Journal of Development Economics

Contract Design for Selection and Scale: Evidence from a City-Wide Automated Demand Response Program

--Manuscript Draft--

Manuscript Number:	DEVEC-D-26-00275R2
Article Type:	Registered Report Stage 1: Proposal
Section/Category:	Experimental Papers, credit, insurance
Keywords:	Demand response; Technology Adoption; contract design; selection; electricity markets; Field Experiment
Corresponding Author:	Zeynep Gurguc Queen Mary University of London LONDON, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND
First Author:	Teevrat Garg
Order of Authors:	Teevrat Garg
	Greer Gosnell
	Zeynep Gurguc
	Shefali Khanna
	Jorge Lemus
	Guillermo Marshall
	Ralf Martin
Abstract:	Many policy programs rely on voluntary participation, yet their effectiveness depends jointly on scale (enrollment) and selection (whether those who enroll are the ones who generate the greatest social value). When participation requires a costly investment and benefits are realized through repeated actions, expanding enrollment may dilute cost-effectiveness if marginal participants rarely follow through, while screening for high-value participants may limit aggregate impact. This paper studies how incentive design navigates this tradeoff in the context of automated residential demand response. We propose a large-scale field experiment in Mumbai that evaluates adoption and compliance incentives for an automated demand response program delivered through smart switches installed on power-intensive household appliances such as air conditioners, where program value is generated through repeated, short-duration switch-off events that households can override. Using device telemetry and high-frequency electricity consumption data, we measure program enrollment, retention, compliance, and delivered demand response, and compare cost-effectiveness across incentive designs. The experiment provides evidence on how to design voluntary programs that both attract the right participants and scale efficiently when participation decisions are shaped by private information.
Response to Reviewers:	We would like to thank the reviewers and the editor for their careful consideration of our proposal.

Dr Zeynep Gürgüç
Lecturer, School of Business
and Management
Queen Mary University of London
London E1 4NS UK
z.gurguc@qmul.ac.uk

1 July 2026

Journal of Development Economics
Registered Report Stage 1: Proposal

Dear Professor Yang,

On behalf of my co-authors, I would like to kindly ask you to consider the revised manuscript for "Contract Design for Selection and Scale: Evidence from a City-Wide Automated Demand Response Program" for the registered report proposal stage.

We are grateful to both of the reviewers for their invaluable and careful feedback, which has substantially strengthened the paper. Their comments helped us sharpen the theoretical framing, and the empirical methods. We hope the revisions meet the reviewers' satisfaction and thank them again for their time and expertise.

Kind regards,

Zeynep Gürgüç

Title page with author details

Contract Design for Selection and Scale:
Evidence from a City-Wide Automated Demand Response Program

Teevrat Garg* Greer Gosnell † Zeynep Gürgüç‡ Shefali Khanna§
Jorge Lemus¶ Guillermo Marshall‖ Ralf Martin**

July 1, 2026

Abstract

Many policy programs rely on voluntary participation, yet their effectiveness depends jointly on scale (enrollment) and selection (whether those who enroll are the ones who generate the greatest social value). When participation requires a costly investment and benefits are realized through repeated actions, expanding enrollment may dilute cost-effectiveness if marginal participants rarely follow through, while screening for high-value participants may limit aggregate impact. This paper studies how incentive design navigates this tradeoff in the context of automated residential demand response. We propose a large-scale field experiment in Mumbai that evaluates adoption and compliance incentives for an automated demand response program delivered through smart switches installed on power-intensive household appliances such as air conditioners, where program value is generated through repeated, short-duration switch-off events that households can override. Using device telemetry and high-frequency electricity consumption data, we measure program enrollment, retention, compliance, and delivered demand response, and compare cost-effectiveness across incentive designs. The experiment provides evidence on how to design voluntary programs that both attract the right participants and scale efficiently when participation decisions are shaped by private information.

Keywords: demand response, technology adoption, contract design, selection, electricity markets, field experiment

JEL Codes: L94, Q41, D12, D82

Study pre-registration: This study will be pre-registered at the AEA RCT Registry prior to the first invitation wave in July 2026. The registration ID will be provided on completion.

Proposed Timeline: Our data collection activities in Mumbai will last from July 2026 to July 2027.

*School of Global Policy and Strategy, UC San Diego; teevrat@ucsd.edu
†Giving Green; greer.gosnell@givinggreen.earth
‡School of Business and Management, Queen Mary University of London. z.gurguc@qmul.ac.uk
§London School of Economics. S.Khanna13@lse.ac.uk
¶Department of Economics, University of Illinois; jalemus@illinois.edu
‖Sauder School of Business, University of British Columbia; guillermo.marshall@sauder.ubc.ca
**Imperial College Business School, IFC, LSE and CEPR.r.martin@imperial.ac.uk

Click here to access/download
Editable source files
PAUSE_Demand_JDE_Revised_July.zip

Teevrat Garg: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Writing – original draft; Writing – review & editing.

Greer Gosnell: Conceptualization; Investigation; Methodology; Software; Writing – review & editing.

Zeynep Gurguc: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Software; Writing – original draft; Writing – review & editing.

Shefali Khanna: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Writing – review & editing.

Jorge Lemus: Conceptualization; Formal analysis; Investigation; Methodology; Writing – review & editing.

Guillermo Marshall: Conceptualization; Formal analysis; Investigation; Methodology.

Ralf Martin: Conceptualization; Formal analysis; Investigation; Methodology; Software.

Teevrat Garg, Greer Gosnell and Zeynep Gurguc are co-founders of Pow-dr Technologies Ltd., a social enterprise that developed the technology used to facilitate this research. All other authors declare no competing interests.

Contract Design for Selection and Scale:

Evidence from a City-Wide Automated Demand Response Program

July 1, 2026

Abstract

Many policy programs rely on voluntary participation, yet their effectiveness depends jointly on scale (enrollment) and selection (whether those who enroll are the ones who generate the greatest social value). When participation requires a costly investment and benefits are realized through repeated actions, expanding enrollment may dilute cost-effectiveness if marginal participants rarely follow through, while screening for high-value participants may limit aggregate impact. This paper studies how incentive design navigates this tradeoff in the context of automated residential demand response. We propose a large-scale field experiment in Mumbai that evaluates adoption and compliance incentives for an automated demand response program delivered through smart switches installed on power-intensive household appliances such as air conditioners, where program value is generated through repeated, short-duration switch-off events that households can override. Using device telemetry and high-frequency electricity consumption data, we measure program enrollment, retention, compliance, and delivered demand response, and compare cost-effectiveness across incentive designs. The experiment provides evidence on how to design voluntary programs that both attract the right participants and scale efficiently when participation decisions are shaped by private information.

