Journal of Development Economics

Energy Conservation under Repeated Contracts and Contests: Registered Report

--Manuscript Draft--

Manuscript Number:	DEVEC-D-26-00069R1
Article Type:	Registered Report Stage 1: Proposal
Section/Category:	Experimental Papers, credit, insurance
Keywords:	Electric Utilities; Demand-Side Management; Consumer Economics; Contests; Contracts
Corresponding Author:	Chi Ta University of Texas at Austin Lyndon B Johnson School of Public Affairs UNITED STATES OF AMERICA
First Author:	Chi Ta
Order of Authors:	Chi Ta
	Teevrat Garg
	Jorge Lemus
	Guillermo Marshall
Abstract:	We propose a field experiment to evaluate whether performance-based incentives for residential energy conservation are effective when repeated and deployed at scale. In our prior study with EVNHANOI, we found that both contests (rewards based on relative performance) and contracts (rewards based on absolute savings) reduced electricity use by 7–9%, with contests delivering similar savings at lower cost. In summer 2026, we will revisit the cost-effectiveness of contracts and contests when a) implemented across both an opt-in population and a random set of households to assess the role of self-selection, b) when households are repeatedly exposed to these incentives, and c) when incentive budgets are fixed over larger populations. To answer these questions, we will randomly assign over six hundred thousand households in Hanoi to one of two enrollment regimes, either opt-in or default, and then cross-randomize them to one of six incentive schemes: contracts offering $2 for 5%, 10%, or 15% electricity reductions, contests awarding a $40 prize in contests of 50 or 250 households, or a control group receiving no incentives. Additionally, households will be cross-randomized to either two or three sequential 2-week intervention periods. Using smart meter data, we will measure electricity consumption to assess both immediate effects and dynamic behavioral responses over time.
Response to Reviewers:	

Energy Conservation under Repeated Contracts and Contests: Registered Report

April 15, 2026

Abstract

We propose a field experiment to evaluate whether performance-based incentives for residential energy conservation are effective when repeated and deployed at scale. In our prior study with EVNHANOI, we found that both contests (rewards based on relative performance) and contracts (rewards based on absolute savings) reduced electricity use by 7–9%, with contests delivering similar savings at lower cost. In summer 2026, we will revisit the cost-effectiveness of contracts and contests when a) implemented across both an opt-in population and a random set of households to assess the role of self-selection, b) when households are repeatedly exposed to these incentives, and c) when incentive budgets are fixed over larger populations. To answer these questions, we will randomly assign over six hundred thousand households in Hanoi to one of two enrollment regimes, either opt-in or default, and then cross-randomize them to one of six incentive schemes: contracts offering \$2 for 5%, 10%, or 15% electricity reductions, contests awarding a \$40 prize in contests of 50 or 250 households, or a control group receiving no incentives. Additionally, households will be cross-randomized to either two or three sequential 2-week intervention periods. Using smart meter data, we will measure electricity consumption to assess both immediate effects and dynamic behavioral responses over time.

JEL Codes: L94, Q41, D12

[1]We are grateful to the King Climate Action Initiative of the Jameel Poverty Action Lab as well as the Deep Decarbonization Initiative at UC San Diego for generous funding to support this research.

Table of Contents

1 Introduction **1**

2 Context of the Experiment **4**

3 Research Design **6**

3.1 Participants 6

3.2 Design 6

4 Data **9**

4.1 Variables 9

4.2 Balance Checks 10

5 Analysis Plan **11**

5.1 Primary Hypotheses 12

5.2 Heterogeneity Analysis 13

5.3 Inference and Robustness 13

6 Previous Results **13**

7 Power Calculations **15**

8 Proposed Timeline **16**

9 Administrative Information **16**

1 Introduction

Electricity consumption in rapidly growing urban areas in low- and middle-income countries (LMICs) is placing increasing strain on power systems. This is especially true during peak summer months, when demand spikes due to cooling needs and electricity supply systems must rely on costly and carbon-intensive generation sources. In these settings, electric utilities often confront a dual mandate: deliver reliable service while keeping costs and emissions in check. However, utilities face limited options to meet surging demand in the short run, especially in contexts where adding generation or storage capacity is expensive, politically difficult, or environmentally undesirable. Against this backdrop, demand-side management (DSM)—interventions that encourage households to reduce electricity use—has emerged as a promising alternative (Auffhammer et al., 2008; Borenstein, 2012).

The economics of DSM is conceptually straightforward. By shifting or reducing electricity demand, especially during peak periods, utilities can avoid costly spot market purchases, delay capacity and distribution infrastructure expansion, and reduce carbon and particulate emissions (Callaway et al., 2018; Graff Zivin and Neidell, 2018). However, implementing DSM programs at scale in LMICs presents both logistical and behavioral challenges. One central issue is how to design incentive schemes that are cost-effective, easily communicable, and capable of sustaining behavioral change across a large number of households and across repeated deployments. While tiered pricing, social comparison nudges, and direct subsidies have all been employed (Allcott and Mullainathan, 2010; Allcott, 2011; Allcott and Kessler, 2019), there is growing interest in performance-based financial incentives, particularly in the form of either contracts or contests.

