Choosing Who Chooses:

Selection-Driven Targeting in Energy Rebate Programs*

Takanori Ida[1], Takunori Ishihara[2], Koichiro Ito[3], Daido Kido[1],

Toru Kitagawa[4], Shosei Sakaguchi[5], and Shusaku Sasaki[6]

[1]Kyoto University, [2]Kyoto University of Advanced Science, [3]University of Chicago and NBER, [4]Brown University and UCL, [5]University of Tokyo, [6]Osaka University

This version: September 6, 2022.

Abstract

We develop an optimal policy assignment rule that integrates two distinctive approaches commonly used in economics—targeting by *observables* and targeting through *self-selection*. Our method can be used with experimental or quasi-experimental data to identify who should be treated, untreated, and self-select to achieve a policymaker's objective. Applying this method to a randomized controlled trial on a residential energy rebate program, we find that targeting that optimally exploits both observable data and self-selection outperforms conventional targeting for a utilitarian welfare function as well as welfare functions that balance the equity-efficiency trade-off. We highlight that the Local Average Treatment Effect (LATE) framework (Imbens and Angrist, 1994) can be used to investigate the mechanism behind our approach. By estimating several key LATEs based on the random variation created by our experiment, we demonstrate how our method allows policymakers to identify whose self-selection would be valuable and harmful to social welfare.

*Ida and Kido: Graduate School of Economics, Kyoto University, Yoshida, Sakyo, Kyoto 606-8501, Japan (e-mails: ida@econ.kyoto-u.ac.jp and daido.kido@gmail.com). Ishihara: Faculty of Economics and Business Administration, Kyoto University of Advanced Science, 18 Yamanouchi Gotanda, Ukyo, Kyoto 615-8577, Japan (e-mail: ishihara.takunori@kuas.ac.jp). Ito: Harris School of Public Policy, University of Chicago, 1307 East 60th St., Chicago, IL 60637 (e-mail: ito@uchicago.edu). Kitagawa: Department of Economics, Brown University, 64 Waterman St., Providence, RI 02912 (e-mail: toru_kitagawa@brown.edu). Sakaguchi: Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan (e-mail: sakaguchi@e.u-tokyo.ac.jp). Sasaki: Faculty of Economics, Tohoku Gakuin University, 1-3-1 Tsuchitoi, Aoba-ku, Sendai, Miyagi 980-8511, Japan (ssasaki.econ@gmail.com). We would like to thank Keisuke Ito for exceptional research assistance and Judson Boomhower, Severin Borenstein, Fiona Burlig, Mark Jacobsen, Ryan Kellogg, Louis Preonas, Wolfram Schlenker, Frank Wolak, and seminar participants at UC Berkeley, the Global Energy Challenge Conference at University of Chicago, and the Center on Japanese Economy and Business (CJEB) at Columbia Business School for their helpful comments. We thank the Japanese Ministry of Environment for their collaboration for this study. Kitagawa and Sakaguchi gratefully acknowledge financial support from ERC grant (number 715940) and the ESRC Centre for Microdata Methods and Practice (CeMMAP) (grant number RES-589-28-0001). Ito gratefully acknowledge support from Research Institute of Economy, Trade and Industry, and note that this project was conducted as part of a research project "Empirical Research on Energy and Environmental Economics".

1 Introduction

Targeting has become a central question in economics and policy design. When policymakers face budget constraints, identifying those who should be treated is critical to maximizing policy impacts. Advances in machine learning and econometric methods have led to a surge in research on targeting in many policy domains, including job training programs (Kitagawa and Tetenov, 2018), social safety net programs (Finkelstein and Notowidigdo, 2019; Deshpande and Li, 2019), energy efficiency programs (Burlig, Knittel, Rapson, Reguant, and Wolfram, 2020), behavioral nudges for electricity conservation (Knittel and Stolper, 2021), and dynamic electricity pricing (Ito, Ida, and Takana, forthcoming).

Economists generally consider two distinctive approaches to the design of effective targeting. The first approach is based on *observable characteristics*. In this approach, policymakers use individuals' observable data to explore optimal targeting (Kitagawa and Tetenov, 2018; Athey and Wager, 2021). The second approach is based on *self-selection*. In this approach, policymakers consider individuals' self-selection as valuable information to target certain individual types (Heckman and Vytlacil, 2005; Heckman, 2010; Alatas, Purnamasari, Wai-Poi, Banerjee, Olken, and Hanna, 2016; Ito, Ida, and Takana, forthcoming).

A priori, which approach is desirable for policymakers is unclear. For example, referring to the two distinctive approaches above as "planner's decisions" and "laissez-faire," Manski (2013) summarizes,

> *"The bottom line is that one should be skeptical of broad assertions that individuals are better informed than planners and hence make better decisions. Of course, skepticism of such assertions does not imply that planning is more effective than laissez-faire. Their relative merits depend on the particulars of the choice problem."*
>
> —Charles F. Manski, *Public Policy in an Uncertain World*

A common view in the literature, reflected in this quote, is that the appropriate approach depends on the context, and therefore, researchers and policymakers need to decide which to use on a case-by-case basis.

In this study, we develop an optimal policy assignment rule that systematically integrates these two distinctive approaches commonly used in economics. Consider a treatment from which the social welfare gains are heterogeneous across individuals and can be positive, negative, or zero, depending on who takes the treatment. Our idea is that policymakers can leverage both of the *observable* and *unobservable* information by identifying three types of individuals based on their observable characteristics: i) individuals who should

be untreated, ii) those who should be treated, and iii) those who should choose by themselves whether to receive the treatment. Once these individual types are identified, policymakers can design a targeting policy that takes advantage of observed and unobserved heterogeneity in the treatment effect.

We begin by formulating this idea by characterizing a social planner's optimal policy assignment problem following the statistical treatment choice literature (Manski, 2004). Thereafter, we highlight that the Local Average Treatment Effect (LATE) framework (Imbens and Angrist, 1994) can be used to investigate the mechanism behind our approach. When individuals have an option to take a treatment, we can define two individual types. *Takers* are individuals who would take the treatment and *non-takers* are those who would not take the treatment if they could choose whether to receive the treatment. We demonstrate that the planner's decision rule can be characterized by the conditional LATEs for takers and non-takers as well as the conditional average treatment effect (ATE), all conditional on individuals' observable characteristics.

Thereafter, we show that all of the optimal policy assignment rule, LATEs for takers and non-takers, and the ATE can be identified and estimated by a randomized controlled trial (RCT) or a quasi-experiment with three randomly-assigned groups: an untreated group, a treated group, and a self-selection group in which individuals choose whether to take the treatment. To estimate the optimal policy assignment, we use the empirical welfare maximization (EWM) method developed by Kitagawa and Tetenov (2018) with policy trees (Zhou, Athey, and Wager, forthcoming). Further, we demonstarte that the conventional estimation strategy for the LATE (Imbens and Angrist, 1994) can be applied to the three randomly-assigned groups to estimate the LATEs for takers and non-takers.

The theoretical framework described above clarifies what variation has to be generated by an RCT or quasi-experiment to estimate the optimal policy assignment. With this insight, we designed an RCT on a residential electricity rebate program and implemented a field experiment in collaboration with the Japanese Ministry of Environment. The policy goal of the rebate program is to incentivize energy conservation in peak demand hours when the marginal cost of electricity tends to be substantially higher than the time-invariant residential electricity price. In our context, the social welfare gain from this rebate program can be heterogeneous across individuals and can be positive, negative, or zero given the fact that there was a per-household implementation cost. This implies that optimal targeting could improve the social welfare gain from this program.

We randomly assigned households to an untreated group, a treated group, and a self-selection group to generate data for our empirical analysis. Using the data from this RCT, we estimate the optimal policy

assignment, the ATE, and LATEs for takers and non-takers. We begin by presenting substantial heterogeneity in the rebate program's causal impact on peak-hour electricity usage. Customers with different values of household income, self-efficacy in energy conservation, the number of people usually at home on weekdays, and electricity usage in the pre-experimental period responded differently to the treatments. This finding suggests that the optimal policy assignment is likely to differ across household types.

Building upon this insight, we use our framework to quantify the program's social welfare gain for each of the five policies: i) all consumers get untreated, ii) all consumers get treated, iii) all households self-select, iv) optimal targeting without self-selection (*selection-absent targeting*), and v) optimal targeting with self-selection (*selection-driven targeting*). Our findings suggest that although the conventional targeting (selection-absent targeting) outperforms non-targeting policies, the selection-driven targeting substantially improves welfare relative to the selection-absent targeting. The optimal assignment suggests that 24% of households should be untreated, 31% should be treated, and 45% should be assigned to self-selection. The selection-driven targeting would provide an additional 47% of social welfare gain from the rebate program relative to the selection-absent targeting.

We then use the LATE framework described above to investigate the mechanism behind our optimal policy assignment. Given the random assignment in our field experiment, we are able to estimate the LATEs for takers and non-takers conditional on observables. This implies that we can estimate these LATEs for each of the three groups obtained by the optimal assignment rule. Consider households who would be assigned to the self-selection group by the optimal assignment rule. For these households, we find that the LATE for takers is positive and large, and the LATE for non-takers is negative. Hence, self-selection is useful for the planner to sort customers in this group to get treated or untreated by their choice. In contrast, these two LATEs for those who are not assigned to the self-selection group suggest that allowing self-selection hurts social welfare because the planner can obtain higher social welfare gains by assigning them to either compulsory treatment or compulsory un-treatment.

Our main empirical analysis is based on the conventional utilitarian social welfare function. That is, the planner's goal is to optimize the efficiency of the policy, ignoring the equity implications. We emphasize that our framework is not restricted to the utilitarian social welfare function. To shed light on this point, we consider a social welfare function that balances the equity-efficiency trade-off. We use a framework developed by Saez (2002) and used by Allcott, Lockwood, and Taubinsky (2019) and Lockwood (2020). In this framework, the planner can include Pareto weights into a social welfare function to balance the

equity-efficiency trade-off. We demonstrate that our method can quantify the optimal targeting for different degrees of redistribution goals. With this method, the planner can improve the equity of the policy at the cost of having a lower efficiency gain. We find that the selection-driven targeting still outperforms the selection-absent targeting even if we take into account redistribution goals.

Related literature and our contributions—Our study is related to three strands of the literature. First, many recent studies in economics have explored targeting based either on "observables" or "unobservables" through self-selection. Along with the papers cited earlier in this introduction, recent studies on targeting based on individuals' observable characteristics include Johnson, Levine, and Toffel (2020); Murakami, Shimada, Ushifusa, and Ida (forthcoming); Cagala, Glogowsky, Rincke, and Strittmatter (2021); Christensen, Francisco, Myers, Shao, and Souza (2021); Gerarden and Yang (2021) and studies on targeting based on self-selection include Dynarski, Libassi, Michelmore, and Owen (2018); Lieber and Lockwood (2019); Unrath (2021); Waldinger (2021). However, we are not aware of any existing study that builds an algorithm that systematically integrates these two distinctive targeting approaches to maximize a policy's social welfare gain.

Second, our econometric framework builds on the growing statistical treatment choice literature. Generally assuming discrete characteristics, earlier studies in this literature (Manski, 2004; Dehejia, 2005; Hirano and Porter, 2009; Stoye, 2009, 2012; Chamberlain, 2011; Tetenov, 2012, among others) formulate estimation of a treatment assignment rule as a statistical decision problem. The empirical welfare maximization approach proposed by Kitagawa and Tetenov (2018) estimates a treatment assignment rule by maximizing the in-sample empirical welfare criterion over a class of assignment rules. This approach can accommodate multi-armed treatment assignment and a rich set of household characteristics, including continuous characteristics, as is the case for our empirical application. We employ a class of tree partitions considered in Athey and Wager (2021) and Zhou, Athey, and Wager (forthcoming) as our class of policy rules. Finally, building on the LATE framework by Imbens and Angrist (1994), we demonstrate that the newly-defined estimators, the LATEs for *takers* and *non-takers*, can be used to investigate the mechanism behind the optimal policy assignment in the presence of self-selection. These LATE estimands can be viewed as the complier's average treatment effects under a multi-valued discrete instrument, which indexes the three arms randomly assigned in the experiment.

Third, the medical statistics literature has studied hybrid sampling designs that combine randomization and treatment choice by patients. See, e.g., Janevic, Janz, Dodge, Lin, Pan, Sinco, and Clark (2003), Long,

Little, and Lin (2008), and references therein. In the medical literature, the sampling process used in our experiment is referred to as "a doubly randomized preference trial" (Rücker, 1989). An example of a clinical trial that implements a doubly randomized preference design is the Woman Take Pride study analyzed in Janevic, Janz, Dodge, Lin, Pan, Sinco, and Clark (2003). These studies focus on assessing whether letting patients choose their own treatment can have a direct causal effect on their health status beyond the causal effect of the treatment itself. See Knox, Yamamoto, Baum, and Berinsky (2019) for partial identification analysis in such a context and an application to political science. Double randomized preference trials have received less attention in economics. The only study we are aware of is Bhattacharya (2013), which uses double randomization between randomized control trials and planner's allocation to assess the efficiency of the planner's treatment allocations. To our knowledge, no work has analyzed double randomized preference trial data to integrate targeting by observable characteristics and targeting through self-selection.

2 Conceptual Framework

In this section, we present a theoretical framework of optimal policy assignment in the presence of self-selection. We begin by formulating an optimal policy assignment problem in Section 2.1. In Section 2.2, we present that the Local Average Treatment Effect (LATE) framework (Imbens and Angrist, 1994) can be used to investigate the mechanism behind our approach. In Section 2.3, we describe how to empirically estimate the optimal policy assignment and LATEs using data from an RCT and the EWM method.

2.1 Optimal Policy Assignment in the Presence of Self-Selection

Consider a planner who wishes to introduce a policy intervention (program) to a population of interest. Instead of the uniform assignment over the entire population, the planner is interested in targeted assignment for heterogeneous individuals. A novel feature of our setting is that the planner can control not only who is compulsorily exposed to the program but also who is given an option to opt-in to the program. Interpreting an individual's take-up of the program as their exposure to the treatment, the planner's goal is therefore to assign each individual in the population to one of the three arms: *compulsorily treated* (indexed as T), *compulsorily untreated* (indexed as U), and *self-selection* (indexed as S). An individual assigned to T or U is exposed to or excluded from the program with no opt-out or opt-in option, whereas an individual assigned to S chooses whether to take it up by themselves.

