Energy & Buildings 278 (2023) 112627

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Human decision making during eco-feedback intervention in smart and connected energy-aware communities

Huijeong Kim [a,b], Ilias Bilionis [a,c], Panagiota Karava [a,b,*], James E. Braun [a,b,c]

[a] Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA
[b] Center for High Performance Buildings, Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN, USA
[c] School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA

A R T I C L E I N F O

Article history:
Received 15 July 2022
Revised 26 September 2022
Accepted 31 October 2022
Available online 4 November 2022

Keywords:
Occupant behavior
Human-building interaction
Bayesian modeling
Multi-unit residential building
Smart building
Eco-feedback
Decision-making
Utility theory

A B S T R A C T

Heating and cooling (HC) energy use is responsible for about 42% of the total annual energy consumption of the average household in the U.S and it is significantly affected by residents' energy-related behavior. In this paper, our goal is to realize a new paradigm for energy-aware communities that leverages smart eco-feedback devices and social games to engage residents in understanding and reducing their home HC energy use. Towards this goal, we present a new sociotechnical modeling approach based on utility theory to reveal causal effects in human decision-making and infer attributes affecting the thermostat adjustment behavior of each household during an eco-feedback intervention. Our approach 1) is based on a utility model that quantifies residents' preferences over indoor temperatures given decision attributes related to their thermal environment and eco-feedback design and 2) incorporates latent parameters that determine the unique behavioral characteristics of each household. For parameter learning, we develop a hierarchical Bayesian model with non-centered parameterization calibrated using field data collected from a multi-unit residential community located in Fort Wayne, IN. The dataset comprises two parts; a baseline period without any behavioral intervention, and an intervention period, where personalized eco-feedback and social games are deployed through resident engagement devices with smart thermostat control capabilities, including a wall-mounted tablet and smart speaker. Through the model calibration, we quantify the impact of the eco-feedback on households' thermostat-adjustment behaviors. We propose that the utility model developed in this work can serve as the foundation for analyzing resident behavior in connected residential communities with eco-feedback energy-saving programs.

1. Introduction

The residential sector is responsible for about 22 % of total U.S. energy consumption and about 42 % of the households' annual energy is spent on heating and cooling (HC) on average [1]. To address opportunities for heating/cooling energy savings, recent studies examined behavior-based strategies to motivate residents' energy conservation behavior [2,3]. The impact of the behavior-based strategies is expected to be significant for multi-family residential buildings in the U.S. as about 88 % of the residents are renters who pay their own utility bills and are often prohibited from renovating their apartments [4]. Unlike previous research with a primary emphasis on retrofits, such strategies are reported to be cost-effective without requiring physical building upgrades [5,6]. For example, efficient setpoint control is reported to reduce residential HC energy consumption by 30 % [7]. This is particularly important for low-income residents who spend a larger portion of their income (i.e., about 16 %) on home energy costs than average-income households (i.e., 4 %) [8]. In fact, about 40 % of the total energy costs for low-income residents in the U.S. are associated with their HC use [8]. Within this context, engaging residents in understanding and reducing their home HC energy use by providing eco-feedback has emerged as a cost-effective approach for energy savings.

Numerous behavior-based strategies have been implemented in eco-feedback systems to support residents' decision-making processes and motivate their engagement. Previous studies adopted information-based intervention strategies such as historical comparisons and recommendations, as a tool for affecting short-term behavioral changes and energy savings [9,10]. However, several studies have shown that engagement and savings may decay over time and reach a plateau in long-term interventions [11–13]. To overcome this, researchers have investigated the power of norms,

* Corresponding author.
E-mail address: pkarava@purdue.edu (P. Karava).

https://doi.org/10.1016/j.enbuild.2022.112627

i.e., standards of appropriate behavior, in shaping energy use [14–16]. In particular, descriptive norms which simply compare the energy use of an individual or household to the energy use of other households turned out to reduce the energy consumption of above-average energy consumers by 6 % or more, depending on the comparison group [11,14,17,18], even though many individuals are not aware of the influence of descriptive norms on their behavior [19]. For example, a simple bar graph presenting a resident's energy consumption next to those of his or her neighbors may be sufficient to generate these behavioral changes [11]. Also, a social proof message indicating that "99 % of people in your community reported turning off their lights when not in use" may generate a significant increase in energy-conserving behaviors as well [20]. This type of normative feedback has also been tested in the context of affordable public housing [10,21–24] and shown to reduce energy consumption by more than 10 % [25]. Norms can also play a key role in social games where various gamification strategies (e.g., competition, cooperation, mixture of these) are used for increasing user engagement [26–36]. Specifically, the social proof of neighbors' energy-conserving behaviors can increase social conformity and play a key role in increasing the cooperation rate in goal-oriented collaborative social games [26,34,37,38]. Such energy games could improve residents' energy-related collective behaviors in subsidized affordable housing communities resulting in significant cost savings [39,40].

However, research on energy conservation and norms has shown that responses to feedback [41,42] and normative messages [43] vary to a certain extent among individuals. Specifically, individual responses may vary with the type of intervention strategy [44], season, and residents' thermal preferences [45]. To better understand what works best for individual users, previous studies pursued large-scale surveys to acquire household information and exploit their characteristics (such as socioeconomic status) to construct household-specific feedback recommendations for energy savings [46,47]. However, the efficacy of this type of approach has not been well validated through field deployment, and the relative impact of different feedback treatments was poorly explained due to the limitations in designing a controlled experiment in real-world testbeds [46–48]. To address this issue, recent studies emphasized the importance of modeling the human decision-making process, as an approach to eliciting human decision preferences that affect behaviors during the eco-feedback intervention [49–52].

For modeling human decision-making, a classical decision theory is one of the widely adopted methods for modeling the rationale behind human behavior [53]. This theory explores the decision makers' behavior and choices by formulating their preferences over a set of competing decision alternatives [54]. The preferences can be represented by a utility function, which maps each choice to a scalar that quantifies the outcome of the choice given the attributes that affect the decisions. In this theory, the decision-making process is considered as a maximization of the expected utility that enables rational decision-makers to make choices that would increase their level of satisfaction [54,55]. Such approaches have been gaining attention in the human–computer-interaction context for user modeling [56,57] and can be leveraged to quantify resident decision preferences while further evaluating the impact of the feedback in eco-feedback intervention studies [52].

In this paper, our goal is to realize a new paradigm for energy-aware communities that leverages smart eco-feedback devices and social games to engage residents in understanding and reducing their home HC energy use. Towards this goal, we present a new sociotechnical modeling approach based on utility theory that reveals causal effects in human decision-making and infers attributes affecting the thermostat responses of each household during an eco-feedback intervention. We calibrate the model using field data collected from a multi-unit residential community located in Fort Wayne, IN that is comprised of two parts; a baseline period without any behavioral intervention, and an intervention period, where personalized eco-feedback and social games are deployed through resident engagement devices with smart thermostat control capabilities, including a wall-mounted tablet and smart speaker. Through the model calibration, we quantify the impact of the eco-feedback on households' thermostat-adjustment behaviors.

The rest of the paper is as follows. In Section 2, we present the overview of our field study and summarize key findings that provide the basis for the decision-making model presented in Section 3. The model structure is introduced along with the calibration process in Section 3 while Section 4 outlines the results from the model fitting and the parameter learning. Section 5 summarizes the main conclusions of this study. The conceptual framework of the study is presented in Fig. 1.

2. Research testbed and field study overview

To examine the impact of eco-feedback on resident behaviors, we conducted a field study in a multi-family residential community from December 1st, 2019 to February 13th, 2022 (Section 2.1). The field data comprises two periods; a baseline period without any behavioral intervention and an intervention period where personalized eco-feedback and social games were presented through a wall-mounted tablet and Amazon, Alexa with smart thermostat control capability (Section 2.2). The technology deployment and the resident engagement process are described in Section 2.3 followed by the key findings observed from the field study presented in Section 2.4.

2.1. Research testbed and sensors

Our test-bed is an affordable housing community located in Fort Wayne, Indiana (Fig. 2). The community includes a pair of two-story multi-family apartment buildings occupied by low-income residents whose utility bills are subsidized. Each building consists of 22 apartment units (14, 2-bed and 8, 3-bed units), where 10 units are located on the 1st floor along with a leasing office, and 12 units are located on the 2nd floor. On each floor, all the corners are occupied by 3-bed units while 2-bed units are located along the hallway.

The apartment units are equipped with a WiFi-enabled smart thermostat (i.e., Ecobee4[1]) that controls the apartment's dedicated electric heat pump and air-handling unit. The thermostats measure indoor temperature, relative humidity, and occupancy while storing the user-selected system settings (e.g., heat, cool, off), and temperature setpoint (i.e., heating and cooling). In the apartment units, we also installed sub-circuit smart power meters (i.e., GreenEye Monitor[2]) to measure disaggregated electricity energy usage (heat pump, air handler, plug load, water heater, etc.) of the household. All sensor data are collected through each vendor's web-based application program interface (API) in 5-minute intervals. Weather information, such as outdoor air temperature and relative humidity, is downloaded from National Oceanic and Atmospheric Administration (NOAA[3]) every hour, which provides public weather station data from Fort Wayne International Airport [58]. For interventions, we installed eco-feedback devices including a wall-mounted tablet and Amazon Echo Dot in each apartment unit. These devices allowed

[1] https://www.ecobee.com.
[2] https://www.brultech.com/greeneye.
[3] https://www.ncei.noaa.gov/access/weather-climate-tools.