Keywords: demand response, technology adoption, contract design, selection, electricity markets, field experiment

JEL Codes: L94, Q41, D12, D82

Study pre-registration: This study will be pre-registered at the AEA RCT Registry prior to the first invitation wave in July 2026. The registration ID will be provided on completion.

Proposed Timeline: Our data collection activities in Mumbai will last from July 2026 to July 2027.

1 Introduction

Many policy programs rely on voluntary enrollment and sustained participation to deliver social benefits. Whether the objective is improved health, higher educational attainment, cleaner air, or lower carbon emissions, policymakers typically confront two related design problems. The first is scale: effectively encouraging enough people to enroll so that the program meaningfully moves aggregate outcomes. The second is selection: getting the right people to enroll, in the sense that participants are those for whom the program generates the largest net social value. When participation is voluntary, enrollment choices often reflect private information about costs, constraints, and expected gains, thereby driving selection. Whether selection is favorable or adverse is an empirical question. We study this question in the context of an automated residential demand-response program, in which households permit a utility to remotely cycle selected high-consumption appliances during peak hours in exchange for enrollment and compliance incentives. Individuals most willing to enroll may face low participation costs because they can connect appliances that are infrequently used or easily interruptible (e.g. secondary or rarely operated room air conditioning units), which would imply limited marginal social benefit from participation. In contrast, households with the greatest potential to deliver peak demand reductions may perceive barriers higher due to limited attention or setup frictions, reducing their likelihood of enrolling, even though conditional on participation they may comply more fully with automated control and generate high marginal social value.

A natural implication is that program design must be evaluated not only by the average effect among those who enroll, but also by how program rules shape who enrolls and how participants behave after enrolling. In many settings, policymakers can move both margins using a combination of adoption incentives, ongoing performance incentives, and participation requirements that change the time, hassle, or perceived risk of participation. A growing empirical literature shows that incentive menus and choice architectures can shift both selection and outcomes in domains ranging from physical activity and mindfulness programs to clean fuel adoption (Abubakari et al., 2024; Adjerid et al., 2022; Dizon-Ross and Zucker, 2025; Woerner et al., 2025). Related work shows that program targeting can materially affect downstream behavior and program effectiveness (Beaman et al., 2023; Jack, 2013). These results underscore a central measurement challenge: differences in program performance across designs may reflect who enrolls, what incentives enrollees face, or both. Our design therefore builds in variation that holds realized incentives fixed while changing what households know at the enrollment stage, allowing us to separate selection effects from incentive effects.

This scale-selection tradeoff is particularly salient for demand response (DR) in electricity markets. DR programs can reduce peak demand, lower procurement costs, and reduce emissions by shifting load away from periods when supply is scarce or carbon intensive (Parrish et al., 2019). In India, these benefits are especially valuable given rapid demand growth, continued reliance on fossil fuels, and the system-level challenge of integrating variable renewable energy while maintaining reliability (Debnath et al., 2021). Yet, scaling DR to residential customers requires overcoming two practical obstacles. First, many DR approaches impose ongoing attention and decision-making burdens on households, which can limit effectiveness (Delmas et al., 2013; Jessoe and Rapson, 2014). Second, automated approaches that reduce cognitive burden often require hardware adoption and continued connectivity, and may raise concerns about privacy, control, and hassle costs (Bailey et al., 2025; Blonz et al., 2025; Coutellier et al., 2020; Khanna et al., 2025). In other words, residential DR is a setting where both scale and selection are first-order: device installation has fixed costs, and installing devices on households unlikely to contribute to ongoing program objectives wastes program resources and lowers cost-effectiveness.

This pre-results plan describes a field experiment in Mumbai in summer 2026 that tests adoption and compliance incentive designs for automated demand response delivered through smart switches installed on high power consuming appliances such as air conditioners. Working with Tata Power's communication channels, we will invite customers from an eligibility pool defined using administrative consumption data (we will exclude users in the bottom and top 5% of the consumption distribution). We will install approximately 4,500 switches (one per household) for which we will observe frequent device-level outcomes. Where available, we will pair observed device-level consumption data with high-frequency smart-meter data to determine the extent to which switch-off events result in demand leakage. The design varies enrollment bonuses, retention bonuses, and event-level compliance incentives. Additionally, it randomizes invitation coverage within feeders or substations (e.g., 50%/100% of eligible customers), allowing us to study how program performance and implementation costs change as enrollment density increases within local distribution networks.

The experiment focuses on three questions:

1. **Program enrollment and targeting:** Which incentive structures attract households most likely to generate DR value (i.e., enroll, remain online, and comply)?
2. **Selection versus incentives:** How much of observed variation in delivered demand response is due to who enrolls (selection) versus how enrolled households are incentivized to behave (incentive effects)?
3. **Scaling and saturation:** How does program performance and implementation cost change as invitation coverage increases within distribution feeders?

This experiment contributes to three strands of economics research: selection into voluntary programs, incentive design under repeated performance, and the economics of scaling policy interventions within networked infrastructure.

First, the design provides unusually clean evidence on the role of selection into voluntary programs. A central challenge in the empirical literature on screening and self-targeting is that different program designs typically change both who participates and the incentives participants face after enrollment, making it difficult to disentangle selection effects from treatment effects. In our experiment, we create paired arms in which households face identical post-enrollment incentives and event structures, but differ in what they know at the time of enrollment about future retention or performance bonuses. This structure allows differences in downstream demand response outcomes across these arms to be attributed to selection rather than differential incentives after adoption. In this respect, our design complements work on allocation under private information (Beaman et al., 2023; Jack, 2013; Oliva et al., 2020), and speaks directly to the energy literature showing that welfare gains depend critically on who opts into programs when households have private information about costs and benefits (Ito et al., 2023). We extend this line of work to a setting with two dimensions of private information: an upfront adoption cost shaped by hassle, attention, and trust, and an ongoing compliance cost shaped by the household's disutility from foregone appliance use. In Appendix A we show that independent randomization of adoption and compliance incentives, combined with randomization of information timing, separates selection on the adoption margin from selection on the joint adoption–compliance type.

Second, the experiment brings ideas from the screening and menu design literature into the context of automated residential demand response with real adoption decisions and repeated compliance opportunities. Prior experimental work shows that allowing agents to choose among incentive contracts can meaningfully alter participation and outcomes in domains such as physical activity, mindfulness, and clean fuel adoption and payments for ecosystem services (Abubakari et al., 2024; Adjerid et al., 2022; Dizon-Ross and Zucker, 2025; Jack and Jayachandran, 2019; Woerner et al., 2025). We extend this literature by studying screening in a setting where participation requires installing a physical device, program benefits accrue through repeated, low-salience compliance decisions over time, and the principal's objective is naturally expressed in terms of cost per unit of delivered service (kWh curtailed during events) rather than solely in terms of average effects among compliers. This is particularly important because program costs are incurred upfront, while benefits depend on the cumulative, realized load reductions delivered across many events, rather than on average behavioral responses among compliers. Our design allows us to assess whether screening contracts improve cost-effectiveness by reallocating scarce devices away from low-value participants, who enroll at a low cost by infrequently used appliances, and toward households with higher abatement potential that face greater enrollment frictions but exhibit stronger compliance once enrolled, rather than simply changing behavior conditional on participation.

