Incentivized learning and attention-driven treatment effects: a field experiment on energy conservation

Marco Castillo* Ragan Petrie†‡

December 2023

Abstract

We investigate how incentives affect learning when attention is multidimensional. Households are provided high-frequency information on gas usage and/or monetary incentives to reduce energy consumption. Information coupled with incentives leads to lower consumption, and information without incentives leads to higher consumption. Higher consumption persists a year later for those who did not receive incentives. Both groups accessed the same information technology to learn preferences and costs for warmer/colder indoor temperatures yet have different durable treatment effects. Incentives focused learning on cost, rather than comfort - those offered incentives explored colder house temperatures, while those without incentives tried a warmer house. Objective, real-time information can produce opposite behavior, as incentives affect learning.

Keywords: Learning, selective attention, incentives, field experiment, energy usage

JEL codes: D91, D12, C93

*Castillo: Department of Economics, Texas A&M University; Melbourne Institute; IZA; marco.castillo@tamu.edu

†Petrie: Department of Economics, Texas A&M University; Melbourne Institute; rpetrie@tamu.edu

‡Castillo acknowledges the financial support of the National Science Foundation (Award 1522876). We grateful for the support of our partners, ENSTAR Natural Gas and the Alaska Housing Finance Corporation (AHFC), without whom the study would not be possible. The study is registered at the AEA RCT Registry (AEARCTR-0003938) and has human subjects approval from Texas A&M University (IRB2018-1356M). Castillo and Petrie declare no conflicts of interest.

1 Introduction

Incentives drive attention (Chetty et al., 2009; Farhi and Gabaix, 2020; Rees-Jones and Taubinsky, 2020; Bordalo et al., 2013; Handel and Schwartzstein, 2018; Hastings and Shapiro, 2013), but this might be away from preferred alternatives. We present field experiment evidence consistent with incentives driving attention towards cost, rather than comfort, in the context of energy conservation. Consumers provided with real-time, high-frequency energy usage data and a monetary incentive to reduce consumption learn to consume less energy, while consumers with usage data, but no incentive, learn to consume more. A significant gap in behavior between these groups is detectable a year after the experiment. We use process and administrative data to show that this finding is due to differences in experimentation with household temperature settings across experimental groups. Those without conservation incentives were more likely to try out warmer temperature settings, while those with incentives mainly kept their house cold. Both groups had access to the same ex-ante information technology to discern preferences and costs for warmth, but they learned differently during the study. This learning persists a year later.

Imperfect information limits the ability of individuals to make optimal decisions and the effectiveness of policies (Stigler, 1961; Leland, 1968; Sandmo, 1971; Gabaix, 2014; Farhi and Gabaix, 2020; Chetty et al., 2009; Taubinsky and Rees-Jones, 2018; Jessoe and Rapson, 2014; Larcom et al., 2017). Providing relevant information could improve decision-making. However, when attention is endogenous, information may be acquired differently and features of the choice set, i.e. prices, can affect how attention is allocated (Bordalo et al., 2013; Hastings and Shapiro, 2013; Handel and Schwartzstein, 2018). Individuals might pay attention to different aspects of a decision problem (Hanna et al., 2014), leading to persistent differences in behavior and outcomes. Whether due to friction or mental gaps (Handel and Schwartzstein, 2018), there is no guarantee that the same information will have the same effects, and thus has implications for policy design and targeting. Our results are consistent with a model of multidimensional attention whereby incentives focus learning in one dimension, at the expense of learning in others.

We use a field experiment in the context of energy conservation to explore how learning and behavior can take different pathways in the presence of relevant high-frequency information on usage and monetary incentives. We only provide usage information so we can focus on studying how objective information on behavior affects decisions. Other information, such as social comparison or suggested actions, is omitted. Experimental research on household energy usage has shown that access to easy-to-use information can significantly increase price responsiveness (e.g. Jessoe and Rapson, 2014; Harding and Lamarche, 2016) and access to more frequent information on usage can affect consumption (e.g. Houde et al., 2013; Gans et al., 2013; Martin and Rivers, 2018; Asmare et al., 2021).[1]

Economic theory is agnostic about whether information would increase or decrease energy consumption and whether it might produce different directional effects when interacted with incentives. Consumers might overestimate the marginal cost of a one degree change in home temperature.[2] Information showing the cost is lower than anticipated could lead to consumers increasing energy usage. There is a growing number of products that give consumers access to energy usage information and tools to manage usage, suggesting information is helpful for conservation, but this might not always be the case.[3]