The planner's goal is to optimize a social welfare criterion by assigning individuals to these three arms. Following the statistical treatment choice literature (Manski, 2004), we specify the planner's social welfare criterion to be the sum of individuals' welfare contributions. An individual's welfare contribution is a known function of the individual's response to being assigned to arm T, U, or S, and the per-person cost of the treatment. An individual's welfare contribution may not correspond to their utility. Hence, if an individual is assigned to S, their utility maximizing decision may not correspond to the choice that maximizes the planner's objective.

Let W_T, W_U, and W_S denote the potential welfare contributions that would be realized if an individual were assigned to T, U, and S.

We assume that the planner observes a pre-treatment characteristic vector for each individual $x \in \mathcal{X}$, where $\mathcal{X}$ denotes the support of the characteristics. Depending on these observable characteristics, the planner assigns each individual to one of the three arms. Let $G_T \subseteq \mathcal{X}$ denote a set of the pre-treatment characteristics x such that any individual whose x belongs to G_T is assigned to T. Similarly, let G_U and G_S denote sets of the pre-treatment characteristics x such that the individuals with $x \in G_U$ are assigned to U and individuals with $x \in G_S$ are assigned to S, respectively.

We call a partition $G := (G_T, G_U, G_S)$ an *assignment policy*. G describes how individuals are assigned to arms according to their observable characteristics x. The realized welfare contribution after assignment for an individual with characteristics x is either W_T, W_U, or W_S depending on $x \in G_T$, $x \in G_U$, or $x \in G_S$. Hence, her welfare contribution under the policy G can be written as

$$\sum_{j \in \{T,U,S\}} W_j \cdot 1\{x \in G_j\}. \tag{1}$$

Viewing individual characteristics and their potential welfare contributions as random variables, the average welfare contribution under assignment policy G can be written as

$$\mathcal{W}(G) \equiv E\left[\sum_{j \in \{T,U,S\}} W_j \cdot 1\{X \in G_j\}\right], \tag{2}$$

where the expectation is with respect to (W_T, W_U, W_S, X).

We define $\mathcal{W}(G)$ as our social welfare function. The social welfare function depends on the assignment policy G through the post-assignment distribution of individual welfare contributions, which can be manip-

ulated by changing which individuals are assigned to which arms. This form of social welfare is standard in the statistical treatment choice literature. W_j is not restricted to any specific functional form. Therefore, the planner can choose an appropriate social welfare function, including utilitarian and non-utilitarian welfare functions. In our empirical analysis, we highlight this point in Section 6.

The planner's objective is to find the optimal assignment policy G^* that maximizes the social welfare $\mathcal{W}(G)$ over a set of possible assignment policies. If the planner can implement any assignment policy, this set of assignment policies corresponds to the set of measurable partitions of $\mathcal{X}$. Accordingly, G^* can be defined by

$$G^* \in \arg\max_{G \in \widetilde{\mathcal{G}}} \mathcal{W}(G), \tag{3}$$

where $\widetilde{\mathcal{G}} := \{G = (G_T, G_U, G_S) : G \text{ is a measurable partition of } \mathcal{X}\}$.

It is desirable that individuals with characteristics x be assigned to an arm that provides the largest conditional mean welfare contribution among $\{E[W_j|x] : j \in \{T, U, S\}\}$. In the absence of a self-selection treatment arm, the planner's assignment policy is to allocate them to either T or U. The optimal choice is then determined by comparing $E[W_T|x]$ and $E[W_U|x]$. That is, an optimal assignment policy exploits only heterogeneity in the average welfare contribution conditional on observable characteristics x, which can be assessed by the planner prior to assignment. We use $G^\dagger$ to denote this sub-optimal policy assignment and call it *the selection-absent targeting*.

Once individuals are permitted to self-select into treatment, social welfare can be improved beyond the level attained by the selection-absent targeting. This is because an individual may possess private information, which drives or helps predict their response to the treatment, and choose whether to receive treatment based on it. Importantly, there can be significant heterogeneity in the usefulness of self-selection for the planner's objective. Individuals with some values of x choose by themselves the treatment that is optimal in terms of the social welfare. In contrast, individuals with other values of x may choose treatment that does not improve social welfare. Thus, an optimal assignment policy that identifies who should be assigned to S along with T and U could further improve welfare. We use G^* to denote this optimal policy assignment and call it *the selection-driven targeting*. In this case, the planner allocate individuals with x to either T, U, or S by comparing $E[W_T|x]$, $E[W_U|x]$, and $E[W_S|x]$.

2.2 Using the LATE Framework to Investigate the Mechanism

In this section, we present a simple model that clarifies how the optimal assignment policy G^* assigns T, U, and S to individuals in accordance with individual observable characteristics x. We highlight that the framework of the Local Average Treatment Effect (LATE) developed by Imbens and Angrist (1994) can be used to uncover the mechanism behind our approach.

Let $D_S \in \{0, 1\}$ denote the individual's take-up of treatment when assigned to S. The choice D_S may depend on both observable characteristics X and unobservable characteristics (i.e., private information). We define the LATEs for *takers* and *non-takers* as follows, which will be useful statistics to characterize the mechanism of optimal policy assignment.

Definition 2.1. *(The LATEs for takers and non-takers) Let $D_S \in \{0, 1\}$ denote an individual's treatment take-up when they self-select into treatment and (W_T, W_U) denote the treated and untreated potential outcomes. We define the LATE for takers by $E[W_T - W_U | D_S = 1]$ and the LATE for non-takers by $E[W_T - W_U | D_S = 0]$.*[1]

Additionally, we make the following assumption, which is not required for the validity of our method in Section 2.1 but useful to investigate the mechanism.

$$W_S = W_T \cdot 1\{D_S = 1\} + W_U \cdot 1\{D_S = 0\}. \tag{4}$$

An implicit assumption here is that an individual's response to the treatment is the same irrespective of whether they self-select themselves or are assigned to it by the planner. This is similar to the exclusion restriction for instrumental variables, with an indicator for assignment to the self-selection treatment corresponding to an instrumental variable.[2]

We use $p_1(x) = P(D_S = 1|x)$ and $p_0(x) = P(D_S = 0|x)$ to denote the probability of take-up

[1] An alternative way to define $E[W_T - W_U|D_S = 1]$ and $E[W_T - W_U|D_S = 0]$ is to use the average treatment effects on the treated (ATT) and the average treatment effects on the untreated (ATU). $E[W_T - W_U|D_S = 1]$ is the ATT for individuals assigned to group S and $E[W_T - W_U|D_S = 0]$ is the ATU for individuals assigned to group S. In our context, these terms could create confusion because there is another ATT for those assigned to the compulsory treatment group (T) and another ATU for those assigned to the compulsory untreated group (U). To avoid this confusion, we use the terms defined in definition 2.1.

[2] This exclusion restriction is not required for the validity of our method in Section 4 because what the planner needs to do is to compare $E[W_T|x]$, $E[W_U|x]$, and $E[W_S|x]$, and the planner does not need to decompose $E[W_S|x]$.

conditional on x. Using equation (4), $E[W_j|x]$ can be decomposed by,

$$E[W_j|x] = \begin{cases} p_1(x) \cdot E[W_T|D_S = 1, x] + p_0(x) \cdot E[W_U|D_S = 0, x] & \text{if } j = S \\ p_1(x) \cdot E[W_j|D_S = 1, x] + p_0(x) \cdot E[W_j|D_S = 0, x] & \text{if } j \in \{T, U\}. \end{cases} \tag{5}$$

We can use equation (5) to investigate how the planner ranks the three assignments (T, U, S) for individuals with x. First, consider what condition makes the planner prefer S over U. Equation (5) implies that $E[W_S - W_U|x] = p_1(x) \cdot E[W_T - W_U|D_S = 1, x]$. Assuming $p_1(x) > 0$, $E[W_S - W_U|x] \geq 0$ if only if $E[W_T - W_U|D_S = 1, x] \geq 0$. That is, the LATE for takers has to be greater than or equals to 0. Second, consider what condition makes the planner prefer S over T. Equation (5) implies that $E[W_S - W_T|x] = p_0(x) \cdot E[W_U - W_T|D_S = 0, x]$. Assuming $p_0(x) > 0$, $E[W_S - W_T|x] \geq 0$ if only if $E[W_T - W_U|D_S = 0, x] \leq 0$. That is, the LATE for non-takers has to be less than or equals to 0.

Finally, the condition that makes the planner prefer T over U is trivial such that $E[W_T - W_U|x] \geq 0$. Combining the three conditions, we can characterize the optimal assignment policy G^* as defined in equation (3) that has the form $G^* = (G^*_T, G^*_U, G^*_S)$ with

$$\begin{aligned} G^*_T &= \{x \in \mathcal{X} : E[W_T - W_U|x] \geq 0 \text{ and } E[W_T - W_U|D_S = 0, x] > 0\}, \\ G^*_U &= \{x \in \mathcal{X} : E[W_T - W_U|x] < 0 \text{ and } E[W_T - W_U|D_S = 1, x] < 0\}, \\ G^*_S &= \{x \in \mathcal{X} : E[W_T - W_U|D_S = 1, x] \geq 0 \text{ and } E[W_T - W_U|D_S = 0, x] \leq 0\}. \end{aligned} \tag{6}$$

Equation (6) implies that the key statistics that characterize the optimal assignment mechanism are the conditional ATE ($E[W_T - W_U|x]$), the conditional LATE for takers ($E[W_T - W_U|D_S = 1, x)$, and the conditional LATE for non-takers ($E[W_T - W_U|D_S = 0, x]$).

Figure 1 illustrates how an optimal policy G^* partitions customers, using the two-dimensional characteristic space $\mathcal{X}$ as an example. Among the six subspaces from A to F, $CATE(x) := E[W_T - W_U|x]$ (i.e., the conditional ATE) is non-negative only in B, C, and D; $CATE_T(x) := E[W_T - W_U|D_S = T, x]$ (i.e., the conditional LATE for takers) is non-negative only in C, D, and E; $CATE_{NT}(x) := E[W_T - W_U|D_S = U, x]$ (i.e., the conditional LATE for non-takers) is non-negative only in A, B, and C. Therefore, according to the optimal policy characterization (6), B and C are assigned to G^*_T; A and F are assigned to G^*_U; and D and E are assigned to G^*_S.

[Figure 1 about here]

2.3 Estimation

In this section, we describe how data from an RCT allows us to estimate the optimal policy assignment (G^*) presented in Section 2.1 and LATEs for takers and non-takers described in Section 2.2.

To estimate G^*, we use the EWM method in Kitagawa and Tetenov (2018). Let the RCT data be a size n random sample of (W_i, Z_i, X_i), where $Z_i \in \{T, U, S\}$ is individual i's randomly-assigned treatment arm, W_i is their observed outcome (welfare contribution), and X_i are their observable pre-treatment characteristics. Letting $\{W_{T,i}, W_{U,i}, W_{S,i}\}$ denote potential outcomes for individual i, the observed outcome W_i is subject to $W_i = \sum_{j\in\{T,U,S\}} W_{j,i} 1\{Z_i = j\}$. We assume that $\{W_{T,i}, W_{U,i}, W_{S,i}, X_i\}_{i=1,\ldots,n}$ are independently and identically distributed as $\{W_T, W_U, W_S, X\}$.

Using the RCT data and a class $\mathcal{G}$ of policies G, the EWM method estimates an optimal policy G^* by maximizing the empirical analogue of the social welfare function over $\mathcal{G}$:

$$\begin{aligned} \hat{G}^* &\in \arg\max_{G\in\mathcal{G}} \widehat{\mathcal{W}}(G), \\ \widehat{\mathcal{W}}(G) &\equiv \frac{1}{n}\sum_{i=1}^{n} \sum_{j\in\{T,U,S\}} \left(\frac{W_i \cdot 1\{Z_i = j\}}{P(Z_i = j|X_i)} \cdot 1\{X_i \in G_j\} \right), \end{aligned} \tag{7}$$

where $\widehat{\mathcal{W}}(G)$ is an empirical welfare function of G that produces an unbiased estimate of the population social welfare $\mathcal{W}(G)$. Observations are weighted by the inverse of the propensity scores, $P(Z_i = j|X_i)$, which are known from the RCT design.

The EWM approach is model-free: It does not require any assumptions or a functional form specification for the potential outcome distributions. However, the class of policies $\mathcal{G}$ must be specified, taking into account any feasibility constraints for assignment policies. If the class $\mathcal{G}$ is too rich, the EWM solution $\hat{G}_{EWM}$ will overfit the RCT data, and the social welfare attained by the estimated policy falls.

We use a class of decision trees (Breiman, Friedman, Olshen, and Stone, 2017) as $\mathcal{G}$. The main reasons for this choice are the ease of interpretation of the decision tree based assignment policies and the availability of partition search algorithms from the classification tree literature. To illustrate the interpretation of a decision tree-based assignment policy, Figure 2 presents an example decision tree of depth 2 for a two-dimensional $\mathcal{X}$. By traversing a tree from its top node to a bottom node, we map from x to one of the tree

assignment options. This tree structure generates a partition of the characteristic space $\mathcal{X}$ as in Figures 2 (b). A decision tree of depth 2 partitions $\mathcal{X}$ into four subspaces, with individuals whose x belongs to each subspace assigned to one of the three options. Generally, a decision tree of depth L partitions $\mathcal{X}$ into 2^L subspaces.

[Figure 2 about here]

A decision tree of depth L comprises two components: (i) a set of inequalities allocated to the nodes in the top $L-1$ layers and (ii) a set of options allocated to the terminal nodes. Thus searching an optimal decision tree of depth L corresponds to searching for an optimal combination of inequalities in the nodes in the top $L-1$ layers and an assignment option for each terminal node. For the example depth 2 decision tree in Figure A.1, searching for the optimal tree is equivalent to optimally choosing an X for each node in the first and second layers (i.e., triplet of indices $(j,k,l) \in \{1,\ldots,K\}$ of the elements of X where K denotes the dimension of X), threshold values (a_1, a_2, a_3) for these same nodes, and an assignment option $(\text{opt1},\ldots,\text{opt4}) \in \{T,U,S\}^4$ for each of the bottom nodes. Learning an optimal decision tree of depth L by the EWM method corresponds to finding a tree partition that maximizes the empirical welfare function $\mathcal{W}(\hat{G})$ over a class $\mathcal{G}$ of decision trees of depth L. The complexity of the policy class $\mathcal{G}$ can be controlled by fixing the depth of possible decision trees (see, e.g., Zhou, Athey, and Wager (forthcoming)). In our empirical application, we estimate equations (7) in Section 4.