Third, the study makes a scaling contribution. By randomizing invitation coverage at the feeder or substation level (e.g., inviting 50%, or 100% of eligible customers within a feeder) while maintaining individual-level randomization into incentive arms, the experiment generates exogenous variation in enrollment density within local distribution networks. This saturation-style design enables measurement of how aggregate demand response, enrollment rates, and implementation costs evolve as outreach scales, and whether program benefits or costs exhibit non-linearities with respect to coverage. The approach is closely related in spirit to experimental work on electrification that uses large-scale rollout variation to study impacts and implementation at policy-relevant scale (Lee et al., 2020), but applied here to the deployment of demand-side flexibility. Beyond these individual contributions, a further contribution of this paper is to study selection, compliance, and scale jointly within a single experimental framework, enabling a decomposition of program cost-effectiveness into its margin-specific drivers that no subset of the literature, studied in isolation, can deliver.

Relevance to Development Economics

Improving the efficacy and reliability of electricity grids is of central interest to development economics, and many papers published in this Journal study precisely this question—from the industrial consequences of electricity provision and outages (Cole et al., 2018; Fisher-Vanden et al., 2015; Kassem, 2024; Mensah, 2024; Rud, 2012) to the labor-market and structural-change effects of rural electrification (Fetter and Usmani, 2024). Our paper sits within this literature and addresses a policy lever that has become increasingly central to grid performance in rapidly-urbanizing developing economies: the design of voluntary demand-response programs, which utilities rely on to manage peak load in the absence of cost-reflective pricing. India's government planning body has explicitly flagged this as a priority: NITI Aayog (2026) identify dynamic tariffs and demand-response protocols as core infrastructure for a modern grid, and note that consumer engagement with demand response remains weak because households and businesses underestimate their influence on consumption patterns.

Our contribution also speaks to a broader class of development-economics settings in which adoption of a welfare-improving technology is free or heavily subsidized but subsequent *use* remains a voluntary behavioral margin. This pattern recurs across canonical applications in public health and environment, for example, insecticide-treated bed nets (Cohen and Dupas, 2010; Dupas, 2014), point-of-use water treatment (Ashraf et al., 2010), and clean cookstoves (Berkouwer and Dean, 2022), where a central open question is whether subsidies that raise adoption also sustain the post-adoption usage on which welfare gains depend. Our experimental design is directly portable to those contexts: by separating the selection induced by enrollment incentives from the selection induced by use incentives, and by varying the timing at which information about future use incentives is revealed, the design isolates the two margins that the existing take-up literature typically conflates. The high-frequency telemetric outcome data available in our setting further provide an unusually clean measure of post-adoption behavior, one rarely observed in classic take-up studies.

Several structural features of the developing-country context shape the mechanisms we study. Residential tariffs are typically cross-subsidized and disconnected from marginal supply costs (McRae, 2015), meaning that price signals alone do not elicit peak-shifting behavior and that voluntary incentive-based programs become a first-order policy instrument rather than a supplemental one. Supply itself is chronically constrained: Indian households face frequent outages that impose large welfare costs (Allcott et al., 2016; Khanna and Rowe, 2024), so the social value of peak-load reductions is substantially higher than in settings where the counterfactual to DR is merely higher wholesale prices. And baseline trust between utilities and residential customers is lower than in the settings that dominate the existing DR literature, with implications for which households self-select into remote-actuation programs and how they respond to information about future incentives. Each of these features creates scope for selection and incentive responses that would be absent or muted in a comparable program in a high-income setting.

2 Context and Setting

2.1 Geography and utility context

The study will be conducted in Mumbai in partnership with Tata Power, one of the city's primary electricity distributors. Tata Power's service territory contains 179 primarily residential distribution feeders encompassing 1054 substations serving 212,138 residential customers.

2.2 Technology

The intervention uses smart switches installed by licensed electricians on high power consuming household appliances, such as air conditioners (ACs). Each participating household receives one smart switch installed on the selected appliance. The switch enables:

- remote switch-off events of 30 minutes duration;
- device telemetry (online/offline status, switch control logs, and measurement of energy consumption);
- a household-facing app interface for notifications and override.

2.3 Data environment

We will obtain:

- **Smart meter data at 30 minute intervals** for the subset of customers with smart meters (approximately 74%)
- **Billing-cycle (monthly) consumption** for all customers;
- Switch telemetry, energy consumption via the switch and event logs for all installed households.

3 Research Design

3.1 Overview

The experiment includes three layers of randomization:

- **Feeder-level invitation intensity randomization (saturation):** Each feeder is assigned to an invitation coverage level: 50%, or 100% of eligible customers invited.
- **Household-level randomization to incentive arms:** Among invited households, random assignment to one of seven main arms (plus within-arm information timing sub-randomizations, described below).

This method of randomization allows us to examine the effects of randomized upfront program incentive information on whether households enroll (i.e. opt into the program and undergo smart switch installation), as well as whether they stay connected for at least two and eight weeks. Additionally, we will examine the effects of incentives and selection on the override rate for switch-off events. We will also translate this into total DR value based on when and how much electricity reduction is achieved. Finally, we will analyze spillovers, where we will track whether neighbors express an interest in signing up for alerts for next year's program one month, three months, and six months after their neighbors have enrolled in the program.

3.2 Eligibility and sampling frame

Eligibility. We first exclude very low-usage and very high-usage customers (i.e., customers for whom the fixed costs of installation are unlikely to be cost-effective, and customers likely outside the study's operational reach) by removing consumer substations that only serve commercial entities or serve less than 0.5% of maximum number of households served by a consumer substation. We will further finetune eligibility criteria prior to the experiment using 12 months of pre-period utility data. We will document the final thresholds in the public pre-analysis plan before treatment rollout.

Sampling frame. The sampling frame consists of residential customers in the utility's Mumbai service area who:

- meet the (to-be-finalized) consumption-based eligibility criteria using the prior 12 months;
- have an appliance (such as AC) suitable for installation (verified during enrollment/installation);
- can support home WiFi connectivity sufficient for the device to remain online.

3.3 Feeder-level invitation intensity (saturation)

Let f index feeders. Each feeder is randomly assigned to an invitation saturation level $s_f \in \{0.50, 1\}$, stratified by substation (and other feeder characteristics if available, e.g., baseline load, number of households, or smart meter penetration). Within feeder f, we randomly select a fraction s_f of eligible customers to receive invitations.

