# Scaling Low-Cost Digital Interventions: Lessons from an Energy Conservation Experiment*

Teevrat Garg† Minkyong Ko‡ Eric Lewis§ Chi Ta¶

December 26, 2025

**Abstract**

Digital tools hold promise for scaling energy conservation by giving households real-time information about their electricity use and costs. Yet whether such app-based interventions can meaningfully reduce consumption depends on users' engagement. We conduct a natural field experiment on a random sample of 45,000 electricity customers in Hanoi, Vietnam, that tested two mobile-app interventions built on the utility's smart-meter platform. One treatment ("price salience") displayed each household's current marginal price tier and consumption to date; the other ("billing salience") showed consumption and bill to date. Across the full sample, neither intervention reduced electricity use on average, and we can rule out effects as small as one percent. To understand this precise null, we examine engagement with the app and find no effects on the extensive margin, and only limited responses on the intensive margin. Among households that already engage with the app, the price-salience treatment modestly increased engagement and led to small consumption declines late in the billing cycle, when marginal prices rise mechanically under the nonlinear tariff. These results underscore both the promise and limits of digital behavioral tools for demand management – while low-cost app integrations can inform attentive users, engagement does not necessarily scale with delivery, limiting the ability of such interventions to automatically generate population-level energy savings.

*We are grateful to the International Growth Center as well as the Deep Decarbonization Initiative at UC San Diego for generous funding to support this research. We thank John Leiman for excellent research assistance. We are grateful for feedback from many seminars and conferences that have helped improve our paper. Any remaining errors are our own. This research received IRB approval from UC San Diego (# 802829) and Virginia Tech (#22-127). The trial was registered on AEA's social science registry as AEARCTR-0012178. Authors names appear in alphabetical order and authorship is equally shared.

†School of Global Policy and Strategy, UC San Diego; teevrat@ucsd.edu

‡Center for International Studies, Yonsei University; minkyong@yonsei.ac.kr

§Brigham Young University; eric.lewis@byu.edu

¶LBJ School of Public Affairs, University of Texas, Austin; chi.ta@austin.utexas.edu

# 1 Introduction

Digital technologies offer an unprecedented opportunity to scale behavioral and information interventions by reaching large populations at very low marginal cost. Once developed, digital platforms can automatically deliver information and reminders at scale, making them attractive tools for improving household decision-making in energy, health, and finance. However, their effectiveness ultimately hinges on whether users engage with the information provided. Many digital programs depend on active engagement and sustained attention – features that are often uneven across users. More recently, economists have emphasized that many interventions exhibit substantial attenuation when scaled, not because the underlying mechanism is absent, but because adoption and adherence fall outside controlled experimental settings (Al-Ubaydli et al., 2017; DellaVigna and Linos, 2022; Vivalt, 2020). In the context of electricity use, this raises an important question: can low-cost digital tools that communicate real-time consumption and pricing information meaningfully change behavior, or do limits to user engagement undermine their potential to deliver energy conservation at scale?

A growing literature shows that households facing nonlinear electricity tariffs often fail to respond to marginal prices, weakening the effectiveness of non-linear price schedules designed to manage demand or promote efficiency (Ito, 2014; Jessoe and Rapson, 2014; Borenstein, 2009). For example, households may mistakenly perceive the marginal price as the price they must pay for all consumption units, resulting in underconsumption and significant welfare losses (Shaffer, 2020).[1] One explanation is that marginal prices are not salient: consumers may not know when they cross into higher price tiers, may find rate structures too complex, or may lack timely feedback about their consumption. Salient information interventions such as in-home display (IHD) devices that provide real-time usage and pricing information have been shown to increase responsiveness to marginal prices (Jessoe and Rapson, 2014). However, IHD devices remain expensive and difficult to deploy at scale, particularly in low- and middle-income countries. A promising alternative is to integrate increasingly available smart-meter data into mobile applications that automatically communicate real-time usage and pricing information to consumers. Whether these low-cost, app-based interventions can replicate the behavioral effects of more intensive technologies – and for whom – remains an open empirical question.

We address this question through a pre-registered natural field experiment on a random

[1] Related evidence shows that consumers often under-attend to energy costs even when those costs are salient and economically meaningful, such as in appliance purchase decisions (Houde and Myers, 2021).

sample of 45,000 residential customers of the public electricity utility in Hanoi, Vietnam.[2] These households were randomly selected from the population that had already downloaded and registered on the utility's mobile application – approximately 30 percent of the utility's more than two million residential customers at the time of the experiment. Residential consumers face nonlinear tariffs with six steeply increasing price tiers, yet many households remain unaware of when they cross into higher marginal rates. Working with the utility's smart-meter data platform, we designed two information interventions delivered through its existing mobile app. The first treatment ("price salience") displayed each household's current marginal price tier and usage to date on a real-time graph of the tariff schedule. The second ("billing salience") displayed cumulative consumption and bill to date using a dynamic visual gauge.

Across the full sample, we find that neither treatment significantly reduced electricity consumption on average. With sufficient statistical power to detect effects as small as one percent, we can rule out meaningful average energy savings. To understand this precise null, we examine whether the interventions succeeded in scaling engagement on the app. We find no effect on engagement on the extensive margin: treated households are no more likely to open the app than control households. The price-salience treatment modestly increases engagement on the intensive margin. Consistent with this pattern, we find suggestive evidence of modest consumption reductions only among households who regularly check the app and primarily late in the billing cycle, when marginal prices rise mechanically under the nonlinear tariff.

**Contributions.** This paper makes two central contributions. First, we provide a policy-relevant evaluation of app-based information interventions for residential electricity conservation at scale. A central insight of the science-of-scaling literature is that evidence should be generated in the population and delivery environment that would be used under real-world implementation, as many interventions attenuate once scaled beyond tightly controlled settings (Al-Ubaydli et al., 2017; DellaVigna and Linos, 2022; Vivalt, 2020; List, 2024; Fowlie and Meeks, 2021). Our experimental design explicitly reflects this concern. We randomly sample 45,000 households from the relevant population of customers—those who had already downloaded and registered on the utility's mobile application, which at the time of the experiment comprised approximately 30 percent of the utility's more than two million residential customers. This sampling frame contrasts with much of the prior

[2]The final analysis follows the experimental design, sampling strategy, outcome definitions, and estimation approach specified in the pre-analysis plan discussed in Garg et al. (2023). We discuss details on necessary deviations from the pre-analysis plan arising from implementation constraints and data availability in Appendix Section A.3.

evidence on energy feedback and information, which often relies on pilots, dedicated hardware deployments, or opt-in participation (Houde et al., 2013; Schultz et al., 2015; Geelen et al., 2019; Tiefenbeck et al., 2018; Burkhardt et al., 2023). In addition, by studying information that is delivered through an existing utility platform, we complement field experiments that combine information with pricing incentives (Wolak, 2011; Kahn and Wolak, 2013; Stojanovski et al., 2020). Across this scale-relevant sample, we find no evidence that providing real-time information about marginal prices or cumulative bills reduces electricity consumption on average, and we can rule out effects of one percent or larger.[3]

Second, we diagnose why these interventions fail to generate population-level impacts when delivered through a scalable digital platform. While mobile applications allow information to be delivered widely at low marginal cost, they do not ensure attention or engagement. We show that the interventions do not increase engagement on the extensive margin—households are no more likely to open the app at all. Any engagement effects operate only on the intensive margin among households that already use the app regularly, consistent with models of limited attention and infrequent optimization (Sims, 2003; Reis, 2006; Chetty et al., 2009). As a consequence, reductions in energy use are necessarily confined to a subset of engaged users. Consistent with this mechanism, we find that modest consumption reductions, when present, occur only among households that routinely check their app and primarily late in the billing cycle, when marginal prices become most salient. This timing aligns with evidence that attention to energy costs is episodic and shaped by billing and other salient moments (Gilbert and Graff Zivin, 2014; Sexton, 2015; Wichman, 2017; Singhal, 2024; Prest, 2020).

Our findings highlight a central challenge for digital demand-side interventions emphasized in the energy efficiency literature: low marginal delivery cost does not imply low realized savings. By combining a scale-relevant experimental sample with high-frequency consumption data and direct measures of engagement, we show that understanding both outcomes and engagement is essential for diagnosing when – and why – digital behavioral interventions fail to scale.

[3] We also contribute to the broader literature on nonlinear tariffs and perceived marginal incentives by testing whether making marginal costs and tariff kinks more transparent changes consumption behavior (Ito, 2014; Borenstein, 2009; Shaffer, 2020; Jessoe and Rapson, 2014).

## 2 Experimental Design and Data

We conduct our study in Hanoi, the capital and second-largest city of Vietnam. Vietnam, a lower-middle-income country in Southeast Asia, has experienced rapid economic growth, with GDP increasing by 6-8% per year over the past decade. This growth has been accompanied by a sharp rise in energy demand: according to the International Energy Agency, Vietnam's electricity consumption increased from roughly 30 TWh in the early 2000s to over 240 TWh by 2022, an almost eightfold increase. Rising household incomes and a growing middle class have driven greater ownership of electricity-intensive appliances, particularly air conditioning. The country's hot and humid climate, particularly during the summer months from May to September, further increases electricity demand. Hydropower accounts for over a third of Vietnam's electricity, but summer droughts often cause shortages, making it difficult for the utility to meet rising demand.

As is common for electricity utilities operating as local monopolies, retail electricity prices in Vietnam are fixed by regulation, limiting the use of dynamic pricing and requiring utilities to rely on non-price mechanisms commonly referred to as demand-side management (DSM) to overcome supply constraints during peak summer months. EVNHanoi, Hanoi's state-owned electricity utility, has introduced a variety of initiatives to encourage energy conservation in peak months, including campaigns via television, radio, community meetings, news outlets, and home visits. At the same time, Vietnam is advancing digital transformation to improve efficiency and cut costs. Smart meters are now installed at 100% of all 2.8 million households in Hanoi, and nearly 30% of customers use EVNHanoi's mobile app to track their electricity use, giving the utility more effective tools to manage demand and engage directly with consumers.