[1] Many studies combine different types of information (i.e. usage feedback, audits, energy-savings tips, ways to monitor or control usage, social comparison to neighbors, etc.). There is a robust literature that provides social comparison information to consumers and examines its effect on energy consumption (e.g. Schultz et al., 2007; Allcott, 2011; Allcott and Rogers, 2014; Costa and Kahn, 2013; Ferraro and Price, 2013; Burkhardt et al., 2019; Byrne et al., 2018; Fang et al., 2023). Meta-analyses that combine studies using all types of information interventions suggest an energy consumption reduction of 5-7% (Nemati and Penn, 2020; Delmas et al., 2013).

[2] Attari et al. (2010) document the public's misperceptions of energy consumption and savings.

[3] Programmable thermostats and Nest thermostats are two popular examples. In a well-powered field experiment, Brandon et al. (2021) find smart thermostats yield a null effect on energy consumption because humans undo the energy savings programmed by engineers. Buchanan et al. (2015) argues that energy usage feedback may not be welfare-enhancing.

To systematically examine the effects of information and incentives on learning, we use a 2×2 experimental design. One arm varies whether or not participants receive high-frequency information on usage via a web portal. This is crossed with another arm that varies whether or not participants receive monetary incentives to reduce energy consumption. The design allows us to see the effect of information, with and without incentives, to understand if learning is different in the presence of incentives. The process and administrative data we collected allow for an examination of how differential learning explains behavior.

The field experiment was conducted in the city of Anchorage, Alaska during the winter months and focuses on gas consumption for home heating. In Anchorage, households receive information on monthly gas consumption at the end of the billing cycle. High-frequency usage information was novel. In the information treatments, gas consumption and indoor temperature were read by an installed device and made available to participants via a web portal in real time. The portal displayed energy usage and costs in five different panels. One panel displayed energy cost and inside and outside temperature to help participants learn how energy costs depend on this difference.[4] The web portal was designed by us to track usage and provides a measure of the demand for information.

The incentive arm of the field experiment was implemented as a monetary bonus. For every 1% reduction in gas usage during the study period, relative to the same period in the prior year, was awarded \$10. The bonus was capped at \$100, or a 10% reduction in usage, to mitigate extreme reductions in home temperature.

We partnered with ENSTAR, the company that provides gas service to the Anchorage Bowl area, and the Alaska Housing Finance Corporation (AHFC) to implement the field experiment. ENSTAR contacted its customers by email and invited them to participate in a study on energy usage. Interested individuals completed

[4] Engineering models of indoor heating show an almost linear relationship between energy use and the gap between outside and inside temperature (ASHRAE, 2017).

an online signup survey that was used to determine study eligibility. All eligible households were randomized into one of the four treatments, and 550 household participated in the study. ENSTAR provided administrative data on monthly gas usage prior to, during and after the intervention. We use these data to test for, and reject, selection effects into the study and attrition bias.[5] Not all gas readers were successfully installed in households, so we use assignment to treatment as an instrument and estimate difference-in-difference effects using treatment-on-the-treated, as well as intent-to-treat. We investigate distributional, average and median effects. Results are robust across all analyses.

We have several key findings. First, we document large heterogeneity in the marginal cost to heat a home, reflecting variation in home energy efficiency. Uncertainty of these costs is a necessary condition for information to have any value to a household. Second, access to high-frequency information on gas usage decreased energy consumption when coupled with monetary incentives but increased energy consumption when incentives were absent. This is confirmed by significant differences in the distributions, average and median of year-over-year change in consumption. Third, the additive effect of information to incentives is null. If information helps consumers learn costs or better monitor consumption, we would have expected the additive effect of information to be negative. This hints that learning was different with and without incentives. Fourth, those who focused learning on comfort, and thus used more energy, during the experiment continued to use more energy a year later.

How do we understand our findings? Both groups with access to high-frequency information on usage had the potential to learn the same during the study, but they did not. We find no evidence this is due to inattention. Instead, the result is consistent with participants focusing on different aspects of the decision, i.e. cost and comfort,

[5]This is a natural concern since our field experiment is what Harrison and List (2004) call a *framed field experiment*. In this context, households select to be participants in the study.

rather than not paying attention at all, and is predicted by a behavioral model of attention (Farhi and Gabaix, 2020). Indeed, the divergent consumption behavior across the two groups can be explained by different exploration of home temperature settings. Those with incentives to conserve were more likely to choose low home temperature settings during the study. But those with no incentives were more likely to keep the house warmer, and this behavior persist a year later. This highlights the durable effect of learning about comfort on future behavior, even after information on usage is no longer available.