This "randomized saturation" layer is designed to enable:

- measurement of spillovers and peer effects;
- estimation of how operational costs scale with program intensity;
- estimation of feeder-level impacts on aggregate load (where data permit).

3.4 Household-level randomization

Randomization into incentive arms occurs at the individual household level. Household-level assignment will be stratified by feeder/substation and baseline household consumption bins, ensuring balance across arms within operationally meaningful strata.

3.5 Recruitment and enrollment process

Communication channels. Invitations will be delivered using Tata Power's customer communication tools (e.g., SMS/WhatsApp/email/app notifications).

Batching. Invitations will be sent in batches (tentatively 20,000 households per wave) to manage budgets and installation capacity. Households will have up to 2–3 weeks to respond. Non-responders are not reassigned; subsequent waves draw new households consistent with the feeder's saturation assignment.

Unique enrollment codes. Each invited household receives a unique enrollment code matched to a customer ID, preventing cross-household code sharing.

Adaptive invitation volumes (implementation rule). We will begin with equal invitation counts per incentive arm. If take-up differs substantially across arms, we will adjust the number of invitations sent to each arm in subsequent waves using a pre-specified rule based only on observed enrollment/installation rates (not on consumption or outcomes). The objective is to reach the target of 4,500 installations overall, while maintaining adequate sample sizes in each arm for analysis. The exact adaptive rule will be documented prior to rollout (e.g., arm-level invitation weights proportional to the inverse of the cumulative installation rate).

3.6 Installation and baseline measurement

Installation. Enrolled households receive one smart switch installed by a licensed electrician on an appliance of their choice.

Baseline survey. We will administer a short in-person survey at installation. Survey content will include basic household demographics, appliance information, and self-reported preferences about comfort and scheduling.

3.7 Demand response events and schedules

Event frequency and timing. We will schedule DR events subject to:

- **Event duration:** 30 minutes.
- **Event window:** 8am–midnight.
- **Frequency caps:** no more than 2 events/day and no more than 10 events/week.

Predictable rotating schedule (default): All households begin on a predictable rotating schedule. Households are informed one day in advance of predictable schedule events. Event timings will be randomized across households.

Opt-out and overrides: Households can opt out on a per-event basis. The primary measure of compliance is whether an event is not overridden. Households may opt out via the app; they may reconnect and re-enter the program at any time.

4 Treatment Arms

4.1 Notation and general principles

All incentive amounts are denominated in INR and parameterized here:

- A: adoption/activation incentive (paid after meeting a 2-week online requirement);
- R: additional retention bonus (paid after meeting an 8-week online requirement);
- p: piece-rate per non-overridden event;
- B: weekly bonus for meeting compliance threshold;

Final numerical values will be set prior to experiment launch depending on the wholesale electricity price in the market.

In Appendix A, we provide a theoretical framework for our experimental design.

4.2 Online requirements for adoption/retention payments

For arms with adoption/retention incentives:

- **2-week activation criterion:** device online >22 hours/day for at least 12 out of 14 days after installation.
- **8-week retention criterion:** device online >22 hours/day for at least 50 out of 56 days after installation.

4.3 Incentive arms and within-arm information timing

Table 1 summarizes the incentive arms. Several arms include within-arm sub-randomizations that vary whether an additional bonus is *revealed at sign-up* versus *revealed only after enrollment*. These sub-randomizations are central for distinguishing selection from incentive effects.

Arm	Name	Description
1	Control (Free installation; no incentives)	Free installation. DR events occur. No adoption, retention, or event-compliance payments.
2	Adoption incentive only (baseline)	Earn A if activation criterion is met (2-week online requirement). DR events occur, but no payments tied to compliance.
2b	Late-reveal retention offer (within-arm)	A randomized subset of Arm 2 receives an *additional* offer after enrollment: earn R if the 8-week retention criterion is met. This offer is not disclosed at sign-up.
3	Adoption + known retention bonus	Earn A if activation criterion is met, and earn R if the 8-week retention criterion is met. Both are disclosed at sign-up.
4	Piece-rate only (baseline)	Earn p for each non-overridden DR event. No adoption/retention payments.
4b	Late-reveal weekly bonus (within-arm)	A randomized subset of Arm 4 receives an additional offer after enrollment: earn B in weeks where non-overrides ≥ 0.8 of scheduled events. Offer not disclosed at sign-up.
5	Piece-rate + known weekly bonus	Earn p per non-overridden event and earn B in weeks where non-overrides ≥ 0.8 of scheduled events. Disclosed at sign-up.
6	Adoption incentive + piece-rate	Earn A if activation criterion is met and earn p per non-overridden event.

Table 1: Treatment arms (incentive amounts in INR; values set pre-launch). All arms receive DR events; incentives vary by arm.

4.4 Primary design contrasts

Key comparisons are organized into four pre-specified hypothesis domains.

Adoption domain: We test the effect of adoption incentives by comparing Arm 2 vs Arm 1 and Arm 6 vs Arm 4. Both contrasts isolate the effect of the adoption incentive A.

Retention domain: We test the effect of known retention bonuses (selection) Arm 3 vs Arm 2b (effect of known retention bonus R) and Arm 5 vs Arm 4b (effect of known weekly bonus B). Both contrasts isolate the selection effect of revealing bonus at sign-up rather than post-enrollment.

Compliance and DR domains: We test the effect of piece rate incentives on ongoing compliance or DR value by comparing Arm 4 vs Arm 1 and Arm 6 vs Arm 2. Both contrasts isolate the effect of the piece-rate incentive p, the first against pure control, and the second on top of an adoption incentive.

5 Outcomes and Measurement

5.1 Primary outcomes

Enrollment.

- Responds and consents.
- Installation completed (switch installed on the appliance).

Retention.

- Online minutes per day; indicator for meeting the activation (2-week) and retention (8-week) criteria.
- Time-to-offline (survival-type metrics).

Compliance.

- Event-level indicator: non-overridden event.
- Weekly compliance rate; indicator for meeting the weekly threshold 0.8.
- Opt-out behavior (app actions; WiFi disconnects).

DR value.

- Event-level kWh reduced during the 30-minute event window.
- Peak load reduction (aggregate and household-level).

5.2 Exploratory outcomes

- Spillovers and peer effects: neighbors' interest in the program at 1, 3 and 6 months post-enrollment.
- Total electricity consumption: measured daily for smart-metered households; monthly for all).

- Load shifting: evidence of demand leakage to pre/post event windows, where data allow.
- Engagement metrics: message opens/clicks, app logins, notification interactions.
- Program costs: installation, incentives paid, and outreach costs: cost-effectiveness measured as INR per kWh reduced.
- Grid reliability indicators: outage incidence and feeder-level demand.