Contracts are individually targeted incentives that reward households if their energy savings exceed pre-specified thresholds. They are straightforward to understand and implement, and are widely used in DSM programs globally (Ito, 2015; Neveu and Sherlock, 2016; Fowlie et al., 2021). Contests, in contrast, offer rewards based on relative performance—households compete against each other for a fixed prize, with only the top performer in a group receiving a reward. The performance of contests has been investigated in various settings (Gross, 2020; Bhattacharya, 2021; Lemus and Marshall, 2021; Ta, 2024) and their effects on effort relative to contracts are ambiguous (Lazear and Rosen, 1981; Green and Stokey, 1983; Garg et al., 2025). In the context of electricity conservation, recent evidence suggests that both mechanisms can reduce household electricity use, and that contests may achieve comparable savings at lower cost (Garg et al., 2025). Yet, the existing evidence concerns one-time interventions or relatively selected participant pools.[1]

This project proposes a field experiment with EVNHANOI, the electricity utility serving Hanoi, Vietnam, to address these questions. The experiment is motivated by a previous randomized controlled trial that we conducted with EVNHANOI during the summer of 2023. In that study, both contracts and contests reduced electricity consumption by roughly 7 to 9 percent relative to a control group, while contests were more cost-effective (Garg et al., 2025). Those results established that performance-based incentives can meaningfully reduce residential electricity demand in this setting.

Encouraged by these findings, EVNHANOI expressed interest in deploying demand-side management programs at scale. However, several critical questions remain unanswered before such programs can be adopted as a regular policy tool. First, to what extent do results from the initial experiment scale when applied to a randomly selected sample of households rather than an opt-in sample? Second, utilities typically deploy conservation programs repeatedly (often annually or seasonally); what happens when these incentives are offered repeatedly? Specifically, do households that do not win or earn a reward in the first round become discouraged and reduce effort in subsequent rounds? Or are such schemes robust to repeated exposure? Finally, scaling these programs requires operating under fixed incentive budgets, which mechanically lowers expected payouts per household as the number of participants increases. Whether contests and contracts continue to perform well under these constraints is an open empirical question. Addressing these issues is essential for understanding not only whether such incentives work, but whether they can be sustained and scaled as part of real-world electricity policy.

Our field experiment is designed around these three questions. The first design feature is variation in *enrollment regime*. Some households are assigned to an *opt-in* regime, in which they receive an invitation to participate and are included only if they affirmatively enroll. Other households are assigned to a *default* regime, in which they are enrolled automatically. This comparison allows us to study how the performance of conservation incentives differs between a self-selected participant pool and a broader policy-relevant population.

The second design feature is variation in *incentive mechanism*. Within each enrollment regime, households are assigned either to a contract treatment, a contest treatment, or a control group. In the opt-in regime, the design focuses on a parsimonious set of arms: a 10 percent contract, a 50-household contest, and a control group. In the default regime, the design is richer and varies both contract difficulty and contest scale. Specifically, households may face contracts with 5, 10, or 15 percent savings thresholds, contests with 50-household

[1] Learning about the performance of repeated interventions has been studied, for example, by Ito et al. (2018) in the context of an energy conservation program involving price changes during high-demand times as well as moral suasion.

or 250-household groups, or no financial incentive. This structure allows us to compare the performance of contracts and contests across a range of practical program designs.

The third design feature is variation in *repeated exposure*. Selected treatment arms last for two rounds, while others last for three rounds. Each round is a two-week conservation period during the summer. This aspect of the design allows us to test whether treatment effects persist, weaken, or strengthen with repetition. Repeated exposure is especially important for contests, since households that do not win early on may become discouraged and reduce effort in later rounds. By comparing two-round and three-round treatment arms, we can evaluate whether repeated incentives remain behaviorally effective and whether persistence differs across mechanisms.

These design features allow the experiment to contribute to the literature and to policy along three margins. First, the paper speaks to the external validity of behavioral and incentive-based energy conservation programs by comparing effects in opt-in and default-enrolled populations. Second, it contributes to the literature on dynamic incentives by studying repeated exposure rather than a one-time intervention. Third, it provides evidence on the cost-effectiveness of alternative performance-based mechanisms when programs are scaled and expected payouts are compressed.