[50] N. Awalgaonkar, I. Bilionis, X. Liu, P. Karava, A. Tzempelikos, Learning Personalized Thermal Preferences via Bayesian Active Learning with Unimodality Constraints, (2019). http://arxiv.org/abs/1903.09094.
[51] J. Xiong, S. Lee, P. Karava, A. Tzempelikos, Personalized visual satisfaction profiles from comparative preferences using Bayesian inference, Energy Procedia 122 (2017) 547–552, https://doi.org/10.1016/j.egypro.2017.07.407.
[52] L.J. Ratliff, M. Jin, I.C. Konstantakopoulos, C. Spanos, S.S. Sastry, Social Game for Building Energy Efficiency: Incentive Design, (2014). http://escholarship.org/content/qt3ww649ck/qt3ww649ck.pdf.
[53] J.O. Berger, Statistical Decision Theory and Bayesian Analysis, 2nd ed.,., Springer Science & Business Media, 2013.
[54] J. Von Neumann, O. Morgenstern, Theory of games and economic behavior, Princet. Univ, Press, 1955.
[55] Detlof V. Winterfeldt, Gregory W. Fischer, MULTI-ATTRIBUTE UTILITY THEORY: MODELS AND ASSESSMENT PROCEDURES, Natl. Tech. Inf. Serv. (1973).
[56] S.J. Payne, A. Howes, Adaptive Interaction: A Utility Maximization Approach to Understanding Human Interaction with Technology, 1st ed.,., Morgan & Claypool Publishers, 2013.
[57] A. Jameson, B. Berendt, S. Gabrielli, F. Cena, C. Gena, F. Vernero, K. Reinecke, Choice Architecture for Human-Computer Interaction, 2013. http://www.nowpublishers.com/articles/foundations-and-trends-in-humancomputer-interaction/HCI-028.
[58] National Oceanic and Atmospheric Administration (NOAA), Climate Data Online Data Tools (https://www.ncdc.noaa.gov/cdo-web/datatools), (2020).
[59] H. Kim, S. Ham, M. Promann, H. Devarapalli, G. Bihani, T. Ringenberg, V. Kwarteng, I. Bilionis, J.E. Braun, J.T. Rayz, L. Raymond, T. Reimer, P. Karava, MySmartE – An eco-feedback and gaming platform to promote energy conserving thermostat-adjustment behaviors in multi-unit residential buildings, Build. Environ. 221 (2022).
[60] H.A. Simon, Bounded rationality, in: Util. Probab., Palgrave Macmillan, London, 1990: pp. 15–8.
[61] R. McElreath, Statistical Rethinking, Stat. Rethink. (2018), https://doi.org/10.1201/9781315372495.
[62] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin, Bayesian Data Analysis, n.d.
[63] M. Betancourt, M. Girolami, Hamiltonian Monte Carlo for Hierarchical Models, Curr. Trends Bayesian Methodol. Appl. (2015) 79–101, https://doi.org/10.1201/b18502-5.
[64] M. Betancourt, Hierarchical Modeling, (2020). https://betanalpha.github.io/assets/case_studies/hierarchical_modeling.html#1_Modeling_Heterogeneity.
[65] C.M. Bishop, Pattern Recognition and Machine Learning, New York, NY (2006), https://doi.org/10.1007/978-3-642-46890-2_1.
[66] J. Salvatier, T.V. Wiecki, C. Fonnesbeck, Probabilistic programming in Python using PyMC3, PeerJ Comput. Sci. 2016 (2016) 1–24, https://doi.org/10.7717/peerj-cs.55.
[67] A. Gelman, J. Hill, Data Analysis Using Regression and Multilevel/Hierarchical Models, Cambridge University Press, 2007.
[68] H. Liu, Y. Wu, B. Li, Y. Cheng, R. Yao, Seasonal variation of thermal sensations in residential buildings in the Hot Summer and Cold Winter zone of China, Energy Build. 140 (2017) 9–18, https://doi.org/10.1016/j.enbuild.2017.01.066.
[69] T. Dietz, Altruism, self-interest, and energy consumption, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 1654–1655, https://doi.org/10.1073/pnas.1423686112.
[70] C. Alós-Ferrer, S. Hügelschäfer, J. Li, Inertia and Decision Making, Front. Psychol. 7 (2016) 1–9, https://doi.org/10.3389/fpsyg.2016.00169.
[71] T. Parkinson, R. de Dear, G. Brager, Nudging the adaptive thermal comfort model, Energy Build. 206 (2020), https://doi.org/10.1016/j.enbuild.2019.109559.

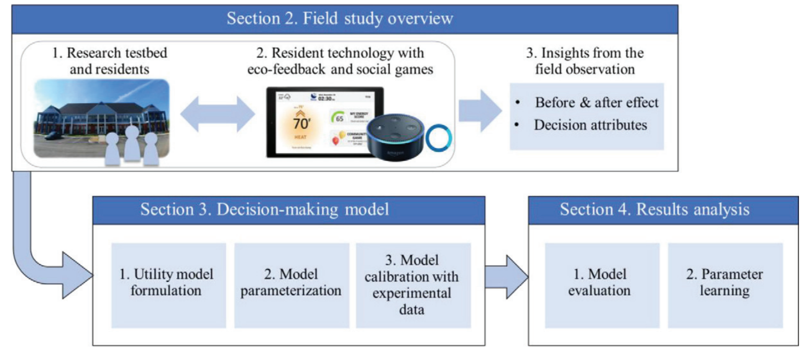

Fig. 1. Conceptual framework of the study.

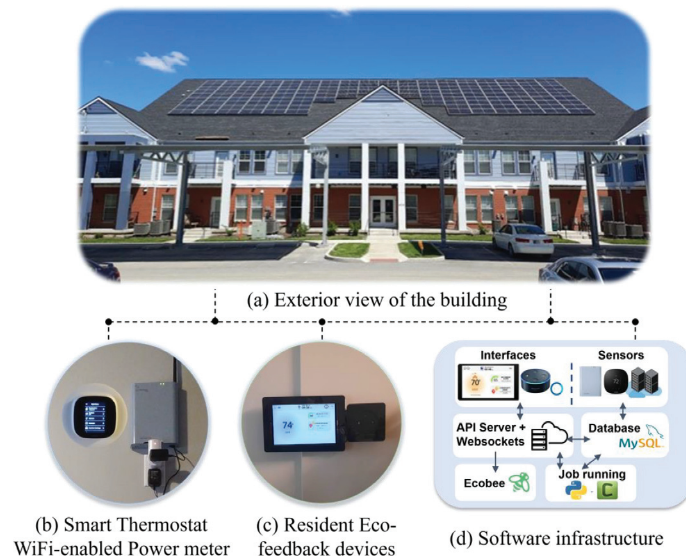

Fig. 2. Overview of the research test-bed.

users to access our resident technology that provides personalized eco-feedback and social games discussed in detail in the next section.

2.2. Resident technology

In each household, we implemented a cloud-based web application, MySmartE [59], which is an eco-feedback and gaming platform designed to help residents effectively adjust their thermostats and reduce their HC energy usage. The app is accessible through a tablet display and Amazon Alexa that allow residents to interact with three major features: 1) personalized eco-feedback that provides a user-centric metric and custom action recommendations based on each household's thermostat usage; 2) a collaborative social energy game for community-level energy-saving initiatives, and 3) smart thermostat control capability provided through a visual (wall-mounted tablet) and voice control (Alexa) user interface.

First, on the 'My energy score' page, the app displays personalized feedback such as an energy score and a daily action recommendation (Fig. 3 (b)). The energy score is designed to evaluate residents' efforts in reducing HC energy usage, with the highest score of 100 denoting the best, and 0 denoting the worst possible performance. To derive the score, the app uses thermostat data from each household collected during the current calendar week and evaluates the setpoints selected in each timestamp using an algorithm presented in [59]. The score is updated every day from Monday to Sunday with the past 7 days of trajectories to inform users about their progress throughout the week. To help residents increase their scores, the app provides a personalized tip that informs them about the best action that could have resulted in the highest increase in their energy score. This tip is presented with an action button (i.e., 'Apply Tip to Schedule') that allows residents to directly implement their tip on the thermostat schedule. The app also enables users to apply the tip verbally or query their current information using the voice user interface (VUI) accessible from the Amazon Echo Dot.

Second, the app allows residents to participate in the social game from the 'Community game' page (Fig. 3(c)). This game is designed to promote community-level energy savings by incentivizing subsidized housing residents in achieving a shared goal. For instance, the residents get a communal mission to increase the community's average energy score to be above or equal to the pre-defined weekly goal. When the community achieves the goal, residents who met the minimum participation requirement become eligible for the game reward (i.e., keeping the score above or equal to the participation threshold by the end of the game week). Each game constitutes four rounds of weekly games where the new game starts every Monday and ends on Sunday with the results announcement. The game information such as daily status, results, or rules is accessible through both the visual interface and the VUI.