Now, we demonstrate that the LATE for takers ($E[W_T - W_U|D_s = 1, x]$) and non-takers ($E[W_T - W_U|D_S = 0, x]$) can be also identified and estimated by the RCT data. Under the assumptions that the experimental assignment Z is randomly assigned and that equation (4) holds, our identification strategy is the same as that of Imbens and Angrist (1994).[3] In what follows, we denote the observed take-up by $D \in \{0,1\}$, which obeys $D = 1\{Z = T\} + 1\{Z = S, D_S = 1\}$. Furthermore, we suppress the dependence on x for ease of notation, although all expectations are taken conditional on x.

First, we discuss the identification and estimation of the LATE for takers. As illustrated in Section 2.2, the ITT between S and U (i.e., $E[W_S - W_U]$) equals to $p_1 \cdot E[W_T - W_U|D_S = 1]$. Then, the experimental variation of S and U allows us to identify the ITT and p_1 by $E[W|Z = S] - E[W|Z = U]$ and $P(D =$

[3] In general, the identification of LATE requires monotonicity assumption. In our application, this assumption is automatically satisfied by the nature of arms. In fact, if we define $D_T \in \{0,1\}$ and $D_U \in \{0,1\}$ as the individual's potential take-up of treatment when assigned to T and U, it always holds that $1 \equiv D_T \geq D_S \geq D_U \equiv 0$ since non-compliance is not allowed under T and U.

$1|Z = S)$, respectively. As a result, the LATE for takers can be identified by

$$E\left[W_T - W_U | D_S = 1\right] = \frac{E\left[W|Z = S\right] - E\left[W|Z = U\right]}{P(D = 1|Z = S)}. \tag{8}$$

This identification result is simply the application of the conventional LATE framework to experimental groups S and U. Thus, we can estimate this LATE by running the instrumental variable (IV) estimation using data from two groups ($Z \in \{S, U\}$) with the randomly-assigned Z as an instrument for take-up D.

Similarly, the ITT between T and S (i.e., $E[W_T - W_S]$) can be written as $p_0 \cdot E[W_T - W_U | D_S = 0]$. Then, the experimental variation of T and S allows us to identify the ITT and p_0 by $E[W|Z = T] - E[W|Z = S]$ and $P(D = 0|Z = S)$, respectively. As a result, the LATE for non-takers can be identified by

$$E\left[W_T - W_U | D_S = 0\right] = \frac{E\left[W|Z = T\right] - E\left[W|Z = S\right]}{P(D = 0|Z = S)}. \tag{9}$$

As compared to the case of LATE for takers, this result can be regarded as the application of conventional LATE framework to experimental groups T and S. In our empirical application, we estimate equations (8) and (9) in Section 5.1.

3 Field Experiment and Data

The framework in Section 2 highlighted that data from an RCT can be used to estimate the optimal policy assignment in the presence of self-selection. In this section, we describe how we designed and implemented such an RCT in the context of a residential energy rebate program in Japan. Section 3.1 provides an overview of the field experiment. Section 3.2 presents summary statistics and balance test results. Section 3.3 examines heterogeneity in the treatment effects and self-selection. Substantial heterogeneity found in this analysis suggests that optimal targeting that we empirically investigate in Section 4 could substantially improve the policy outcome.

3.1 Field Experiment

We conducted our field experiment in the summer of 2020 in collaboration with the Ministry of the Environment, Government of Japan in the Kansai (around Osaka) and Chubu (around Nagoya) regions of Japan. To include a broad set of households, we invited customers in these regions both by letter and email

with a participation reward with 2000 JPY ($\approx$ 20 USD , given 1 ¢ $\approx$ 1 JPY in the summer of 2020). A total of 4446 customers pre-registered for the experiment. Non-residential customers, those who canceled their electricity contracts in the middle of the experiment, and those who have incomplete high-frequency electricity usage data, were excluded. This left us with 3870 residential customers. That is, our experiment was an RCT for households who agreed to participate in the experiment, which is common in the literature of residential electricity demand (Wolak, 2006, 2011; Ito, Ida, and Takana, forthcoming). Therefore, the external validity of the sample is an important question, which we explore in Section 3.2.[4]

We randomly assigned the 3870 households to one of the following three groups: untreated group (U), treated group (T), and self-selection group (S).[5]

Untreated group (U): 1577 customers did not participate in the rebate program.

Treated group (T): 1486 customers participated in the rebate program.

Selection group (S): 807 customers were asked to choose whether they intended to participate in the rebate program.

The rebate program in our experiment is called the "peak-time rebate" (PTR) program (Wolak, 2011). The fundamental inefficiency in electricity markets in many countries is that residential electricity prices do not fully reflect the time-varying marginal cost of electricity. In peak hours, the time-invariant residential price tends to be too low relative to time-variant marginal cost. This creates a text-book example of short-run deadweight loss. The goal of peak-time rebate programs is to lower this deadweight loss by setting the rebate incentive close to the marginal cost.

It is well known that the PTR is likely to be inferior to simply setting the hourly prices equal to c for two reasons (Wolak, 2011). First, although the PTR provides customers a marginal incentive to reduce their electricity usage, it does not "penalize" the customers if they increase their usage. Second, if a policymaker does not carefully set the "baseline usage", which we discuss below, customers may have an incentive to manipulate their baseline usage to obtain a rebate. This is why economists tend to consider that dynamic pricing is a better option than the PTR (Wolak, 2011; Ito, Ida, and Takana, forthcoming).

However, it is often difficult or impossible for policymakers to implement dynamic pricing for broad population because of political feasibility. The PTR is politically favorable in many countries because

[4]We obtained an Institutional Review Board approval from the ethics committee of the Inter-Graduate School Program for Sustainable Development and Survivable Societies, Kyoto University. Furthermore, We registered the experiment in the AEA RCT Registry (Ida, Ishihara, Kido, and Sasaki, 2020).

[5]The random assignment process was designed such that U: T: S= 2: 2: 1. A relatively large number of households were assigned to the U and T groups in consideration that the data for these groups might be used in future studies in this project.

participating customers do not lose money, although they have a marginal incentive to conserve. This is why the Japanese government, the partner of our experiment, intended to study the PTR rather than dynamic pricing in this research project, and in fact implemented a similar policy in reality in the summer of 2022.[6]

The objective of our PTR was to reduce residential electricity consumption in the system peak hours (between 1 pm and 5 pm) during the week of August 24 to 30, 2020. To prevent customers from manipulating their baseline usage, we did not tell them how the baseline was calculated until August. The baseline usage is each customer's average electricity usage during the peak hours from July 1 to 31. During the treatment week (from August 24 to 30), customers who enrolled in the rebate program received a rebate that was equal to the energy conservation during the peak hours relative to the baseline (kWh) times 100 JPY per 1kW. Customers who enrolled in the program were notified about the information about the treatment week, peak hours, and reward calculation procedure in the beginning of August.

Customers in the selection group (S) were asked to send an email or a prepaid post card during the two-week period from July 31 to August 11 if they intended to participate in the rebate program. The take-up rate was 37.17%, which was rather higher than those for Critical Peak Pricing (CPP) in previous studies.[7] As mentioned above, the PTR never make consumers pay more, unlike the CPP treatment, which may have contributed to the higher take-up rate. At the same time, because the PTR would not make any participating household worse off financially, the take-up rate was far from 100%, which could imply that there were non-financial reasons for a relatively low take-up, including inertia to participate in a new program.

3.2 Data and Summary Statistics

Our primary data is household-level electricity consumption over a 30-minute interval. We collected this data in the pre-experimental period (from 1 Jul, 2020 to 31 Jul, 2020) and in the experimental period (from 24 Aug, 2020 to 30 Aug, 2020). Moreover, we conducted a survey before the experiment to collect a variety of household characteristics.

Table 1 presents summary statistics and balance check. Columns 1, 2, and 3 present the sample average by the randomly-assigned group ($Z = \{U, T, S\}$) with the standard deviation in brackets. Columns 4 to 6

[6] In June 2022, the Japanese government announced the launch of a electricity rebate program called "Setsuden Point Program" to address the expected shortage of electricity supply in the summer of 2022. Similar to the rebate program studied in our field experiment, the Setsuden Point rebate program provided a rebate for electricity customers if they reduced their electricity usage. Likewise, California used similar electricity rebate programs in the California electricity crisis in 2000-2002 and subsequent years (Reiss and White, 2008; Ito, 2015).

[7] The take-up rate for the CPP was 20% in Fowlie, Wolfram, Baylis, Spurlock, Todd-Blick, and Cappers (2021) and 16% (without a take-up incentive) and 31% (with a financial incentive for take-up) in (Ito, Ida, and Takana, forthcoming).

report the difference in sample means with the standard error in parentheses. The first three variables are electricity usage (watt hour per 30-minute) in peak hours (from 1 pm to 5 pm), pre-peak hours (from 10 am to 1 pm), and post-peak hours (from 5 pm to 8 pm). The following three variables are from the survey. “Number of people at home" is the number of household members usually at home on weekdays. The survey also asked the self-efficacy in energy conservation using the 5-point Likert scale, in which higher scores imply higher self-efficacy. The household income is reported in 1000 JPY.

[Table 1 about here]

Columns 4, 5, and 6 suggest that all variables are balanced between the randomly-assigned groups, except for the household income. We do not observe statistically significant difference between the U and S groups, and the T and S groups. However, we do observe a statistically significant difference between the U and T groups (p-value is 0.021). In our empirical analysis, we include household fixed effects to consider the influence of their time-invariant characteristics.

As we described above, our experiment was an RCT for households who agreed to participate in the experiment, and therefore, the external validity of the sample is an important question. To investigate this point, we collected data from a random sample of 2070 customers who resided in the experimental locations but did not participate in the experiment. Table A.1 suggests that the experimental sample has higher sample averages in the monthly electricity usage, number of people at home on weekdays, self-efficacy in energy conservation, and household income.

3.3 Heterogeneity in the Program’s Impact on Peak-Hour Electricity Usage

The rebate program aimed at incentivizing energy conservation in peak hours. A key variable in our social welfare function is, therefore, electricity usage in peak hours as we present in Section 4.1. Before we proceed to the analysis of social welfare gains in Section 4, this section provides a simple analysis on heterogeneity in the rebate program’s impact on peak-hour electricity usage.

Consider estimating the intention-to-treat (ITT) of the randomly-assigned groups $Z = \{T, S\}$ relative to $Z = U$ by the OLS with an estimating equation,

$$y_{it} = \beta_T T_{it} + \beta_S S_{it} + \lambda_i + \theta_t + \epsilon_{it}, \tag{10}$$

where y_{it} is the natural log of electricity usage for household i in a 30-minute interval t. We include data from the pre-experimental period and experimental period.[8] A dummy variable T_{it} equals one if household i is in group T and t is in the treatment period. A dummy variable S_{it} equals one if household i is in group S and t is in the treatment period. We include household fixed effects λ_i and time fixed effects θ_t for each 30-minute interval to control for time-specific shocks such as weather. Given that $Z = \{U, T, S\}$ is randomly assigned, β_T and β_S provides the ITT of $Z = \{T, S\}$ relative to $Z = U$. Because there is no self-selection in group T, β is also the average treatment effect (ATE) of T relative to U. We cluster standard errors at the household level.

In Table 2, we report the estimation results of equation (10). We begin by demonstrating the ITT for the entire sample in Column 1. The compulsory treatment (T) resulted in a reduction in peak-hour electricity usage by 0.097 log points (9.2%). The self-selection treatment (S) induced a reduction by 0.052 log points (5.1%). The p-value of the difference in these two ITTs is 0.088.[9]

[Table 2 about here]

Along with these overall program impacts, an important question for our analysis is whether there is substantial heterogeneity in the effects. If different household types respond to T and S differently, the optimal targeting policy presented in Section 2 could enhance the welfare gain from the policy. We investigate this question in the remaining columns of Table 2.[10] Each pair of columns splits the sample into two groups: those with a below median value of a particular variable and those with an above median value.

We find evidence of rich heterogeneity in the program's impact. In columns 2 and 3, for instance, we split customers by peak-hour electricity usage relative to pre-peak hour usage, based on data in the pre-experimental period. For households with lower values of this variable, we find that $\hat{\beta}_T = -0.108$ and $\hat{\beta}_S = -0.022$, and the p-value for the difference is 0.013. In contrast, for households with higher values of this variable, $\hat{\beta}_T = -0.079$ and $\hat{\beta}_S = -0.073$, and the p-value for the difference is 0.88. That is, in terms of the ITT for peak-hour electricity usage, T provides a larger reduction than S for a subgroup, but this is not the case for other groups. We find similar heterogeneity when we split the sample based on the number

[8] Because of randomization, the pre-experimental data is not necessary for obtaining the consistent estimator. The primary benefit of including the pre-experimental data is that the inclusion of household fixed effects can substantially increase the precision of the estimates because residential electricity usage tends to form a significant part of household-specific time-invariant variation.

[9] For the self-selection group, we can use the ITT and take-up rate to obtain the local average treatment effect for takers (LATE), which is -0.14. This LATE is larger (in absolute value) than the ATE obtained by the T group, suggesting the possibility of selection on welfare gains similar to the findings in Ito, Ida, and Takana (forthcoming).

[10] This table presents results on the covariates selected for estimating the optimal policy in Section 4.2.

of people at home, the self-efficacy in energy conservation, and household income.

This heterogeneity in the program's impact on peak-hour electricity usage implies that optimal targeting is likely to enhance the welfare gain from the policy. However, although the simple analysis in Table 2 is useful, there are two caveats in this analysis. First, these ITTs are not equivalent to the social welfare gains, and therefore, do not necessarily provide full information to rank T and S. For example, these ITTs are related to but do not directly measure consumer surplus or social surplus. Furthermore, the cost of the policy (i.e., the implementation cost per participating household) is not included. Second, the true heterogeneity can be more complex as covariates may have nonlinear and interaction effects on the program's impact. For this reason, we conduct more comprehensive analysis based on a machine-learning method (the decision tree) in Section 4.