The design allows us to test several hypotheses of interest. First, we assess whether contest-based incentives continue to outperform contracts in terms of cost-effectiveness and energy savings in a large sample drawn at random from the population of interest. Second, by exposing households to the same incentives in two and three consecutive rounds, we test whether conservation responses persist over time or attenuate due to discouragement, particularly among households that do not earn rewards in the first round. If effort declines more in contests than in contracts among non-recipients, this would suggest that rank-order mechanisms may underperform when used repeatedly. Third, we examine whether the generosity of the incentive (i.e., contest prize size or contract reward amount) modulates both initial participation and subsequent retention or engagement. In contests, larger group sizes mechanically reduce winning probabilities, while in contracts, higher savings thresholds reduce the likelihood of earning a reward. Comparing behavioral responses across these margins allows us to assess whether relative-performance incentives remain effective as they are scaled, and whether they outperform absolute-performance contracts when both mechanisms are constrained by realistic budget limits.

2 Context of the Experiment

Vietnam's electricity sector is undergoing a transformation driven by rapid economic growth, rising residential demand, and national commitments to reduce emissions. In Hanoi—the capital and second-largest city—these dynamics are especially pronounced. As households enter the middle class, they increasingly adopt energy-intensive appliances such as air conditioners and washing machines, leading to sharp increases in electricity use during the hot summer months (Gertler et al., 2016). This seasonal demand surge strains grid infrastructure and raises the risk of blackouts and service disruptions.

EVNHANOI, the city's state-owned electric utility, relies heavily on hydropower generation. As a result, unlike solar-dominated systems where peak supply and peak demand are misaligned at the daily level, Hanoi's power shortages emerge from a different mismatch. Hydropower generation is highly seasonal and dependent on rainfall. Reservoirs are often depleted by mid-summer, which coincides with the hottest weeks of the year. This makes electricity shortages in Hanoi a *week-of-year* problem—rooted in the cumulative seasonal imbalance between supply and demand—rather than a *time-of-day* problem. Peak reductions during this critical seasonal window are therefore of particular value for grid reliability and emissions mitigation.

To address this challenge, EVNHANOI has begun exploring demand-side management (DSM) strategies. While dynamic pricing and energy efficiency subsidies are still in early stages, the utility has invested in digital infrastructure, including widespread installation of smart meters and a mobile app that allows households to track their daily usage. These tools provide a platform to test behavioral and incentive-based DSM interventions.

In a 2023 field experiment, we partnered with EVNHANOI to evaluate two such interventions: individual contracts (which reward households for reducing consumption relative to their own past use) and contests (which reward the household with the greatest relative savings within a group). That study found that both mechanisms led to significant energy savings—7 to 9 percent during the hottest month of the year—with contests delivering these savings at roughly half the cost of contracts (Garg et al., 2025). The contest design was particularly appealing to utility managers because it fixed total payouts ex ante, independent of the number of households meeting consumption targets.

Encouraged by these findings, EVNHANOI is considering scaling such programs. However, the earlier experiment left several important questions unresolved. First, that study relied heavily on households willing to engage with the program, so it remains unclear whether similar treatment effects would arise in a broader population enrolled by default. Second, the earlier intervention was designed as a one-time program. For policy purposes, however, utilities are often interested in deploying conservation incentives repeatedly over the course of a summer or across years. Repeated exposure may generate persistence in conservation behavior, but it may also weaken responses if households become discouraged after failing to earn a reward or win a contest. Third, the utility is interested not only in whether incentives work, but also in how their performance changes as programs expand and expected payouts per household decline.

The current experiment is designed to address these questions in a policy-relevant setting. The study compares two enrollment regimes. In the opt-in regime, households are first invited to participate, and only those that opt in are assigned to treatment or control groups. In the default regime, households are enrolled automatically and assigned directly to treatment or control groups. This distinction allows us to compare the performance of the same broad classes of incentives in a self-selected sample and in a much larger default-enrolled sample.

Within these enrollment regimes, households are assigned to contract treatments, contest treatments, or a control group. In the opt-in regime, the design focuses on a relatively small set of arms: a 10 percent contract, a 50-household contest, and a control group, with some treated households exposed for two rounds and others for three rounds. The reason is that opt-in take-up is typically small, and we expect around 1 to 2 percent of the households invited to participate to opt in.

In the default regime, the design is richer, implementing several contract thresholds, contest size, and treatment duration. In particular, the default regime includes contracts with 5, 10, and 15 percent savings thresholds, contests with 50-household and 250-household groups, selected three-round treatment arms, and a large control group. This structure allows the experiment to shed light on both the extensive-margin question of participation and the intensive-margin question of how treatment performance changes with incentive design and repeated exposure.

The setting offers three main advantages for answering these questions. First, smart-meter data provide precise daily measures of household electricity consumption before, during, and after the intervention. Second, EVNHANOI can implement treatment assignment and communication at scale using its existing administrative and messaging systems. Third, the utility's interest in practical DSM policy makes the experiment directly relevant for real-world program design. More broadly, because EVNHANOI is a major public utility operating in a rapidly growing urban environment, the results may inform the design of conservation programs in similar settings where summer demand peaks create operational and environmental challenges.