The combination of rapid demand growth and ongoing digital transformation makes Hanoi a particularly useful setting for evaluating digital demand-side interventions. Mobile platforms allow utilities to deliver personalized information at very low marginal cost and to potentially do so at scale. Once deployed, they can automatically send information and reminders, helping households make more informed decisions about energy use. However, the success of these interventions ultimately depends on whether people engage with the guidance provided.

A key part of managing energy use is Vietnam's nationwide nonlinear pricing system, which has six tiers based on consumption within a billing cycle. This structure is designed to encourage careful electricity use, allowing lower-income households to access electricity at affordable rates while assigning higher costs to excessive consumption. However, as is common in other settings, households typically do not react to marginal prices of electricy - the common perception is that most customers do not fully understand the pricing system, creating an important role for effective communication and education.

To address this, our experiment leveraged the EVNHanoi mobile app to make pricing and billing more salient for households. The EVNHanoi app, which already allowed customers to click on a tab on the homescreen to track daily electricity consumption, was enhanced to display estimated electricity prices for each unit of consumption and projected billing information. This added functionality allowed households to see how their usage translated into costs more frequently, providing a clearer link between behavior and financial impact. By integrating this pricing and billing information directly into the app, the study aimed to test whether making consumption costs more visible could improve energy decision-making.

## 2.1 Sampling

We obtain monthly consumption data for 700,000 households from the electric utility, covering the period from August 2022 to July 2023. These households represent the full population of all app users at the time of the study. To ensure data quality, we excluded households with more than 15% missing daily data from May 2023 to July 2023 and those with missing data for any month over a twelve-month span.[4] We explicitly exclude households that had participated in other experiments (Garg et al., 2025).

Our study focuses on evaluating the impact of marginal price and billing salience on energy consumption. In our experimental context, the highest price tier applies when monthly energy use exceeds 400 kWh (see Figure A.1). Conversely, the lowest price tier is applicable when monthly consumption is at or below 50 kWh. As such, households consuming significantly more than 400 kWh/month or significantly less than 50 kWh/month will not be affected by our marginal price intervention. Thus, we further limit our sample to households with monthly energy usage ranging from 10 to 1000 kWh.[5] Incorporating these parameters, we draw our sample from a population of about 400,000 households. Our power

[4] Typically, 2-5% of households have missing data due to idiosyncratic meter malfunctions or other technical problems. Households with an unusually high rate of missing data are often those whose utility services have been disconnected.

[5] Less than 1% of all 700,000 app users consume less than 10 kWh monthly, and only 2% consume 1,000 kWh or more monthly. Our pre-analysis plan has further details on sampling (Garg et al., 2023).

calculations indicate that 10,000 households per treatment group are sufficient to detect a small 1% effect. We therefore used a total sample size of 45,000 households, with 15,000 assigned to each treatment or control arm. The experiment was done on 45,000 households. However, our final sample has 44,997 households: We drop one household due to a missing household ID variable and drop two households due to conflicting consumption data.[6] Importantly, our results generalize to the broader population of app users at the time of the study in Hanoi as our sample was selected randomly from the final frame of approximately 400,000 households.[7]

## 2.2 Treatment Assignment

We randomly assigned 45,000 households evenly into a control group and two treatment groups. Using the app, all households can check their current daily usage, review daily consumption from the past week, and monitor their monthly usage throughout the current calendar year. The treatment groups received additional features: one group received real-time pricing information, while the other received real-time billing updates. Notably, these information treatments were presented prominently at the top of the homepage of the app, as shown in Figure A.2.

**Control Group**: Households in this group did not have access to any interactive displays for daily electricity prices or updated bills. Figure A.2A shows the app interface of the control group. The figure displays the app's homepage. From this page, customers can access various features, including online bill payment, management of their electricity service contract, daily energy usage tracking, and advertisements for incentive programs that allow users to earn and spend points through the app. Clicking on any icon or advertisement directs the user to the corresponding page within the app.

**Price Information Treatment**: Households in this treatment group received, in addition to all the information provided to the control group, details about their estimated electricity prices for each unit of consumption in their current billing cycle via the app's interactive display. Figure A.2B presents an example of the interactive display, visually representing the nonlinear pricing system, as compared to the original interface shown in Figure A.2A.

[6] Specifically, our consumption data from the technology partner lists each of these households in multiple rows with conflicting consumption data in each row.

[7] Our sample size was constrained by both fixed and per-user costs associated with working with the technology partner, including user-interface development as well as data security and management requirements.

The graph in Figure A.2B highlights the primary intervention, allowing customers to visualize their current energy consumption along with the corresponding marginal price on a real-time six-tier graph. The design's objective is to help participants better understand this pricing structure.

**Bill Information Treatment**: In this treatment, similar to the price-information treatment, households received information regarding their estimated, up-to-date electricity usage and bill each day of their current billing cycle through the app's interactive display. Figure A.2C showcases the interactive interface for this treatment. Households in this group viewed a semicircular gauge on their app that visually represents energy consumption. As energy usage increases over time, the gauge fills up and changes color from blue (indicating low consumption compared to the previous month) to red (indicating high consumption relative to the previous month).

## 2.3 Experiment Timeline and Key Events

Figure 1 shows information on key dates in our study and highlights several important events that took place during the sample period. In the pre-period, electricity prices were raised once, on May 4, 2023. A second increase occurred on November 9, 2023, during the mid-period. Figure A.1 presents changes in the nonlinear pricing structure faced by consumers during our study period. In the post-period, starting after February 29, 2025, the utility adjusted billing schedules so that all households shared the same start and end dates aligned with the calendar month.

Our initial electricity consumption data begins in August, 2022, and our pre-treatment period extends from August 2022 to end of September, 2023. The experiment was initially launched on October 1, 2023. However, the treatment was temporarily suspended on October 20, 2023, at the utility's request, as they planned to raise electricity prices on November 9, 2023. The treatment was reintroduced on December 30, 2023, and continued until May 30, 2024. Based on this timeline, we define the pre-period as August 1, 2022, the start of our consumption data, to September 30, 2023, the day before the initial rollout. The mid-period runs from October 1 to December 29, 2023, covering the gap between the initial launch and the full resumption. The post-period begins on December 30, 2023, when the rollout resumed fully.

In our analysis, our main specifications exclude the mid-period (October 1-December 29, 2023) as this was when the treatment was temporarily rolled out and then discontinued in October 2023. In some robustness checks, we include all dates in the sample, treating observations from October 1-December 29, 2023 as treated dates.

### 2.4 Data and Experimental Balance

The primary outcome variable is daily household electricity consumption from our sample of 44,997 households, tracked from August 1, 2022, to June 30, 2024. In addition, we gather billing data from participating households, which includes information on electricity consumption and billing dates. We also collect temperature data, since temperature changes affect electricity use as households adjust their air conditioners or heaters. Figure A.3 shows average daily consumption and feels-like temperature during the sample period. Average daily consumption ranges from 6.5 to nearly 16 kWh. Consumption is generally lowest during the winter months (December to February), when temperatures are coolest, and highest during the summer months (May to August), when temperatures peak. Additionally, we collected app usage data from households in our sample between November 6, 2023, and June 4, 2024. Although the app usage data fully cover the post-treatment period, they unfortunately do not include the pretreatment period.[8] Appendix Table A.1 presents summary statistics for all variables in the dataset, measured at the day by household level.

We assessed the balance between treatment and control groups using historical household electricity consumption data. Specifically, we examined average daily consumption for each month from August 2022 to December 2023, prior to the intervention. Appendix Tables A.2, A.3, and A.4 present the results of our balance checks, showing no noticeable differences in historical electricity consumption patterns between the control and treatment groups.

## 3 No Aggregate Effect on Electricity Consumption

We begin by estimating the average effect of the two information interventions on household electricity consumption. Our primary outcome is daily electricity use at the house-

[8]This limitation arises from budget constraints and data protection requirements: the technology partner did not routinely track or store individual-level app usage data. Under our data-sharing agreement, such data were collected solely for study participants and only during the study period, making app usage data available only for the post-treatment phase.

hold level, measured using high-frequency smart meter data. To estimate average treatment effects, we estimate the following difference-in-differences specification:

$$f(y_{it}) = \sum_{k=1}^{2} \beta_k \cdot \mathbb{1}\{treatment_i = k\} \times post_t + \gamma_t + \delta_i + \varepsilon_{it}, \tag{1}$$

where $y_{it}$ denotes daily electricity consumption of household $i$ on day $t$, $post_t$ is an indicator equal to one for dates on or after December 30, 2023 (when the intervention was fully rolled out), and $k = 1$ and $k = 2$ index the price-salience and billing-salience treatments, respectively. The specification includes household fixed effects $\delta_i$ and date fixed effects $\gamma_t$, so identification comes from within-household changes in consumption following the intervention, relative to the control group. Standard errors are clustered at the household level.

We estimate equation (1) using several specifications, including the inverse hyperbolic sine transformation, the $\log(1 + y_{it})$ transformation, a Poisson specification, and consumption in levels. Our main specification uses the inverse hyperbolic sine transformation, which accommodates zero values, reduces sensitivity to outliers, and improves statistical precision in the presence of skewed consumption data. Our finding of a null result is robust to these different specifications.

Table 1 reports the estimated average treatment effects. Across all specifications, we find no evidence that either the price-salience treatment or the billing-salience treatment reduces electricity consumption on average. Estimated coefficients are small in magnitude and statistically indistinguishable from zero. In some specifications, the point estimate for the price-salience treatment is positive, while the billing-salience treatment yields slightly negative estimates, but a test of equality between the two treatment coefficients fails to reject the null at conventional significance levels.