Lastly, the app integrates smart thermostat control features. The 'Thermostat panel' in the main interface (Fig. 3 (a)) allows users to update their weekly thermostat schedule to operate the home HC system based on the user-programmed occupancy schedule and the setpoints for specific occupancy modes. The panel also enables users to manually set up their thermostat by selecting a desired system function, setpoint temperature, and the duration to hold the setpoint as well as an option to override the setpoint to their thermostat schedule. Furthermore, the thermostat panel includes an 'Eco-Away' feature so users can efficiently setback or up the setpoint before they leave their home. All the features are controllable not only through the visual interface but also through the VUI that serves as a functional alternative and provides a hands-free option for residents to control thermostats.

To facilitate field experiments with the resident technology, we developed a new software infrastructure that administers multi-modal communications among various software elements (Fig. 2 (d)). This includes a MySQL database for storing multiple data sources such as real-time user interactions with visual and voice user interfaces, sensor readings from the Ecobee thermostat and power meter, and computation results from the feedback algorithms. In addition, a Google Cloud-based API server was built to enable real-time communications among the web and voice applications, connected devices, and a database using a unified API. This API enables our web app to replicate Ecobee functionalities in real-time by supporting data synchronization via polling every 5 s and on-demand WebSockets connections.

2.3. Technology deployment and resident engagement

Our field study was conducted for about two years from December 1st, 2019 to February 13th, 2022. In the first year (from December 1st, 2019 to December 31st, 2020), we collected baseline data to observe residents' seasonal thermostat usage in the absence of any behavioral intervention. Then from January 1st, 2021, we deployed resident technology in each apartment unit and began the eco-feedback intervention. In collaboration with the housing developer and the facility manager of the community, we recruited the residents of 41 households to participate in our study. The residents who agreed to participate received a one-time sign-up payment and a quarterly participation payment. The initial payment was for installing feedback devices and consenting to the intervention while the reoccurring payment was provided in compensation for the technology adoption, any technical inconvenience that may arise during the interventions, and occasional surveys and user interviews that we conducted to collect user feedback. In each unit, we installed the tablet and Echo Dot on the wall located near the living room and dining area. Then, we assisted residents to become familiar with our technology by holding virtual onboarding sessions that guided them on how to use the MySmartE app on the tablet and Alexa.

When the intervention began, the residents were given access to the 'My energy score' page and the thermostat control features in the app. As a default, each household received customized feedback from the 'My energy score' page throughout the entire intervention. The 'Community game' page was only activated during the heating and cooling seasons (i.e., heating: from February 1st to March 28th, 2021 and from December 6th, 2021 to February 13th, 2022, cooling: from June 7th to August 1st, 2021). During the first heating season game, we focused on software improvement based on the initial data collected from the community. After that, the residents were able to participate in the fully operational cooling season community game with the enhanced app usability. At the beginning of the game, 36 out of 41 units participated in the studies as two units opted-out and three units had technical issues with heat pumps and smart devices. Due to a network issue that occurred during the intervention, about 26 units were participating by the end of the game in February 2022. Our field study was designed and conducted in compliance with Purdue's Institutional Review Board (IRB Protocol #: 1702018811) to ensure that all human subjects research was conducted ethically.

Throughout the intervention, residents received daily popups on the tablet display about their energy score, tips, or the game status. During the game season, the participating units received weekly game flyers that summarized the social proof information (e.g., how many neighbors increased the score this week or applied tips) and the new updates on the goals and rewards.

2.4. Experimental results

This section presents field experiment results, where the effect of feedback intervention on resident behaviors was analyzed by comparing the thermostat data collected during the baseline and intervention periods. For the comparison, we first categorized the data based on the data collection periods: 1) baseline data collected before the installation of eco-feedback devices (from December 1st, 2019 to December 31st, 2020) and 2) intervention data affected by the presence of eco-feedback devices (from January 1st, 2021 to February 13th, 2022). Then the intervention data were grouped into individual feedback and community game data based on the pages activated during the intervention. For instance, the individual feedback refers to the data observed when only the 'My energy score' page was available to the residents, while community game data were collected when the 'Community game' page was activated in addition to the 'My energy score' page during the game seasons. For the analysis, we filtered the households who participated in the intervention study for both the heating and cooling seasons and screened out the units that lost baseline data due to network issues, resulting in a final count of 10 households.

Acknowledgment

This work was funded by the National Science Foundation under grant number 1737591 and supported by the Big Ideas Challenge program at Purdue University.