This follow-up study seeks to determine whether the effectiveness of contests survives broader enrollment, repeated implementation, and larger-scale deployment. The new randomized controlled trial involves a broader sample of six hundred thousand households. By implementing multiple consecutive rounds of incentives, we aim to evaluate not only the replicability of our prior results but also the behavioral persistence of effort under repeated incentives in a hydropower-dependent grid facing seasonal shortages.

3 Research Design

3.1 Participants

The study population will consist of EVNHANOI residential customers ("households") that meet a pre-treatment eligibility criteria based on their daily electricity consumption during the year preceding the target start date. A household is eligible if it satisfies all of the following conditions:

1. Less than 20% of their daily consumption observations are equal to 0 kWh.
2. Fewer than 10% of daily consumption observations are missing. These data are not truly "missing," but due to IT system glitches, are recovered later on. While this is not an issue for data analysis, it precludes including households in our randomization.
3. Average daily consumption (during the entire year) is between 1 kWh and 24 kWh.
4. Average daily consumption during the reference period (i.e., the same period as the proposed experiment period but one year prior) is between 1 kWh and 30 kWh.

These eligibility criteria are similar to those used in our previous study (Garg et al., 2025), and are intended to limit the variance in the outcome variable of interest by excluding households in the thin (but long) upper tail of the distribution of electricity consumption. Households with extremely high and volatile electricity consumption would add noise to the analysis.

3.2 Design

Let the full study population consist of N households, with household-level monthly electricity consumption observed before and after the intervention. At the beginning of the experiment, we will randomly assign six hundred thousand households to two enrollment regimes: an *opt-in regime* and a *default regime*. Specifically, about 400,000 households will be assigned to the opt-in regime and 200,000 to the default regime. These group sizes are determined based on power calculations, budget constraints, and expected take-up rates, as discussed later. The two regimes are designed to compare the performance of conservation incentives in a self-selected sample of participants and in a broader sample of households enrolled by default.

A first group of size N_1 (the "opt-in" regime) receives an early invitation text message in May 2026, before the intervention, stating that an energy conservation program will run in the summer of 2026 and asking whether the household would like to participate. A second group of size $N_2 = N - N_1$ (the "default" regime) receives no early text.

The purpose of this first randomization is to allow us to compare the effect of incentives on i) a set of households that opt into the study and ii) a random set of households. In other words, this comparison will help us gauge the importance of selection in understanding the impacts of incentives on energy conservation. Among households in the opt-in regime, responses will be recorded in three categories: "Yes", "No", and "No response". These responses reveal the willingness to engage with the program. All households that do not respond "Yes" will be excluded from the study.

In principle, the experiment would follow a $2 \times 6 \times 2$ factorial-like design across three dimensions: enrollment regime, incentive scheme, and repetition scheme. While not all possible combinations are used due to budget restrictions, this structure organizes the study and allows comparisons across key factors. Households in the "opt-in" enrollment regime that choose to participate are randomly assigned at the household level to the relevant incentive and repetition schemes once the opt-in period is complete. In contrast, households in the default regime are randomly assigned an incentive and repetition scheme immediately after the enrollment regime assignment, as no opt-in period is required. This will allow us to estimate intention-to-treat effects for a policy-relevant population of interest, including households that would not ordinarily enroll in energy conservation programs.

The possible combinations of enrollment regime, incentive scheme, and repetition scheme are given by:

1. **Dimension 1: Enrollment regime**
    - **Opt-in regime:** Households receive a text in May 2026 inviting them to participate. Responses are recorded as {Yes, No, No-Response}. Those who respond "Yes" will remain in the study.
    - **Default regime:** Households are defaulted into the study.
2. **Dimension 2: Incentive Scheme**
    - **Incentive 1: Contract (5%)** — All participants who reduce their electricity consumption by 5 percent relative to the same period in the previous year will receive a reward of \$2 (USD).

- **Incentive 2: Contract (10%)** — All participants who reduce their electricity consumption by 10 percent relative to the same period in the previous year will receive a reward of $2 (USD).
- **Incentive 3: Contract (15%)** — All participants who reduce their electricity consumption by 15 percent relative to the same period in the previous year will receive a reward of $2 (USD).
- **Incentive 4: Contest (50)** — Participants will be placed into a 50-household contest, and a single prize of $40 will be awarded to the household that achieves the highest percentage reduction in electricity consumption compared to the same period in the previous year.
- **Incentive 5: Contest (250)** — Participants will be placed into a 250-household contest, and a single prize of $40 will be awarded to the household that achieves the highest percentage reduction in electricity consumption compared to the same period in the previous year.
- **Incentive 6: Control Group** — This group will not participate in a contest or individual contract.

3. **Dimension 3: Repetition Scheme**

    - **3-round group**: Households in this group are invited to participate in a sequence of three 2-week incentive periods, starting on dates June 15, July 15, and August 15.
    - **2-round group**: Households in this group are invited to participate in a sequence of two 2-week incentive periods, starting on dates July 15 and August 15.