In our setting, information does not have a main effect (Muralidharan et al., 2019), but rather interacts with incentives via differential learning to produce both an increase and decrease in usage. That is, information does not always move behavior in the same direction. Our experimental design, which fully interacts information and incentives, uncovers this finding and adds to the related literature (Battalio et al., 1979; Jessoe and Rapson, 2014).[6] Our treatments are run in the same population and thus allow for balance on observables and differences in pre-existing levels of uncertainty or tolerance to it.[7]

Through insight into individuals' demand for information, via the web portal, we provide a more complete picture of the effect of information on decisions. To our knowledge, this has not been studied in the context of energy usage.[8] Information theories predict that the demand for information should be increasing in its impact. We test this with our process data on web portal usage. A telltale sign

[6]Given our focus on learning with and without incentives, our experimental design is closest to Battalio et al. (1979) who have a sample size 100 households across five conditions. Our study sample size is large enough to allow detection of a failed replication, as suggested by Simonsohn (2015) who proposes that a replication should use a sample size two and a half times the original. Our sample size exceeds that threshold.

[7]The literature on precautionary saving Leland (1968) suggests that individuals might consume less due to income uncertainty. This has an analogue in energy consumption if individuals are uncertain of the costs associated with different actions.

[8]Jessoe and Rapson (2014) collected survey responses on the frequency of use of the device installed in the households. They show that more frequent reported use is associated with higher price responsiveness as predicted by theory. We build on their approach by obtaining field measures of demand for information from process data.

that information can have both positive and negative effects on conservation is that the gap in behavior across treatments is increasing in attention. This is what we find. Our field measures of attention provide an illustration of how predictions of rational inattention models (Mackowiak et al., forthcoming) can be tested and relate to the burgeoning literature on attention and costly information acquisition (Caplin, 2016; Gabaix, 2019). Our results suggest caution is needed in extrapolating results from information interventions without an understanding of information acquisition or demand.

Finally, we contextualize our results and how they generalize to other settings by considering the SANS conditions (List, 2020).[9] In terms of selection, our sample includes residents of single-family homes with broadband service, a dedicated meter and at least one year in residence in Anchorage, Alaska. Anchorage has long winters and would be most comparable in energy usage to locations that rely on air conditioning to keep cool during long summers. In terms of attrition, compliance with treatment was incomplete, and this is accounted for in our analysis. Our main outcome variable on energy usage is from administrative data and covers all study participants. In terms of naturalness, participants made decisions over energy usage in a setting in which they naturally engage. Finally, the experiment was designed based on scaling, with consideration of cost and availability of high-frequency gas readers and installation. Our key results are WAVE2 insights (List, 2020), since we replicate Battalio et al. (1979). This is discussed further in Section 7.

The paper is organized as follows. Section 2 provides a behavioral decision framework. Section 3 describes the experiment and Section 4 the field setting. Section 5 presents results on heterogeneous energy costs, intent-to-treat and treatment-on-the-treated effects on energy usage. Section 6 discusses mechanisms, and Section 7 concludes.

[9]SANS is for Selection, Attrition, Naturalness and Scaling.

2 Behavioral decision framework

Before we describe the field experiment design, we present a simple model of endogenous attention and energy consumption. This frames the decision problem faced by individuals with multidimensional attention on learning their preferences for a warm house and the cost to provide warmth. The model follows the approach proposed by Farhi and Gabaix (2020) who analyze welfare implications of behavioral biases and captures singular testable implications that models without endogenous attention do not predict. A full description of the model is in Appendix A.

We distinguish between perceived utility u_s and real utility u. Let the real utility of household temperature τ given income y be equal to $u(\tau, c) = \frac{\theta}{1-1/e}\left(\frac{\tau}{\theta}\right)^{1-1/e} + y - p\tau$ where p is the marginal cost of an extra degree of warmth. c is consumption and substituted out with the budget constraint, $y = c + p\tau$. Parameter θ captures heterogeneous preferences for warmth and parameter e represents the price elasticity. We assume a constant elasticity representation for clarity.