Importantly, the absence of statistically significant effects is not driven by imprecise measurement or limited statistical power. Our analysis is based on a balanced panel of 44,997 households observed daily over a 23-month period, yielding more than 30 million household-day observations. Consumption is measured directly via smart meters, minimizing classical measurement error. Based on pre-registered power calculations, this design allows us to rule out average treatment effects of one percent or larger.[9] Furthermore, the 95% confidence intervals implied by our estimates in Table 1 allow us to reject that the treat-

[9] Details of the power calculations are provided in the pre-analysis plan (Garg et al., 2023).

ment causes a significant change in consumption. For example, the 95% interval for the price salience treatment effect using the Poisson specification is (-0.0039, 0.0101), allowing us to reject a -1% change in consumption and just barely including a +1.01% increase in consumption. We also find no evidence of bunching at tier thresholds in either the pre- or post-treatment periods, further suggesting that the interventions did not meaningfully alter consumption behavior at the aggregate level. Our findings, therefore, indicate that providing frequent, real-time information about marginal prices or cumulative bills through a mobile application does not reduce household electricity consumption on average, even when delivered at scale.

# 4 Limited Engagement and the Failure to Scale Attention

A natural explanation for the absence of aggregate consumption effects is that households may not engage with the information provided through the mobile app. In this section, we examine whether the two treatments affected households' engagement with the app, focusing separately on the extensive-margin (whether households ever check the app during the study period) and the intensive-margin (how frequently they check it).

## 4.1 Measuring App Engagement

We measure engagement using administrative app-usage data, which record whether a household accesses the app on each day. Because app usage data are available beginning November 6, 2023, our analysis focuses on two periods: a mid-period from November 6 to December 29, 2023 (a period after October 2023 during which the treatment was temporarily introduced and then withdrawn), and a post-period from December 30, 2023 onward (when the treatment was fully active). For each household and period, we construct two measures of engagement: (i) an indicator for whether the household checks the app at least once during the period (extensive margin), and (ii) the fraction of days during the period on which the household checks the app (intensive margin).

In our design, treatment consists solely of information displayed on the app screen. As a result, treatment assignment is revealed only after an individual opens the app. Consequently, assignment to treatment should not affect engagement absent other channels. Consistent with this, we find no effect on the extensive margin – the share of those who ever check the app is not statistically different in the treatment and control groups.[10] On the intensive margin, however, it is possible for the treatment to affect app engagement as different types of information may spur households to check the app more or less frequently.

### 4.2 Average Effects on Engagement

To estimate treatment effects on app engagement, we estimate the following household-level regression separately for the mid- and post-periods:

$$y_i^p = \alpha + \beta_1 T1_i + \beta_2 T2_i + \varepsilon_i, \tag{2}$$

where $y_i^p$ denotes an engagement outcome for household $i$ in period $p \in \{\text{mid}, \text{post}\}$, $T1_i$ and $T2_i$ are indicators for assignment to the price-salience and billing-salience treatments, respectively, and $\varepsilon_i$ is an error term. We estimate equation (2) using ordinary least squares.

Table 2 reports the results. Across both the mid- and post-periods, we find no evidence that either treatment affects engagement on the extensive margin. In the post-period, approximately 56 percent of households check the app at least once, and this share does not differ statistically between the control and treatment groups. We find similar null effects on extensive-margin engagement during the mid-period. Because checking the app is the only way households can observe their treatment status, these results indicate that the interventions do not impact whether those who have already downloaded and registered on the app engage at all with it during the study period.

In contrast, we find modest effects on the intensive margin. Column 2 of Table 2 shows that households assigned to the price-salience treatment check the app slightly more frequently during the post-period. Relative to a control-group mean of 8.7 percent of days, households in the price-salience treatment check the app on approximately 9.1 percent of days, a difference that is statistically significant at the five-percent level. We find a similar increase in app-checking frequency during the mid-period (column 4), suggesting that the earlier temporary rollout of the price information treatment in October 2023 impacted app engagement during November and December 2023 when the treatment was temporarily discontinued. In contrast, the billing-salience treatment does not significantly affect en-

[10] Indeed, finding any statistically significant treatment effects on the extensive margin would suggest a lack of balance in treatment assignment.

gagement on either margin in either period.

## 4.3 Treatment Effects by Billing-Cycle Position

To better understand when households engage with the app, we examine daily patterns of app usage. Figure A.4a plots the fraction of households that check the app on each day, separately by treatment status. App usage is highly cyclical, with pronounced spikes at the beginning and end of billing cycles. These spikes become especially sharp after March 1, 2024, when the utility aligned all households' billing cycles with the calendar month. A noticeable increase in app usage also occurs around the November 9, 2023 price increase.

Figure A.4b further shows that these patterns align closely with households' billing-cycle start dates prior to March 2024. Households whose billing cycles begin earlier in the month exhibit earlier peaks in app checking, while those with later cycle start dates peak correspondingly later. These patterns indicate that attention to the app is naturally concentrated around billing events. We formalize these patterns using daily household-level regressions that allow treatment effects to vary by position within the billing cycle. Specifically, we estimate:

$$CheckApp_{it} = \alpha + \sum_{g \in \{\text{begin},\text{middle},\text{end}\}} (\beta_{1g} T1_i + \beta_{2g} T2_i)\, \mathbb{1}\{CyclePos_{it} = g\} + \gamma_t + \varepsilon_{it}, \quad (3)$$

where $CheckApp_{it} = 1$ if household $i$ checks the app on day $t$, $\mathbb{1}\{CyclePos_{it} = g\}$ denotes indicators for whether day $t$ falls at the beginning, middle, or end of the billing cycle, and $\gamma_t$ are date fixed effects. Standard errors are clustered at the household level. In our regression specification, we allow for different ways of dividing the billing cycle into beginning, middle, and end: Specifically, we define the beginning as days 1-X of the billing cycle, the end as the last X days of the billing cycle, and allow values of X ranging from 2 to 5.

Table A.7 reports the results. Consistent with the graphical evidence, overall app checking is significantly higher at the beginning and end of the billing cycle than in the middle. The price-salience treatment increases the probability of app checking particularly during the middle and end of the billing cycle, with effects on the order of 0.3 to 0.4 percentage points. This is consistent with the idea that the information about one's consumption on the marginal price schedule is most valuable when one is more likely to have left the first pricing tier – i.e., after the first few days of the billing cycle. In contrast, the estimated effects of the billing-salience treatment are smaller and statistically insignificant across all cycle positions.

# 5 Limited Consumption Responses Among App Users

Sections 3 and 4 show that the information treatments do not reduce electricity consumption on average and do not increase app engagement on the extensive margin. In this section, we examine whether consumption responses emerge conditional on app usage and whether such responses vary systematically over the billing cycle.

## 5.1 Heterogeneity by App Checking

We first allow treatment effects to differ by whether a household checks the app during the post-treatment period. Let $CA_i$ be an indicator equal to one if household $i$ checks the app at least once during the post-period (December 30, 2023 and later). We estimate the following specification:

$$\begin{aligned} f(y_{it}) = {} & \beta_1 \cdot CA_i \cdot post_t + \beta_2 \cdot T1_i \cdot post_t + \beta_3 \cdot T2_i \cdot post_t \\ & + \beta_4 \cdot T1_i \cdot CA_i \cdot post_t + \beta_5 \cdot T2_i \cdot CA_i \cdot post_t + \gamma_t + \delta_i + \varepsilon_{it}, \end{aligned} \tag{4}$$

where $y_{it}$ denotes daily electricity consumption, $T1_i$ and $T2_i$ are indicators for the price-salience and billing-salience treatments, $post_t$ indicates the post-treatment period, $\delta_i$ are household fixed effects, and $\gamma_t$ are date fixed effects. As in the main specifications, we focus on the inverse hyperbolic sine transformation of consumption.

Appendix Table A.8 reports the results. We find no evidence that either treatment affects consumption for households that do not check the app. For households that do check the app, the interaction between the price-salience treatment and app checking is statistically significant in some specifications. However, joint tests of the total effect of the price-salience treatment for app-checking households fail to reject the null that the treatment has no effect on consumption.[11] Overall, these results indicate that selection into app checking is strongly correlated with consumption outcomes, but that treatment-induced changes in consumption conditional on app usage are not robustly distinguishable from zero.

## 5.2 Heterogeneity by Day of Billing Cycle

We next examine whether treatment effects vary over the billing cycle. This analysis is motivated by the structure of the nonlinear tariff and by the information displayed in the app, which changes mechanically as consumption accumulates within a billing cycle. To allow for flexible heterogeneity, we estimate the following specification:

$$f(y_{it}) = \sum_{d=1}^{31} \sum_{p\in\{\text{pre},\text{post}\}} \sum_{r\in\{\text{control},T1,T2\}} \beta_{d,p,r} \cdot \mathbb{1}\{d,p,r\}_{it} + \gamma \cdot t + w_t + \delta_i + \varepsilon_{it}, \tag{5}$$

where $d$ indexes the day of the billing cycle, $w_t$ captures interactions between weather deciles and day-of-week indicators, $\delta_i$ are household fixed effects, and other terms are defined as above. We exclude the mid-period (October 1–December 29, 2023) from this analysis because the treatment was temporarily rolled out and rolled back during this window.

We summarize the day-specific treatment effects using the following difference-in-differences estimands:

$$DID_{d,\text{T1 vs C}} = \beta_{d,\text{post},T1} - \beta_{d,\text{pre},T1} - (\beta_{d,\text{post,control}} - \beta_{d,\text{pre,control}}), \tag{6}$$

$$DID_{d,\text{T2 vs C}} = \beta_{d,\text{post},T2} - \beta_{d,\text{pre},T2} - (\beta_{d,\text{post,control}} - \beta_{d,\text{pre,control}}). \tag{7}$$

Figure 2 shows estimated treatment effects for the price information treatment. For the full sample and for households that do not check the app, estimated effects are close to zero throughout the billing cycle. For households that check the app, we observe suggestive evidence of reduced consumption during the second half of the billing cycle, particularly between days 15 and 20. However, confidence intervals widen toward the end of the cycle, and the estimated effects are not statistically distinguishable from zero.