4 Optimal Assignment Policy and Welfare Gains

In this section, we apply the framework we developed in Section 2 to our experimental data. In our framework, the planner's objective is to find the optimal policy assignment rule $G^* = (G_T^*, G_U^*, G_S^*)$ that maximizes the expected welfare gain $\mathcal{W}(G)$ in equation (2). We begin by defining $\mathcal{W}(G)$ in our empirical context in Section 4.1, describe exogenous parameters and estimation details in Section 4.2, and report the estimation results in Section 4.3.

4.1 Construction of the Social Welfare Criterion

We use p and c to denote the price and marginal cost of electricity. In peak hours, the time-invariant residential price p tends to be too low relative to c. The goal of peak-time rebate programs is to reduce welfare loss from this economic inefficiency by setting the rebate incentive equal to c.

Consider a household that takes the rebate program. We use Y_U and Y_T to denote the potential untreated and treated outcomes of electricity consumption. We assume a locally-linear demand curve for electricity usage. Then, the short-run social welfare gain from this program is be written by $\frac{1}{2}(p-c)(Y_T-Y_U)$. Further, we consider that the reduction in consumption creates an additional long-run social welfare gain as it saves the cost of power plant investments. We denote this long-run gain by $\delta(Y_T - Y_U)$, where δ is the price per kW in the capacity market. Finally, the participation to the rebate program incurs an implementation cost per customer by a.

Then, for each $j \in \{T, U, S\}$, the social welfare gain from the rebate program can be written by,

$$W_j := b \cdot (Y_j - Y_U) - a \cdot 1\{D_j = 1\}, \tag{11}$$

where $b = \frac{1}{2}(p - c) + \delta$, Y_j is the potential outcome of electricity usage for $j \in \{T, U, S\}$, and $D_j = 1$ if the consumer is treated. Equation (11) implies that $W_j = 0$ if the consumer is untreated ($j = U$).

As presented in Section 2, the planner's objective is determining the optimal policy assignment rule $G^* = (G_T^*, G_U^*, G_S^*)$ that maximizes $\mathcal{W}(G) \equiv E\left[\sum_{j\in\{T,U,S\}} W_j \cdot 1\{X \in G_j\}\right]$. By inserting equation (11) to this objective function, $-b \cdot Y_U$ becomes a constant and does not depend on the policy assignment G. Therefore, to find G^*, we can maximize $\mathcal{W}(G)$ by replacing W_j with $\tilde{W}_j := b \cdot Y_j - a \cdot 1\{D_i = 1\}$. As described in Section 2.3, we use the EWM to estimate the optimal policy by maximizing the following objective function over a class of policies $\mathcal{G}$: $\frac{1}{n}\sum_{i=1}^{n}\sum_{j\in\{T,U,S\}} \frac{W_i \cdot 1\{Z_i=j\}}{P(Z_i=j|X_i)} \cdot 1\{X_i \in G_j\}$, where i indicates each household, n is the sample size, $W_i = b \cdot Y_i - a \cdot 1\{D_i = 1\}$, Y_i is the observed electricity usage for household i, $D_i = 1$ if household i is treated, and $Z_i \in \{T, U, S\}$ is the randomly-assigned group in our experiment.

4.2 Estimation Details

Equation (11) includes four exogenous parameters: p, c, a, and δ. We use data from the Japanese electricity market during our experimental period to set the values for these parameters. p is the unit price of electricity. We set to $p = 25$ JPY/kWh, approximately the regulated price of electricity in Japan, which is independent of the time of a day.[11] c is the marginal cost of production for electricity. We specify $c = 125$ JPY/kWh, so that the difference between p and c is equal to the per kWh rebate. The wholesale price of electricity sometimes soars during peak hours such as summer afternoons or winter evenings, reflecting supply constraints. In the past, the wholesale price has occasionally exceeded 100 JPY/kWh in summer afternoons.[12]

[11] In Japan, until April 1 2016 household electricity was supplied by local power companies and retail prices were regulated. Since then, entry into the retail electricity industry has been fully liberalized, allowing all households to freely choose their price menu. However, as a transitional measure, the regulated price for households is being maintained for the time being, and is set at approximately 25 JPY/kWh regardless of the time of day.

[12] The wholesale electricity market, where the power generation sector and the retail sector trade electricity, is operated by the Japan Electric Power Exchange (JEPX). Most trading takes place in the "day-ahead market" where both sectors trade electricity on the day before the actual demand period. Trading results are disclosed, and we confirm that the price exceeded 100 JPY/kWh on July, 25, 2018. Moreover, the price has exceeded even 125 JPY/kWh. For example, the price reached 250 JPY/kWh on January, 15, 2021.

Parameter a represents the administrative cost of implementing our energy saving program. This cost comprises several items, including the installation cost of the Home Energy Management System (HEMS) required to participate. In 2016, the Japanese government estimated the cost of implementing a demand reduction program, including the installation cost of HEMS, to be 291.1 JPY per household per season (Ida and Ushifusa, 2017).[13] We use this as the value of the administrative cost.

Parameter δ represents the long-term benefits of a unit reduction in energy consumption. Here, we consider the effect of a unit reduction on the capacity market, where future supply capacity is traded between the power generation and retail sectors. In Japan, the capacity market was established in 2020, with the first auction held at that time. In that auction, the Japanese government provided a reference price 9,425 JPY/kW to bidders, which we use as the value for δ.

To estimate the optimal policy G^*, we need to solve the optimization problem with the objective function in Section 4.1. To do so, we specify the policy class to be the class of decision trees of a depth of 6. We select five variables among candidates to be used in constructing the decision trees. These are Peak - Pre-peak, Peak - Post-peak, household income, the number of residents in the same housing unit, and a measure of the households' self-efficacy in energy conservation. These variables are selected based on their ability to predict electricity consumption and the conditional average treatment effects. Specifically, we select these variables by running two off-the-shelf machine learning algorithms, lasso and random forest, with all the available covariates and assessing the importance of each variable. When using lasso to assess importance, we regress Y_i on all the available covariates with a l_1-penalization term. We order variables in terms of importance by increasing the penalization parameter step-wise and checking which variables remain selected for large penalization parameter values. When using random forest, we estimate the conditional average treatment effects using the causal forest algorithm of (Wager and Athey, 2018) with all available covariates included. We use the frequency with which a variable is used to split nodes as a measure of its importance. These selected variables are those that appeared on the lists of important variables produced by both methods.

We use the decision tree at depth 6, and exactly maximize the empirical welfare criterion by applying the exhaustive search algorithm of Zhou, Athey, and Wager (forthcoming). An important technical detail of

[13] We do not include the installation cost for a smart meter in the administrative cost. Since the Great East Japan Earthquake of March 11, 2011 and the accident at the Fukushima Daiichi Nuclear Power Plant, the Japanese government has stipulated that smart meters should be installed in all homes by the end of the decade, so this cost is "sunk" in that it will be paid regardless of whether a demand reduction program is implemented or not.

the EWM estimation is that the optimized empirical welfare value from the estimation will be an upwardly biased estimate of the true welfare attained by the estimated policy. This is known as the winner's bias (see, e.g., Andrews, Kitagawa, and McCloskey, 2019), and is caused by using the same data twice: once to learn the policy and once to infer the policy's welfare.[14] To control for the winner's bias in our point estimates and confidence intervals, we create artificial test data by fitting a causal forest Wager and Athey (2018) to run regressions of the outcome onto all the covariates, and generating data with permuted regression residuals. We include detailed explanations about our estimation procedure in Appendices A.2 and A.3.

4.3 Results of the Optimal Policy Assignment

We estimate the optimal policy assignment that maximizes social welfare based on the algorithm we described in equation (7) in Section 2. We compare five alternative policies: 1) assigning everyone to U, 2) assigning everyone to T, 3) assigning everyone to S, 4) the selection-absent targeting $G^{\dagger}$, and 5) the selection-driven targeting G^{*}.

In Table 3, we present the welfare performances of three benchmark policies without targeting (100% U, 100% T, and 100% S) followed by the suboptimal and optimal targeting policies ($G^{\dagger}$ and G^{*}). For each policy, we estimate the ITT of the welfare gain in JPY per household per season.[15]

[Table 3 about here]

We find that the 100% T policy induces a welfare gain of 120.7 per consumer, but the effect is not statistically significant. The 100% S policy results in a welfare gain by 180.6 per consumer and is marginally significant at p-value $= 0.107$. These results suggest that without targeting, we cannot reject that the policy's net welfare gain can be zero.

Recall that our policy intervention induces both cost (from the implementation cost) and benefit (from the energy conservation), and therefore, the net welfare gain from a consumer can be positive, negative, or zero. This implies that we could be able to increase the policy performance by targeting policies, $G^{\dagger}$ and G^{*}. Results in Table 3 suggest that the selection-absent targeting ($G^{\dagger}$) attains a welfare gain by 387.8 per consumer. Our algorithm identifies that 52.4% of consumers should be treated, and 47.6% of them should be untreated.

[14] The estimation and inference procedures proposed by Andrews, Kitagawa, and McCloskey (2019) cannot be directly applied to decision tree based policies because the number of candidate policies is infinite.

[15] The ITT is equivalent to the ATE when a policy does not allow for self-selection.

Furthermore, we find that the selection-driven targeting (G^*) results in a welfare gain of 553.7 per consumer. With this policy, our algorithm identifies that 31.4% of consumers should be treated, 23.9% of them should be untreated, and 44.7% of them should self-select.

In Table 4, we statistically test the null hypothesis that a policy's welfare gain is larger than another policy's welfare gain. The 100% S generates a lager welfare gain than the 100% T policy, but the difference is not statistically significant (p-value is 0.29). Both of our targeting policies ($G^\dagger$ and G^*) generate statistically larger welfare gains than non-targeting policies. Finally, we find that the selection-driven targeting (G^*) results in a 43% ($= 553.7/387.8 - 1$) larger welfare gain than the selection-absent targeting ($G^\dagger$), and the difference is statistically significant at p-value of 0.003 .

[Table 4 about here]

Table 5 presents the covariates distribution by the optimal policy assignment group $\hat{G}^* = (\hat{G}^*_U, \hat{G}^*_T, \hat{G}^*_S)$. Columns 1, 2 and 3 show the mean and standard deviation by group, and columns 4, 5 and 6 show the difference between the means and its standard errors. For example, the means of household income indicate that higher-income households are more likely to be assigned to U rather than T or S. Similarly, the means of self-efficacy in energy conservation suggest that households with lower efficacy in energy conservation are more likely to be assigned to U.

[Table 5 about here]

5 Mechanism Behind the Optimal Policy Assignment

In this section, we investigate the mechanism behind the optimal policy assignment G^*. To do so, we analyze the LATEs for takers and non-takers in Section 5.1 and the counterfactual ITTs in Section 5.2.

5.1 Using the LATE Framework to Uncover the Mechanism

As presented in Section 2.2, an advantage of our research design is that we can identify both of the LATE for *takers* ($E[W_T - W_U|D_S = 1]$) and the LATE for *non-takers* ($E[W_T - W_U|D_S = 0]$). In this section, we demonstrate that these two LATEs can be used to examine the mechanism behind the selection-driven targeting.

Recall that in our empirical context, we define takers and non-takers in the following way with the notations used in Section 2.1. If a consumer is assigned to the self-selection group (S), the consumer has a binary choice between getting treated or untreated. We use $D_S = \{0, 1\}$ to denote this potential outcome. That is, $D_S = 1$ (takers) means that the consumer would take the treatment if she is assigned to S, and $D_S = 0$ (non-takers) means that she would not take the treatment if she is assigned to S.

As shown in equation (8) in Section 2.2, we can use the conventional LATE framework by Imbens and Angrist (1994) to demonstrate that $E\left[W_T - W_U | D_S = 1\right] = \frac{E[W|Z=S] - E[W|Z=U]}{P(D=1|Z=S)}$, where $Z = \{S, U\}$ is randomly assigned in our RCT, $Z = S$ is the selection group, $Z = U$ is the untreated group, and D is the observed treatment take-up for those who were assigned to $Z = S$. The numerator of the right hand side of the equation is the difference in the ITTs between groups S and U, and the denominator is the take-up rate in groups S. Therefore, the sample analogue of this equation can be estimated from our experimental data.[16] A unique feature of our research design is that we have a randomly-assigned compulsory treatment group ($Z = T$) along with groups $Z = \{S, U\}$. As presented in equation (8) in Section 2.2, we can use two groups $Z = \{S, T\}$ to estimate the LATE for non-takers by $E\left[W_T - W_U | D_S = 0\right] = \frac{E[W|Z=T] - E[W|Z=S]}{P(D=0|Z=S)}$.

Note that the LATEs for takers and non-takers can be estimated conditional on X because the randomization of $Z = (U, T, S)$ holds given X. This implies that we can estimate these LATEs by customer types based on X. Now consider the optimal assignment rule computed by the selection-driven targeting policy, $\hat{G}^* = (\hat{G}^*_U, \hat{G}^*_T, \hat{G}^*_S)$. This policy divides customers into three groups based on their observables: those who should be untreated ($X \in \hat{G}^*_U$), those who should be treated ($X \in \hat{G}^*_T$), and those who should self-select ($X \in \hat{G}^*_S$).

In Figure 3, we estimate equations (8) and (9) for these three groups, $\hat{G}^*_U$, $\hat{G}^*_T$, and $\hat{G}^*_S$. For those who are assigned to the selection group ($\hat{G}^*_S$), the LATE for takers is 2062 and the LATE for non-takers is -823. This implies that self-selection is a useful tool for this group to let customers sort into the treatment choice that is in line with the planner's objective.

[Figure 3 about here]

By contrast, if we allow self-selection for customers in $\hat{G}^*_U$, it would decrease welfare because the LATE for takers in this group is -2335. Similarly, if we allow self-selection in $\hat{G}^*_T$, it would lower welfare because

[16] Equation (8) implies that the LATE for takers is equivalent to the LATE for compliers when we consider two groups with a binary instrument $Z = \{U, S\}$. This also implies that we can use the conventional IV estimation to estimate equation (8) given the regular assumptions required for the LATE.

self-selection would make non-takers out of the treatment even though their LATE is positive and large at 1019. Therefore, the LATE for takers and non-takers presented in Figure 3 highlights how our algorithm chooses who should get treated, untreated, and choose to get treated by themselves.