6 Data

6.1 Administrative and telemetry data

We will compile:

- Customer roster: feeder/substation, contact info availability, tariffs (if relevant).
- Billing data for all customers (monthly kWh at minimum), for at least 12 months pre-period.
- Smart meter data (30-minute intervals) for smart-metered customers, for at least 12 months pre-period through the experiment.
- Switch telemetry: online/offline, control actions, event execution logs.
- Event schedule logs: time, lead time, window type (predictable vs individualized), and override indicators.

6.2 Weather and calendar

We will merge local weather (temperature, humidity) and calendar controls (day-of-week, holidays).

6.3 Survey data

We will collect brief household and appliance information at installation (in person).

6.4 Data quality

Device connectivity is central to both compliance measurement and outcome data quality. We classify offline periods using a three-way rule based on 30-minute heartbeat signals logged by the device firmware. A device that goes offline during a scheduled DR event is coded as non-compliant. A device that goes offline outside DR event windows is flagged but excluded from compliance calculations. Grid outages are identified using feeder-level SCADA data from Tata Power, and all device-hours affected by confirmed outages are excluded entirely from the analysis. For smart meter data, we remove outliers above the 95th percentile of half-hourly consumption as standard practice.

The primary data quality challenge identified in prior work with these devices is intermittent Wi-Fi connectivity, which affected a subset of pilot households. Households will be alerted via SMS and email if their device remains offline continuously for 48 hours or longer, and will be guided on how to reconnect if they wish to continue participating. A separate API malfunction in July 2023 caused a temporary gap in smart switch data during the pilot, which is documented in Khanna et al. (2024). When devices are connected, however, the switch telemetry provides highly accurate, continuous measurement of appliance-level energy consumption, logged at 5-minute intervals directly from the device. The pilot also confirmed that device-level and meter-level consumption estimates are closely aligned — the smart switch accounts for approximately 13% of total household consumption on average — which supports using device telemetry as the primary compliance measure even for households without smart meters. We additionally pre-specify a robustness check that restricts the estimation sample to households with device connectivity above 70%.

7 Empirical Strategy

7.1 Notation

Let i index households, $f(i)$ their feeder, and t time (days or 5-minute intervals). Let Z_i denote assigned treatment arm. Let S_f denote feeder invitation saturation.

7.2 Adoption and installation

We estimate intention-to-treat (ITT) effects among invited households:

$$Adopt_i = \alpha + \sum_{k \neq 1} \beta_k \cdot \mathbf{1}\{Z_i = k\} + \gamma_{strata(i)} + \varepsilon_i, \tag{1}$$

where Arm 1 (control) is omitted, and $\gamma_{strata(i)}$ includes feeder/substation strata fixed effects. Standard errors are clustered at feeder level, reflecting the potential for correlated shocks within feeders due to shared infrastructure and local weather conditions. We report heteroskedasticity-robust standard errors as a robustness check.

7.3 Retention

We will estimate analogous models for meeting the 2-week and 8-week online criteria, and for continuous measures (e.g., online hours/day).

7.4 Compliance and DR value

For event-level outcomes, we estimate:

$$Y_{ie} = \alpha + \sum_{k \neq 1} \beta_k \cdot \mathbf{1}\{Z_i = k\} + \delta_e + \theta_i + \varepsilon_{ie}, \tag{2}$$

where e indexes events, Y_{ie} is either (i) non-override, or (ii) event-level kWh reduction, δ_e includes event-date/time fixed effects and θ_i includes individual fixed effects. Standard errors clustered at the household level. For high-frequency (5-minute) consumption, we will also estimate event-study specifications around event start and end times.

7.5 Selection versus incentive effects via information timing

We will explicitly compare pairs with identical ex post incentive schedules but different information timing:

- Arm 3 vs Arm 2b (known vs late-reveal retention bonus).
- Arm 5 vs Arm 4b (known vs late-reveal weekly bonus).

These contrasts identify how *anticipation at enrollment* changes take-up and downstream outcomes.

7.6 Feeder-level saturation effects

We will analyze feeder-level outcomes (e.g., number enrolled, number installed, total kWh reduced in feeder) as a function of saturation assignment:

$$\overline{Y}_f = \alpha + \pi_{50}\mathbf{1}\{S_f = 0.50\} + X_f'\Gamma + u_f, \tag{3}$$

with standard errors clustered at the feeder level (or randomization-inference p-values as a robustness check).

7.7 Multiple hypothesis testing

We will define families for (i) adoption/installation, (ii) retention, (iii) compliance, and (iv) DR value. We will report both unadjusted p-values and family-wise adjusted p-values (Holm), specified in the final pre-analysis plan.

8 Power Calculations

The experiment will aim to install 4,500 switches (one per household). Given the high-frequency (5-minute) data via the switch is available for a majority of participants, we expect relatively precise estimates for event-level load reductions and compliance. We look at minimum detectable effects for initial enrollment rate, the retention rates at 2 weeks and 8 weeks, and event-level kWh reduction among smart-metered households.

Assuming a very low initial adoption of the program in the population, we assess whether different scenarios for the increase in enrollment rate (e.g. 0.1% to 1%,2% and 5%, or 1% to 2% and 5%) are detectable with at least power of 80%. We focus on two primary adoption contrasts that isolate the effect of adoption incentives (Arm 2 versus Arm 1 and Arm 6 versus Arm 4). The same approach is valid to investigate the effect of known retention/weekly bonus (Arm 3 vs Arm 2b and Arm 5 vs Arm 4b), as well as piece rate incentives (Arm 4 vs Arm 1 and Arm 6 vs Arm 2). To account for multiple hypothesis tests within each set of two primary comparisons, we apply a Bonferroni correction and use a one-sided significance level of 0.025 for each test.[1] We also check MDE and power given a sample size of 500 per comparison group (Figure 1).

alpha	power	# Baseline	# Comparison	delta	p1	p2
.025	.8	1,088	1,088	.009	.001	.01
.025	.8	459	459	.019	.001	.02
.025	.8	167	167	.049	.001	.05
.025	.8	2,456	2,456	.01	.01	.02
.025	.8	302	302	.04	.01	.05

Table 2: Enrollment rates - sample size analysis

Our previous findings indicate a 30% dropout rate, hence 70% retention; therefore, we also check MDE and power given a sample size of 500 per comparison group (Figure 2). Main comparisons for this analysis include checking the effect of retention bonus (Arm 3 and Arm 2b vs Arm 2), as well as checking the effect of weekly bonus (Arm 5 vs 4, and Arm 4b vs 4). To account for multiple hypothesis tests within each set of two primary comparisons, we again apply a Bonferroni correction and use a one-sided significance level of 0.025, and find that we are able to detect approximately 10% change from the baseline retention rate.