Consumers may perceive their preferences for warmth and the marginal cost of warmth as different to what they are, i.e. $\theta_s \neq \theta$, and $p_s \neq p$. For given parameters θ_s and p_s, perceived utility is defined by $u_s(\tau, c) = \frac{\theta_s}{1-1/e}\left(\frac{\tau}{\theta_s}\right)^{1-1/e} + y - p_s\tau$. We represent utility this way because we do not know ex-ante if consumers in the experiment will learn their true preferences or the true cost of consumption.

To allow for endogenous attention and learning, we define a_θ as the amount of attention a consumer devotes to learning her preferences and define a_p as the amount of attention a consumer devotes to learning about price. We assume that $a_i \in [0, 1], i = \theta, p$. We also assume that attention is costly and equal to $\kappa(a_\theta + a_p)$. We make this assumption for three reasons. First, as shown by Gabaix (2014), this function allows for a corner solution in which an individual pays no attention. Second, as shown in the empirical analysis, we do not have separate proxies for these two types of attention. Last, the revealed preference intuition we present is independent of this latter assumption.

We assume that perceived parameters are a linear combination of real parameters and default values and that attention is multidimensional. In particular, perceived preference is $\theta_s(a_\theta) = \theta_d + a_\theta(\theta - \theta_d)$, where a_θ is the attention placed on preferences, θ_d is a default preference and θ is the real preference parameter. The perceived price is $p_s(a_p) = p_d + a_p(p - p_d)$, where a_p is the attention placed on price, p_d is a default price and p is the real price. If the consumer is fully attentive in both domains, i.e. $a_\theta, a_p = 1$, they respond to real preferences (θ) and price (p). We let $\tau_s(\theta_s(a_\theta), p_s(a_p)) = \theta_s p_s^{-e}$ be the optimal household temperature choice given perceived preferences and price.

In the case of new information or changes in the cost of attention, consumers allocate attention to preferences and price to maximize real utility knowing that their own behavior will follow perceived parameters rather than real parameters. The intuition is that consumers will develop attention-constrained optimal choices. Given an amount of attention (a_θ, a_p), a behavioral-constrained individual will choose according to $(\theta_s(a_\theta), p_s(a_p))$, the perceived parameters, not the real ones. So, an omniscient ego will maximize real utility in anticipation of constrained decisions. In particular, attention will be allocated to solve:

$$max_{a_\theta, a_p} \theta^{1/e} \frac{1}{1-1/e} \tau_s(\theta_s(a_\theta), p_s(a_p))^{1-1/e} - p\tau_s(\theta_s(a_\theta), p_s(a_p)) - \kappa(a_\theta + a_p) \qquad (1)$$

Equation (1) shows that the consumer chooses the amount of attention to place on learning preferences and price by maximizing real utility under the assumption that later decisions will be made based on perceived parameters not real parameters, i.e. according to $\tau_s(\theta_s(a_\theta), p_s(a_p))$. Default parameters θ_d and p_d are assumed to be the result of previous attention allocation decisions. We do not assume that these default parameters are correct, but rather the individual's model of the world evolves as circumstance change.

The solution to this maximization problem produces several insights. First, in the case of over-consumption, a consumer will place more attention on the price of warmth if that price increases.[10] At the same time, a price increase can crowd out attention placed on learning one's own preferences for warmth. Second, in the case of under-consumption, a consumer will place more attention on preferences for warmth and less on price. There is less incentive to update information on cost that might lead to even lower consumption. While the model is stark, in that the consumer might pay attention to one dimension and not the other, it highlights the potential differential effect of access to information when attention is multidimensional.

2.1 Hypotheses

Our experiment manipulates p via monetary incentives to reduce gas consumption and provides high-frequency information on usage to affect learning and, therefore, attention, κ. The main predictions from the model for the field experiment are as follows.

Hypothesis 1: Attention is driven by treatment incentives. Manipulating p focuses attention on learning about price and leads to a reduction in energy usage. Without a change in p, attention could be devoted to learning about preferences and price, and thus, energy usage could increase without incentives. In other words, in the same high-frequency information environment, learning can differ with and without incentives. We note that a reduction in energy usage would occur with a change in p, even without high-frequency information, but nothing would be learned about the price of warmth.