Figure 3 presents analogous results for the bill information treatment. Estimated effects are

[11] Specifically, we test that $\beta_2 + \beta_4 = 0$ and $\beta_3 + \beta_5 = 0$.

generally small and unstable across billing-cycle days. In some specifications, we observe marginally significant effects for the full sample, but these patterns do not persist among households that check the app. Formal $F$-tests fail to reject the null that all day-specific treatment effects are jointly equal to zero.

Overall, we find no statistically significant evidence of heterogeneous treatment effects by app usage or billing-cycle position. While some patterns are suggestive of behavioral responses to marginal price information among app-checking households, these effects are limited in magnitude and not robust across specifications, consistent with the absence of aggregate treatment effects.

# 6 Conclusion

Digital platforms are often viewed as a promising pathway for scaling behavioral interventions, since information can be delivered to large populations at near-zero marginal cost. Whether such interventions can generate meaningful behavioral change at scale, however, depends critically on attention and engagement. We study this question using a large-scale randomized experiment with residential electricity customers in Hanoi, Vietnam, in which two information interventions were embedded directly into the utility's existing mobile application and powered by smart-meter data. Importantly, our experimental sample of 45,000 households was drawn at random from the population of customers in the city who were already using the app (approximately 30% of the households in the city). By randomizing within the population that would plausibly be exposed under real-world scale-up, our estimates speak directly to the effects that would be realized if the intervention were rolled out broadly within the platform, consistent with recent work emphasizing the importance of sampling from the relevant population of interest when evaluating scalability (List, 2024).

Across this population, we find no evidence that providing real-time information about marginal prices or cumulative bills reduces electricity consumption on average. Estimated effects are tightly centered around zero, and our confidence intervals allow us to rule out average consumption changes of one percent or larger. The absence of aggregate effects is not driven by a lack of statistical power or measurement error, but instead reflects limited engagement with the intervention itself. While the price-salience treatment modestly increases app usage intensity among households that already check the app, neither intervention increases engagement on the extensive margin. As a result, attention does not scale with delivery: the intervention reaches many households mechanically but meaningfully informs only a subset of already attentive users.

Consistent with this interpretation, we find that behavioral responses, when they occur, are conditional and temporally concentrated. Modest reductions in electricity consumption appear only among households that actively check the app and primarily late in the billing cycle, when marginal prices rise mechanically under the nonlinear tariff and the information becomes most decision-relevant. These patterns are consistent with standard models of limited attention and salience, but they are not robust or widespread enough to generate detectable changes in average consumption. Interpreting these conditional effects as evidence of a scalable intervention would therefore be misleading: they illustrate a mechanism without delivering population-level impact.

Our findings highlight a fundamental constraint on scaling digital behavioral interventions. Low marginal delivery cost does not imply low attention cost. Even when information is well designed, economically relevant, and delivered through an existing platform, limited and uneven engagement can sharply attenuate its aggregate impact. In this setting, app-based information changes behavior for some users some of the time, but fails to generate broad energy savings.

Our results point to several directions for future research. First, what design features or institutional complements can reliably increase attention at scale, rather than merely intensifying engagement among existing users? Second, how should policymakers evaluate platform-based interventions when treatment exposure is endogenous to user attention, even under randomized assignment? Third, under what conditions can digital tools substitute for higher-touch technologies such as in-home displays, and when should they instead be viewed as complements? Addressing these questions is essential for understanding when – and whether – digital platforms can fulfill their promise as scalable tools for demand-side management and other behavioral policies.

# 7 References

Al-Ubaydli, O., J. A. List, D. LoRe, and D. Suskind (2017). Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives 31*(4), 125–144.

Borenstein, S. (2009). To what electricity price do consumers respond? residential demand elasticity under increasing-block pricing. *Mimeo*.

Burkhardt, J., K. Gillingham, and P. Kopalle (2023). Field experimental evidence on the effect of pricing on residential electricity conservation. *Management Science 69*(12), 7675–7702.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *American Economic Review 99*(4), 1145–1177.

DellaVigna, S. and E. Linos (2022). Rcts to scale: Comprehensive evidence from two nudge units. *Econometrica 90*(1), 81–116.

Fowlie, M. and R. Meeks (2021). The economics of energy efficiency in developing countries. *Review of Environmental Economics and Policy 15*(2), 238–260.

Garg, T., M. Ko, and C. Ta (2023). Enhanced salience of nonlinear pricing and energy conservation. *AEA RCT Registry*. https://doi.org/10.1257/rct.12178-1.0.

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

Geelen, D., D. Keyson, S. U. Boess, and H. W. van Dijk (2019). Saving energy by supporting energy-efficient behavior: The effects of a smartphone feedback application in households. *Energy Efficiency 12*, 1635–1660.

Gilbert, B. and J. Graff Zivin (2014). Dynamic salience with intermittent billing: Evidence from smart electricity meters. *Journal of Economic Behavior & Organization 107*, 176–190.

Houde, S. and E. Myers (2021). Are consumers attentive to local energy costs? evidence from the appliance market. *Journal of Public Economics 201*, 104480.

Houde, S., A. Todd, A. Sudarshan, J. A. Flora, and K. C. Armel (2013). Real-time feedback and electricity consumption: A field experiment assessing the potential for savings and persistence. *The Energy Journal 34*(1), 87–102.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review 104*(2), 537–563.

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

Kahn, M. E. and F. A. Wolak (2013). Using information to improve the effectiveness of nonlinear pricing: Evidence from a field experiment. Working paper, Stanford University.

List, J. A. (2024). Optimally generate policy-based evidence before scaling. *Nature 626*(7999), 491–499.

Prest, B. (2020). Peaking interest: How awareness drives the effectiveness of time-of-use electricity pricing. *Journal of the Association of Environmental and Resource Economists 7*(1), 103–143.

Reis, R. (2006). Inattentive consumers. *Journal of Monetary Economics 53*(8), 1761–1800.

Schultz, P. W., M. Estrada, J. Schmitt, R. Dachel, and R. Ragazzi (2015). Using in-home displays to provide smart meter feedback about household electricity consumption. *Energy 90*, 351–358.

Sexton, S. (2015). Automatic bill payment and salience effects: Evidence from electricity consumption. *Review of Economics and Statistics 97*(2), 229–241.

Shaffer, B. (2020). Misunderstanding nonlinear prices: Evidence from a natural experiment on residential electricity demand. *American Economic Journal: Economic Policy 12*(3), 433–461.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Singhal, P. (2024). Inform me when it matters: Cost salience, energy consumption, and efficiency investments. *Energy Economics 133*, 107484.

Stojanovski, O., G. W. Leslie, F. A. Wolak, J. E. H. Wong, and M. C. Thurber (2020). Increasing the energy cognizance of electricity consumers in mexico: Results from a field experiment. *Journal of Environmental Economics and Management 102*, 102323.

Tiefenbeck, V., L. Goette, K. Degen, V. Tasic, E. Fleisch, R. Lalive, and T. Staake (2018). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management Science 64*(3), 1458–1476.

Vivalt, E. (2020). How much can we generalize from impact evaluations? *Journal of the European Economic Association 18*(6), 3045–3089.

Wichman, C. (2017). Information provision and consumer behavior: A natural experiment in billing frequency. *Journal of Public Economics 152*, 13–33.

Wolak, F. A. (2011). Do residential customers respond to hourly prices? evidence from a dynamic pricing experiment. *American Economic Review 101*(3), 83–87.