5.2 Counterfactual Intention-to-Treat Analysis

Another way to investigate the mechanism behind our algorithm is to estimate the counterfactual ITTs separately for each of the three groups $\hat{G}^*_U$, $\hat{G}^*_T$, and $\hat{G}^*_S$. For instance, consider customers in $\hat{G}^*_U$, who should be untreated according to the optimal assignment rule. Note that within this group, our experiment provided a random variation of $Z = (U, T, S)$. Therefore, we can use this variation to estimate three ITTs as if they were assigned to U, T, and S. Similarly, we can estimate these three ITTs for customers in $\hat{G}^*_T$ and those in $\hat{G}^*_S$.

Table 6 presents these counterfactual ITTs for each of $\hat{G}^*_U$, $\hat{G}^*_T$, and $\hat{G}^*_S$. Results for customers in $\hat{G}^*_U$ imply that their ITTs would be negative (-905.4 and -923.0) if they were assigned to T and S, respectively. That is, the welfare gains is maximized if they are assigned to U. Similarly, results for customers in $\hat{G}^*_T$ and $\hat{G}^*_S$ suggest that their welfare gains are maximized when they are assigned to T and S, respectively.

[Table 6 about here]

Hence, the policy assignment computed by $\hat{G}^*$ coincides with the arm that attains the highest welfare gain for every subgroup of $\{X \in \hat{G}^*_j\}$, $j \in \{U, T, S\}$. In other words, the estimated policy $\hat{G}^*$ captures households' heterogeneous responses by using both observed and unobserved characteristics to maximize social welfare. This result is true by construction because our algorithm finds the optimal assignment by maximizing the ITT of the welfare gain. Yet, Table 6 is useful to visualize the mechanism by observing that we indeed see the optimal ITTs in the diagonal line.

6 Welfare Maximization with Redistribution

In Section 4, we presented the results based on the utilitarian welfare function. However, our method is not necessarily restricted to a conventional utilitarian framework. Rather, one can apply our method to any welfare function most appropriate for a policy goal. In this section, we shed light on this point by considering a policy goal that balances the equity-efficiency trade-off.

Table 7 presents the redistribution implications of the optimal utilitarian policy presented in Section 4. The efficiency gain (553.7 in the table) is the ITT of the welfare gain of the optimal policy (the selection-driven targeting) based on the utilitarian welfare function. As presented in Table 4, this targeting policy ($\hat{G}^*$) maximizes the utilitarian welfare gain compared to other policy options.

[Table 7 about here]

However, Table 7 suggests that this policy may create a concern for equity. We compare the average rebate that would be distributed to consumers across the household income distribution. We find that the optimal targeting policy would distribute more rebates to higher income households. That is, although this targeting maximizes the efficiency gain from the policy, it may not be appealing to policymakers who weigh on redistribution implications.

To address this equity concern, we consider a social welfare function that balances the equity-efficiency trade-off by using a framework developed by Saez (2002) and used by Allcott, Lockwood, and Taubinsky (2019) and Lockwood (2020). Consider Pareto weights for a household: $w = h^{-\nu}$ where h is household income and ν is a scalar parameter that represents a policymaker's preference for redistribution. With this specification, a higher ν implies a stronger preference for redistribution, $\nu = \infty$ corresponds to the Rawlsian criterion, and $\nu = 0$ corresponds to utilitarianism. When $\nu = 1$, it approximately corresponds to the weight that would arise under logarithmic utility from income. We modify our utilitarian social welfare function by using this weight for each consumer's surplus from the policy intervention. In Appendix A.4, we show that the social welfare function with this Pareto weight can be written by:

$$W_j = \tilde{b} \cdot (Y_j - Y_U) - (1 - w)R_j - a \cdot 1\{D_j = 1\}, \tag{12}$$

for $j \in \{T, U, S\}$, where $\tilde{b} = \frac{2-w}{2}(p - c) + \delta$, $w = h^{-\nu}/H$ is the normalized weight, and H is the sum of $h^{-\nu}$ over all households. R_j is the potential amount of rebate for arm j and it is defined by $R_j = (c - p) \cdot \max\{Y_{\text{base}} - Y_j, 0\} \cdot 1\{D_j = 1\}$, where Y_{base} denotes the baseline consumption of rebate payment for the household. When $\nu = 0$, w is equal to 1 for all households, and hence, the equation (12) becomes the utilitarian welfare function in (11).[17]

[17]The potential amount of rebate R_j does not appear in the utilitarian weight because it is just a lump-sum transfer between consumers and producers. However, once we allow differential weight for each consumer's surplus, R_j is not cancelled out in the welfare function.

In Table 7, we present the results with $\nu = 1$ and $\nu = 2$ below the result for the utilitarian welfare function, which is equivalent to the case with $\nu = 0$. We find that the welfare function with $\nu = 2$ is able to roughly equalize the average rebate distributed to households across the income distribution. However, as we expect, the efficiency gain (i.e., the welfare gain evaluated based on the utilitarian welfare function) is compromised as we increase the weight on the preference for redistribution.

These results indicate that policymakers can choose the welfare function that most appropriately balances the equity-efficiency trade-off and apply our method to such welfare functions. We also present in Table A.2 that the selection-driven targeting maximizes the welfare function with all values of ν that are considered in our analysis.

7 Conclusion

We develop an optimal policy assignment rule that systematically integrates two distinctive approaches commonly used in the literature—targeting by "observables" and targeting through "self-selection." Our method identifies who should be treated, untreated, and self-select into a treatment to maximize a policy's social welfare gain. To generate data required to estimate optimal policy assignment, we designed a randomized controlled trial for a residential energy rebate program. We show that targeting that leverages information on both observables and self-selection outperforms conventional targeting for a standard utilitarian welfare function as well as welfare functions that balance the equity-efficiency trade-off. Finally, we highlight that the LATE framework (Imbens and Angrist, 1994) can be used to uncover the mechanism behind our approach. We introduce new estimators, the LATEs for *takers* and *non-takers*, to demonstrate how our method identifies whose self-selection is useful and harmful for the planner to maximize social welfare.

References

ALATAS, V., R. PURNAMASARI, M. WAI-POI, A. BANERJEE, B. A. OLKEN, AND R. HANNA (2016): "Self-targeting: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 124, 371–427.

ALLCOTT, H., B. B. LOCKWOOD, AND D. TAUBINSKY (2019): "Regressive sin taxes, with an application to the optimal soda tax," *The Quarterly Journal of Economics*, 134, 1557–1626.

ANDREWS, I. S., T. KITAGAWA, AND A. MCCLOSKEY (2019): "Inference on Winners," *NBER working paper*.

ATHEY, S. AND S. WAGER (2021): "Efficient policy learning with observational data," *Econometrica*, 89, 133–161.

BERTSIMAS, D. AND J. DUNN (2017): "Optimal classification trees," *Machine Learning*, 106, 1039–1082.

BHATTACHARYA, D. (2013): "Evaluating Treatment Protocols by Combining Experimental and Observational Data," *Journal of Econometrics*, 173, 160–174.

BREIMAN, L., J. H. FRIEDMAN, R. A. OLSHEN, AND C. J. STONE (2017): *Classification and regression trees*, Routledge.

BURLIG, F., C. KNITTEL, D. RAPSON, M. REGUANT, AND C. WOLFRAM (2020): "Machine learning from schools about energy efficiency," *Journal of the Association of Environmental and Resource Economists*, 7, 1181–1217.

CAGALA, T., U. GLOGOWSKY, J. RINCKE, AND A. STRITTMATTER (2021): "Optimal Targeting in Fundraising: A Causal Machine-Learning Approach," *arXiv preprint arXiv:2103.10251*.

CHAMBERLAIN, G. (2011): "Bayesian aspects of treatment choice," in *The Oxford Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. van Dijk, Oxford University Press, 11–39.

CHRISTENSEN, P., P. FRANCISCO, E. MYERS, H. SHAO, AND M. SOUZA (2021): "Energy Efficiency Can Deliver for Climate Policy: Evidence from Machine Learning-Based Targeting," *Working Paper*.

DEHEJIA (2005): "Program evaluation as a decision problem," *Journal of Econometrics*, 125, 141–173.

DESHPANDE, M. AND Y. LI (2019): "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, 11, 213–248.

DYNARSKI, S., C. LIBASSI, K. MICHELMORE, AND S. OWEN (2018): "Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students," *NBER working paper*.

FINKELSTEIN, A. AND M. J. NOTOWIDIGDO (2019): "Take-up and Targeting: Experimental Evidence from SNAP," *Quarterly Journal of Economics*, 134, 1505–1556.

FOWLIE, M., C. WOLFRAM, P. BAYLIS, C. A. SPURLOCK, A. TODD-BLICK, AND P. CAPPERS (2021): "Default Effects And Follow-On Behaviour: Evidence From An Electricity Pricing Program," *The Review of Economic Studies*, 88, 2886–2934.

FRIEDBERG, R., J. TIBSHIRANI, S. ATHEY, AND S. WAGER (2021): "Local Linear Forests," *Journal of Computational and Graphical Statistics*, 30, 503–517.

GERARDEN, T. D. AND M. YANG (2021): "Using Targeting to Optimize Program Design: Evidence from an Energy Conservation Experiment," *Working Paper*.

HECKMAN, J. J. (2010): "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, 48, 356–398.

HECKMAN, J. J. AND E. VYTLACIL (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73, 669–738.

HIRANO, K. AND J. R. PORTER (2009): "Asymptotics for statistical treatment rules," *Econometrica*, 77, 1683–1701.

IDA, T., T. ISHIHARA, D. KIDO, AND S. SASAKI (2020): "A field experiment using rebates and machine learnings to promote energy-saving behavior," *AEA RCT Registry, No.6139.*

IDA, T. AND Y. USHIFUSA (2017): "Cost-Benefit Analysis of Price-based Residential Demand Response," *Proceedings of the Japan Joint Automatic Control Conference*, 60, 304–307.

IMBENS, G. W. AND J. D. ANGRIST (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

ITO, K. (2015): "Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program," *American Economic Journal: Economic Policy*, 7, 209–37.

ITO, K., T. IDA, AND M. TAKANA (forthcoming): "Selection on Welfare Gains: Experimental Evidence from Electricity Plan Choice," *American Economic Review*.

JANEVIC, M. R., N. K. JANZ, J. A. DODGE, X. LIN, W. PAN, B. R. SINCO, AND N. M. CLARK (2003): "The role of choice in health education intervention trials: a review and case study," *Social Science and Medicine*, 56, 1581–1594.

JOHNSON, M. S., D. I. LEVINE, AND M. W. TOFFEL (2020): "Improving regulatory effectiveness through better targeting: Evidence from OSHA," *Harvard Business School Technology & Operations Mgt. Unit Working Paper*.

KITAGAWA, T. AND A. TETENOV (2018): "Who should be treated? Empirical welfare maximization methods for treatment choice," *Econometrica*, 86, 591–616.

KNITTEL, C. R. AND S. STOLPER (2021): "Machine Learning about Treatment Effect Heterogeneity: The Case of Household Energy Use," *AEA Papers and Proceedings*, 111, 440–44.

KNOX, D., T. YAMAMOTO, M. A. BAUM, AND A. J. BERINSKY (2019): "Design, Identification, and Sensitivity Analysis for Patient Preference Trials," *Journal of the American Statistical Association*, 114, 1532–1546.

LIEBER, E. M. AND L. M. LOCKWOOD (2019): "Targeting with in-kind transfers: Evidence from Medicaid home care," *American Economic Review*, 109, 1461–85.

LOCKWOOD, B. B. (2020): "Optimal income taxation with present bias," *American Economic Journal: Economic Policy*, 12, 298–327.

LONG, Q., R. LITTLE, AND X. LIN (2008): "Causal Inference in Hybrid Intervention Trials Involving Treatment Choice," *Journal of the American Statistical Association*, 103, 474–484.

MANSKI, C. (2013): *Public Policy in an Uncertain World*, Cambridge, MA: Harvard University Press.

MANSKI, C. F. (2004): "Statistical treatment rules for heterogeneous populations," *Econometrica*, 72, 1221–1246.

MURAKAMI, K., H. SHIMADA, Y. USHIFUSA, AND T. IDA (forthcoming): "Heterogeneous Treatment Effects of Nudge and Rebate: Causal Machine Learning in a Field Experiment on Electricity Conservation," *International Economic Review*.

REISS, P. C. AND M. W. WHITE (2008): "What changes energy consumption? Prices and public pressures." *RAND Journal of Economics*, 39, 636–663.

RÜCKER, G. (1989): "A Two-Stage Trial Design for Testing Treatment, Self-Selection, and Treatment Preference Effects," *Statistics in Medicine*, 8, 477–485.

SAEZ, E. (2002): "Optimal income transfer programs: intensive versus extensive labor supply responses," *The Quarterly Journal of Economics*, 117, 1039–1073.

STOYE, J. (2009): "Minimax regret treatment choice with finite samples," *Journal of Econometrics*, 151, 70–81.

——— (2012): "Minimax regret treatment choice with covariates or with limited validity of experiments," *Journal of Econometrics*, 166, 138–156.

TETENOV, A. (2012): "Statistical treatment choice based on asymmetric minimax regret criteria," *Journal of Econometrics*, 166, 157–165.

UNRATH, M. (2021): "Targeting, Screening, and Retention: Evidence from California's Food Stamps Program," *Working paper*.

WAGER, S. AND S. ATHEY (2018): "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113, 1228–1242.

WALDINGER, D. (2021): "Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation," *American Economic Review*, 111, 2660–96.