Hypothesis 2: After high-frequency information and monetary incentives (p) are

[10]Over-consumption means that at real utility the marginal utility of consumption is smaller than marginal cost. Under-consumption means that at real utility the marginal utility of consumption is larger than marginal cost.

removed, behavior learned with incentives could revert to old patterns in the long-run, but behavior learned without incentives persists. This occurs because, when there are incentives to reduce energy usage, the consumer does not learn how to improve comfort. However, those who have no incentive to pay attention to price may instead pay attention to comfort. Such learning would persist.

Hypothesis 3: If attention is unidimensional, by contrast, providing high-frequency information has the same directional learning effect, with or without monetary incentives. Learning about cost and preferences for warmth would be the same in both treatments.[11] This would be consistent with Jessoe and Rapson (2014) where real-time information on energy usage increased price responsiveness.

Hypothesis 4: Treatment effects should be larger among those paying more attention (a). The model comparative statics show that treatment effects increase as more attention is paid to learning preferences and price.

Finally, we note that endogenous attention can manifest as a change in the distribution of consumption. Thus, our analysis focuses on average changes in consumption, as well as changes in the distribution.[12]

[11] Gabaix (2014); Farhi and Gabaix (2020) show that price responsiveness will increase with reductions in attention costs.

[12] In the case of constant price elasticity, and absent endogeneous attention, we expect a change in the log of indoor temperature ($\Delta\tau$) to be a function of a change in the log of prices (Δp), i.e., $\Delta \log \tau_i = \epsilon_i \Delta p$. This suggests that treatment effects manifest as a shift in the distribution of indoor temperature and costs. In the case of endogenous attention, consumers react to changes in perceived prices Δp^s, in particular, $\Delta p^s = \Delta p + \Delta m(\Delta p, \epsilon_i)$, where $m(\Delta p, \epsilon_i)$ is the level of attention to prices if $p^s = m(p, \epsilon_i)p$., which are in turn a function of ϵ_i and real prices. In this context as well, treatment effects could have distributional effects on consumption (see Taubinsky and Rees-Jones, 2018; Farhi and Gabaix, 2020).

3 Field experiment

3.1 Design

The field experiment is designed to test the effects of high-frequency, energy-usage information and conservation incentives on energy consumption of natural gas.[13] All study materials are in Appendix B. The four treatments are based on a 2×2 factorial design that varies access to information and incentives. The design allows us to see the effects of incentives absent information and information absent incentives. The Control group has no information and no incentives, Incentive Only has incentives, Information Only has information, and Information and Incentive has information and incentives.

The high-frequency usage information is displayed in real-time via an online dashboard. The dashboard shows five panels of information: instantaneous usage (CCF) displayed in 5-minute intervals, cumulative usage (CCF), 5-minute interval usage displayed in cost ($), cumulative usage in cost ($) and a panel that shows instantaneous usage, indoor temperature and outdoor temperature. The dashboard was designed to provide the user relevant information on usage and cost without guiding learning in any particular direction. The user can click to change the time frame displayed (i.e. day, week, month, year) and date range for all panels.[14] Interactions with the dashboard are recorded through logins and clicks on dashboard panels. A self-installed device in the participant's home reads the gas meter and indoor temperature in 5-minute intervals.

The incentive to reduce gas usage is based on a year-over-year decrease in consumption. For every 1% reduction in usage during a four-week period, relative to the same period in the previous year, a participant is paid $10. Payment is capped at

[13] The study is registered at the AEA RCT Registry (AEARCTR-0003938) and has human subjects approval from Texas A&M University (IRB2018-1356M).

[14] The dashboard is personalized with the user's previous year energy consumption. This is displayed if the user clicks month or year.

$100 (or a 10% reduction in usage). The incentives were calibrated based on evidence that a one-degree reduction in indoor home temperature would reduce consumption by 2-3% and that price responsiveness is low (ASHRAE, 2017). These design elements limit the incentive for an extreme reduction in indoor home temperature yet provide scope for differential response across participants.

3.2 Implementation

The field experiment was conducted in Anchorage, Alaska in the winter of 2019. We partnered with ENSTAR, the company that provides gas service to the Anchorage Bowl area, and the Alaska Housing Finance Corporation (AHFC).[15] Figure C.1 provides a graphical representation of the construction of the sample.