The incentive-scheme dimension will allow us to compare the effectiveness of each incentive (the various contracts, contests, and controls) within each incentive period across groups. The repetition scheme will allow us to measure whether the effectiveness of a particular incentive (the various contracts, contests, and controls) changes with repetition. Staggering the start of the incentive periods allows us to measure the "repetition" effect while accounting for temporal variation, such as weather, since treatment effects may differ across months and not solely reflect repeated exposure. However, because the interval between treatments is only two weeks, we do not expect time effects to be very large. Households will not be informed of subsequent incentive periods until the new incentive period is due to commence.

Due to selection, households in the opt-in regime are not directly comparable to those in the default regime. But measuring the impact of the different incentives on energy consumption across groups (opt-in versus default) will reveal the role of selection on the effects of the

Group	Rounds	Incentive	Reward/Round ($)	N	Win Prob.	Total Payout ($)	Avg. Payout / Round ($)
O1	2-round	Contract (10%)	2	750	0.25	750	0.50
O2	3-round	Contract (10%)	2	750	0.25	1125	0.50
O3	2-round	Contest (50)	40	750	0.02	1200	0.80
O4	3-round	Contest (50)	40	750	0.02	1800	0.80
OC	NA	Control	0	750	0.00	0	0.00
			Total:	**3750**		**4875**	

Table 1: Summary of All Treatment Groups and Expected Payouts in Opt-in Regime

incentives. Learning about the role of selection is relevant to the question of whether it is more cost-effective to target incentives to households that choose to opt in.

Tables 1 and 2 show the summary of all treatment groups and expected payouts in the opt-in and default regimes, respectively.[2] The expected probabilities of winning are based on findings from our previous study. The sample sizes are determined based on power calculations, data from previous experiments, the project's budget, and an expected invitation response (take-up) rate of about 1% in the *opt-in regime*. We expect roughly 51,750 households to participate across the various incentive groups and approximately 152,000 households in the control group.

During each 2-week incentive period, we plan to send households reminders about the program. The tentative reminder schedule within each incentive period is outlined below:

- Day 3: A reminder that the contest/contract has started.
- Day 7: A message noting that the first week is complete, with one more week to go.
- Day 10: A reminder that only a few days remain in the contest/contract period.
- Day 14: Notification that the contest/contract has ended. The message will include: "The contest/contract has ended. Winners will be notified soon."

4 Data

4.1 Variables

Electricity Consumption: The main variable of interest is daily electricity consumption at the household level. This variable is obtained from the utility company, which mea-

[2]All incentive schemes include two rounds; a third round will be included for some incentive schemes (e.g., D3 and D6) depending on the availability of funding.

Group	Rounds	Incentive	Reward/Round ($)	N	Win Prob.	Total Payout ($)	Avg. Payout / Round ($)
D1	2-round	Contract (5%)	2	3000	0.40	4800	0.80
D2	2-round	Contract (10%)	2	6000	0.25	6000	0.50
D3*	3-round	Contract (10%)	2	6000	0.25	9000	0.50
D4	2-round	Contract (15%)	2	6000	0.20	4800	0.40
D5	2-round	Contest (50)	40	6000	0.02	9600	0.80
D6*	3-round	Contest (50)	40	6000	0.02	14400	0.80
D7	2-round	Contest (250)	40	15000	0.004	4800	0.16
DC	NA	Control	0	152000	0.00	0	0.00
			Total:	**200000**		**53400**	

Table 2: Summary of All Treatment Groups and Expected Payouts in Default Regime

sures electricity consumption with smart meters installed in every home.[3] We collect daily consumption data for each household for a period of at least twelve months prior to the intervention, throughout the duration of the experiment, and for at least twelve months following the conclusion of the interventions. This extended data collection period allows us to assess both the immediate and long-term effects of the treatments on household energy use.

Weather Data: Weather plays a significant role in influencing a household's electricity consumption and their likelihood of winning a prize in a contract or contest. As a result, we gather daily air temperature data for Hanoi from Visual Crossings. This dataset encompasses the air temperature variable, along with a "feels like" temperature variable, which takes into account temperature and humidity to provide a more accurate representation of the perceived outdoor temperature. We utilize these data to study heterogeneous responses by weather conditions on a given day.

Generation and Transmission Data: To assess the cost-effectiveness and welfare impacts, we also obtain anonymized administrative generation and transmission data from the utility, allowing us to quantify the benefits of energy savings in terms of reduced energy production, carbon emissions, and the prevention of blackouts and system failures.

4.2 Balance Checks

We will check balance between the treatment and control groups using data on the historical electricity use of households. Specifically, we will use the daily electricity consumption at the household level computed month by month for the 12 months prior to our experiment. We will check balance in two steps.

[3] To ensure data security and privacy protection, the utility company provides the data in anonymized form as a condition of access.