Finally, given the results of our previous study with smart switches which indicate a 69% reduction in device-level electricity use (Khanna et al., 2025), we seek to detect a reduction (per household device in our treatment groups compared to our control groups) of at least

[1] We use a Fisher's exact test given the near zero values of potential enrollment rates.

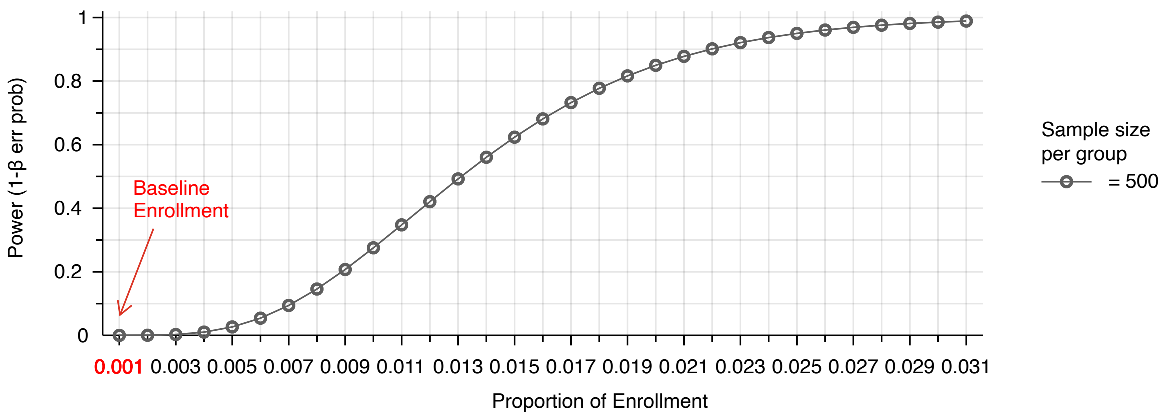

Figure 1: Minimum detectable effect for enrollment

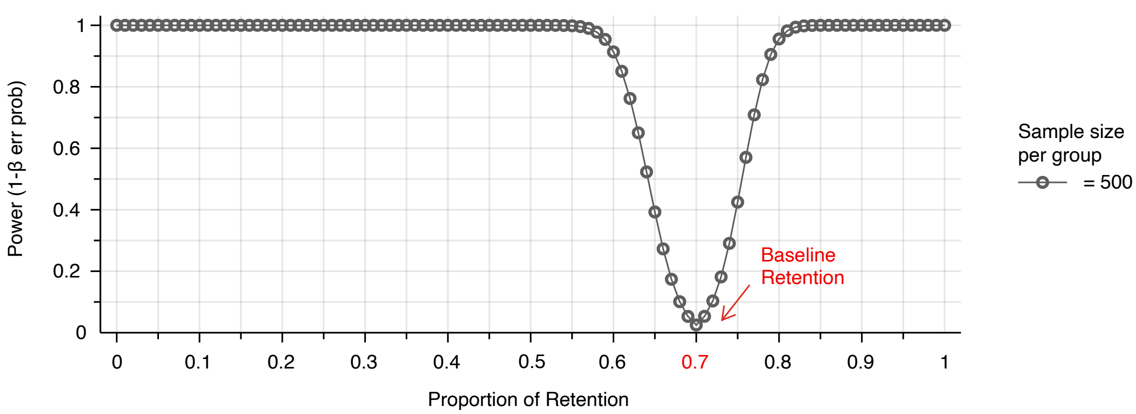

Figure 2: Minimum detectable effect for retention

50%. The standard deviation in device-level electricity consumption in our pilots is 180Wh. We choose 0.05 significance level and a power of at least 80%. We adjust our desired significance level of 0.05 for running two simultaneous hypothesis tests by compensating with Bonferroni correction and utilize 0.025. We use a on e-sided test for the effect of incentives compared to their corresponding control group and find a sample size of 567 per group.

9 Timeline

- **Pre-experiment data assembly:** at least 12 months prior to first invitation wave.
- **Invitation waves:** rolling waves, 2–3 week response windows.
- **Installations:** ongoing until 4,500 switches installed.
- **DR events:** begin after installation; continue for at least 4 months (potentially longer).

10 Administrative Information

Ethics. The study has received ethics approval from UCSD, and IFMR. Amendments to existing IRB approvals are currently ongoing to accommodate the latest design. Participants can opt out of any event and may exit the study at any time.

Funding. This work is supported by Solutions and Advancements through Research for Water and Air (SARWA.GR.2025.05), London School of Economics and Political Science (GRSF 112552.N.010.2332.1043), International Growth Centre (IND-25289), Private Enterprise Development in Low-Income Countries, Centre for Economic Policy and Research (ESGC01 101006).

Pre-registration. We will register a detailed pre-analysis plan before outcome data are analyzed.

Data security. All administrative and telemetry data will be stored on secure servers with access restricted to authorized researchers.

Acknowledgements. We are grateful to Ankit Kumar, Srinish Muthukrishnan, and Utkarsh Agarwal for their research support on this project.

References

S. Abubakari, K. Asante, M. Daouda, B. K. Jack, D. Jack, F. Malagutti, and P. Oliva. Targeting subsidies through price menus: Menu design and evidence from clean fuels. Technical report, Working paper, 2024.

I. Adjerid, G. Loewenstein, R. Purta, and A. Striegel. Gain-loss incentives and physical activity: the role of choice and wearable health tools. *Management Science*, 68(4):2642–2667, 2022.

H. Allcott, A. Collard-Wexler, and S. D. O'Connell. How do electricity shortages affect industry? Evidence from India. *American Economic Review*, 106(3):587–624, 2016. doi: 10.1257/aer.20140389.

N. Ashraf, J. Berry, and J. M. Shapiro. Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review*, 100(5):2383–2413, 2010. doi: 10.1257/aer.100.5.2383.

M. R. Bailey, D. P. Brown, B. C. Shaffer, and F. A. Wolak. Take the load off: Time and technology as determinants of electricity demand response. Technical report, National Bureau of Economic Research, 2025.

L. Beaman, D. Karlan, B. Thuysbaert, and C. Udry. Selection into credit markets: Evidence from agriculture in mali. *Econometrica*, 91(5):1595–1627, 2023.

S. B. Berkouwer and J. T. Dean. Credit, attention, and externalities in the adoption of energy efficient technologies by low-income households. *American Economic Review*, 112 (10):3291–3330, 2022. doi: 10.1257/aer.20210766.

J. Blonz, K. Palmer, C. J. Wichman, and D. C. Wietelman. Smart thermostats, automation, and time-varying prices. *American Economic Journal: Applied Economics*, 17(1):90–125, 2025.

J. Cohen and P. Dupas. Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *Quarterly Journal of Economics*, 125(1):1–45, 2010. doi: 10.1162/qjec.2010.125.1.1.