References

[1] EIA, Residential Energy Consumption Survey (RECS), (2015).
[2] J. Froehlich, L. Findlater, J. Landay, C. Science, The Design of Eco-Feedback Technology, (2010).
[3] C. Wilson, H. Dowlatabadi, Models of Decision Making and Residential Energy Use, Annu. Rev. Environ. Resour. 32 (2007) 169–203, https://doi.org/10.1146/annurev.energy.32.053006.141137.
[4] S. Samarripas, Dan York, Lauren Ross, More Savings for More Residents: Progress in Multifamily Housing Energy Efficiency, (2017). https://www.aceee.org/sites/default/files/publications/researchreports/u1702.pdf.
[5] V. Tiefenbeck, A. Wörner, S. Schöb, E. Fleisch, T. Staake, Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives, Nat. Energy 4 (2019) 35–41, https://doi.org/10.1038/s41560-018-0282-1.
[6] K. Steemers, G.Y. Yun, Household energy consumption: A study of the role of occupants, Build. Res. Inf. 37 (2009) 625–637, https://doi.org/10.1080/09613210903186661.
[7] J. Rotondo, R. Johnson, N. Gonzalez, A. Waranowski, C. Badger, N. Lange, E. Goldman, R. Foster, Overview of Existing and Future Residential Use Cases for Connected Thermostats, (2016).
[8] J.F. Eisenberg, Weatherization Assistance Program Technical Memorandum Background Data and Statistics On Low-Income Energy Use and Burdens, (2014).
[9] G. Peschiera, J.E. Taylor, The impact of peer network position on electricity consumption in building occupant networks utilizing energy feedback systems, Energy Build. 49 (2012) 584–590, https://doi.org/10.1016/j.enbuild.2012.03.011.
[10] S.W. Ham, P. Karava, I. Bilionis, J. Braun, Real-time model for unit-level heating and cooling energy prediction in multi-family residential housing, J. Build. Perform. Simul. 14 (2021) 420–445, https://doi.org/10.1080/19401493.2021.1968495.
[11] H. Allcott, T. Rogers, The Short-Run and Long-Run Effects of Behavioral Interventions: Experiment Evidence From Energy Conservation, Am. Econ. Rev. 104 (2012) 3003–3037, https://doi.org/10.1257/aer.104.10.3003.
[12] H.E. Moore, J. Boldero, Designing interventions that last: A classification of environmental behaviors in relation to the activities, costs, and effort involved for adoption and maintenance, Front. Psychol. 8 (2017), https://doi.org/10.3389/fpsyg.2017.01874.
[13] P. Candio, K.B. Pouwels, D. Meads, A.J. Hill, L. Bojke, C. Williams, Modelling decay in effectiveness for evaluation of behaviour change interventions: a tutorial for public health economists, Eur. J. Heal. Econ. (2021), https://doi.org/10.1007/s10198-021-01417-7.
[14] I. Ayres, S. Raseman, A. Shih, Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage, J. Law, Econ, Organ. 29 (2013) 992–1022, https://doi.org/10.1093/jleo/ews020.
[15] A. Laskey, O. Kavazovic, Opower, XRDS Crossroads, ACM Mag. Students. 17 (2011) 47–51, https://doi.org/10.1145/1961678.1961687.
[16] K.S. Wolske, K.T. Gillingham, P.W. Schultz, Peer influence on household energy behaviours, Nat. Energy 5 (2020), https://doi.org/10.1038/s41560-019-0541-9.
[17] Schultz et al., The Constructive, Destructive, and Reconstructive Power of Social Norms, Psycological Sci. 18 (2007) 429–434.
[18] R.K. Jain, J.E. Taylor, G. Peschiera, Assessing eco-feedback interface usage and design to drive energy efficiency in buildings, Energy Build. 48 (2012) 8–17, https://doi.org/10.1016/j.enbuild.2011.12.033.
[19] J.M. Nolan, P.W. Schultz, R.B. Cialdini, N.J. Goldstein, V. Griskevicius, Normative social influence is underdetected, Personal. Soc. Psychol. Bull. 34 (2008) 913–923, https://doi.org/10.1177/0146167208316691.
[20] K. McGonigal, The Willpower Instinct: How Self-Control Works, Why It Matters, and What You Can Do to Get More of It, Penguin, 2011.
[21] S. Woo Ham, P. Karava, Identifying Peer Groups in a Multifamily Residential Building for Eco-Feedback Design, Int. High Perform. Build. Conf. (2018) Paper 319. https://docs.lib.purdue.edu/ihpbc/319.
[22] K. Song, K. Anderson, S.H. Lee, An energy-cyber-physical system for personalized normative messaging interventions: Identification and classification of behavioral reference groups, Appl. Energy 260 (2020), https://doi.org/10.1016/j.apenergy.2019.114237.
[23] P. Petkov, F. Köbler, M. Foth, H. Krcmar, Motivating domestic energy conservation through comparative, community-based feedback in mobile and social media, (2012) 21. doi:10.1145/2103354.2103358.
[24] R.K. Jain, J.E. Taylor, P.J. Culligan, Investigating the impact eco-feedback information representation has on building occupant energy consumption behavior and savings, Energy Build. 64 (2013) 408–414, https://doi.org/10.1016/j.enbuild.2013.05.011.
[25] P. Dolan, R.D. Metcalfe, Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation, SSRN Electron. J. (2015), https://doi.org/10.2139/ssrn.2589269.
[26] D. Wemyss, R. Castri, F. Cellina, V. De Luca, E. Lobsiger-Kägi, V. Carabias, Examining community-level collaborative vs. competitive approaches to enhance household electricity-saving behavior, Energ. Effi. 11 (2018) 2057–2075, https://doi.org/10.1007/s12053-018-9691-z.
[27] I.C. Konstantakopoulos, A.R. Barkan, S. He, T. Veeravalli, H. Liu, C. Spanos, A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure, Appl. Energy 237 (2019) 810–821, https://doi.org/10.1016/j.apenergy.2018.12.065.
[28] H.P. Das, I.C. Konstantakopoulos, A.B. Manasawala, T. Veeravalli, H. Liu, C.J. Spanos, Segmentation Analysis in Human Centric Cyber-Physical Systems using Graphical Lasso, (2018). http://arxiv.org/abs/1810.10533.
[29] J. Novak, M. Melenhorst, I. Micheel, C. Pasini, P. Fraternali, A.E. Rizzoli, Integrating behavioural change and gamified incentive modelling for stimulating water saving, Environ. Model. Softw. 102 (2018) 120–137, https://doi.org/10.1016/j.envsoft.2017.11.038.
[30] P. Ponce, A. Meier, J.I. Méndez, T. Peffer, A. Molina, O. Mata, Tailored gamification and serious game framework based on fuzzy logic for saving energy in connected thermostats, J. Clean. Prod. 262 (2020), https://doi.org/10.1016/j.jclepro.2020.121167.
[31] S. Chadoulos, I. Koutsopoulos, G.C. Polyzos, Mobile apps meet the smart energy grid: A survey on consumer engagement and Machine Learning applications, IEEE Access 8 (2020), https://doi.org/10.1109/ACCESS.2020.3042758.
[32] B. Xia, H. Ming, K.Y. Lee, Y. Li, Y. Zhou, S. Bansal, S. Shakkottai, L. Xie, EnergyCoupon: A case study on incentive-based demand response in smart grid, E-Energy 2017 - Proc. 8th Int. Conf. Futur. Energy Syst. (2017) 80–90. doi:10.1145/3077839.3077862.
[33] D. Foster, S. Lawson, M. Blythe, P. Cairns, Wattsup?, in: Proc. 6th Nord. Conf. Human-Computer Interact. Extending Boundaries - Nord. '10, ACM Press, New York, New York, USA, 2010: p. 178. doi:10.1145/1868914.1868938.
[34] C. Grevet, J. Mankoff, S.D. Anderson, Design and evaluation of a social visualization aimed at encouraging sustainable behavior, Proc. Annu. Hawaii Int. Conf. Syst. Sci. (2010) 1–8, https://doi.org/10.1109/HICSS.2010.135.
[35] J.E. Petersen, V. Shunturov, K. Janda, G. Platt, K. Weinberger, Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives, Int. J. Sustain. High. Educ. 8 (2007) 16–33, https://doi.org/10.1108/14676370710717562.
[36] P. Fraternali, S. Herrera, J. Novak, M. Melenhorst, D. Tzovaras, S. Krinidis, A.E. Rizzoli, C. Rottondi, F. Cellina, EnCOMPASS - An integrative approach to behavioural change for energy saving, GIoTS 2017 - Glob. Internet Things Summit, Proc. (2017) 1–6. doi:10.1109/GIOTS.2017.8016256.
[37] V. De Luca, R. Castri, The Social Power Game: A Smart Application for Sharing Energy-Saving Behaviours in the City, Proc. AVI 2014 Work. Foster. Smart Energy Appl. through Adv. Vis. Interfaces. (2014) 1–4. doi:10.1145/2598153.2602224.
[38] A. Järvinen, Game design for social networks: Interaction design for playful dispositions, Proc. 2009 ACM SIGGRAPH Symp. Video Games, Sandbox '09. 1 (2009) 95–102.
[39] O. Sachs, V. Tiefenbeck, C. Duvier, A. Qin, K. Cheney, C. Akers, K. Roth, Field Evaluation of Programmable Thermostats, (2012).
[40] K. Lusson, SMART THERMOSTATS :Assessing Their Value in Low-Income Weatherization Programs, (2020).
[41] K. Anderson, S.H. Lee, An empirically grounded model for simulating normative energy use feedback interventions, Appl. Energy 173 (2016) 272–282, https://doi.org/10.1016/j.apenergy.2016.04.063.
[42] A. Khosrowpour, R.K. Jain, J.E. Taylor, G. Peschiera, J. Chen, R. Gulbinas, A review of occupant energy feedback research: Opportunities for methodological fusion at the intersection of experimentation, analytics, surveys and simulation, Appl. Energy 218 (2018) 304–316, https://doi.org/10.1016/j.apenergy.2018.02.148.
[43] O.I. Asensio, M.A. Delmas, Nonprice incentives and energy conservation, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) E510–E515, https://doi.org/10.1073/pnas.1401880112.
[44] B. Morschheuser, J. Hamari, A. Maedche, Cooperation or competition – When do people contribute more? A field experiment on gamification of crowdsourcing, Int. J. Hum Comput Stud. 127 (2019) 7–24, https://doi.org/10.1016/j.ijhcs.2018.10.001.
[45] B. Huchuk, W. O'Brien, S. Sanner, A longitudinal study of thermostat behaviors based on climate, seasonal, and energy price considerations using connected thermostat data, Build. Environ. 139 (2018) 199–210, https://doi.org/10.1016/j.buildenv.2018.05.003.
[46] Q. Han, I. Nieuwenhijsen, B. de Vries, E. Blokhuis, W. Schaefer, Intervention strategy to stimulate energy-saving behavior of local residents, Energy Policy. 52 (2013) 706–715, https://doi.org/10.1016/j.enpol.2012.10.031.
[47] A. Rinaldi, M. Schweiker, F. Iannone, On uses of energy in buildings: Extracting influencing factors of occupant behaviour by means of a questionnaire survey, Energy Build. 168 (2018) 298–308, https://doi.org/10.1016/j.enbuild.2018.03.045.
[48] L. Pereira, N. Nunes, Understanding the practical issues of deploying energy monitoring and eco-feedback technology in the wild: Lesson learned from three long-term deployments, Energy Rep. 6 (2020) 94–106, https://doi.org/10.1016/j.egyr.2019.11.025.
[49] X. Liu, S. Lee, I. Bilionis, P. Karava, J. Joe, S.A. Sadeghi, A user-interactive system for smart thermal environment control in office buildings, Appl. Energy 298 (2021), https://doi.org/10.1016/j.apenergy.2021.117005.

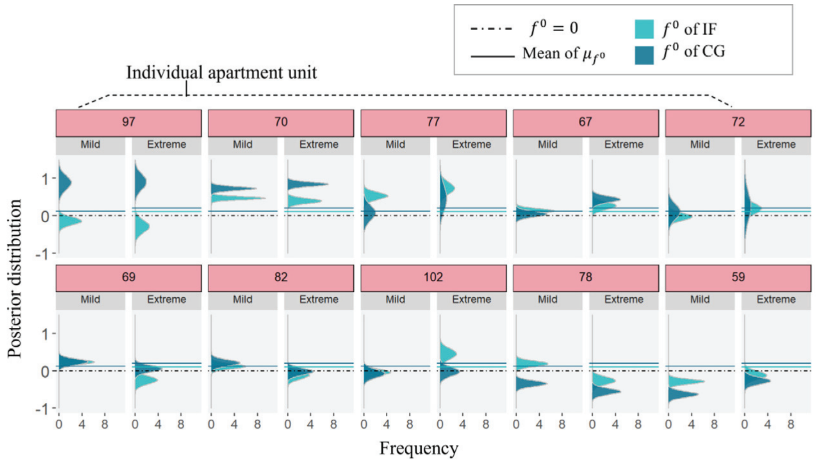

Fig. 12. Posterior distribution of f_0 for heating season. The overall impact of the intervention was smaller than that observed during the cooling season.

data observed in Fig. 4. Other apartment units such as 69, 78, and 70 have their f^0 located slightly lower than the population mean for all feedback types during the mild weather, which represents a smaller impact of the intervention compared to their neighbors. In addition, one can infer the preferences of individual households' toward certain feedback types by comparing f_0 estimated for different treatments. For example, the impact of the CG was higher than the IF for households 77 and 78 while the opposite trend was observed for 72, 59, and 82.