Enrollment Regime Step We consider the daily electricity consumption at the household level computed month by month for the 12 months prior to our experiment. For each of these variables, we will run the following specification:

$$y_i = \alpha + \text{opt in}_i \cdot \psi + \varepsilon_i,$$

where opt in_i is an indicator that takes the value one if household i was assigned to the opt-in regime. In our balance analysis, we will report estimates for the coefficients ψ (one for each variable listed above) and their standard errors.

Treatment Group Within each enrollment regime, and for each of the variables listed above, we will run the following specification:

$$y_i = \alpha + \sum_k 1\{\text{treatment}_i = k\}\beta_k + \varepsilon_i,$$

where treatment_i is a variable indicating the treatment assignment of household i. The regression includes indicators for all treatment groups except for the control group (the omitted category). In our balance analysis, we will report estimates for the coefficients $\{\beta_k\}$, their standard errors, and the p-value from a joint test of statistical significance of all coefficients on the treatments indicators (i.e., a test where $H_0 : \beta_1 = \beta_2 = \cdots = \beta_K = 0$) for every variable listed above.

In the case of the opt-in enrollment regime households, the balance analysis will be conducted for households that choose to opt into the study.

5 Analysis Plan

We will estimate the impact of repeated incentive treatments on household electricity consumption separately for the households in the opt-in and default enrollment regimes.

For the households in each enrollment regime e, we will use the following household-day level specification:

$$y_{it} = \alpha + \sum_{r=1}^{3}\sum_{k=1}^{K} \beta^e_{k,r} \cdot \text{Treat}_{ik} \cdot \text{Round r}_{it} + \gamma_i + \delta_t + \varepsilon_{it}, \tag{1}$$

where y_{it} is daily electricity consumption (in kWh) for household i on day t, Treat_{ik} is an indicator for assignment to treatment group k (including control, contest variants, and contract variants), Round r_{it} is an indicator that takes the value one during the round $r \in$

$\{1, 2, 3\}$ incentive period faced by household i, γ_i are household fixed effects, and δ_t are day fixed effects that control for time-varying factors such as temperature or calendar effects.[4] The omitted category in this equation is the control group.

Our parameters of interest are $\beta^e_{k,r}$, which measure the average impact of intervention k in round r on electricity consumption for households in enrollment regime e.

Standard errors will be clustered at the household level. We will also estimate alternative versions of Equation 1 using the natural logarithm or the inverse hyperbolic sine transformation of the electricity consumption (in kWh) as the dependent variable. All of these econometric models will make use of daily household-level electricity consumption data from before, during, and after the experimental period.

We will also estimate versions of Equation 1 where we allow for time-varying treatment effects (i.e., event study design), to measure whether these incentives induce energy conservation beyond the experimental period.

5.1 Primary Hypotheses

1. **Effectiveness of Incentives (Round 1):** We test whether households in any treatment group reduce electricity use during the first intervention period relative to the control group.

2. **Persistence and Discouragement (Rounds 2 and 3):** We test whether treatment effects persist or attenuate in the second and third rounds, with specific attention to differences in performance between those who did and did not win or earn rewards in Round 1.

3. **Relative Performance of Mechanisms:** We compare contests to contracts on both behavioral and cost-effectiveness dimensions, using average energy saved and reward payouts.

4. **Role of Selection:** We compare whether the response to incentives is greater among households who opt in to the study relative to households who were defaulted into the study.

As discussed, we will estimate treatment effects separately for each round and compare changes in consumption between rounds within households to test for effort persistence or discouragement effects.

[4] δ_t also captures experimental round fixed effects.

5.2 Heterogeneity Analysis

We will explore heterogeneity in treatment effects along key household characteristics and the outcomes of the first-round interventions, including:

i) Baseline electricity consumption (above vs. below median)

ii) Baseline variance in electricity consumption (above vs. below median)

iii) Electricity consumption in round 1

iv) Winning/earning a reward in round 1

v) Electricity consumption in round 2

vi) Winning/earning a reward in round 2

vii) Weather

We will include interaction terms between these variables and treatment indicators.

5.3 Inference and Robustness

All regressions will report robust standard errors clustered at the household level. To address potential imbalances in covariates across treatment arms, we will control for baseline electricity use (from the same month in the previous year) and household-level averages during the 30 days preceding the experiment, if the specification does not include household fixed effects.

We will also conduct robustness checks using alternative outcome definitions, including log consumption and winsorized consumption at the top 1%. Additional specifications will test sensitivity to outliers and alternative baselines (e.g., consumption relative to the 30-day moving average instead of year-over-year).

Finally, we will conduct cost-effectiveness calculations for each treatment arm, defined as the monetary cost per kWh saved and per metric ton of CO_2 abated, using emission factors provided by EVNHANOI for summer peak generation.