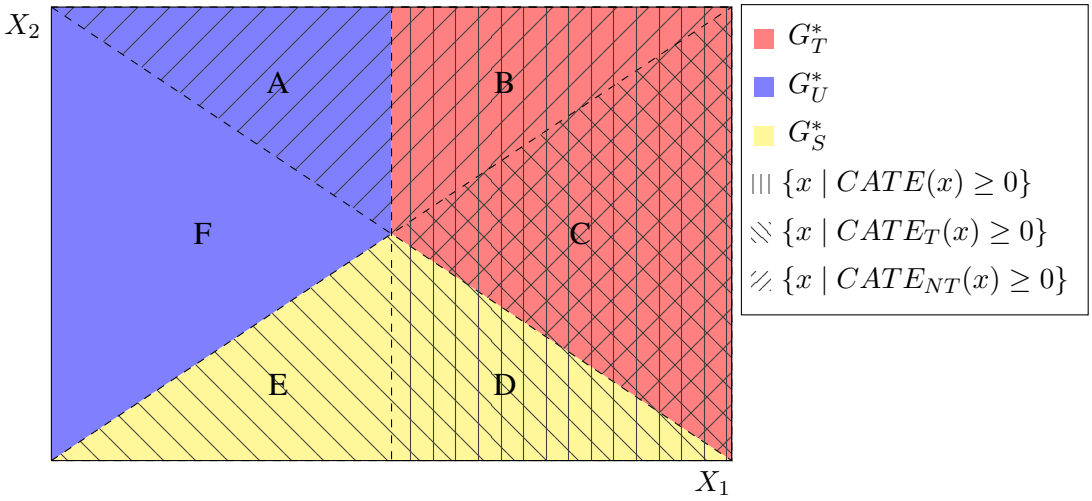
WOLAK, F. A. (2006): "Residential Customer Response to Real-Time Pricing: the Anaheim Critical-Peak Pricing Experiment," *Working Paper*.

——— (2011): "Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment," *The American Economic Review*, 101, 83–87.

ZHOU, Z., S. ATHEY, AND S. WAGER (forthcoming): "Offline multi-action policy learning: Generalization and optimization," *Operations Research*.

Figures

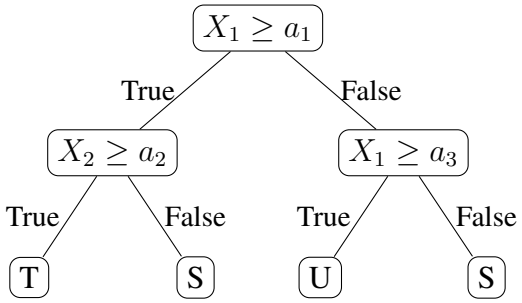
Figure 1: Example of Optimal Policy Assignment G^*



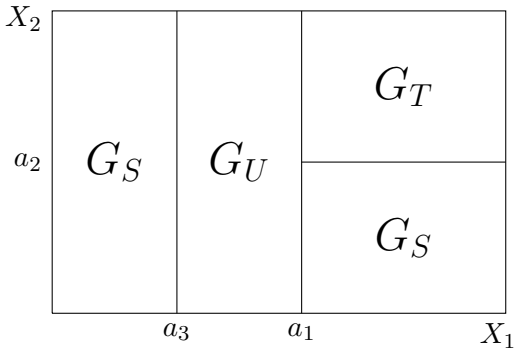
Notes: This figure illustrates how an optimal policy rule G^* described in equation (6) partitions customers, using the two-dimensional characteristic space $\mathcal{X}$ as an example. Let $CATE(x) := E[W_T - W_U|x]$, $CATE_T(x) := E[W_T - W_U|D_S = T, x]$ (i.e., the LATE for takers), and $CATE_{NT}(x) := E[W_T - W_U|D_S = U, x]$ (i.e., the LATE for non-takers). Among the six subspaces from A to F, $CATE(x)$ is non-negative only in B, C, and D; $CATE_T(x)$ is non-negative only in C, D, and E; $CATE_{NT}(x)$ is non-negative only in A, B, and C. Therefore, according to the optimal policy characterization (6), B and C are assigned to G_T^*; A and F are assigned to G_U^*; and D and E are assigned to G_S^*.

Figure 2: Illustration of the Decision Tree

(a) Decision tree with depth-2

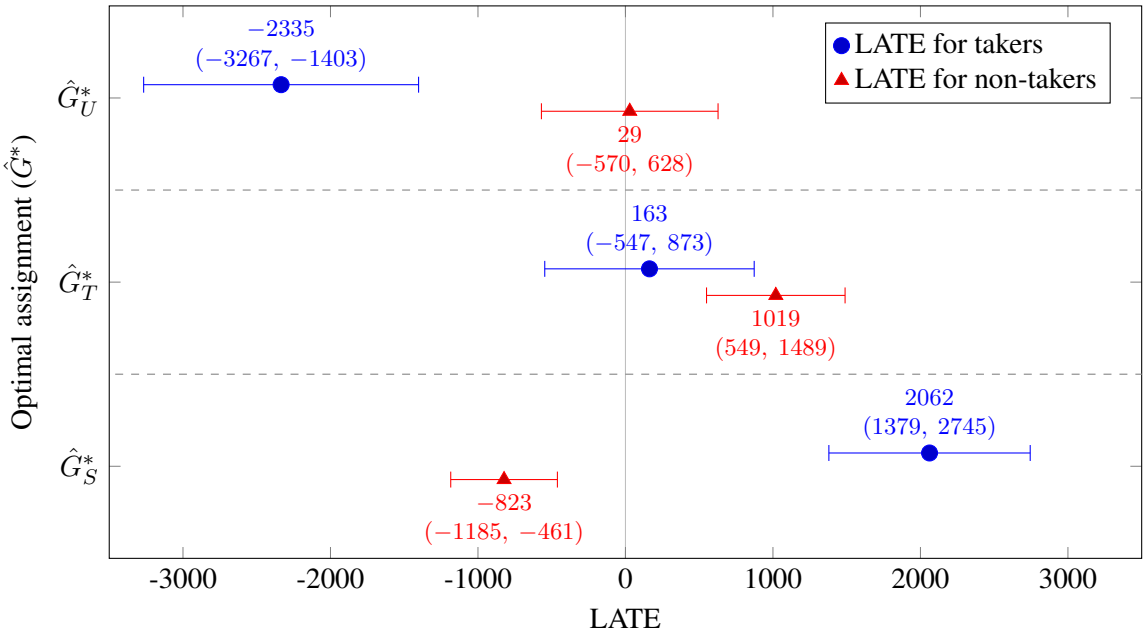


(b) Partition of $\mathcal{X}$ by the decision tree



Notes: This figure illustrates the decision tree described in Section 2.3 by using an example case with two variables and depth-2. Note that in our empirical analysis, we use more than two variables and depth deeper than 2, so this figure is only for illustration. Policy assignment T, U, S corresponds to the treated, untreated, and selection groups.

Figure 3: Mechanism Behind the Algorithm: The LATEs for Takers and Non-Takers



Notes: This figure shows the estimation results in Section 5.1. For each of the three groups in the optimal assignment ($x \in \hat{G}^*_U, x \in \hat{G}^*_T, x \in \hat{G}^*_S$), we estimate the LATE for *takers* ($E[W_T - W_U|D_S = 1]$) and the LATE for *non-takers* ($E[W_T - W_U|D_S = 0]$) to investigate the mechanism behind the optimal assignment. In the figure, we show the point estimates with the 95% confidence intervals. For example, for those who are assigned to the selection group ($\hat{G}^*_S$), the LATE for takers is 2062 and the LATE for non-takers is -823. This implies that self-selection is a useful tool for this group to let customers sort into the treatment choice that is in line with the planner's objective. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

Tables

Table 1: Summary Statistics and Balance Check

	Sample mean by group [standard deviation]			Difference in sample means (standard error)		
	Untreated ($Z = U$)	Treated ($Z = T$)	Selection ($Z = S$)	U vs. T	U vs. S	T vs. S
Peak hour usage (Wh)	192	190	189	2.57	2.87	0.29
	[141]	[138]	[134]	(5.03)	(5.91)	(5.93)
Pre-peak hour usage (Wh)	179	176	180	3.79	−1.11	−4.89
	[137]	[135]	[142]	(4.92)	(6.07)	(6.11)
Post-peak hour usage (Wh)	299	297	293	1.94	6.02	4.08
	[175]	[171]	[174]	(6.26)	(7.54)	(7.56)
Number of people at home	2.48	2.44	2.47	0.04	0.01	−0.03
	[1.24]	[1.24]	[1.27]	(0.04)	(0.05)	(0.06)
Self-efficacy in energy	3.45	3.46	3.49	−0.01	−0.04	−0.02
conservation (1-5 scale)	[0.85]	[0.85]	[0.83]	(0.03)	(0.04)	(0.04)
Household income	645	613	637	31.69	8.45	−23.23
(JPY 10,000)	[399]	[362]	[391]	(13.75)	(17.06)	(16.67)
All electric	0.32	0.31	0.30	0.01	0.02	0.00
	[0.47]	[0.46]	[0.46]	(0.02)	(0.02)	(0.02)
Number of air conditioners	3.14	3.11	3.08	0.03	0.05	0.02
	[1.69]	[1.71]	[1.67]	(0.06)	(0.07)	(0.07)
Number of fans	2.80	2.73	2.77	0.07	0.04	−0.04
	[1.63]	[1.63]	[1.56]	(0.06)	(0.07)	(0.07)
Number of household members	2.76	2.73	2.75	0.04	0.01	−0.03
	[1.27]	[1.27]	[1.28]	(0.05)	(0.06)	(0.06)
Total living area (m^2)	107.29	105.51	103.42	1.78	3.87	2.09
	[48.57]	[49.61]	[46.14]	(1.78)	(2.03)	(2.07)

Notes: Columns 1-3 present the sample mean and standard deviations in blackets for the pre-experiment consumption data and demographic variables by randomly-assigned group: untreated ($Z = U$), treated ($Z = T$), and selection ($Z = S$). Columns 4-6 show the difference in the sample means with the standard error of the difference in parentheses. The number of households are 1,577 (U), 1,486 (T), and 807 (S). "Number of people at home" is the average number of people at home from 17:00 to 21:00 on weekdays. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

Table 2: Intention-to-Treat Estimates

		Peak hour usage − Pre-peak hour usage (in pre-experiment)		Peak hour usage − Post-peak hour usage (in pre-experiment)	
	All	Low	High	Low	High
Treated group ($Z = T$)	-0.097	-0.108	-0.079	-0.089	-0.094
	(0.021)	(0.028)	(0.031)	(0.030)	(0.028)
Selection group ($Z = S$)	-0.052	-0.022	-0.073	-0.070	-0.023
	(0.027)	(0.034)	(0.041)	(0.037)	(0.037)
Number of customers	3,870	1,935	1,935	1,937	1,933
Number of observations	1,176,480	588,240	588,240	589,152	587,328
p-value (T = S)	0.088	0.013	0.880	0.595	0.047
Take-up rate in group S	37.2%	36.9%	37.4%	39.9%	34.7%

	Number of people at home		Self-efficacy		Household income	
	Low	High	Low	High	Low	High
Treated group ($Z = T$)	-0.096	-0.098	-0.134	-0.057	-0.071	-0.125
	(0.027)	(0.034)	(0.028)	(0.031)	(0.028)	(0.031)
Selection group ($Z = S$)	-0.022	-0.094	-0.036	-0.072	-0.036	-0.060
	(0.034)	(0.042)	(0.035)	(0.040)	(0.038)	(0.037)
Number of customers	2,245	1,625	1,967	1,903	2,036	1,834
Number of observations	682,480	494,000	597,968	578,512	618,944	557,536
p-value (T = S)	0.020	0.934	0.004	0.715	0.336	0.094
Take-up rate in group S	37.6%	36.6%	33.8%	40.6%	34.8%	39.7%

Notes: This table shows the estimation results for equation (10) using the full-sample (the first column of the upper panel) or sub-samples (the remaining columns). The dependent variable is the log of household-level electricity consumption over a 30-minute interval. We include household fixed effects and time fixed effects for each 30-minute interval. The standard errors are clustered at the household level to adjust for serial correlation. To investigate the heterogeneity of the treatment effects, we focused on the five variables selected for estimating the optimal policy in Section 4.2 and divided the sample into five sets of two sub-groups. For the five different variables, the first sub-group includes households who are below the median of this variable and the second includes those who are above the median. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

Table 3: Welfare Gains from Each Policy

Policy	Welfare Gain	Share of customers in each arm		
		G_U	G_T	G_S
100% untreated	0 (—)	100%	0.0%	0.0%
100% treated	120.7 (98.8)	0.0%	100%	0.0%
100% self-selection	180.6 (112.1)	0.0%	0.0%	100%
Selection-absent targeting ($\hat{G}^{\dagger}$)	387.8 (55.7)	47.6%	52.4%	0.0%
Selection-driven targeting ($\hat{G}^{*}$)	553.7 (68.0)	23.9%	31.4%	44.7%

Notes: This table summarizes characteristics of three benchmark policies (100% untreated, 100% treated, and 100% self-selection), selection-absent targeting ($\hat{G}^{\dagger}$), and selection-driven targeting ($\hat{G}^{*}$). The column titled "Welfare Gain" shows the estimated ITT of welfare gain in JPY per household per season, with its standard error in parentheses. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

Table 4: Comparisons of Alternative Policies

	Difference in Welfare Gains	p-value
100% self-selection vs. 100% treated	59.9 (110.0)	0.293
Selection-absent targeting ($\hat{G}^{\dagger}$) vs. 100% treated	267.2 (99.7)	0.004
Selection-absent targeting ($\hat{G}^{\dagger}$) vs. 100% self-selection	207.3 (116.9)	0.038
Selection-driven targeting ($\hat{G}^{*}$) vs. 100% treated	433.0 (106.8)	0.000
Selection-driven targeting ($\hat{G}^{*}$) vs. 100% self-selection	373.1 (113.3)	0.000
Selection-driven targeting ($\hat{G}^{*}$) vs. Selection-absent targeting ($\hat{G}^{\dagger}$)	165.8 (61.1)	0.003

Notes: This table compares welfare gains from each policy. For each row, the column "Difference in Welfare Gains" shows the estimated welfare gain of the policy on the left-hand side (W_L) relative to the policy on the right-hand side (W_R) in JPY per household per season, with its standard error in parenthesis. The column "p-value" gives the p-value for the null hypothesis: $H_0 : W_L \geq W_R$.