During the heating season, the population mean of f^0 was estimated above but near the 0 lines, indicating that the overall impact of the intervention was smaller than that observed during the cooling season (Fig. 12). For mild weather, the marginal impact of the IF and the CG was almost identical while the impact of the community game was slightly higher during the colder weather. Regardless of the small marginal impact, we observed a more diverse range of the feedback impact during the heating season than that observed in the cooling season. For instance, units 97 and 70 distinctively improved their behavior while playing the CG while some other units showed the opposite impact. For example, household 59 had a negative intervention impact ($f^0 < 0$) while receiving the IF and participating in the CG for all types of weather.

5. Conclusions

In this study, we introduced a sociotechnical modeling approach based on utility theory to reveal causal effects in human decision-making and infer attributes affecting the thermostat responses of each household during the eco-feedback intervention. A utility model was developed to determine the unique behavioral characteristics of each household. The model structure and parameters were designed based on field data collected from a multi-family residential community with an eco-feedback gaming platform deployed in each apartment. For parameter learning, a hierarchical Bayesian model was designed with a non-centered parameterization and it wascalibrated with field data collected for about two years. The calibration results revealed that the estimates of the resident baseline temperature preferences aligned with the room air temperatures that the residents maintained during the baseline period. Our model also captured the trend that the households' temperature preference varies at the different ranges of the outdoor air temperature. In addition, the impact of the eco-feedback was quantified by adopting a latent variable and the estimation results showed that all types of feedback interventions had a positive impact on energy-conserving behaviors during the cooling season, increasing the room air temperatures of most households from the baseline. The overall impact of the intervention was smaller during the heating season than that observed during the cooling season. Nonetheless, a wider range of feedback impact was observed in the heating season than in the cooling season. One thing to note is that such diversity in the impact could have been caused by various factors that determine the decision-characteristics of residents (e.g., seasonal thermal preference [45,68], pro-environmental attitude [3], attitude toward personal interest or collaboration [69], decision inertia [70], adaptative thermal comfort [71], etc). To analyze the detailed causal effects of such factors on human decision-making, extensive data collection needs to be conducted with increased population size in future work. Towards this direction, we propose that the utility model may become a generalized approach for analyzing resident behavior in connected residential communities with eco-feedback energy-saving programs.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

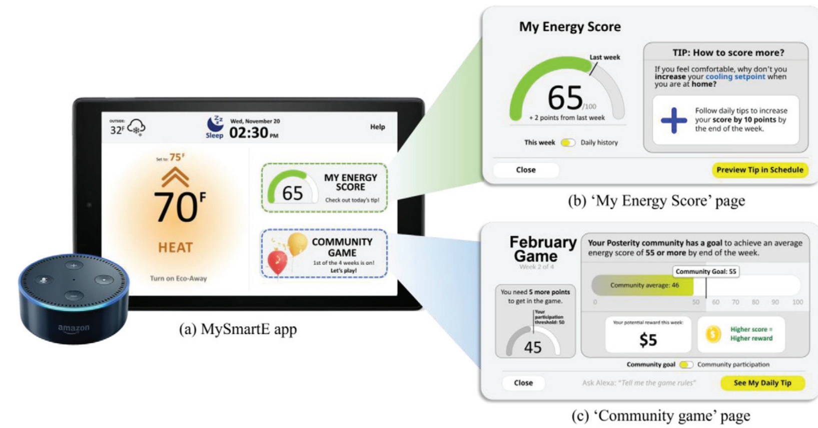

Fig. 3. User interface of MySmartE app.

Next, we categorized heating and cooling season data based on the prevailing HC operational mode observed in the community [59]. From the data, the households' heat pump system was mainly operating in cooling mode when the daily average outdoor air temperature was higher than or equal to 18℃, while the heating mode was prevailing when the temperature was lower than or equal to 7℃. We labeled the heating and cooling days based on these outdoor air temperature thresholds and filtered the thermostat data collected from each household during their occupied hours. Then, we compared the room air temperatures of each household observed during the baseline period to those observed in the individual feedback and community game periods under similar outdoor air conditions. For each data set, the hourly room air temperatures were grouped by the outdoor air condition bin defined based on the range of the dry-bulb temperature and wet-bulb temperature. For each bin, we resampled the hourly indoor temperature data to compare the properties of the three different datasets under the same sample size [59]. Finally, the bootstrap medians were calculated from each bin and are illustrated in Fig. 4 for each household with 95 % credible intervals.

Fig. 4 (a) and (b) show the room air temperatures that each household selected during the baseline and intervention periods of the cooling and heating seasons. In the graphs, the baseline data are presented in the red box while the intervention data collected during the individual feedback (IF) and the community game (CG) experiments are highlighted in green and blue boxes, respectively. At a glance, one can notice that the changes in the household room air temperature are more distinct in the cooling season compared to the heating season. During the cooling season, 90 % of the households increased their indoor temperature by $1 \sim 2$℃ after receiving either IF or CG treatments from the baseline. Considering that this is a weather normalized comparison, we can infer that such changes were caused mainly by the improvement in residents' thermostat adjustment behaviors resulting in an increase in their usual cooling setpoints. On the other hand, the impact of the intervention was mitigated during the heating season as about only 30 % of the households improved their behavior by decreasing their household temperature (e.g., by lowering the heating setpoint or turning off the HC), while 30 % showed reversed behavior (i.e., increased temperature), and the rest of the households maintained their baseline temperature during the intervention period.

Additionally, the magnitude of the temperature changes varied for different households and the outdoor air condition due to the heterogeneity of resident preferences. Specifically, the households maintained a diverse range of baseline temperatures based on their thermal preferences without concern about utility bills. For instance, the baseline temperature ranged from 20 to 26℃ and from 20 to 28℃ during the cooling and heating seasons, respectively. Moreover, the baseline temperature preference also seemed to vary by the range of the outdoor air temperature within the same season. We also observed that the impact of the IF and CG varied by different household and weather conditions. Based on these observations, we assume that resident behavior change depends on the type of feedback, which may affect their level of motivation to take action, and their preferences about the thermal environment in their home.

3. Modeling methodology

In this section, we present the development of the utility model that quantifies the households' preferences over indoor temperatures given the decision attributes related to their thermal environment and the feedback information. In the following sub-sections, we present the model details, parameterization, and calibration process using the experimental data.

3.1. Household utility model

This section presents the utility model that mathematically describes residents' HC-related decision-making during the baseline and intervention period. We designed the model based on the main experimental observations presented in Section 2.4.

First, we define the decision variables for household i that are related to residents' HC behavior. In our field experiment, residents can take various actions to maintain their desired thermal environment. For instance, residents utilize a mix of various thermostat control features in the MySmartE app to program thermostat schedules, temporally hold temperature setpoints, or change the

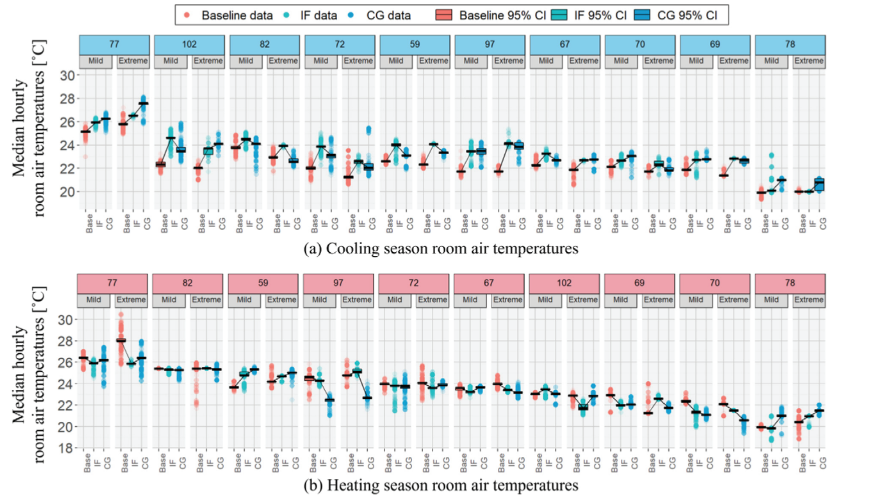

Fig. 4. Room air temperature before and after the intervention in each unit. Two major trends are observed: 1) the changes in the household room air temperature are more distinct in the cooling season compared to the heating season and 2) the magnitude of the temperature changes vary for different units and outdoor air conditions due to the heterogeneity of resident preferences.

mode of the HC system setting. Considering that these actions lead to a change in the indoor air temperature, let us define the decision variable for each household i to be the average room air temperature T_i^r, which is the composite outcome of residents' thermostat-adjustment behaviors.

