The Pre-Analysis Plan for
"A field experiment on energy conservation with manipulation"

Date: 15th August 2025

Takanori Ida*   Takunori Ishihara†   Daido Kido‡   Shusaku Sasaki§

Hypotheses Setting

We propose three main hypotheses to be tested in our experiment. While Hypotheses 1 and 2 help establish the existence and heterogeneity of strategic responses, the primary objective of our study is to assess how such responses affect policy design by using Empirical Welfare Maximization (EWM) and other policy learning methods (Athey & Wager, 2021; Kitagawa & Tetenov, 2018; Ida et al., 2022; Mbakop & Tabord-Meehan, 2021).

Hypothesis 1 (Strategic Response to Future Treatment Assignment Rules):
When the treatment assignment in period 2 is conditional on behavior in period 1 and the rule is disclosed in advance, individuals will strategically adjust their behavior in period 1 to receive the treatment in period 2.

According to a microeconomic model, if individuals are informed of a future assignment rule (e.g., assignment of a rebate program in period 2 is based on meeting a conservation threshold in period 1), they will weigh the utility from receiving the treatment in period 2 against the disutility of changing behavior in period 1. If the net utility gain is positive, we expect individuals to act strategically by adjusting their behavior in period 1. Conversely, if the period 2 assignment is independent of period 1 behavior, such incentives for strategic adjustment disappear.

This implies that observable behavior in period 1—electricity consumption—will systematically differ depending on whether or not eligibility for treatment in period 2 is conditional on period 1 behavior.

* Kyoto University, ida@econ.kyoto-u.ac.jp
† Kyoto University of Advanced Science, ishihara.takunori@kuas.ac.jp
‡ Otaru University of Commerce, dkido@res.otaru-uc.ac.jp
§ The University of Osaka, ssasaki.econ@cider.osaka-u.ac.jp

In our electricity rebate experiment, the treatment consists of a financial incentive based on electricity conservation. We examine a rule in which households are eligible for the rebate program in period 2 only if their electricity reduction rate in period 1 exceeds a predetermined threshold (e.g., 10%).

If this eligibility rule is disclosed in advance, households with a strong preference for receiving the rebate in period 2 may reduce their electricity usage in period 1—even in the absence of a contemporaneous incentive—in order to qualify.

Identifying such strategic behavior is a necessary first step in evaluating whether policy performance can be improved by explicitly accounting for these anticipatory responses using methods such as Empirical Welfare Maximization (EWM).

Hypothesis 2 (Heterogeneity in Strategic Behavior):
We expect there to be substantial heterogeneity in individuals' strategic responses to future treatment rules.

The decision to act strategically depends on the tradeoff between the utility gain from receiving the treatment and the disutility from adjusting behavior. This tradeoff probably varies across individuals due to differences in preferences, constraints, or baseline consumption. Some individuals will find it worthwhile to act strategically; others will not.

Accordingly, we expect variation in the magnitude and direction of behavioral adjustment in period 1 in response to the disclosure of an eligibility rule for period 2 treatment.

In our setting, some households may find it easy to meet the required electricity reduction thresholds—10%, 30%, or 50%, depending on the treatment group—while others may find it costly or infeasible. Even when the eligibility rule is commonly known, we therefore expect to observe heterogeneous responses in first-period electricity usage. We furthermore expect that these heterogeneous responses will be systematically related to observable household characteristics.

Hypothesis 3 (Policy Performance under Strategic Behavior):
When individuals respond strategically to future assignment rules, the optimal targeting policy that accounts for such behavior will differ substantially—both in form and in performance—from one that does not.

Recent work in econometrics has developed methods for estimating optimal treatment assignment rules, typically assuming no strategic response to the rules. However, in many real-world settings, policymakers must disclose allocation rules in advance, potentially inducing strategic responses. If such responses are present and ignored in policy design, performance may suffer. Conversely, explicitly accounting for strategic behavior in targeting rule design can lead to improved policy outcomes.

Conversely, explicitly incorporating strategic behavior into the policy design process can yield more effective targeting rules and improve overall welfare.

In our study, we use policy learning methods, such as Empirical Welfare Maximization (EWM), to estimate and compare optimal targeting policies under both strategic and non-strategic behavioral assumptions. By evaluating how these policies differ in structure and performance, we aim to quantify the welfare costs of ignoring strategic behavior and the potential gains from accounting for it.

Empirical Analyses

Behavioral Evidence: Strategic Response in Period 1 (Hypothesis 1)

To test Hypothesis 1, we will apply a difference-in-means estimator to electricity consumption data from the summer of 2025 (period 1). In particular, we will compare average electricity usage between groups where the Period 2 treatment assignment rule depends on Period 1 behavior and those where it does not.

A significant difference in electricity consumption between these two groups would be consistent with the presence of strategic responses to the future assignment rule. This behavioral response forms the basis for evaluating the welfare implications of policy design under strategic behavior, as addressed in Hypothesis 3.

Heterogeneity in Strategic Behavior (Hypothesis 2)

To examine Hypothesis 2, we will apply off-the-shelf machine learning methods—such as causal forests (Wager & Athey, 2018)—to electricity consumption data from period 1. These methods will allow us to examine how the effect of being in an experimental group with a behavior-dependent period 2 assignment rule varies across households. Specifically, we will estimate conditional average treatment effects (CATEs) and examine how these vary with observable household characteristics.

If Hypothesis 2 is correct, we expect to observe substantial heterogeneity in electricity saving behavior: some households are likely to reduce electricity usage significantly in order to qualify for the period 2 rebate, while others will exhibit little or no behavioral adjustment.

Main Analysis: Empirical Welfare Maximization (Hypothesis 3)

To examine Hypothesis 3, we use Empirical Welfare Maximization (EWM) and related methods to estimate optimal targeting policies based on data from all periods. We will focus on comparing two policy rules of particular interest. Our main objective is to compare two policy rules that differ in whether they account for strategic behavior.

The first policy is estimated using data only from the experimental groups in which period 2 treatment eligibility is independent of period 1 electricity usage. This setting corresponds to a conventional optimal targeting rule under the assumption that individuals do not respond strategically—an assumption commonly made in the econometrics literature. Because the assignment rule is fixed and unrelated to earlier behavior, households in these arms have no incentive to adjust their electricity consumption in Period 1.

The second policy is estimated using data that includes experimental groups in which period 2 treatment eligibility does depend on period 1 electricity usage. This design allows us to estimate an optimal targeting rule that incorporates the observed behavioral responses to the eligibility threshold—thus capturing the effects of strategic adaptation.

In addition to these two primary policies, we will compare them to benchmark rules involving no targeting—such as universal treatment (providing the rebate to all households) and universal non-treatment (providing it to none)—to contextualize the welfare implications of optimized targeting.

By comparing the structure and welfare performance of these targeting rules, we aim to assess the extent to which accounting for strategic behavior improves policy effectiveness.

References:

Athey, S., & Wager, S. (2021). Policy learning with observational data. *Econometrica*, *89*(1), 133-161.

Ida, T., Ishihara, T., Ito, K., Kido, D., Kitagawa, T., Sakaguchi, S., and Sasaki, S. (2022). *Choosing who chooses: Selection-driven targeting in energy rebate programs* (No. w30469). National Bureau of Economic Research.

Kitagawa, T., and Tetenov, A. (2018). Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, *86*(2), 591-616.

Mbakop, E., & Tabord-Meehan, M. (2021). Model selection for treatment choice: Penalized welfare maximization. *Econometrica*, *89*(2), 825-848.

Wager, S., and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, *113*(523), 1228-1242.