In early January, ENSTAR sent an email to all of its customers in the city of Anchorage with an email address inviting them to participate in a study on energy usage in Anchorage. In anticipation of possible selection and/or attrition issues, the email invitation randomized the advertised study completion payment, either $40 or $60. To sign up, interested individuals completed an online signup survey that included questions on household characteristics and was used to determine study eligibility. The signup survey was completed by 1,566 individuals, giving a response rate of 4% to emails sent and roughly 22% of emails opened.[16] Those who were offered a study completion fee of $40 or $60 were equally likely to complete the signup survey.

To be eligible for the study, participants needed to have broadband home internet service, a dedicated and new generation gas meter for the residence, resided at the house for at least a year and reside in a single-family home. The first two requirements

[15]The Alaska Housing Finance Corporation (AHFC) oversaw the Home Energy Rebate Program (HERP) in 2008-18, an incentive program that provided households rebates for making home improvements that increased energy efficiency. It is a respected organization in Anchorage specializing in home energy efficiency.

[16]The email campaign did not record whether or not the email was opened. We assume a 17% open rate, which is in the middle of the range of open rates across industries, https://mailchimp.com/resources/email-marketing-benchmarks/. Thus, roughly 7,140 individuals opened the email.

Table 1: Balance across treatments

	Signed-up	Assigned to treatment	Treatments: Control	Information Only	Information and Incentive	Incentive Only	F-test (p-value)
Household size	2.9	3.0	2.9	3.0	3.1	3.0	0.646
# Children	0.9	0.9	0.9	0.9	1.0	0.9	0.998
Years in residence	5.4	6.3	5.8	6.7	6.5	6.2	0.240
Household Income	90,363	97,185	97,735	96,100	100,803	96,851	0.686
Have prog thermostat	72.4	78.4	75.7	78.0	82.5	79.1	0.537
HERP	22.1	35.3	35.4	44.0	40.9	36.8	0.598
Avg temp (M-F, 8am-5pm)	n/a	66.3	66.5	66.6	66.2	66.1	0.761
Avg temp (M-F, 5-10pm)	n/a	68.3	68.6	68.4	68.2	68.1	0.236
Avg temp (M-F, after 10pm)	n/a	66.0	66.2	66.2	66.1	65.8	0.653
Avg temp (Sat-Sun, day)	n/a	68.0	68.2	67.9	67.9	67.9	0.692
Avg temp (Sat-Sun, night)	n/a	66.2	66.3	66.8	66.3	65.8	0.372
Year built	1981	1981	1982	1980	1981	1980	0.605
Home value	352,903	351,273	353,516	365,387	357,376	347,330	0.839
Avg Jan bill (reported)	177.3	186.8	186.1	205.8	192.1	185.9	0.564
Feb-March 2018 bill (admin)	254.8	378.1	384.9	388.7	382.8	363.7	0.516
Obs	1,566	652	181	50	137	182	

Notes: Signed-up is the sample of participants that completed the sign-up survey. Assigned to treatment is the sample of participants that were eligible for the study. Treatments include all participants in the estimation sample (n=550). Household size (adults and children), # children, years in residence, household income and having a programmable thermostat are from the signup survey. Participation in the Alaska Home Energy Rebate Program (HERP) and reported average January gas bill are from the confirmation survey. Year built and home value are from tax records. Feb-March 2018 gas bill is from ENSTAR administrative records. Average temperature in the house (Avg temp) are reported temperature settings from the confirmation survey. The F-test is the joint test that the treatments are jointly zero.

were needed for the gas reader to work properly.[17] A year of residency was needed to be able to implement the incentive treatments, and a single-family home provided more variation in energy usage and scope for changes. There were 652 households that fit the eligibility requirements.[18] All of these households were randomly assigned to one of the four treatments.

Some assigned households dropped out prior to the start of the study, giving an estimation sample of 550 households assigned to: Control (181), Incentive Only (182), Information Only (50) and Information and Incentive (137). Table 1 shows

[17]The ENSTAR gas meter needed to be a new generation model for the reader to work. ENSTAR was in the process of replacing the meters, but not all eligible households had one installed.

[18]Reasons households that completed the signup survey were ineligible are: 41 had a shared meter, 44 did not have broadband internet, 9 did not live in the Anchorage bowl area, 204 had lived less than a year in the house and 251 were not in a single-family home. Another 365 could not be included in the eligible sample because the home did not have the new generation gas meter needed to work with the reader.

the treatments are balanced on covariates.[19] Attrition rates are significantly different across study completion incentives (18.9% if \$40, and 12.3% if \$60), thus, all estimates are adjusted using inverse probability weighting that includes treatment and study completion payment assignment. Both are valid for weighting as they are randomly assigned to households and orthogonal to one another.