Then, we formulate the utility of an individual household, $U^H_{i,k,s,l}$, to represent the residents' preferences over a decision variable T_i^r. We consider the utility as a numerical representation of the level of satisfaction that the residents perceive by choosing a certain temperature T_i^r for season s, type of feedback intervention k, and the weather intensity l. The details of the group variables (i.e., k, s, l) will be further discussed in Section 3.2. In the current setup, we assume the household utility, $U^H_{i,k,s,l}$, to be affected by two additive decision attributes that represent residents' satisfaction with the thermal environment,$U^C_{i,s,l}$, and their satisfaction with the eco-feedback intervention, $U^F_{i,k,s,l}$, as shown in Eq. (1):

$$U^H_{i,k,s,l}(T_i^r) = U^C_{i,s,l}(T_i^r) + U^F_{i,k,s,l}(T_i^r). \tag{1}$$

We model $U^C_{i,s,l}$ based on the assumption that the level of satisfaction decreases if the indoor temperature deviates from the residents' preferred value. Assuming that residents' thermal preference is best reflected in the selected room air temperatures during the baseline period, we design $U^C_{i,s,l}$ to have its maximum value at $T^{rc}_{i,s,l}$ which is the room air temperature that residents in the household i preferred during the baseline period. Considering that resident thermal preference is subject to seasonal variations [45], we group $U^C_{i,s,l}$ and $T^{rc}_{i,s,l}$ by seasonal index s and the weather intensity index l. A plausible choice of $U^C_{i,s,l}$ is a concave quadratic function:

$$U^C_{i,s,l}(T_i^r) = -0.5\left(T_i^r - T^{rc}_{i,s,l}\right)^2, \tag{2}$$

where, $U^C_{i,s,l}$ is a quadratic function of T_i^r that represents the thermal satisfaction of residents in household i, at room air temperature T_i^r. This type of function has been also adopted in [49,52] to represent the thermal comfort level of the occupants in office environments.

Another component of the utility, $U^F_{i,k,s,l}$, is designed to represent residents' satisfaction with adopting pro-environmental behaviors after being motivated by the individual feedback or social game information shared through our eco-feedback devices during the intervention period. In this work, we constrain the residents' pro-environmental behaviors to their attempts to reduce the use of HC by adjusting the thermostat (e.g., applying tips, holding or scheduling energy conservative thermostat setpoints, turning off the HC system). Such attempts are likely to increase the household's energy score, improve its performance in the social game, and eventually increase residents' satisfaction levels. In addition, these lead to an increase in the indoor air temperature during the cooling season or a decreased indoor air temperature during the heating season. With these in mind, we expect $U^F_{i,k,s,l}$ to be a linear function that increases when the room air temperature shifts in the direction that consumes less energy for the heating and cooling seasons. A simple model is:

$$U^F_{i,k,s,l}(T_i^r) = \eta\left(\psi f^0_{i,k,s,l} T_i^r - (1-\psi) f^0_{i,k,s,l} T_i^r\right), \tag{3}$$

where, $U^F_{i,k,s,l}$ as a linear function of T_i^r that is only present during the intervention period. The term η denotes the presence of eco-feedback during the experiments with a binary value (baseline period without any feedback: 0, intervention period when personalized feedback or social games were provided: 1) while ψ represents the type of the season during the intervention (heating: 0, cooling: 1). $f^0_{i,k,s,l}$ is the satisfaction of the residents in household i for a certain type of feedback intervention k during the season s

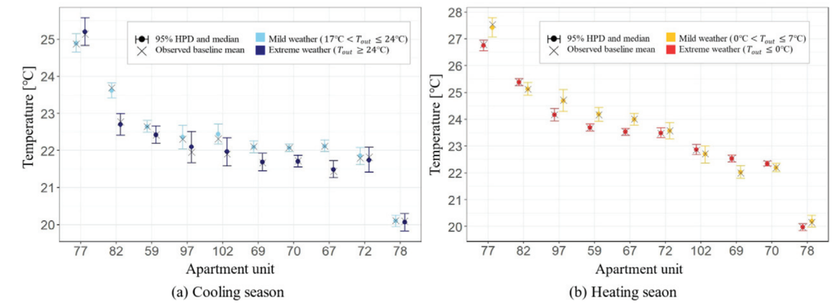

Fig. 10. Estimated T^{rc} of individual households during the cooling and heating seasons. T^{rc} aligns with the average indoor temperature observed during the baseline period.

respectively. As defined in Section 3.1, this parameter represents the satisfaction of the residents in household i for a certain type of feedback intervention, season, and weather intensity. The parameter quantifies the effect of the eco-feedback by estimating the relative changes in the household's room air temperature that occurred after implementing the feedback compared to the baseline T^{rc} (Eq. (7)). From this, we can interpret positive f^0 as the positive effect of feedback treatment that reduced the residents' energy use by increasing their room air temperature during the cooling season or decreasing the temperature during the heating season. On the other hand, negative f^0 represents the reversed effect of feedback that leads to an increase in energy consumption by increasing the room temperature for heating and decreasing it for cooling.

In Fig. 11, we illustrate the f^0 of individual apartments estimated with respect to the type of feedback and the weather intensity. For each apartment unit, we plot the posterior distribution of f^0 with a dotdashed line located at 0 to examine the effect of the feedback. One can notice that the distributions of f^0 are located above 0 in most of the households during the cooling season. This reveals that most of the units increased their room air temperature regardless of the type of feedback provided during the intervention. On the graph, we also highlight the posterior mean of the hyperparameter μ_{f^0} with solid lines to visualize the general impact of the intervention observed in the community for different types of feedback and the weather intensity. Considering that the lines of the IF and CG are located close to each other above the 0 lines, we can infer that IF and CG had a similar level of positive impact on the participants in general. The same trend was observed for both levels of weather intensity. For the individual households, we can also infer how each unit reacted to the feedback relative to their neighbors by comparing the location of the μ_{f^0} lines and the distribution of f^0. For example, household 77 and 102 tend to have higher f^0 than the mean of the μ_{f^0} when it comes to CG. This represents the higher impact of the CG on the households' thermostat-adjustment behavior compared to the IF, which aligns with the

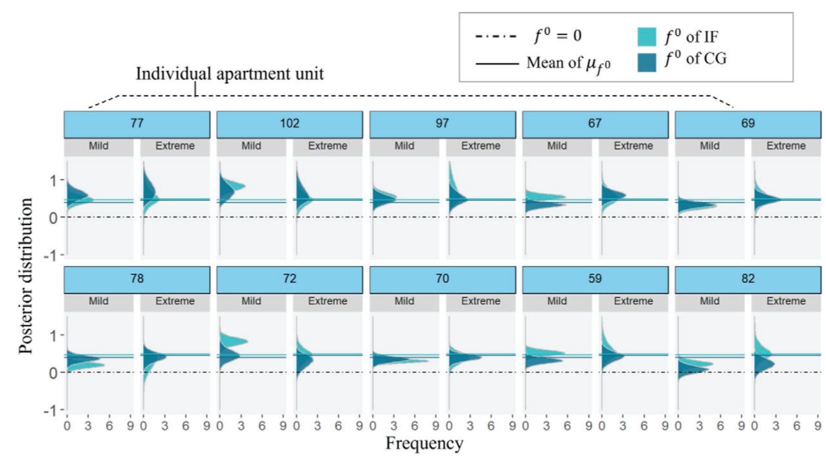

Fig. 11. Posterior distribution of f^0 for cooling season. f^0 are located above 0 in most of the households, indicating the positive impact of the feedback during the intervention periods.

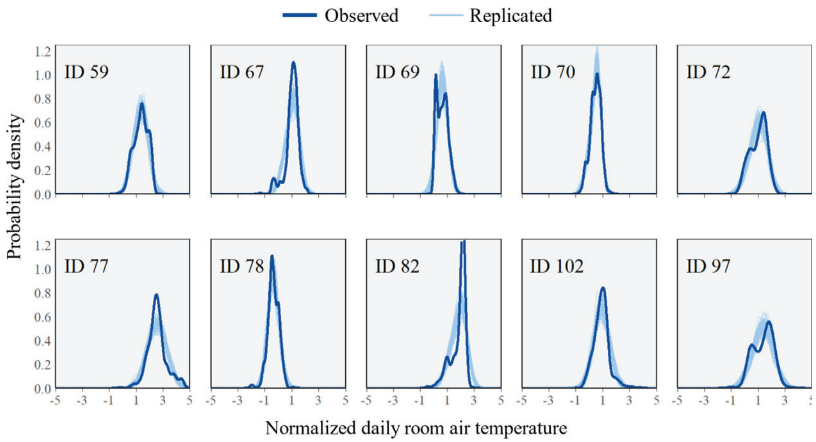

Fig. 8. Household-level posterior predictive checking.

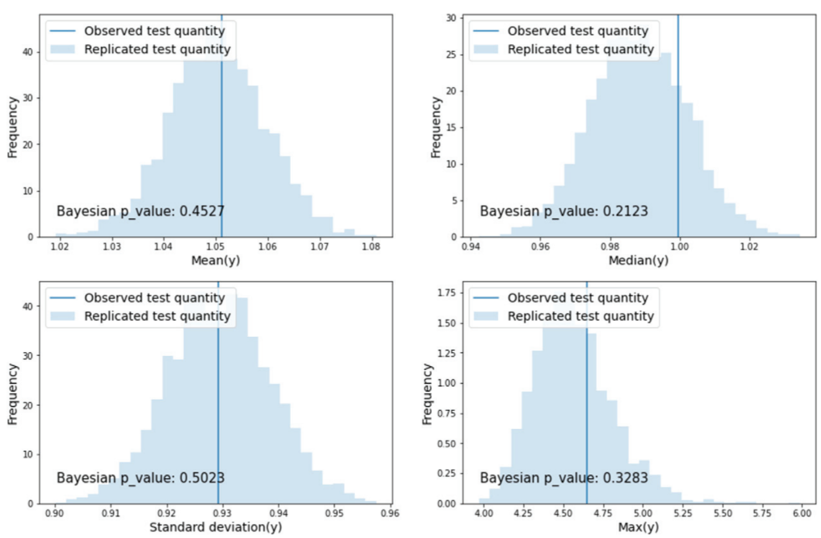

Fig. 9. Bayesian p-values.

most of the households. This supports the validity of our initial assumption that the parameter T^{rc} represents the households' preferred indoor temperature observed during the baseline period. Our model also captures the households' temperature preference trend that varies for different ranges of the outdoor dry-bulb temperature. For instance, the model is able to depict that 70 % of households preferred cooler indoor temperatures when the weather intensity was at a higher level during the cooling season. Similarly, for the heating season, the model depicts that 50 % of the households increased their room air temperature during the colder weather (i.e., higher weather intensity).