# Figures

**Figure 1:** Timeline of key events and data coverage

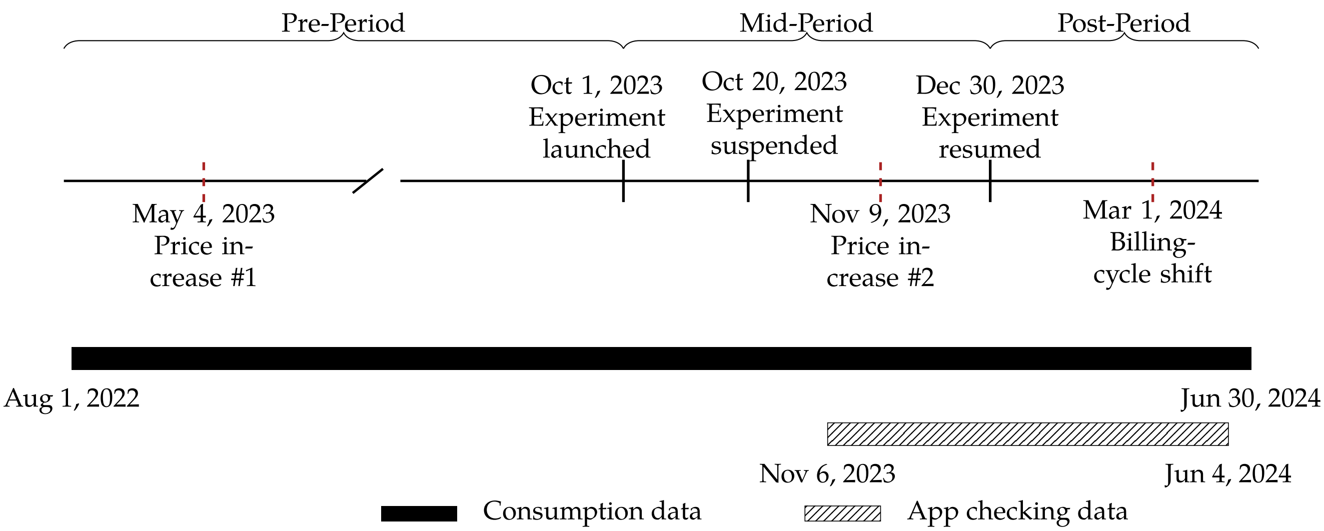
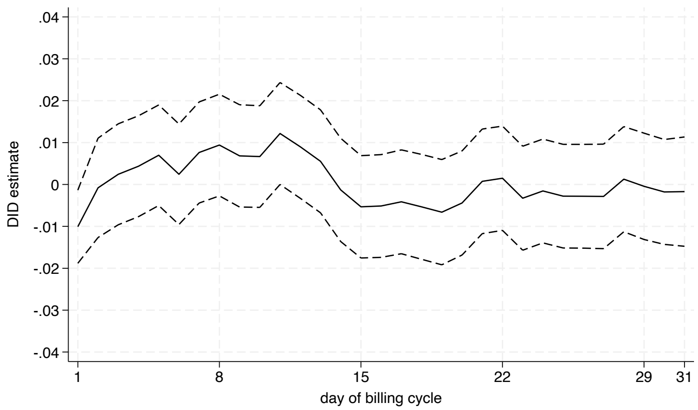
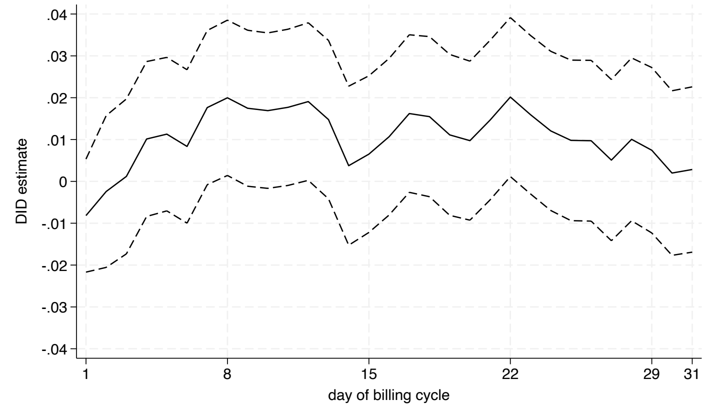


**Figure 2:** Estimated effect of treatment 1 (price treatment) relative to control group. Point estimates of DID treatment effects and 95% confidence intervals from estimates of equation 5.

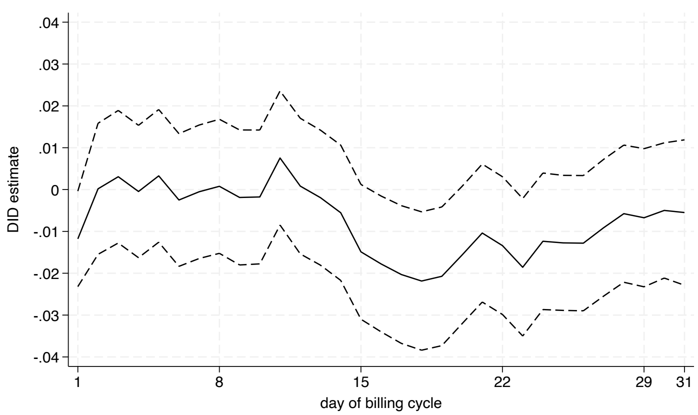
(a) All households:




(b) Households that do not check app in post-period:




(c) Households that check app in post-period:




(d) All control households plus treatment households that check app in post-period:

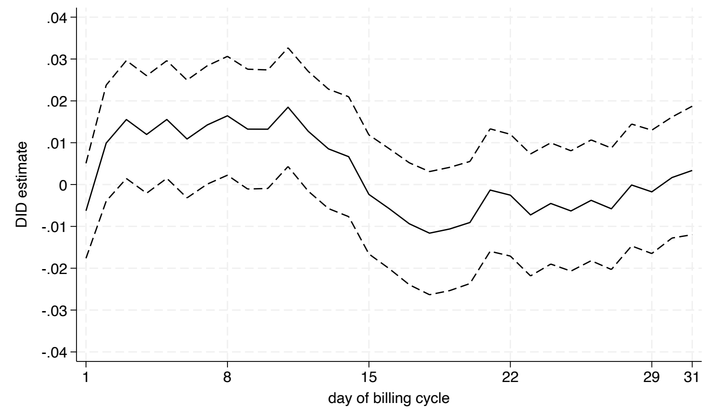
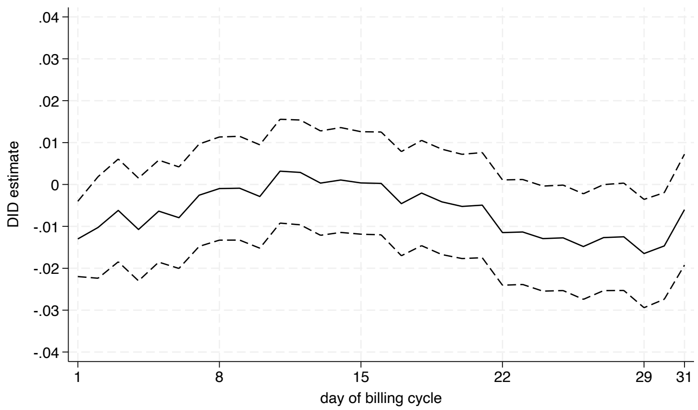
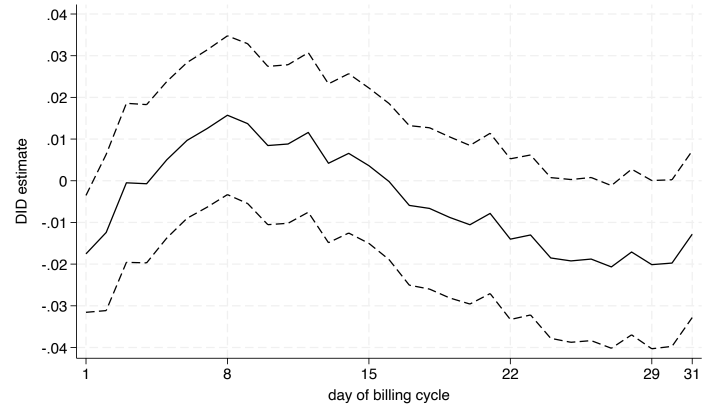


**Figure 3:** Estimated effect of treatment 2 (billing treatment) relative to control group. Point estimates of DID treatment effects and 95% confidence intervals from estimates of equation 5.

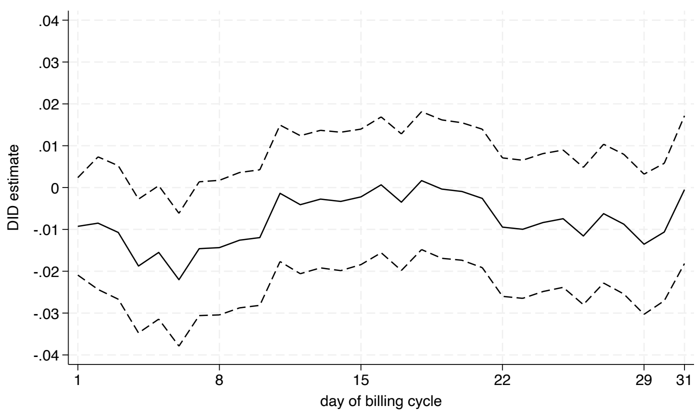
(a) All households:



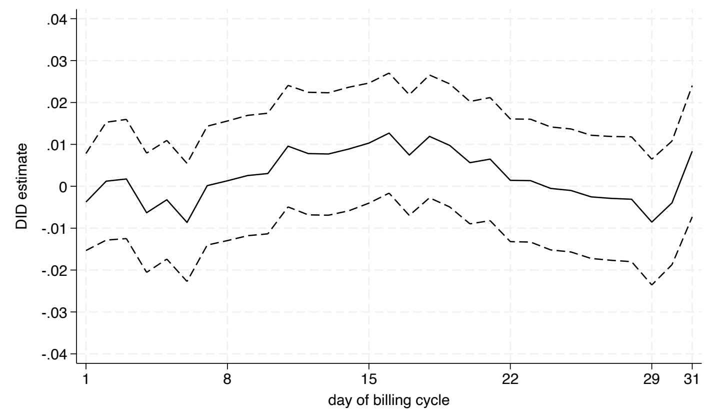

(b) Households that do not check app in post-period:




(c) Households that check app in post-period:




(d) All control households plus treatment households that check app in post-period:



# Tables

**Table 1:** Average Treatment Effects: Consumption

| | (1) IHS | (2) log(1+.) | (3) Poisson | (4) In levels |
|---|---|---|---|---|
| Price information treatment x post | 0.000685 | 0.000837 | 0.00308 | 0.0307 |
| | (0.00420) | (0.00351) | (0.00358) | (0.0347) |
| Billing information treatment x post | -0.00631 | -0.00478 | -0.000225 | -0.00102 |
| | (0.00430) | (0.00358) | (0.00353) | (0.0342) |
| p-value equality | 0.103 | 0.116 | 0.368 | 0.372 |
| R squared | 0.448 | 0.454 | | 0.480 |
| Pseudo R squared | | | 0.298 | |
| Control mean post-period | 2.546 | 2.042 | 9.693 | 9.693 |
| Observations | 26471767 | 26471767 | 26471767 | 26471767 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This regression table using the specification in equation 1. The dependent variable is the inverse hyperbolic sine of consumption (column 1), log(1 + consumption) (column 2), consumption using Poisson regression (column 3), and consumption in levels (column 4). Each specification includes date and household fixed effects. Standard errors are in parentheses and are clustered by household. The row denoted "p-value equality" is the p-value from the test that the coefficient of treatment 1 x post is equal the coefficient of treatment 2 x post.