6 Previous Results

This study builds on a randomized controlled trial we conducted in partnership with EVN-HANOI, the city's exclusive electricity utility, during the summer of 2023. The experiment aimed to evaluate the relative cost-effectiveness of two incentive mechanisms for household energy conservation: *contracts*, which offered fixed rewards for meeting absolute consumption reductions, and *contests*, which rewarded the top performer in each group based on relative performance. The intervention was implemented through EVNHANOI's mobile app and leveraged the city's widespread smart meter infrastructure.

We recruited 11,194 households, primarily through the utility's app, and randomized them into one control group and three treatment arms: two variants of contract schemes and a contest. The experiment ran from July 15 to August 13, 2023. Households in the contract groups were offered tiered payments for reducing electricity consumption by 5%, 10%, or 15% (low-threshold contract), or by 10%, 15%, or 20% (high-threshold contract), compared to their usage in the same period the previous year. Contest participants were grouped into cohorts of 50, with the top saver in each group receiving a prize of approximately $87. Control households received no financial incentives but could monitor their consumption via the app, as could all other groups.

We found that both contracts and contests significantly reduced electricity consumption, with average reductions of 7–9% relative to the control group. These effects were consistent across multiple empirical specifications, including both cross-sectional and within-household analyses. Moreover, we found that these reductions were additional—persisting for at least a week after the end of the experiment—before gradually returning to pre-treatment levels.

While the average energy savings were statistically indistinguishable across the treatment groups, the cost per household differed significantly. On average, contest participants received $1.74 in rewards, compared to $3.14–$3.21 in the contract arms. Thus, contests achieved similar conservation outcomes at roughly half the cost.

We also found evidence that households responded to marginal incentives: those offered a financial reward for reaching a given threshold (e.g., 5%) were more likely to achieve that exact reduction than those who were not. This supports the idea that households fine-tune their effort based on expected payoffs.

Using structural estimates from a model of household energy demand with idiosyncratic and common shocks, we compare the performance of cost-equivalent optimal contracts and contests. The model predicts—and the data confirm—that contests can outperform even optimal contracts when contest size is sufficiently large and expected payments are held constant. This performance gap emerges because contests, unlike contracts, are not sensitive to common shocks (e.g., temperature variation) and allow utilities to fix program budgets ex ante.

Finally, we estimate the marginal abatement cost (MAC) of CO_2 under each mechanism. Contests delivered emissions reductions at $59–76 per ton of CO_2, well below the social cost of carbon. From the utility's perspective, the net cost of conservation under contests was often negative when oil was the marginal electricity source—making the case for these programs even absent explicit climate policy.

These findings demonstrate that contests are not only behaviorally effective but also financially attractive for utilities facing summer peak constraints and emission reduction goals. The current study seeks to test the replicability of these results in a new sample and to examine whether repeated exposure to such incentives results in diminished effort, particularly for contest participants who do not win.

7 Power Calculations

Our previous study, Garg et al. (2025), finds energy conservation effects from both contracts and contests of around 7 percent (and as large as 9 percent), and we cannot reject the null that the effects of contracts and contests are equal. That study included roughly 3,000 participants per control or treatment group. The power calculations below are intentionally more conservative, assuming effect sizes of only 1 to 2.5 percent. This is because we are looking at a random sample of households, not just those who chose to participate in the program, as in our previous study. In addition, we aim to detect a small but meaningful 1 percentage point difference between treatment arms, not only between treatment and control groups. As a result, we expect smaller effect sizes in this study than in the previous one.

We use the following equation to calculate the appropriate sample size for our study:

$$J = \frac{\left(t_{1-\kappa} + t_{\frac{\alpha}{2}}\right)^2}{P\left(1-P\right)} \frac{\sigma^2}{MDE^2} \left(\rho + \frac{1-\rho}{T}\right)$$

where

- J is the sample size
- κ is the probability of correctly rejecting a false null or the power
- α is the probability of a type I error
- $t_{\frac{\alpha}{2}}$ and $t_{1-\kappa}$ are the critical values of t distributions
- MDE or the mean detectable effect is the smallest effect size where an effect can still be detected if there is one
- P is the proportion of the sample that is treated

- σ^2 is the variance of the treatment effect estimator
- ρ is the intracluster correlation coefficient
- T is the length of the experiment in days

We set κ to 0.80, or 80%, and α to 0.05, or 5%, values typically used for these calculations. We use the available data to calculate the variance of the outcome variable and the intraclass correlation coefficient $\rho = 0.566$. For sensitivity, we vary σ^2 from 0.05 to 0.25. Due to budget constraints, we can assign a much larger sample size to the control group than to the treatment groups, and some treatment groups may be relatively larger than others (for example, larger-size contests versus smaller-size contests given the same total payout). Thus, for the purposes of power calculations, we also vary the proportion of the sample assigned to treatment from 10 percent to 50 percent.