M. A. Cole, R. J. R. Elliott, G. Occhiali, and E. Strobl. Power outages and firm performance in Sub-Saharan Africa. *Journal of Development Economics*, 134:150–159, 2018. doi: 10.1016/j.jdeveco.2018.05.003.

Q. Coutellier, G. Gosnell, Z. Gürgüç, R. Martin, and M. Muûls. *Consumer-driven Virtual Power Plants: A Field Experiment on the Adoption and Use of a Prosocial Technology*. Grantham Research Institute on Climate Change and the Environment, 2020.

R. Debnath, V. Mittal, and A. Jindal. A review of challenges from increasing renewable generation in the indian power sector: Way forward for electricity (amendment) bill 2020. *Energy & Environment*, page 0958305X20986246, 2021.

M. A. Delmas, M. Fischlein, and O. I. Asensio. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy*, 61: 729–739, 2013.

R. Dizon-Ross and A. D. Zucker. Mechanism design for personalized policy: A field experiment incentivizing exercise. Technical report, National Bureau of Economic Research, 2025.

P. Dupas. Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica*, 82(1):197–228, 2014. doi: 10.3982/ECTA9508.

T. R. Fetter and F. Usmani. Fracking, farmers, and rural electrification in India. *Journal of Development Economics*, 170:103308, 2024. doi: 10.1016/j.jdeveco.2024.103308.

K. Fisher-Vanden, E. T. Mansur, and Q. J. Wang. Electricity shortages and firm productivity: Evidence from China's industrial firms. *Journal of Development Economics*, 114:172–188, 2015. doi: 10.1016/j.jdeveco.2015.01.002.

K. Ito, T. Ida, and M. Tanaka. Selection on welfare gains: Experimental evidence from electricity plan choice. *American Economic Review*, 113(11):2937–2973, 2023.

B. K. Jack. Private information and the allocation of land use subsidies in malawi. *American Economic Journal: Applied Economics*, 5(3):113–135, 2013.

B. K. Jack and S. Jayachandran. Self-selection into payments for ecosystem services programs. *Proceedings of the National Academy of Sciences*, 116(12):5326–5333, 2019.

K. Jessoe and D. Rapson. Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review*, 104(4):1417–1438, 2014.

D. Kassem. Does electrification cause industrial development? Grid expansion and firm turnover in Indonesia. *Journal of Development Economics*, 167:103234, 2024. doi: 10.1016/j.jdeveco.2023.103234.

S. Khanna and K. Rowe. The long-run value of electricity reliability in India. *Resource and Energy Economics*, 77:101425, 2024. doi: 10.1016/j.reseneeco.2024.101425.

S. Khanna, R. Martin, and M. Muûls. Building virtual power plants: Incentives and automation for demand-side flexibility. *Available at SSRN 5089649*, 2025.

K. Lee, E. Miguel, and C. Wolfram. Experimental evidence on the economics of rural electrification. *Journal of Political Economy*, 128(4):1523–1565, 2020.

S. McRae. Infrastructure quality and the subsidy trap. *American Economic Review*, 105(1): 35–66, 2015. doi: 10.1257/aer.20110572.

J. T. Mensah. Jobs! Electricity shortages and unemployment in Africa. *Journal of Development Economics*, 167:103231, 2024. doi: 10.1016/j.jdeveco.2023.103231.

NITI Aayog. A study report on scenarios towards Viksit Bharat and Net Zero: An overview (Vol. 1). Technical report, Government of India, New Delhi, 2026. URL https://niti.gov.in/publications/division-reports.

P. Oliva, B. K. Jack, S. Bell, E. Mettetal, and C. Severen. Technology adoption under uncertainty: Take-up and subsequent investment in zambia. *Review of Economics and Statistics*, 102(3):617–632, 2020.

B. Parrish, R. Gross, and P. Heptonstall. On demand: Can demand response live up to expectations in managing electricity systems? *Energy Research & Social Science*, 51: 107–118, 2019.

J. P. Rud. Electricity provision and industrial development: Evidence from India. *Journal of Development Economics*, 97(2):352–367, 2012. doi: 10.1016/j.jdeveco.2011.06.010.

A. Woerner, G. Romagnoli, B. M. Probst, N. Bartmann, J. N. Cloughesy, and J. W. Lindemans. Should individuals choose their own incentives? evidence from a mindfulness meditation intervention. *Management Science*, 71(8):7056–7070, 2025.

A Theoretical Framework

We provide a simple model that rationalizes the treatment arms and addresses three issues: (1) how different contracts affect *scale* (enrollment), (2) how they affect *selection* (which households enroll), and (3) why the *timing* of information about future rewards isolates these two effects.

Households are heterogeneous in two dimensions: an *upfront adoption cost* a, which captures time, hassle, scheduling an electrician, app setup, WiFi setup, privacy concerns, and any inconvenience from allowing remote control (interpreted as a *net* cost, so that $a < 0$ is possible for households who value the free device itself); and an *ongoing compliance cost* c, capturing the disutility from not overriding events and maintaining device connectivity. There is a unit mass of households, and household i has private type $\theta_i = (a_i, c_i)$ drawn from a joint distribution $F(a, c)$ with density $f(a, c)$.

The utility wants to install devices on households with low c, because these households are more likely to comply with demand response (DR) events. But households also differ in a, so contracts that expand enrollment may either improve or worsen the composition of enrollees. The experimental arms are designed to learn which contract best balances the tradeoff between scale and selection.

A household that is enrolled and installed may: (a) keep the device online long enough to satisfy the *activation* criterion (2 weeks), (b) keep the device online longer to satisfy the *retention* criterion (8 weeks), and (c) comply with scheduled DR events by not overriding them. There are W weeks in the relevant post-installation horizon and n scheduled DR events per week. A compliant event generates a reduction of ℓ kWh, valued by the utility at v per kWh, so each compliant event generates utility value $v\ell$. The utility pays a fixed installation cost K for each installed household.

A contract is defined as $m = (A_m, R_m, p_m, B_m, d_m^R, d_m^B)$, where $A_m \geq 0$ is the activation payment, $R_m \geq 0$ is the retention bonus, $p_m \geq 0$ is the piece-rate per non-overridden event, $B_m \geq 0$ is a weekly threshold bonus, $d_m^R \in \{0, 1\}$ indicates whether the retention bonus is disclosed at sign-up, and $d_m^B \in \{0, 1\}$ indicates whether the weekly threshold bonus is disclosed at sign-up.