Table 5: Covariate Distribution by Optimally Assigned Group $\hat{G}^*$

	Sample mean by group [standard deviation]			Difference in sample means (standard error)		
	$\hat{G}^*_U$	$\hat{G}^*_T$	$\hat{G}^*_S$	$\hat{G}^*_U$ vs. $\hat{G}^*_T$	$\hat{G}^*_U$ vs. $\hat{G}^*_S$	$\hat{G}^*_T$ vs. $\hat{G}^*_S$
Peak hour usage (Wh)	203	180	191	23.03	11.98	−11.05
	[146]	[136]	[135]	(6.18)	(5.79)	(5.08)
Pre-peak hour usage (Wh)	198	167	175	30.56	23.08	−7.48
	[150]	[133]	[132]	(6.23)	(5.86)	(4.97)
Post-peak hour usage (Wh)	329	255	310	73.22	18.82	−54.40
	[176]	[176]	[164]	(7.67)	(7.00)	(6.41)
Number of people at home	2.87	2.27	2.38	0.60	0.48	−0.11
	[1.34]	[1.32]	[1.08]	(0.06)	(0.05)	(0.05)
Self-efficacy in energy conservation (1-5 scale)	3.30	3.49	3.53	−0.19	−0.23	−0.04
	[1.02]	[0.82]	[0.75]	(0.04)	(0.04)	(0.03)
Household income (JPY 10,000)	787	597	572	190.12	215.11	25.00
	[433]	[397]	[318]	(18.23)	(16.15)	(13.73)
All electric	0.36	0.25	0.33	0.11	0.03	−0.08
	[0.48]	[0.43]	[0.47]	(0.02)	(0.02)	(0.02)
Number of air conditioners	3.41	2.82	3.16	0.58	0.24	−0.34
	[1.72]	[1.66]	[1.67]	(0.07)	(0.07)	(0.06)
Number of fans	2.99	2.58	2.78	0.41	0.20	−0.21
	[1.75]	[1.57]	[1.55]	(0.07)	(0.07)	(0.06)
Number of household members	3.17	2.54	2.67	0.63	0.50	−0.13
	[1.31]	[1.36]	[1.14]	(0.06)	(0.05)	(0.05)
Total living area (m^2)	115.41	97.16	106.73	18.25	8.68	−9.57
	[47.77]	[49.13]	[47.37]	(2.11)	(1.94)	(1.81)

Notes: This table shows the covariate distribution by group based on the optimal policy assignment $\hat{G}^*$. The last column shows the difference in the sample means and its standard errors in parentheses. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

Table 6: Mechanism Behind the Algorithm: Counterfactual Analysis of the ITT

	Consumer types based on the optimal assignment rule $\hat{G}^*$		
	$\hat{G}^*_U$	$\hat{G}^*_T$	$\hat{G}^*_S$
Counterfactual ITT (if assigned to U)	0	0	0
	(——)	(——)	(——)
Counterfactual ITT (if assigned to T)	-905.4	662.5	257.8
	(157.8)	(131.4)	(117.5)
Counterfactual ITT (if assigned to S)	-923.0	67.9	772.5
	(166.7)	(150.9)	(119.7)

Notes: This figure shows the estimation results in Section 5.2. For each of the three groups in the optimal assignment ($x \in \hat{G}^*_U, x \in \hat{G}^*_T, x \in \hat{G}^*_S$), we estimate three counterfactual ITTs: ITT if assigned to U, ITT if assigned to T, and ITT if assigned to S, with the standard errors in parentheses. The table indicates that the policy assignment computed by $\hat{G}^*$ coincides with the arm that attains the highest welfare gain for every subgroup of $\{x \in \hat{G}^*_j\}$, $j \in \{U, T, S\}$. The monetary unit is given as 1 ¢= 1 JPY in the summer of 2020.

Table 7: Incorporating Equity-Efficiency Trade-off

	Efficiency gain	Average rebate by the quartiles of household income			
		[0%,25%]	(25%,50%]	(50%,75%]	(75%,100%]
Utilitarian ($\nu = 0$)	553.7	72.8	93.7	144.1	148.9
	(68.0)	(10.3)	(12.9)	(19.1)	(18.9)
With a redistribution goal ($\nu = 1$)	431.2	77.0	132.3	140.2	116.5
	(69.2)	(13.0)	(17.4)	(18.1)	(17.6)
With a redistribution goal ($\nu = 2$)	366.1	105.2	115.7	109.9	119.2
	(69.3)	(14.8)	(16.5)	(16.2)	(20.8)

Notes: The first column "Efficiency gain" shows the welfare gain from the policy measured by the utilitarian welfare function. Other columns present the average rebate amount in each of the quartile of the income distribution. The utilitarian policy maximizes the efficiency gain but its rebate distributions are regressive. In Section 6, we consider a welfare function with a redistribution goal with a Pareto parameter ν. The policies with $\nu = 1$ and 2 reduce regressivity at the cost of sacrificing the efficiency gain. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

A Online Appendix (Not for Publication)

A.1 The external validity of the experimental sample

We randomly sampled 2070 customers from the target population who did not participate in this experiment, and conducted a similar survey to the one for the experimental sample. The purpose of this was to investigate the external validity of our experimental sample by comparing the mean for each variable between the control group from our experimental sample and this random sample. Columns 1 and 2 of Table A.1 present summary statistics for the untreated group and the random sample. Column 3 presents differences in means, with the standard errors of these differences in parentheses. We observe larger means for four variables in the untreated group than in the random sample, and the differences are statistically significant. Our experimental sample has larger pre-experiment electricity usage per month, a larger number of people at home on weekdays, higher self-efficacy in energy conservation, and higher household income. This implies that our sample includes a larger number of customers who are willing and able to reduce their electricity consumption, which should be taken into consideration when discussing the generalizability of this study.

[Table A.1 about here]

A.2 Optimization of tree

We basically perform tree optimization using the tree search algorithm proposed by Zhou, Athey, and Wager (forthcoming), which is implemented in `policy_tree` function from R package `policytree`. Given a depth of tree, their algorithm compares all possible trees of that depth in terms of an objective function and chooses the optimal one. Therefore, in principle, the exhaustive nature of the algorithm allows us to find the globally optimal tree of that depth. Another approach to find a global optimum is to formulate the tree optimization problem as a Mixed Integer Programming (MIP) and solve the problem using commercial solver. However, Zhou, Athey, and Wager (forthcoming) empirically shows that their approach is faster than MIP approach. Furthermore, their algorithm is essentially different from the greedy optimization procedure employed in tree-based classification methods including CART (Breiman, Friedman, Olshen, and Stone, 2017) in terms of quality of solution. The latter algorithm takes top-down approach: starting from the root node, it determines each split without any consideration of its possible impact of future splits in the

tree. In classification problems, it is known that the MIP approach outperforms CART in terms of 0-1 loss (Bertsimas and Dunn, 2017).

However, in our application, it is not practical to use their algorithm directly to obtain a globally optimal tree of depth greater than 3. According to their analysis of computational complexity, the running time of their algorithm scales on the order of $O((2np)^{L-1}d)$, where n is the sample size, p is the number of covariates, L (≥ 3) is the depth of tree, and d is the number of arms. That is, when other factors are fixed, computation time increases exponentially in tree depth. In our computing environment (AMD Ryzen 9 3950X 3.5GHz), a solution was obtained in about an hour for depth 3, while it was not obtained in a reasonable time for depths greater than 3.

Due to this difficulty in finding a globally optimal tree of depth 6, we adapt an approximate algorithm as follows. Specifically, we first optimize a root tree of depth 3 using the `policy_tree` function for the entire sample. This operation divides the sample into 8 subsamples. Then, for each subsample, another tree of depth 3 is optimized using `policy_tree` function, and the resulting tree is grafted on the root tree. As a result, this procedure generates a tree of depth 6. It does not generally result in a global optimal tree of depth 6. However, it is expected that our approach outperforms the completely greedy approach since our algorithm partly considers the effect of the current split on future splits.

A.3 Artificial Test Data

We describe our method for generating artificial test data in detail. First, to motivate our construction of artificial test data, we briefly discuss why bias in point estimates of welfare gain is introduced when one uses the same data to learn an optimal policy and to estimate its welfare gain. Simply speaking, this occurs due to noise in the observed electricity consumption Y and the observed take-up status D of treatment, which are random components in W. Specifically, the observed electricity consumption Y can be decomposed as the sum of an essential term and noise as follows:

$$\begin{aligned} Y = &\underbrace{E[Y_U|X] \cdot 1\{Z = U\} + E[Y_T|X] \cdot 1\{Z = T\} + E[Y_S|X] \cdot 1\{Z = S\}}_{\text{essential term}} \\ &+ \underbrace{\epsilon_U \cdot 1\{Z = U\} + \epsilon_T \cdot 1\{Z = T\} + \epsilon_S \cdot 1\{Z = S\}}_{\text{noise term}}, \end{aligned}$$

where $\epsilon_j := Y_j - E[Y_j|X]$ for each $j \in \{U, T, S\}$. The observed choice D can be similarly decomposed. While only the essential term is necessary for learning an optimal policy, a learning algorithm inevitably responds to the noise term and overfits the training sample at hand. Therefore, when one evaluates the welfare performance of the estimated policy on the same training sample, the welfare estimate is biased upward because the policy fits the noise term as well. This implies that, if we replace the noise term in the training sample with another independent noise sample, we can eliminate the bias from the estimate of welfare performance.

Motivated by this observation, we generate test data $\{Y_i^{\text{test}}, Z_i, D_i^{\text{test}}, X_i\}_{i=1}^n$ by generating another sample of noise. Here, Y_i^{test} denotes test electricity consumption and D_i^{test} denotes test take-up status of treatment. For Y_i^{test} we use the following procedure to generate artificial data: For $j \in \{T, U, S\}$ and samples $i \in I_j := \{i : Z_i = j\}$,

1. Estimate $E[Y_{j,i}|X_i]$ and calculate residuals $\hat{\epsilon}_i = Y_i - \hat{E}[Y_{j,i}|X_i]$.
2. Estimate $E[\epsilon_{j,i}^2|X_i]$ by regressing $\hat{\epsilon}_i^2$ on X_i and calculate $\hat{\sigma}_i = \sqrt{\hat{E}[\epsilon_{j,i}^2|X_i]}$.
3. Sample $\{\tilde{\epsilon}_i\}_{i \in I_j}$ from the standardized residuals $\{\hat{\epsilon}_i/\hat{\sigma}_i\}_{i \in I_j}$ with replacement and calculate $\epsilon_i^{\text{test}} = \tilde{\epsilon}_i \cdot \hat{\sigma}_i$.
4. Construct $Y_i^{\text{test}} = \hat{E}[Y_{j,i}|X_i] + \epsilon_i^{\text{test}}$.

In the first and second steps, one can apply any nonparametric methods to estimate the conditional means. In our application, we estimate the conditional means using random forests (Friedberg, Tibshirani, Athey, and Wager, 2021; Wager and Athey, 2018).

We generate test take-up status $\{D_i^{\text{test}}\}_{i=1}^n$ by different procedures since take-up status takes values in zero and one. For samples $i \in I_U$ and $i \in I_T$, we set $D_i^{\text{test}} = 0$ and $D_i^{\text{test}} = 1$, respectively. For samples $i \in I_S$,

1. Estimate $P(D_{S,i} = 1|X_i)$ to obtain $\hat{P}(D_{S,i} = 1|X_i)$.
2. Sample $\{D_i^{\text{test}}\}_{i \in I_S}$ according to $D_i^{\text{test}} \sim \hat{P}(D_{S,i} = 1|X_i)$.

As is the case of Y_i^{test}, one can apply any nonparametric methods to estimate the conditional probability. In our application, we estimate that quantity using random forest.

A.4 Social Welfare Function with a Redistribution Goal

This section describes the derivation of each household's potential welfare contribution. First, we focus on the change in each household's utility resulting from the introduction of the rebate program. Under the assumption that the utility function of each household is quasi-linear and the electricity demand function is linear, assigning a household to arm $j \in \{T, U, S\}$ changes its utility as follows

$$CS_j = \frac{c-p}{2} \cdot (Y_j - Y_U) + R_j,$$

where $R_j = (c-p) \cdot \max\{Y_{\text{base}} - Y_j, 0\} \cdot 1\{D_j = 1\}$. The first term represents the change in consumer surplus due to an increase in the price of electricity from p to c, while the second term represents the rebate received by reducing electricity consumption below the baseline consumption Y_{base}.

Then, given the Pareto weights w for a household, it is natural to define the weighted potential welfare contribution by

$$W_j = w \cdot CS_j + \Delta PS_j - a - R_j + \delta \cdot (Y_j - Y_U).$$

The welfare contribution is the sum of five terms. $w \cdot CS_j$ is the consumer surplus weighted with the Pareto weight. ΔPS_j is change of producer surplus, and hence we have $\Delta PS_j = (p-c)(Y_j - Y_U)$. The constant a denotes the cost taken to implement rebate program. Note that this cost does not include the rebate payment, which is instead reflected in $-R_j$. Finally, the term $\delta \cdot (Y_j - Y_U)$ is the long-run gain. Using the concrete expression of CS_j, we can rewrite the welfare contribution as follows

$$W_j = \left(\frac{(2-w)(p-c)}{2} + \delta \right) \cdot (Y_j - Y_U) - (1-w) \cdot R_j - a,$$

which is equation (12). Further, when $w = 1$, the welfare contribution boils down to

$$W_j = \left(\frac{p-c}{2} + \delta \right) (Y_j - Y_U) - a,$$

which in turn corresponds to equation (11).

A.5 Appendix Tables

Table A.1: The external validity of the experimental sample

	Experimental sample in the untreated group	Random sample of population	Difference between sample and population
Monthly electricity usage in July (kWh)	356 [205]	304 [177]	51.86 (6.34)
Number of people at home	2.48 [1.24]	2.31 [1.20]	0.17 (0.04)
Self-efficacy in energy conservation (1-5 Likert scale)	3.45 [0.85]	3.32 [0.97]	0.13 (0.03)
Household income (JPY10,000)	645 [399]	581 [384]	63.83 (13.06)
Number of households	1,577	2,070	

Notes: We randomly sampled 2070 customers from the target population who did not participate in this experiment, and conducted a similar survey to the one for the experimental sample. The purpose of this survey was to investigate the external validity of our experimental sample by comparing the mean for each variable between the control group from our experimental sample and this random sample. Columns 1 and 2 show summary statistics for the untreated group and the random sample. Column 3 presents differences in means, with the standard errors of these differences in parentheses. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020. We observe larger means for four variables in the untreated group than in the random sample, and the differences are statistically significant. Our experimental sample has larger pre-experiment electricity usage per month, a larger number of people at home on weekdays, higher self-efficacy in energy conservation, and higher household income. This implies that our sample includes a larger number of customers who are willing and able to reduce their electricity consumption, which should be taken into consideration when discussing this study's generalizability.