4.3. Impact of the eco-feedback

Fig. 11 and Fig. 12 illustrate posterior distributions of f_0 sampled from our model during the cooling and heating season,

and weather intensity l. We assume that $f^0_{i,k,s,l}$ represents how good the residents feel about increasing their energy score and shifting the indoor temperature in the energy-saving direction. In this sense, the positive sign of the $f^0_{i,k,s,l}$ represents the positive effect of the intervention while the negative sign indicates the reversed effect. With the positive $f^0_{i,k,s,l}$, $U^F_{i,k,s,l}$ will be linearly increasing with respect to the increase in T^r_i during the cooling season while the opposite applies for the heating season.

Finally, the utility of an individual household in Eq. (1) is:

$$U^H_{i,k,s,l}(T^r_i) = -0.5\left(T^r_i - T^{rc}_{i,s,l}\right)^2 + \eta\left(\psi f^0_{i,k,s,l}T^r_i - (1-\psi)f^0_{i,k,s,l}T^r_i\right). \tag{4}$$

Based on utility theory [54], we assume that all residents are utility maximizers and take actions T^{r*}_i that return the maximum utility:

$$T^{r*}_i = \arg\max_{T^r_i} U^H_{i,k,s,l}(T^r_i). \tag{5}$$

Taking the derivative of the Eq. (4) with respect to T_r, we get:

$$\frac{\partial U^H_{i,k,s,l}(T^r_i)}{\partial T^r_i} = -T^r_i + T^{rc}_{i,s,l} + \eta\left(\psi f^0_{i,k,s,l} - (1-\psi)f^0_{i,k,s,l}\right). \tag{6}$$

By setting Eq. (6) equal to zero and solving for T^r_i given the parameters $T^{rc}_{i,s,l}$ and $f^0_{i,k,s,l}$, we can predict the household's decision $T^{r*}_{i,k,s,l}$ in Eq. (4) that returns the maximum utility:

$$\widehat{T}^{r*}_{i,k,s,l} = T^{rc}_{i,s,l} + \eta\left(\psi f^0_{i,k,s,l} - (1-\psi)f^0_{i,k,s,l}\right). \tag{7}$$

This is the basis of the structural equation model and the regression performed in subsequent sections.

3.2. Structural equations model parameterization

In this section, we discuss the structural equations of the household utility and the model parameterization process.

First, the structural causal model of the household utility can be formulated based on the Eq.(7) in Section 3.1, where we derived the estimates of the $T^{r*}_{i,k,s,l}$. From the equation, we can infer that the resident's decision on the indoor temperature can be explained by two additive terms, $T^{rc}_{i,s,l}$ and $\eta\left(\psi f^0_{i,k,s,l} - (1-\psi)f^0_{i,k,s,l}\right)$. The first term is the baseline temperature that the residents would choose for increasing their thermal comfort in the absence of the intervention, while the second term represents the effect of the feedback by quantifying the changes in the household's indoor temperature from the baseline. For instance, $f^0_{i,k,s,l}$ measures the relative temperature increased from $T^{rc}_{i,s,l}$ after receiving the feedback during the cooling season ($\eta = 1, \psi = 1$) and the relative decrease during the heating season. The resident's decision is affected by four variables $f^0_{i,k,s,l}$, $T^{rc}_{i,s,l}$, ψ, and η, where the first two are unobserved variables while the rest are observed from the experimental data. $f^0_{i,k,s,l}$ is subject to the types of feedback intervention k (i.e., personalized feedback, a combination of the personalized feedback and the social game), the type of the season s (i.e., heating, cooling), and the intensity of the weather in each season (i.e., mild, extreme). The baseline preferred temperature, $T^{rc}_{i,s,l}$, depends on the season type s and the intensity level l. The ψ depends on the season data and is thus considered an observed variable ($\psi = 1$ if s = "cooling" and 0 if s = "heating"). The graphical causal structure of Eq.(7) is illustrated in Fig. 5 with the observed and latent variables represented in colored and white nodes, respectively.

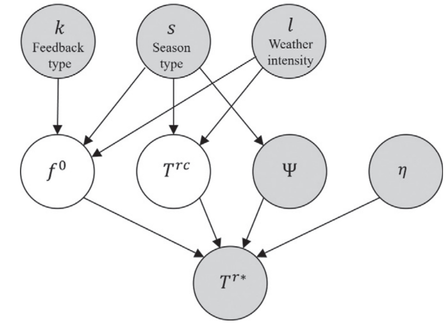

Fig. 5. A structural causal model of an individual household's decision.

With the indoor temperature data (d) collected from each household during the experiment, we can formulate the likelihood of the observed $T^{r*}_{i,k,s,l,d}$ as follows:

$$p\left(T^{r*}_{i,k,s,l,d}|\mu_{T^r i,k,s,l}, {\sigma_{T^r i}}^2\right) = \mathcal{N}\left(T^{r*}_{i,k,s,l,d}|\mu_{T^r i,k,s,l}, {\sigma_{T^r i}}^2\right), \tag{8}$$

$$\mu_{T^r i,k,s,l} = T^{rc}_{i,s,l} + \eta_{i,d}\left(\psi_{i,d} f^0_{i,k,s,l} - (1-\psi_{i,d})f^0_{i,k,s,l}\right), \tag{9}$$

$$\sigma_{T^r i} \sim \mathrm{HalfNormal}(10), \tag{10}$$

where, $\mathcal{N}(\cdot|\mu, \sigma^2)$ denotes the probability density function (PDF) of a univariate Gaussian distribution with mean μ and standard deviation σ. Given Eq.(7), we can model the individual household's likelihood of T^r_i as a Gaussian distribution with the mean $\mu_{T^r i,k,s,l}$ in Eq.(9) and the standard deviation $\sigma_{T^r i}$ specified for each individual household. We use $\sigma_{T^r i}$ to account for the discrepencies between the predicted and the real-world decisions that may arise due to the occupants' limited knowledge or computational capacity [49,60]. To account for such uncertainty and the differences among households, we assigned a weekly informative prior to $\sigma_{T^r i}$ with the standard deviation of 10 as represented in Eq.(10), where HalfNormal(σ) is the PDF of a Half-Normal distribution with the standard deviation σ.

Considering that we do not have strong prior information about the model parameters, we employed a hierarchical Bayesian approach. This allows quantifying the uncertainty in our model without requiring a large amount of data. Also, we formulated the model with a non-centered parameterization to fit the model while accounting for the similar characteristics among all households (i.e., population-level similarity) and the potential variations that may exist in the grouped variables (i.e., inter-group variation) [61,62]. This approach enables us to model the relative deviations of the parameters from the latent population and improve the calibration results by avoiding the so-called "funnel of hell problem" that often arises when fitting hierarchical models [63,64].

To formulate our model, we parameterized $T^{rc}_{i,s,l}$ in a non-centered way using a Normal distribution. The parameter is centered on the population mean $\mu_{T^{rc}s,l}$ with the standard deviation $\sigma_{T^{rc}s,l}$ multiplied by an auxiliary offset variable $\zeta_{T^{rc}i,s,l}$ that accounts for the differences among individual households (Eq.(11)). Then, $f^0_{i,k,s,l}$ is parameterized similarly using a Normal distribution with a population mean $\mu_{f^0 k,s,l}$, standard deviation $\sigma_{f^0 k,s,l}$, and the offset variable $\zeta_{f^0 i,k,s,l}$ (Eq.(12)).

$$T^{rc}_{i,s,l} = \mu_{T^{rc}s,l} + \sigma_{T^{rc}s,l}\zeta_{T^{rc}i,s,l}, \quad (11)$$

$$f^{0}_{i,k,s,l} = \mu_{f^{0}k,s,l} + \sigma_{f^{0}k,s,l}\zeta_{f^{0}i,k,s,l}, \quad (12)$$

where, $\mu_{T^{rc}s,l}$ and $\mu_{f^{0}k,s,l}$ follow Normal distributions (i.e., $\mu_{T^{rc}s,l} \sim \mathcal{N}(0,1)$, $\mu_{f^{0}k,s,l} \sim \mathcal{N}(0,1)$) while $\sigma_{T^{rc}s,l}$ and $\sigma_{f^{0}k,s,l}$ follow Half-Normal distributions (i.e., $\sigma_{T^{rc}s,l} \sim \mathrm{HalfNormal}(1)$, $\sigma_{f^{0}k,s,l} \sim \mathrm{HalfNormal}(1)$). The priors of offset parameters follow a Normal distribution (i.e., $\zeta_{T^{rc}i,s,l} \sim \mathcal{N}(0,1)$, $\zeta_{f^{0}i,k,s,l} \sim \mathcal{N}(0,1)$). The hierarchical structure of the Bayesian model is represented in the plate notation [65] in Fig. 6.