**Table 2:** Impact of Treatment on App Checking

| | (1) Any–post | (2) Frac–post | (3) Any–mid | (4) Frac–mid |
|---|---|---|---|---|
| Price information treatment (T1) | 0.00750 | 0.00452** | 0.00577 | 0.00448** |
| | (0.00573) | (0.00193) | (0.00577) | (0.00195) |
| Billing information treatment (T2) | -0.000326 | 0.00191 | 0.000599 | 0.00194 |
| | (0.00574) | (0.00191) | (0.00577) | (0.00192) |
| Intercept | 0.555*** | 0.0867*** | 0.490*** | 0.0851*** |
| | (0.00406) | (0.00134) | (0.00408) | (0.00135) |
| Observations | 44997 | 44997 | 44997 | 44997 |
| R squared | 0.0000530 | 0.000123 | 0.0000268 | 0.000118 |

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

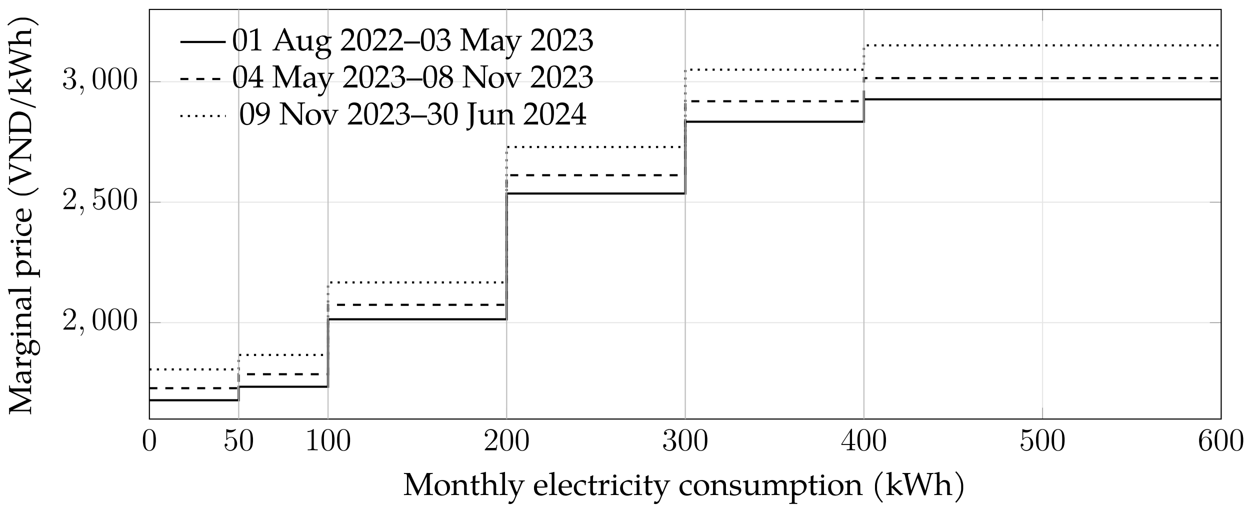
Notes: Regressions at household level. Dependent variables are an indicator for whether ever checked app in the post-period (column 1), fraction of days in post-period household checked app (column 2), whether ever checked app in the mid-period (column 3), and fraction of days in mid-period household checked app (column 4).

# A Appendix

## A.1 Additional Institutional Details

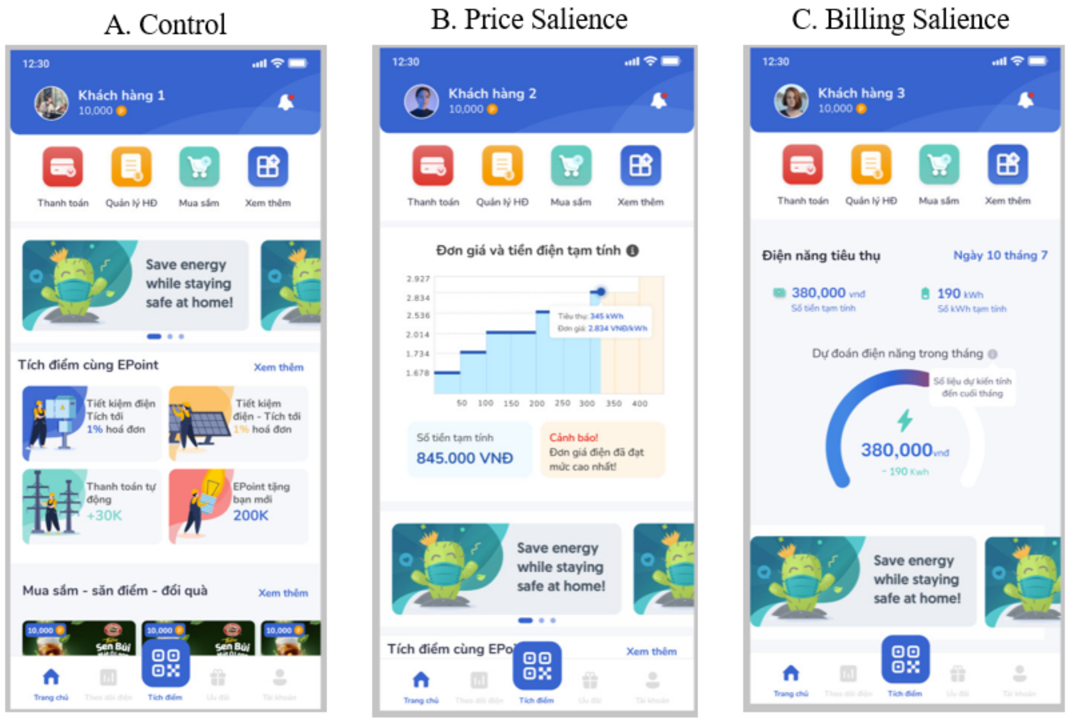
Figure A.1 shows the nonlinear pricing schedule in Hanoi. Figure A.2 shows the app under the control arm (panel A), price information treatment (panel B), and billing information treatment (panel C).

**Figure A.1:** Nonlinear (Increasing Block) Tariff Schedule in Hanoi




*Notes:* Step functions depict the marginal price tier schedule under Vietnam's six-tier increasing block tariff. During the study period, USD 1 ≈ 24,500 VND, so the highest marginal price shown corresponds to ≈ USD 0.128/kWh.

**Figure A.2:** App Interface Variations: Control Group vs. Treatment Groups

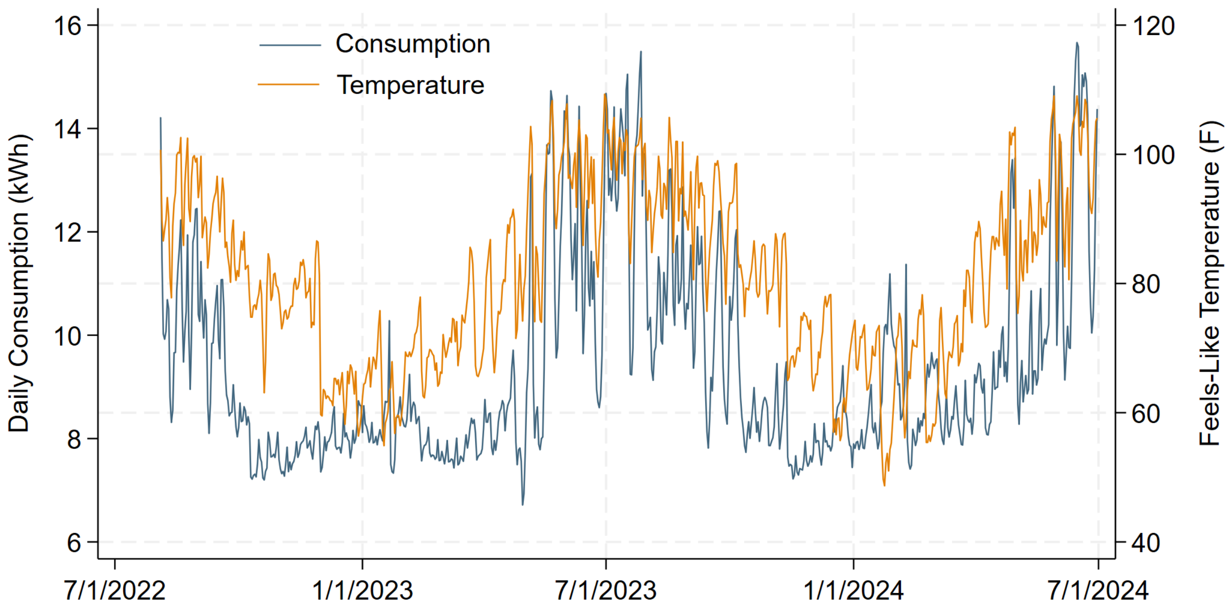



Notes: The control group (Panel A) does not have access to price or usage information on the app's main page. In contrast, the interface for Price Salience Treatment Group (Panel B) has been redesigned for a more visually intuitive understanding of the nonlinear pricing system. The app display for Billing Salience Treatment Group (Panel C) offers users a day-by-day visual of their accumulating energy costs.

**Table A.1:** Summary statistics at the household by day level.

| Variable | Mean | Median | Min | Max | N |
|---|---|---|---|---|---|
| Daily consumption (kWh) | 9.42 | 8.00 | 0.00 | 334.00 | 30,457,445 |
| 1(Checked app at least once in day) | 0.09 | 0.00 | 0.00 | 1.00 | 9,160,416 |
| Mean feels-like temperature (F) | 81.16 | 80.65 | 48.70 | 109.30 | 31,432,800 |

**Figure A.3:** Average Daily Consumption and Feels-Like Temperature during the Sample Period




Notes:This graph displays the average daily electricity consumption (kWh) and feels-like temperature over the sample period from August 1, 2023, to June 30, 2024. The solid blue line represents average daily electricity consumption, while the solid red line indicates the average daily feels-like temperature.