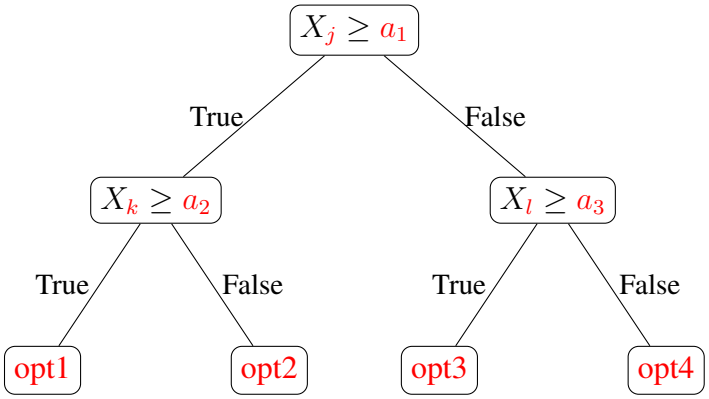
Table A.2: Comparisons of Selection-absent and Selection-driven Targeting in Welfare Function with Redistributional Goal

	Difference in Welfare Gains	p-value
G^* vs. $G^\dagger$ (with $\nu = 1$)	167.1	0.005
	(64.4)	
G^* vs. $G^\dagger$ (with $\nu = 2$)	157.9	0.019
	(75.8)	

Notes: This table compares welfare gains with G^* (selection-driven targeting) and those with $G^\dagger$ (selection-absent targeting) with a redistributional goal. The column "Difference in Welfare Gains" shows the estimated welfare gain of the selection-driven targeting relative to selection-absent targeting in terms of welfare function weighted Pareto weight with parameter $\nu \in \{1, 2\}$, with the standard errors in parentheses. For each row, the column "Difference in Welfare Gains" shows the estimated welfare gain of the policy on the left-hand side (W_L) relative to the policy on the right-hand side (W_R) in JPY per household per season, with its standard error in parenthesis. The column "p-value" gives the p-value for the null hypothesis: $H_0 : W_L \geq W_R$. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

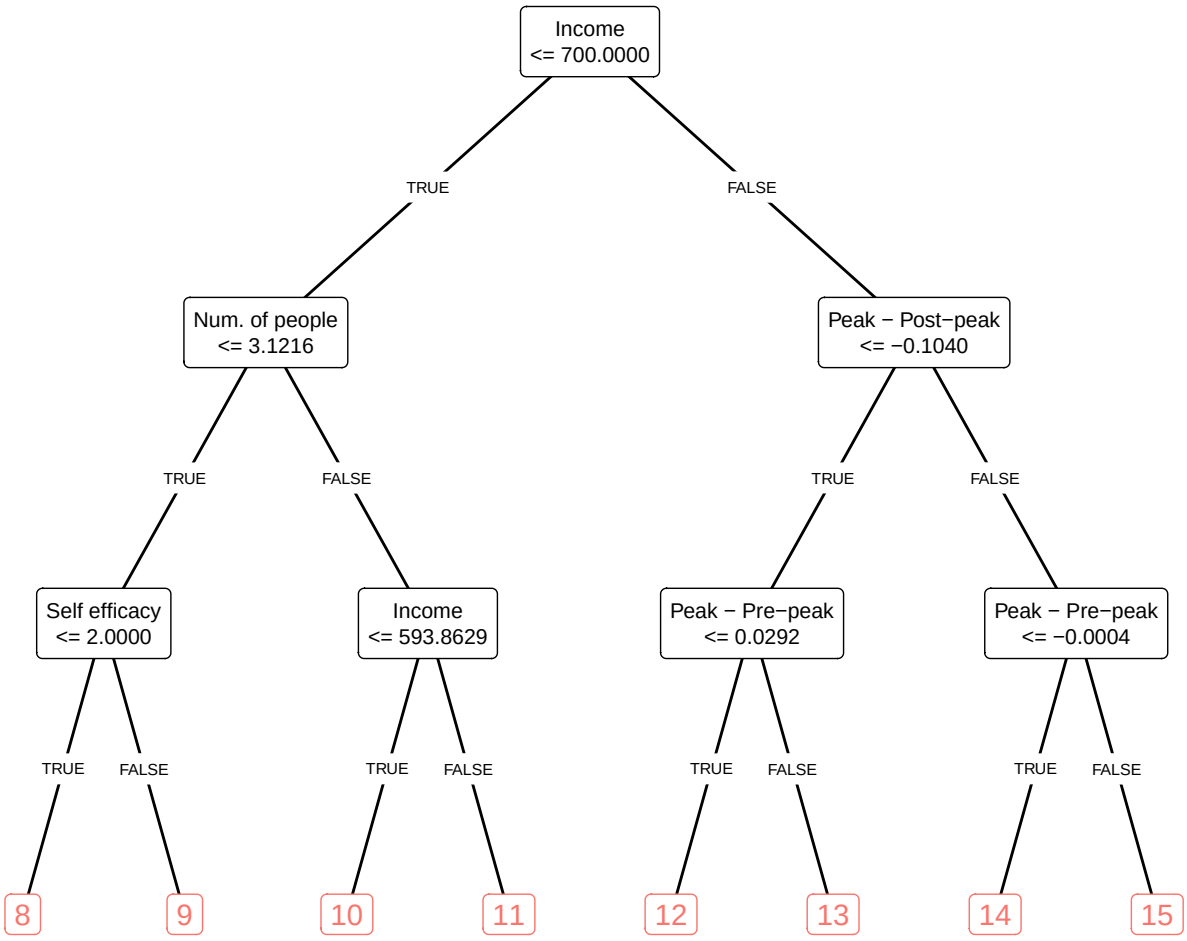
A.6 Appendix Figures

Figure A.1: Decision tree of depth 2



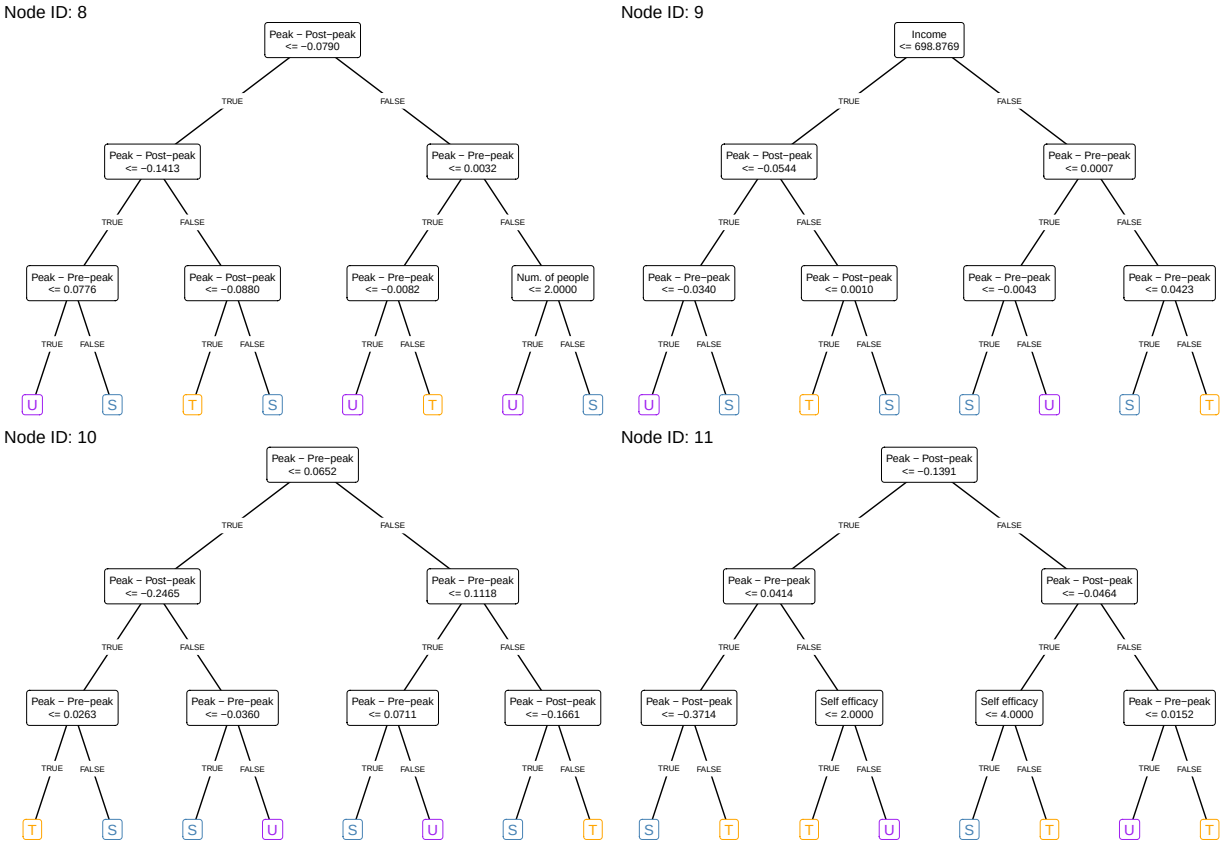
Notes: $(j, k, l) \in \{1, \ldots, K\}^3$, $(a_1, a_2, a_3) \in \mathbb{R}^3$, and $(\text{opt1}, \ldots, \text{opt4}) \in \{T, U, S\}^4$. Searching for the optimal decision tree of depth 2 is equivalent to finding the best combination of indices $(j, k, l) \in \{1, \ldots, K\}^3$ of X and threshold values $(a_1, a_2, a_3) \in \mathbb{R}^3$ in the top 2 layers, and options $(\text{opt1}, \ldots, \text{opt4}) \in \{T, U, S\}^4$ in the bottom layer.

Figure A.2: Optimal Assignment $\hat{G}^*$ (Panel A: From Depth 1 to 3)



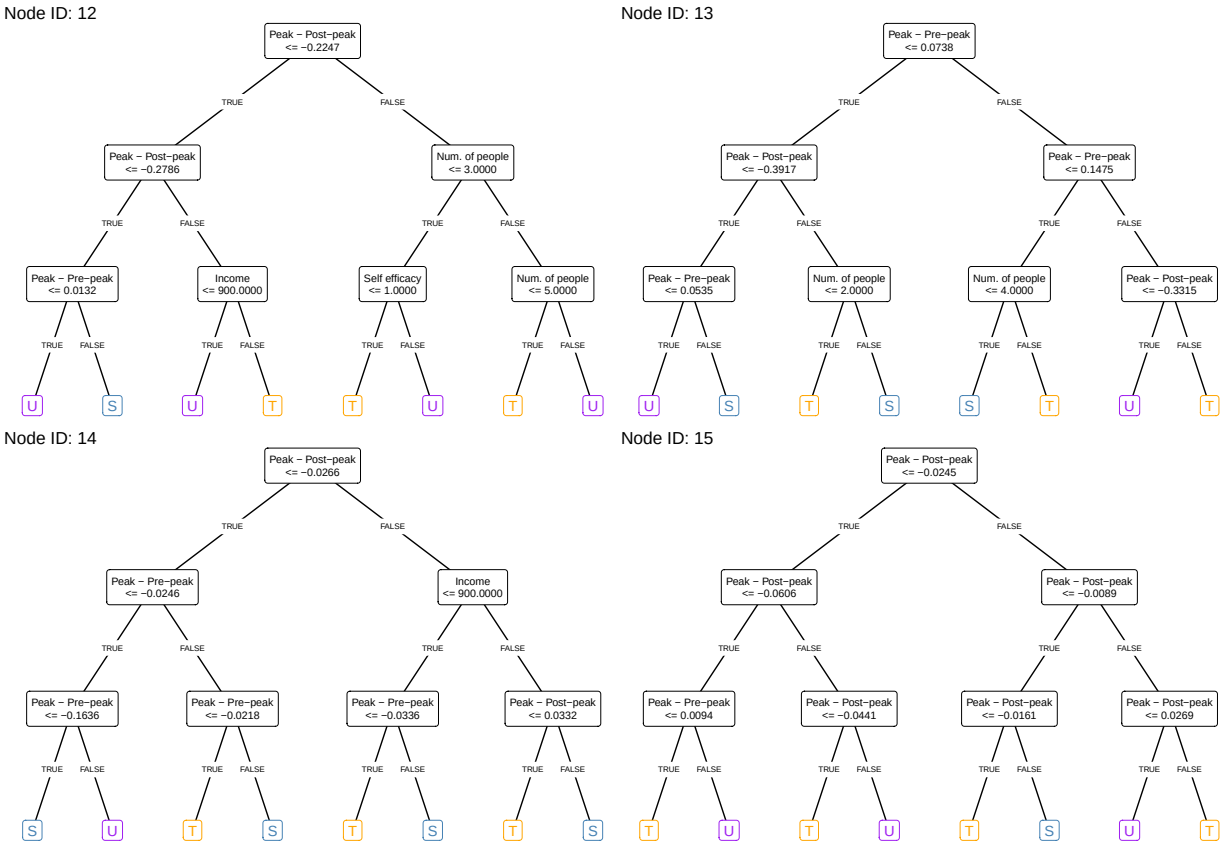
Notes: This figure presents the depth 1 to 3 of the optimal assignment $\hat{G}^*$. Each household answers yes-no questions from its top, and is given one number from 8 to 15. Households given a number less than or equal to 11 go to Panel B below. Other households go to Panel C below. The monetary unit is given as 1 ¢ = 1 JPY in the summer of 2020.

Figure A.2: Optimal Assignment $\hat{G}^*$ (Panel B: From Depth 4 to 6, Left Side)



Notes: This figure presents the left side of depth 4 to 6 of the optimal assignment $\hat{G}^*$. Each household given a number at most 11 in Panel A refers to tree with the same node id. Then, following the yes-no questions from its top, each household assigned to one of U, T, and S.

Figure A.2: Optimal Assignment $\hat{G}^*$ (Panel C: From Depth 4 to 6, Right Side)



Notes: This figure presents the right side of depth 4 to 6 of the optimal assignment $\hat{G}^*$. Each household given a number at least 12 in Panel A refers to tree with the same node id. Then, following the yes-no questions from its top, each household assigned to one of U, T, and S.