Assuming an MDE of 0.01 and the low-variance scenario ($\sigma^2 = 0.05$), the required sample size per treatment group ranges from 2,474 to 4,453, while the required control group size ranges from 22,266 to 4,454 as the treatment proportion P varies from 10% to 50%. In the high-variance scenario ($\sigma^2 = 0.25$), the required sample size per treatment group increases to a range of 12,370 to 22,266, and the required control group size ranges from 111,331 to 22,266 over the same values of treatment proportions.[5]

Similarly, assuming an MDE of 0.025, in the low-variance scenario ($\sigma^2 = 0.05$), the required sample size per treatment group ranges from 396 to 713, and the required control group size ranges from 3,563 to 713 as P varies from 10% to 50%. In the high-variance scenario ($\sigma^2 = 0.25$), the required sample size per treatment group ranges from 1,979 to 3,563, while the required control group size ranges from 17,813 to 3,565 over the same treatment proportions.

8 Proposed Timeline

Table 3 provides a tentative project timeline.

9 Administrative Information

Funding: This work is supported by JPAL K-CAI (grants #KCAI-21-00330 and #KCAI-24-02463).

[5]The high-variance scenario is somewhat unlikely given the precision of our estimates in previous experiments, but we include it to remain conservative.

Table 3: Project Timeline

Task	Start date	Completion date
Randomization	May 2026	May 2026
First-round incentives	June 2026	June 2026
Second-round incentives	July 2026	July 2026
Third-round incentives	August 2026	August 2026

Institutional Review Board (ethics approval): This project was approved by IRB committees at UCSD (IRB #23047) and UBC (BREB #H22-00785).

Pre-registration: The pre-analysis plan will be registered at the AEA RCT registry before outcome data collection begins.

Acknowledgments: We are grateful to the King Climate Action Initiative of the Jameel Poverty Action Lab as well as the Deep Decarbonization Initiative at UC San Diego for generous funding to support this research.

Declaration of interest: None of the authors have any relevant or material financial interests that relate to this research project.

References

H. Allcott. Social norms and energy conservation. *Journal of Public Economics*, 95(9-10): 1082–1095, 2011.

H. Allcott and J. B. Kessler. The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, 11(1):236–276, 2019.

H. Allcott and S. Mullainathan. Behavior and energy policy. *Science*, 327(5970):1204–1205, 2010.

M. Auffhammer, C. Blumstein, and M. Fowlie. Demand-side management and energy efficiency revisited. *The Energy Journal*, 29(3):91–104, 2008.

V. Bhattacharya. An empirical model of r&d procurement contests: An analysis of the dod sbir program. *Econometrica*, 89(5):2189–2224, 2021.

S. Borenstein. The redistributional impact of nonlinear electricity pricing. *American Economic Journal: Economic Policy*, 4(3):56–90, 2012.

D. S. Callaway, M. Fowlie, and G. McCormick. Location, location, location: The variable value of renewable energy and demand-side efficiency resources. *Journal of the Association of Environmental and Resource Economists*, 5(1):39–75, 2018.

M. Fowlie, C. Wolfram, P. Baylis, C. A. Spurlock, A. Todd-Blick, and P. Cappers. Default effects and follow-on behaviour: Evidence from an electricity pricing program. *The Review of Economic Studies*, 88(6):2886–2934, 2021.

T. Garg, J. Lemus, G. Marshall, and C. Ta. A comparison of contests and contracts to deliver cost-effective energy conservation. 2025.

P. J. Gertler, O. Shelef, C. D. Wolfram, and A. Fuchs. The demand for energy-using assets among the world's rising middle classes. *American Economic Review*, 106(6):1366–1401, 2016.

J. Graff Zivin and M. Neidell. Air pollution's hidden impacts. *Science*, 359(6371):39–40, 2018.

J. R. Green and N. L. Stokey. A comparison of tournaments and contracts. *Journal of Political Economy*, 91(3):349–364, 1983.

D. P. Gross. Creativity under fire: The effects of competition on creative production. *Review of Economics and Statistics*, 102(3):583–599, 2020.

K. Ito. Asymmetric incentives in subsidies: Evidence from a large-scale electricity rebate program. *American Economic Journal: Economic Policy*, 7(3):209–237, 2015.

K. Ito, T. Ida, and M. Tanaka. Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, 10(1): 240–267, 2018.

E. P. Lazear and S. Rosen. Rank-order tournaments as optimum labor contracts. *Journal of political Economy*, 89(5):841–864, 1981.

J. Lemus and G. Marshall. Dynamic tournament design: Evidence from prediction contests. *Journal of Political Economy*, 129(2):383–420, 2021.

A. R. Neveu and M. F. Sherlock. An evaluation of tax credits for residential energy efficiency. *Eastern Economic Journal*, 42:63–79, 2016.

C. Ta. Do conservation contests work? an analysis of a large-scale energy competitive rebate program. *Journal of Environmental Economics and Management*, 124:102926, 2024.