**Table A.2:** Balance Analysis: Mean Comparisons Across Control and Treatment Groups

| Month | Control Mean | T1 Coeff. | T1 P-value | T2 Coeff. | T2 P-value | F-test P-value |
|---|---|---|---|---|---|---|
| Aug 2022 | 10.633 | -.03 | .671 | -.006 | .934 | .904 |
| Sep 2022 | 9.659 | -.06 | .357 | -.008 | .903 | .606 |
| Oct 2022 | 7.838 | -.033 | .521 | -.023 | .657 | .806 |
| Nov 2022 | 7.787 | -.046 | .376 | -.016 | .758 | .667 |
| Dec 2022 | 8.033 | -.017 | .764 | .002 | .976 | .935 |
| Jan 2023 | 8.221 | .006 | .925 | .051 | .4 | .655 |
| Feb 2023 | 8.102 | -.034 | .558 | -.009 | .884 | .83 |
| Mar 2023 | 7.795 | -.02 | .71 | -.01 | .852 | .933 |
| Apr 2023 | 8.311 | -.019 | .741 | -.021 | .706 | .919 |
| May 2023 | 10.769 | -.034 | .622 | -.022 | .751 | .883 |
| Jun 2023 | 11.474 | -.013 | .868 | -.005 | .948 | .986 |
| Jul 2023 | 13.156 | -.05 | .565 | .049 | .578 | .529 |
| Aug 2023 | 10.84 | -.031 | .678 | .012 | .872 | .836 |
| Sep 2023 | 10.17 | .021 | .767 | .048 | .498 | .794 |

Notes: Each observation is a household. The analysis is based on the monthly average of daily electricity consumption. Columns 3–6 report OLS coefficients and $p$-values for treatment group indicators. T1 refers to the price salience treatment and T2 refers to the billing salience treatment. T1 P-value and T2 P-value are the p-values from a test that the T1 coefficient and the T2 coefficient are equal to zero, respectively. Column 7 shows the $p$-value from a joint F-test that both the T1 coefficient and the T2 coefficient are equal to zero.

**Table A.3:** Kolmogorov–Smirnov Tests Across Treatment Groups

| Month | Control vs T1 | Control vs T2 | T1 vs T2 |
|---|---|---|---|
| Aug 2022 | .534 | .886 | .456 |
| Sep 2022 | .62 | .933 | .573 |
| Oct 2022 | .866 | .916 | .409 |
| Nov 2022 | .519 | .974 | .518 |
| Dec 2022 | .821 | .866 | .847 |
| Jan 2023 | .885 | .341 | .758 |
| Feb 2023 | .891 | .544 | .742 |
| Mar 2023 | .995 | .233 | .658 |
| Apr 2023 | .794 | .621 | .811 |
| May 2023 | .927 | .55 | .77 |
| Jun 2023 | .75 | .547 | .968 |
| Jul 2023 | .591 | .526 | .347 |
| Aug 2023 | .996 | .838 | .743 |
| Sep 2023 | .812 | .817 | .912 |

Notes: Each observation is a household. All variables are defined at the household level. T1 refers to the price salience treatment and T2 refers to the billing salience treatment. Each cell reports the $p$-value of a Kolmogorov–Smirnov test comparing empirical distributions across the respective treatment groups.

## A.2 Summary Statistics and Balance

Table A.1 shows summary statistics, where the unit of observation is the household by day. Figure A.3 is a graphic showing average consumption and average "feels like" temperature by date. It shows that consumption and temperature are strongly correlated.

Tables A.2, A.3, and A.4 give information about balance using the pre-period data. Table A.2 shows that within each month, there is no difference between the control mean of consumption and each of the treatment means of consumption. Table A.3 does a similar series of balance tests by month but instead uses a Kolmogorov-Smirnov balance test. Table A.4 shows evidence of balance by consumption bin.

Figures A.4a and A.4b show app checking patterns by day of sample period. Figure A.4a graphs app checking by each treatment group, showing that both treatments and control had similar app checking patterns, although the price salience treatment group has slightly higher app checking rates. Figure A.4b groups app checking by when the households billing cycle starts, where day of billing cycle is based on billing cycle before March 2024, when all households billing cycles moved to align with the calendar month. The graph shows that app checking tends to spike around the time that the bill is due.

**Table A.4:** Distribution of Consumption in Tiers

| Tier | Energy usage (kWh) | Control | | Treatment 1 | | Treatment 2 | | Total | |
|---|---|---|---|---|---|---|---|---|---|
| | | Obs. | % | Obs. | % | Obs. | % | Obs. | % |
| 1 | 0-50 | 11,560 | 6.43 | 11,435 | 6.37 | 11,490 | 6.40 | 34,485 | 6.40 |
| 2 | 51-100 | 13,386 | 7.45 | 13,636 | 7.59 | 13,249 | 7.38 | 40,271 | 7.47 |
| 3 | 101-200 | 43,890 | 24.42 | 43,636 | 24.30 | 44,550 | 24.81 | 132,076 | 24.51 |
| 4 | 201-300 | 45,643 | 25.40 | 45,510 | 25.34 | 44,734 | 24.91 | 135,887 | 25.22 |
| 5 | 301-400 | 29,917 | 16.65 | 29,830 | 16.61 | 29,548 | 16.46 | 89,295 | 16.57 |
| 6 | 401-600 | 25,601 | 14.24 | 25,883 | 14.41 | 26,115 | 14.54 | 77,599 | 14.40 |
| 7 | 601+ | 9,727 | 5.41 | 9,638 | 5.37 | 9,870 | 5.50 | 29,235 | 5.43 |

Notes: This table displays the distribution of consumption across tiers after the randomization, showing the number of observations and their respective percentage share within each tier and group. The unit of observation is household by billing cycle. The sample is limited to those whose billing cycle is in the pre-period, with end-of-billing cycle lying between October 1, 2022 and September 30, 2023, ensuing that we observe the full billing cycle's consumption.

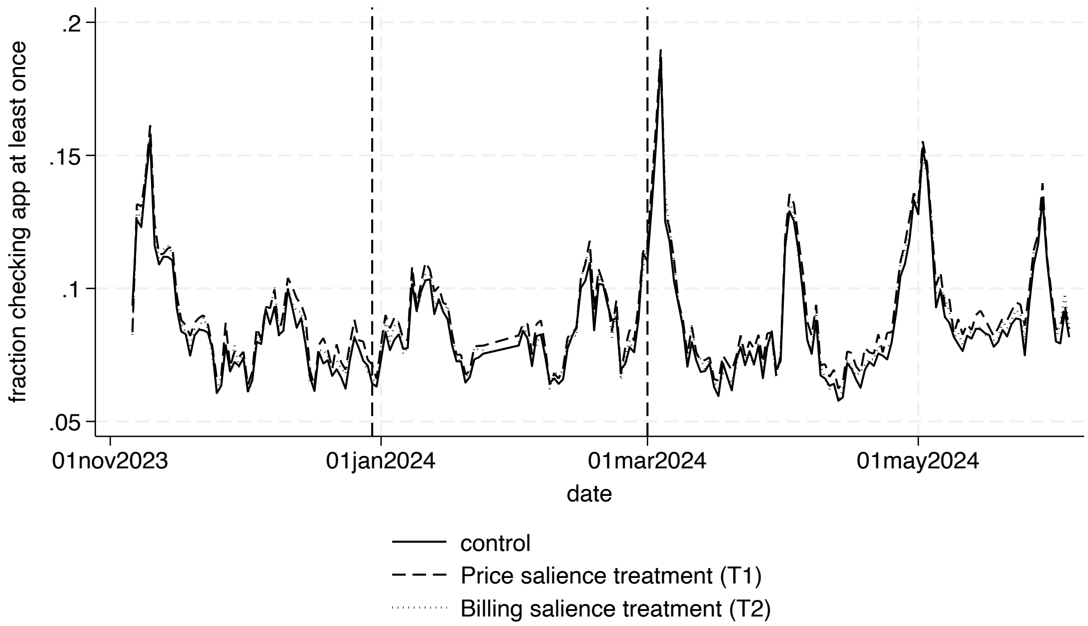
## A.3 Deviations from the Pre-Analysis Plan

The final analysis follows the experimental design, sampling strategy, outcome definitions, and estimation approach specified in the pre-analysis plan. Two deviations arose due to implementation details and data availability. First, the timing of the intervention rollout differed slightly from the period specified in the pre-analysis plan. The pre-analysis plan anticipated that app-based information and accompanying push notifications would be active from September 29 to October 29. In practice, push notifications were sent from October 1 to October 20, after which they were discontinued. These notifications were intended solely to familiarize households with the app interface and pricing displays and were not designed as a distinct behavioral treatment. Consistent with this intent, the analysis does not study push notifications as a separate intervention component.