Online milestones. To earn the activation payment A, the household must remain online long enough to satisfy the 2-week criterion, requiring ongoing effort $\kappa_A c$ with $\kappa_A > 0$. To earn the retention bonus R, the household must satisfy the 8-week criterion, requiring additional effort $\kappa_R c$ with $\kappa_R > \kappa_A$. The net values of these milestone payments are:

$$V_A(c, A) = \max\{A - \kappa_A c,\ 0\}, \qquad V_R(c, R) = \max\{R - \kappa_R c,\ 0\}.$$

Both expressions are decreasing in c, capturing the idea that households with lower ongoing compliance costs are more likely to qualify for milestone payments.

Weekly event compliance. Within a week, the household chooses how many events to comply with. Let $x \in \{0, \ldots, n\}$ denote the number of non-overridden events in a week. The household's weekly utility from event-related incentives is:

$$u(x, c, p, B) = x(p - c) + B \cdot \mathbf{1}\{x \geq \tau n\},$$

where $\tau \in (0, 1)$ is the threshold share required for the weekly bonus (set empirically to 0.8). Because the objective is linear in x except for the threshold bonus, the optimum is attained at one of three points: $x \in \{0,\ \tau n,\ n\}$. Hence the weekly event utility is:

$$V_E(c, p, B) = \max\{0,\ \tau n(p - c) + B,\ n(p - c) + B\},$$

and the corresponding number of compliant events per week is:

$$q(c, p, B) = \begin{cases} n, & \text{if } p \geq c, \\ \tau n, & \text{if } p < c \text{ and } \tau n(p - c) + B \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

This structure formalizes the intended distinction between piece-rates (which reward every compliant event) and weekly bonuses (which create bunching incentives around the weekly threshold).

B Timing and the Household Problem

The experiment's key identification logic comes from the timing of information. At **Stage 0** (sign-up), the household observes the rewards disclosed at sign-up and decides whether to enroll. At **Stage 1** (installation), if the household enrolls, the utility installs the device and the adoption cost a becomes sunk. At **Stage 2** (late-reveal), in late-reveal subarms the household learns about additional rewards after enrollment. At **Stage 3** (behavior), the household chooses whether to remain online and how much to comply with DR events. This timing implies that rewards disclosed only after installation cannot affect selection into enrollment—they can only affect post-enrollment behavior.

The household's value from signing up under contract m is:

$$S_m(c) = V_A(c, A_m) + d_m^R\, V_R(c, R_m) + W\, V_E(c, p_m,\, d_m^B B_m).$$

We assume that late-reveals are surprises to the household, so they do not form expectations about the possibility of late reveals. A household enrolls if and only if $a \leq S_m(c)$, so the enrollment set is

$$\mathcal{H}_m = \{(a, c) : a \leq S_m(c)\}.$$

Once enrolled and installed, the relevant continuation value is:

$$P_m(c) = V_A(c, A_m) + V_R(c, R_m) + W\, V_E(c, p_m, B_m),$$

where rewards absent from a given arm are set to zero. Note that $S_m(c)$ determines *who enrolls*, while $P_m(c)$ determines what *enrolled households* do.

C Principal's Objective

The utility values delivered DR but pays installation and incentive costs. For a household of type (a, c) enrolled under arm m, expected total compliant events are $Q_m(c) = W\, q(c, p_m, B_m)$, and the expected net utility payoff is:

$$\Pi_m = \int_{\mathcal{H}_m} [v\ell\, Q_m(c) - K - T_m(c)]\, dF(a, c),$$

where $T_m(c)$ captures the expected transfers paid to that household (adoption, retention, and event payments). The primary cost-effectiveness objective evaluated by the principal is expected INR per delivered kWh:

$$CE_m = \frac{\displaystyle\int_{\mathcal{H}_m} [K + T_m(c)]\, dF(a, c)}{\displaystyle\int_{\mathcal{H}_m} \ell\, Q_m(c)\, dF(a, c)}.$$

We can empirically measure both the numerator and denominator based on the households in each contract.

We now explain how the treatment arms in the RCT map into the model, to understand the different incentives created by each arm. For compactness, define:

$$\phi_A(c) = V_A(c, A), \quad \phi_R(c) = V_R(c, R), \quad \phi_P(c) = WV_E(c, p, 0), \quad \phi_{PB}(c) = WV_E(c, p, B).$$

The sign-up and post-enrollment values for each experimental arm then map as follows:

Arm	Description	Sign-up value, $S_m(c)$	Post-enrollment value, $P_m(c)$
1	Control	0	0
2	Adoption incentive only	$\phi_A(c)$	$\phi_A(c)$
2b	Late-reveal retention offer	$\phi_A(c)$	$\phi_A(c) + \phi_R(c)$
3	Adoption + known retention bonus	$\phi_A(c) + \phi_R(c)$	$\phi_A(c) + \phi_R(c)$
4	Piece-rate only	$\phi_P(c)$	$\phi_P(c)$
4b	Late-reveal weekly bonus	$\phi_P(c)$	$\phi_{PB}(c)$
5	Piece-rate + known weekly bonus	$\phi_{PB}(c)$	$\phi_{PB}(c)$
6	Adoption incentive + piece-rate	$\phi_A(c) + \phi_P(c)$	$\phi_A(c) + \phi_P(c)$

Table 3: Mapping the Model to the Treatment Arms

Arms 2 and 2b share the exact same sign-up value and thus induce identical selection. Any difference in their downstream retention is a *pure post-enrollment incentive effect*. The same holds for Arms 4 and 4b regarding weekly bonuses. Conversely, Arms 2b and 3 share the same post-enrollment value, so any difference in enrollment composition or compliance between them originates strictly from *selection at sign-up*. The same logic applies to Arms 4b and 5.

Testable Predictions. The model yields some predictions that map directly to the primary design contrasts evaluated in the RCT.

Prediction 1 (Selection via Information Timing): Suppose $V_R(c, R)$ and $V_E(c, p, B) - V_E(c, p, 0)$ are weakly decreasing in c, which holds by construction. Then: (a) revealing the retention bonus at sign-up (Arm 3 vs. 2b) weakly shifts enrollment toward lower-c households; and (b) revealing the weekly bonus at sign-up (Arm 5 vs. 4b) weakly shifts enrollment toward lower-c households. *Intuition:* Low-c households expect to monetize future rewards more easily, so disclosing those rewards before sign-up raises their ex-ante willingness to enroll, improving the quality of the selected participant pool.

Prediction 2 (The Scale-Selection Tradeoff of Upfront Subsidies): An increase in the activation payment A (comparing Arm 2 to Arm 1, or Arm 6 to Arm 4) expands enrollment, but the effect on the average compliance cost among enrollees is ambiguous and depends on the joint distribution $F(a, c)$. If low-a and low-c types are positively associated, adoption subsidies improve both scale and selection. If low-a and high-c types are positively associated—for example, because eager adopters tend to connect secondary, infrequently used appliances—adoption subsidies increase scale but worsen selection, potentially diluting cost-effectiveness CE_m.