3.3. Model calibration using the experimental data

For the model calibration, we use the field data collected from 10 households during the baseline and intervention periods. Our data set $\mathcal{D}$ includes 7 types of daily observations collected from each household: 1) the unique anonymous id indexed for each household (i); 2) the daily averaged indoor air temperature of the household (T^r_i); 3) the season id indexed for the season type (s); 4) the binary season indicator (ψ), 5) weather intensity id indexed based on different levels of outdoor air temperature (l), 6) the information id indexed for different types of feedback presented to residents (k), and 7) the presence of eco-feedback (η).

The data $\mathcal{D}$ was generated by preprocessing the raw data. For instance, the temperature data in 2) were scaled to have a mean of 0. For 3) and 4), we classified the season by following the method in Section 2.4. Based on the data collected in the field, we labeled "cooling season" for the days having a daily average outdoor air temperature higher or equal to 18℃ and "heating season" when the daily average outdoor air temperature was lower or equal to 7℃. The rest of the days that fell in between the two seasons were classified as a shoulder season and filtered out in this process. For 5), we also defined the range of the outdoor air temperatures to classify the data into two groups (i.e., mild and extreme) for each season based on the levels of outdoor air temperature observed during the experiment. The threshold temperatures for dividing the groups were set to 0℃ for the heating season, and 24℃ for the cooling season, to ensure that each group included a similar number of days of the data. In this regard, the weather is labelled as 'extreme' when the daily outdoor air temperatures are below 0℃ during the heating season and above 24℃ during the cooling season. In all other cases the weather data are labelled as 'mild'. The feedback type represents the type of feedback presented to the household during the intervention, which can be classified into individual feedback and the community game (i.e., IF and CG). Finally, we categorized the baseline and intervention periods based on the presence of eco-feedback during the experiment (i.e., baseline:0, intervention:1).

By adopting a fully Bayesian approach with the observed data $\mathcal{D}$, we can infer the set of unobserved variables in our model $\theta = \left(T^{rc}_{i,s,l}, f^{0}_{i,k,s,l}, \mu_{T^{rc}s,l}, \mu_{f^{0}k,s,l}, \sigma_{T^{rc}s,l}, \sigma_{f^{0}k,s,l}, \zeta_{T^{rc}i,s,l}, \zeta_{f^{0}i,k,s,l}, \sigma_{T^{r}i}\right)$ for $i = \{1,\ldots,H\}, s = \{1,\cdots,S\}, k = \{1,\cdots,K\}$, and $l = \{1,\cdots,L\}$, where H, S, K, and L are total number of households ($H = 10$), seasons ($S = 2$), types of feedback ($K = 2$), and weather intensity level ($L = 2$), respectively. We estimated the latent model parameters by sampling from the posterior distribution using a standard Markov chain Monte Carlo (MCMC) method. For this, we implemented the model using a Python package PyMC3 [66]. The four MCMC chains are used for drawing 4000 samples each while dropping 2000 samples for tuning. The convergence of the MCMC samples is diagnosed by calculating Gelman-Rubin statistics ($\hat{R}$) for all variables θ [62]. This assesses the difference between multiple Markov chains by comparing the estimated between-chains and within-chain variances for each model parameter. The decrease in these variances leads to the convergence in the MCMC samples, which is represented with $\hat{R}$ equal to 1. In our model, the $\hat{R}$ converged to 1 for all variables θ, indicating the convergence of the model.

After the calibration, we evaluated the goodness-of-fit of the model using posterior predictive checks [62]. This process allows simulating replicated data under the fitted model and assessing the model fit by comparing the replicated data to the observed data [67]. To be specific, this type of check accounts for uncertainty associated with the estimated parameters, by sampling the replicated data from the joint posterior predictive distributions given all the estimated model parameters θ from:

$$p(y^{rep}|\mathcal{D}) = \int p(y^{rep}|\theta)p(\theta|\mathcal{D})d\theta, \quad (13)$$

where, $p(y^{rep}|\mathcal{D})$ is a posterior predictive distribution of the replicated data (y^{rep}) given the observed data $\mathcal{D}$. This is calculated by marginalizing the distribution of y^{rep} given the estimated parameters θ over the posterior distribution.

4. Model calibration results

4.1. Posterior predictive checks

We evaluated the goodness-of-fit of the model using posterior predictive checks and in this section present the results with 1) graphical posterior predictive checks and 2) Bayesian p-values.

From the posterior predictive simulation, we sampled replicated room air temperature data (y^{rep}) from the joint posterior predictive distributions given all the estimated model parameters θ. Then we conducted a graphical model checking by comparing the distribution of the sampled data to that of the observed data. Fig. 7 shows the results from the visual comparison conducted with 10 sets of replicated data. In all sets of the distributions, the replicated room air temperature of the community visually aligned with that of the observed data. To check if a similar trend is observed for the individual households' data, we generated 100 sets of replicated data and disaggregated each data set by the household id. Then, for each household, we compared the distribution of the replicated and observed data as shown in Fig. 8. For most households, the two data sets aligned with each other in general.

To quantify the discrepancies between the observed data and the replicated data, we calculated Bayesian p-values for a set of test metrics τ (i.e., mean, standard deviation, 25 % quantile, and maximum values). The Bayesian p-values measure the probability that the replicated data could be more extreme than the observed data [62]. Such probability can be calculated using the pre-defined test metrics τ and should be close to 0.5 to ensure a good model fit.

$$p - \text{value} = \Pr(\tau(y^{rep},\theta) \geq \tau(y,\theta)|y). \quad (14)$$

For estimating the p-values, we calculated test quantities over 3000 samples drawn from the posterior predictive distribution. The test quantities are compared to that of the observed data and the resulting Bayesian p-values are represented in Fig. 9. Most of the p-values are close to 0.5, showing a reasonable model fit.

4.2. Inferred model parameters and user characteristics

4.2.1. Baseline temperature preference

To examine the validity of our assumption that the T^{rc} represents households' baseline indoor temperature preference, we compared the estimated parameter with the observed room air temperature from the baseline data as shown in Fig. 10. The figure

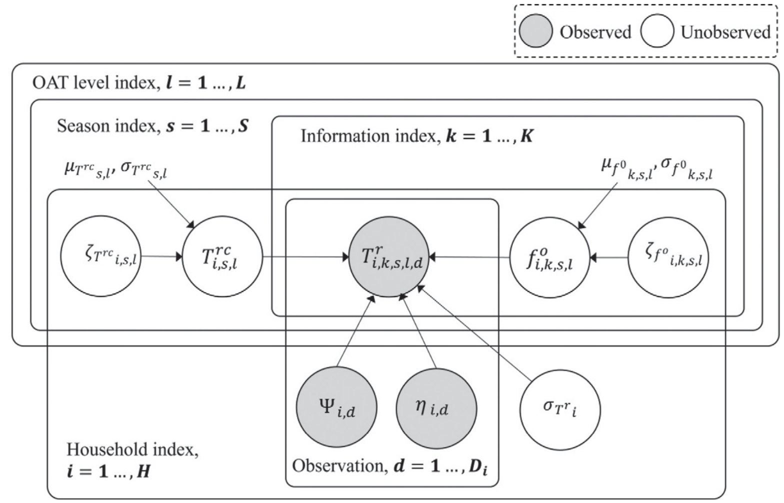

Fig. 6. Plate notation of the hierarchical Bayesian model.

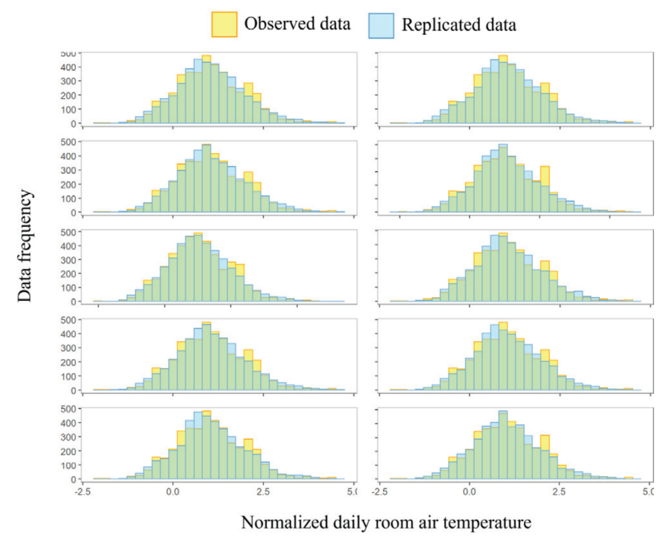

Fig. 7. Community-level graphical posterior predictive checking.

shows individual households' estimated posterior of T^{rc} for different weather intensities during the cooling (Fig. 10(a)) and heating seasons (Fig. 10(b)). For each household, we present the median of the posterior distribution along with the 95 % highest posterior density interval (HPDI). On the plot, we also marked the household's average room air temperature observed during the baseline period. At a glance, the observed data are located within the estimated parameter range and align with the posterior median for