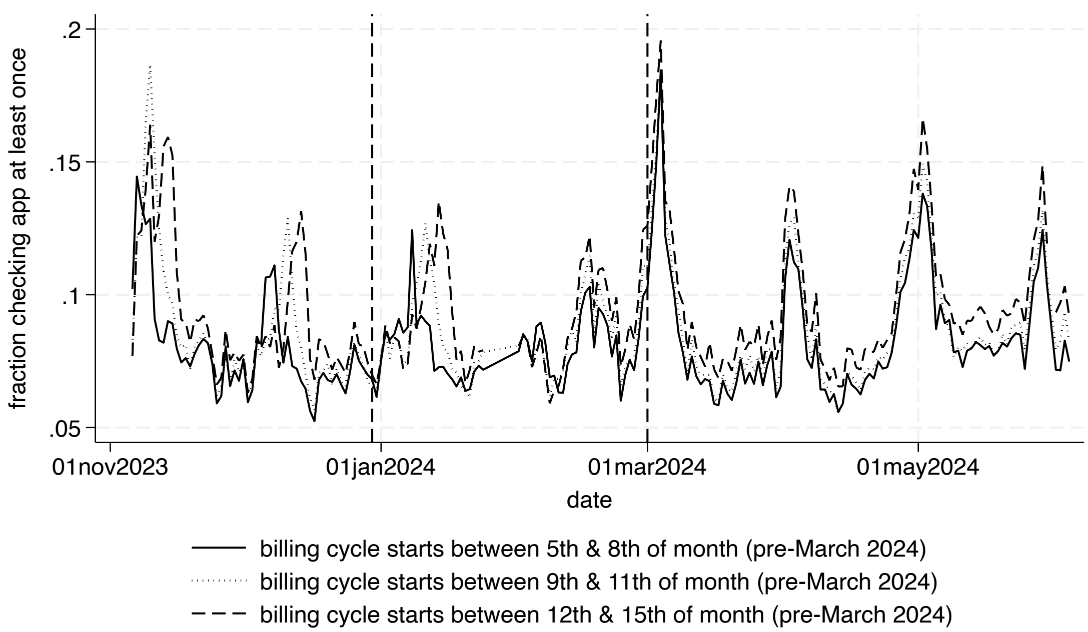
Second, while the pre-analysis plan specified the collection of mobile app usage data, it did not explicitly note that such data could only be collected after participating households were identified and enrolled in the study. App-usage data are available only for the post-treatment period. Consequently, engagement outcomes are analyzed using post-treatment data and are used to study heterogeneity and mechanisms rather than to assess pre-treatment balance.

All other aspects of the analysis were implemented as described in the pre-analysis plan.

**Figure A.4:** App Checking Over Time




(**a**) Fraction of households that check the app by day and treatment status




(**b**) Fraction of households that check the app by day and first day of billing cycle

*Notes:* Panel (a) shows the fraction of households that check the utility's mobile app on each day of the sample period, separately by treatment status. Panel (b) plots app-checking behavior by day of the sample period, grouped by the first day of the household's billing cycle. App usage is highly cyclical, with pronounced spikes at the beginning and end of billing cycles, particularly after March 1, 2024, when billing cycles were aligned to the calendar month.

**Table A.5:** Average Treatment Effects: Compares mid or post with pre-period.

| | (1) IHS | (2) log(1+.) | (3) Poisson | (4) In levels |
|---|---|---|---|---|
| Price information treatment x (mid \| post) | 0.000530 | 0.000641 | 0.00234 | 0.0234 |
| | (0.00371) | (0.00310) | (0.00325) | (0.0302) |
| Billing information treatment x (mid \| post) | -0.00514 | -0.00390 | -0.000623 | -0.00481 |
| | (0.00379) | (0.00315) | (0.00316) | (0.0294) |
| p-value equality | 0.133 | 0.149 | 0.369 | 0.359 |
| R squared | 0.451 | 0.457 | | 0.483 |
| Pseudo R squared | | | 0.297 | |
| Control mean mid and post periods | 2.515 | 2.013 | 9.254 | 9.254 |
| Observations | 30457445 | 30457445 | 30457445 | 30457445 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## A.4 Additional regression results

Tables A.5 and A.6 provide robustness checks for Table 1: In particular, while Table 1 excludes the mid period (Oct 1-Dec 29, 2023) from the analysis, both Table A.5 and Table A.6 include it. Table A.5 examines the treatment effect assuming the same treatment effect for the mid period and the post period. Table A.6 estimates separate treatment effects for the mid period and the post period.

Table A.8 shows a regression using the specification in equation 4, examining heterogeneity in estimated treatment effect as a function of whether the household checks the app at least once during the post period.

**Table A.6:** Average Treatment Effects: Separate treatment effects for mid period vs. post period.

| | (1)<br>IHS | (2)<br>log(1+.) | (3)<br>Poisson | (4)<br>In levels |
|---|---|---|---|---|
| Price information treatment x mid | 0.000176 | 0.000213 | 0.000558 | 0.00844 |
| | (0.00391) | (0.00324) | (0.00343) | (0.0296) |
| Billing information treatment x mid | -0.00278 | -0.00213 | -0.00167 | -0.0130 |
| | (0.00397) | (0.00329) | (0.00329) | (0.0286) |
| Price information treatment x post | 0.000704 | 0.000852 | 0.00310 | 0.0307 |
| | (0.00420) | (0.00351) | (0.00358) | (0.0346) |
| Billing information treatment x post | -0.00630 | -0.00477 | -0.000177 | -0.000775 |
| | (0.00430) | (0.00358) | (0.00353) | (0.0341) |
| p-value equality mid period | 0.455 | 0.477 | 0.514 | 0.469 |
| p-value equality post period | 0.102 | 0.116 | 0.372 | 0.375 |
| R squared | 0.451 | 0.457 | | 0.483 |
| Pseudo R squared | | | 0.297 | |
| Control mean mid period | 2.452 | 1.955 | 8.365 | 8.365 |
| Control mean post period | 2.546 | 2.042 | 9.693 | 9.693 |
| Observations | 30457445 | 30457445 | 30457445 | 30457445 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A.7:** App Checking by Day and Day of Billing Cycle

| | (1) checks app X=2 | (2) checks app X=3 | (3) checks app X=4 | (4) checks app X=5 |
|---|---|---|---|---|
| T1 * 1(days 1-X before next cycle) | 0.00489** | 0.00495** | 0.00476** | 0.00512** |
| | (0.00223) | (0.00215) | (0.00208) | (0.00204) |
| T1 * 1(days 1-X of billing cycle) | 0.00307 | 0.00295 | 0.00314 | 0.00354* |
| | (0.00229) | (0.00218) | (0.00213) | (0.00209) |
| T1 * 1(other days of cycle) | 0.00461** | 0.00468** | 0.00476** | 0.00462** |
| | (0.00188) | (0.00187) | (0.00187) | (0.00186) |
| T2 * 1(days 1-X before next cycle) | 0.00277 | 0.00251 | 0.00246 | 0.00253 |
| | (0.00221) | (0.00212) | (0.00206) | (0.00202) |
| T2 * 1(days 1-X of billing cycle) | 0.00107 | 0.00101 | 0.000885 | 0.00136 |
| | (0.00226) | (0.00215) | (0.00211) | (0.00207) |
| T2 * 1(other days of cycle) | 0.00192 | 0.00196 | 0.00202 | 0.00190 |
| | (0.00185) | (0.00184) | (0.00184) | (0.00183) |
| 1(days 1-X before next cycle) | 0.0262*** | 0.0224*** | 0.0180*** | 0.0144*** |
| | (0.000800) | (0.000706) | (0.000642) | (0.000598) |
| 1(days 1-X of billing cycle) | 0.0335*** | 0.0276*** | 0.0234*** | 0.0201*** |
| | (0.000801) | (0.000682) | (0.000620) | (0.000589) |
| Observations | 9179388 | 9179388 | 9179388 | 9179164 |
| Control group mean | 0.0862 | 0.0862 | 0.0862 | 0.0862 |
| R squared | 0.00612 | 0.00603 | 0.00589 | 0.00580 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Regression results from OLS regression where the dependent variable is an indicator variable for whether a household checks the app on any given date. All regression specifications include date fixed effects. T1 denotes the pricing treatment; T2 denotes the billing treatment. The values of X in the four columns denote window length defining the beginning and end of the billing cycle: The term "days 1-X before next cycle" refers to the end of the billing cycle; the term "days 1-X of billing cycle" refers to the beginning of the billing cycle. Standard errors are clustered by household.

**Table A.8:** Average Treatment Effects by App Checking Status

| | (1) IHS | (2) log(1+.) | (3) Poisson | (4) In levels |
|---|---|---|---|---|
| Checks app x post | 0.0235*** | 0.0195*** | 0.0118** | 0.120** |
| | (0.00607) | (0.00506) | (0.00500) | (0.0478) |
| T1 * post | 0.0108 | 0.00921* | 0.00910 | 0.0852 |
| | (0.00670) | (0.00559) | (0.00594) | (0.0556) |
| T2 * post | -0.00438 | -0.00329 | 0.000575 | 0.00517 |
| | (0.00689) | (0.00573) | (0.00587) | (0.0546) |
| T1 * checks app * post | -0.0183** | -0.0151** | -0.0106 | -0.0983 |
| | (0.00857) | (0.00715) | (0.00741) | (0.0708) |
| T2 * checks app * post | -0.00344 | -0.00265 | -0.00143 | -0.0109 |
| | (0.00877) | (0.00730) | (0.00730) | (0.0696) |
| Constant | 2.551*** | 2.044*** | 2.406*** | 9.553*** |
| | (0.00133) | (0.00111) | (0.00114) | (0.0105) |
| p-value: T1 | 0.162 | 0.185 | 0.740 | 0.764 |
| p-value: T2 | 0.149 | 0.189 | 0.844 | 0.894 |
| R squared | 0.448 | 0.454 | | 0.480 |
| Pseudo R squared | | | 0.298 | |
| Control mean post-period | 2.546 | 2.042 | 9.693 | 9.693 |
| Observations | 26471767 | 26471767 | 26471767 | 26471767 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Notes: Regressions where dependent variable is daily consumption. T1 and T2 are indicators for the price and billing treatments, respectively. Checks app is an indicator for households that check app at least once in the post-period, and post is an indicator variable for dates of Dec 30, 2023 and later. The sample excludes the mid-period (Oct 1-Dec 29, 2023) from the analysis. All specifications include household and date fixed effects. The row "p-value: T1" contains the p-value for a test that the coefficient for T1 * post + the coefficient for T1 * checks app * post is equal to zero. Similarly, the row "p-value: T2" is the p-value from a test that the coefficient for T2 * post + the coefficient for T2 * checks app * post is equal to zero. Standard errors clustered by household.