The number of households assigned to each treatment is not equal. This is for two reasons. First, the high-frequency readers were expensive and we had a limited number.[20] Second, more readers were allocated to the Information and Incentive treatment to make comparison with the Incentive Only treatment more precise. A priori, we anticipated a small difference in treatment effects between the Information and Incentive and Incentive Only treatments. Because the relative treatment sample size is inversely related with the desired detectable effects (i.e., $\frac{n_1}{n_2} = \frac{\delta_2}{\delta_1}$, where n_i's are optimal sample sizes and δ_i's are minimum detectable effects, see List et al. (2011)), a larger sample was devoted to the Information and Incentive treatment, relative to the the Information Only treatment. Ex post, the experiment data confirm our presumption.

Households assigned to the treatments with information were mailed the high-frequency reader to self-install in their house. Instructions were included in the packet, and the research team troubleshot installation issues both on the ground, over email and on the phone. Installation was complete when the participant logged into their account on the online dashboard. Despite our best efforts to install all the readers, some were never installed due to technical issues or participants not setting them up. In the end, 124 readers were installed, 34 (68.0%) in Information only and 90 (65.7%) in Information and Incentive. In the Results section, we report analysis of

[19] Table C.1 shows treatments are balanced on covariates for the 652 households that were originally assigned to treatment. We do not have covariate or energy usage data on the broader ENSTAR population, so we cannot speak to the comparison with our sample.

[20] The cost of each reader was about \$200, including parts and assembly, and our budget allowed for purchase of 200. Some readers did not work, and we were left with 187 usable readers for the study.

intent-to-treat, treatment on the treated and treatment noncompliance. In addition, the analysis includes permutation tests, recommended by Young (2019), to test the robustness of our results, given sample sizes and attrition. All results are robust to these analyses.

The initial study period was 28 days. Gas usage was collected during this period and compared to the same time period the previous year to examine year-over-year changes in gas usage. This period was Feb 1-28 for the Control and Incentive Only conditions. Because of delays getting the reading devices installed, the period is Feb 9-March 8 for the Information Only and Information and Incentive treatments. We adjust for these different periods in the analysis of year-over-year changes.

A second 28-day study period was implemented March 9-April 5 and announced by email on March 8 to all participants. Participants were not aware of this ahead of time. In the incentive treatments, participants were eligible for an additional 28-day incentive period – any incentives earned in this second period would be added to any incentives earned during the first study period. The incentive conditions were the same as the first period but with the comparison of usage to the corresponding previous year period. All participants were reminded of the requirement to complete the endline survey in April to receive the study completion payment ($40 or $60).[21]

After the second study period concluded, participants were sent a final email providing a link to complete the endline survey. Of the 550 participants in our estimation sample, 86.2% completed the endline survey, 89.5% in the Control group, 91.2% in Incentive Only, 72.0% in Information Only and 80.3% in Information and Incentive. Upon completion of the endline survey, participants were paid their study completion payment and, in the incentive treatments, any incentive payments earned via Paypal or an Amazon gift card.

[21] The second 28-day study period was planned as a replication of the first 28-day period. It was not pre-announced to participants to avoid biasing behavior in the first period. We find no different across periods in treatment effects and pool both periods in the analysis.

3.3 Data sources

Our analysis is based on four sources of data.

The first source is three surveys completed by our participants: a signup survey, a confirmation survey and an endline survey. In the signup survey, participants gave consent to be part of the study and allow ENSTAR to share their monthly gas records, and they provided household information. Eligible households were assigned to treatment and asked to complete the confirmation survey to confirm participation in the study, provide payment information and report typical household indoor temperature. In the endline survey, participants reported on any changes to the house during the study period.

The second source of data is from the high-frequency readers and web portal. The readers provided real-time data on gas usage and indoor temperature. The portal recorded clicks and logins. These data are only available for those in the information treatments and only for the study period of Feb 8-April 5. We do not have these data for pre-study or post-study periods. Because Control and Incentive Only households did not have a reader installed in their homes, none of these data are available for these groups.

The third source of data is from ENSTAR gas company. Monthly gas usage data was provided for each household that completed the signup survey. These data are from one year prior to the study, the study period and one year after the study, i.e. Jan 2018 - April 2020. ENSTAR has flat rate pricing, not tiered, so monthly gas usage reflects cost.[22] The fourth source of data is tax records. Participant addresses were matched to tax records to obtain the year the house was built and its assessed value.

