. Automated schedule suggestion, enrollment-time pre-configuration, and direct load control with opt-out are the routes most likely to reach the passive majority, while the *Active Customizer Archetype* requires only validation and reinforcement through rate signals and feedback. These patterns establish schedule taxonomy as a scalable behavioral segmentation tool extractable from existing utility datasets without surveys or additional instrumentation.

In sum, these findings argue for a reorientation from a technology deployment model that assumes behavioral homogeneity to a behavioral segmentation model that characterizes actual scheduling archetypes and designs differentiated interventions accordingly. These archetypes provide an empirically grounded starting point for that reorientation in California and, with appropriate replication, across other smart thermostat markets.

This work could be extended in four directions so as to make the results more actionable. First, archetype-stratified analysis of household response to CAISO anchor events, using a Difference-in-Differences design, would provide the event-level behavioral evidence that cross-sectional profiles cannot capture. Second, LLM-based real-time taxonomy classification could generate archetype-aligned schedule suggestions at enrollment and extend behavioral segmentation to linguistically diverse populations. Third, longitudinal modeling of archetype transitions at the device level would identify the conditions under which households migrate from passive to active scheduling, supplying the causal evidence the present observational design cannot. Fourth, cross-vendor replication, whether through data-sharing agreements with other thermostat manufacturers or through utility DR program telemetry, would test whether the archetypes and the schedule-name signal generalize beyond the Ecobee interface, directly supporting the enrollment screening that state and federal VPP targets require.

Acknowledgments

This work was supported by the Fulbright-García Robles Program (COMEXUS) during a research stay at the University of California, Berkeley, in collaboration with the University of California, Davis, and by the School of Architecture, Art and Design, Tecnologico de Monterrey. We gratefully acknowledge the support and contribution of ecobee and ecobee customers to this research.

Disclosure template

Claude (Anthropic) assisted with drafting and restructuring text across multiple sections. OpenAI Codex assisted with generating and debugging Python code for the clustering pipeline and statistical validation, and R scripts for figures. Grammarly was used for grammar review. The authors reviewed and verified all content, calculations, and citations

References

Beltran, K., Simonson, R., & Abbene, B. (2024). *Behavioral Demand Response – 10,000+ Customers Receive Alerts for Needed Energy-Saving During Peak Demand Events*. 2024 ACEEE Summer Study on Energy Efficiency in Buildings. https://www.aceee.org/sites/default/files/proceedings/ssb24/assets/attachments/20240722160803000_d0a22ddf-4d65-4658-ae2d-dbd3dcfc100d.pdf

Blonz, J., Palmer, K., Wichman, C. J., & Wietelman, D. C. (2025). Smart Thermostats, Automation, and Time-Varying Prices. *American Economic Journal: Applied Economics*, *17*(1), 90–125. https://doi.org/10.1257/app.20210618

Bouveyron, C., Celeux, G., Murphy, T. B., & Raftery, A. E. (2019). *Model-Based Clustering and Classification for Data Science: With Applications in R*. Cambridge University Press.

Bureau, U. C. (2026). *Gazetteer Files*. Census.Gov. https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.html

California Energy Commission. (2023). *Senate Bill 846 Load-Shift Goal Report*. California Energy Commission. https://www.energy.ca.gov/publications/2023/senate-bill-846-load-shift-goal-report

DNV. (2024, March 21). *Forward-looking Smart Thermostat Study. California Public Utilities Commission*. https://www.calmac.org/publications/CPUC_GroupA_Fwdlooking_Tstat_FinalReportES.pdf

Ecobee. (2025). *Donate Your Data | ecobee*. https://www.ecobee.com/en-us/donate-your-data/

Huchuk, B., O'brien, W., & Sanner, S. (2021). Exploring smart thermostat users' schedule override behaviors and the energy consequences. *Science and Technology for the Built Environment*, *27*(2), 195–210. https://doi.org/10.1080/23744731.2020.1814668

Islam, M. R. (2018). Sample size and its role in Central Limit Theorem (CLT). *International Journal of Physics & Mathematics*. https://doi.org/10.31295/ijjpm.v1n1.42

Jung, W., Wang, Z., Hong, T., & Jazizadeh, F. (2023). Smart thermostat data-driven U.S. residential occupancy schedules and development of a U.S. residential occupancy schedule simulator. *Building and Environment*, *243*, 110628. https://doi.org/10.1016/j.buildenv.2023.110628

Kane, M. B., & Sharma, K. (2019). *Data-driven Identification of Occupant Thermostat-Behavior Dynamics* (arXiv:1912.06705). arXiv. https://doi.org/10.48550/arXiv.1912.06705

Li, H., O'Brien, W., Loftness, V., Cochran Hameen, E., & Hong, T. (2025). A critical review of use cases and insights from a large dataset of smart thermostats. *Advances in Applied Energy*, *19*, 100236. https://doi.org/10.1016/j.adapen.2025.100236

Luo, N., & Hong, T. (2022). *Ecobee Donate Your Data 1,000 homes in 2017* [Dataset]. https://doi.org/https://doi.org/10.25584/ecobee/1854924

Meier, A., Ueno, T., Rainer, L., Pritoni, M., Daken, A., & Baldewicz, D. (2019). *What can connected thermostats tell us about American heating and cooling habits?* ECEEE 2019 SUMMER STUDY. https://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2019/4-monitoring-and-evaluation-for-greater-impact/what-can-connected-thermostats-tell-us-about-american-heating-and-cooling-habits/

Murphy, S., Miller, C., Deason, J., Dombrowski, D., & Awuah, P. (2024). *The State of Demand Flexibility Programs and Rates*. Lawrence Berkeley National Laboratory. https://doi.org/10.2172/2429694

Nambiar, C., Rosenberg, S., Peffer, T., Schiavon, S., Brager, G., Zhang, H., & Rees, A. (2024). *Enhancing Participation in Residential Demand Response: Insights from Case Studies Conducted in Alaska and California*. 2024 ACEEE Summer Study on Energy Efficiency in Buildings.

thebigbry1971. (2024, May 19). *Ecobee thermostat interface is awful* [Reddit Post]. R/Ecobee. https://www.reddit.com/r/ecobee/comments/1cvq23u/ecobee_thermostat_interface_is_awful/

U.S. Department of Energy. (2025). *Pathways to Commercial Liftoff: Virtual Power Plants 2025 Update*. Office of Technology Transitions.

Wadsworth, H. M. (1998). *The Handbook of Statistical Methods for Engineers and Scientists*. McGraw-Hill.