[22] https://www.enstarnaturalgas.com/rates-regulatory/

4 Field setting

Anchorage, Alaska has long, cold winters and mild summers. The average annual number of heating degree days is over 10,000, compared to about 4,500 for the lower-48 states (ACRC, 2022; EPA, 2022). Over 93% of households in our sample use natural gas as the primary source of heating. Households in Anchorage spend an average of 4.7% of their income on home energy use, compared to 3.1% in the lower-48 states (Saylor et al., 2008; DOE, 2022). Energy costs are sizeable, and this comes almost entirely from heating in the roughly six months of winter. Our field setting is comparable to other settings where summers are long and hot and energy is used to cool homes.

All the homes that fit the eligibility requirements are single-family. Each household has an average of 0.9 children and a total of three members, i.e. adults and children (Table 1, column 2). Average reported household income is $97,184. This is higher than median household income for Anchorage, $84,928 (Census 2010), and likely reflects that our sample is primarily homeowners (95%) and income in our survey was self-reported in categories. Seventy-eight percent of households have a programmable thermostat, and 35% participated in the Home Energy Rebate Program (HERP). The average home was built around 1981 and has an estimated property value of $351,273. Feb-March 2018 gas bills were on average $378, according to administrative records from ENSTAR.

5 Results

The main analysis focuses on the effects of treatment assignment on gas consumption (intent-to-treat) and causal estimates of the treatment on consumption (treatment on the treated). This analysis allows for examination of the hypotheses generated by the decision model. In the main analysis, we pool the data from both study periods.

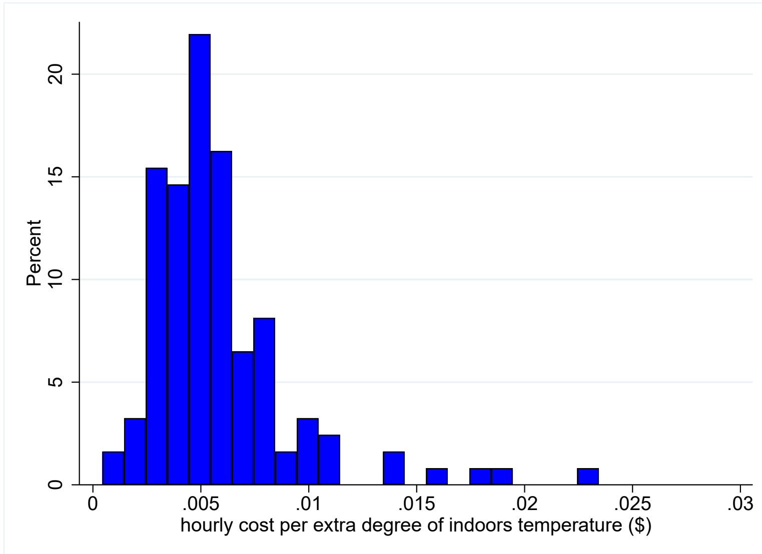
To have an apples-to-apples comparison of household behavior across treatments, we use the same time frame for all households, Feb 1-April 5, to examine energy usage. The results are robust to other pooling criteria and separate period analysis.

5.1 Evidence of heterogeneous energy costs

Before we turn to treatment effects, we first establish that there is scope for participants to learn marginal cost (i.e. p in the decision model in Section 2). To do so, we use data on indoor and outdoor temperature from households in the information treatments to estimate individual-level cost parameters. Figure 1 presents the distribution of these individual-level estimates.[23] Because the estimated cost function is linear (Figure C.5), relative cost savings depend only on the difference between indoor and outdoor temperature. The average hourly cost of an extra indoor degree of warmth, relative to outdoor temperature, is \$0.006 (s.d. \$0.004). This means that, if the indoor temperature is 70 degrees while the outdoor temperature is 30, the cost per hour would be \$0.24 (s.d. \$0.16). In this case, a reduction of the indoor temperature by one degree would reduce the hourly cost by 2.5%.

Figure 1 is based on information from a selected sample: those who have a working reader. To assess the representativeness of these data, we create proxies of the slope of the cost function using administrative data prior to the experiment. The expected cost of gas is $a_i + b_i(\textit{indoor temperature-outdoor temperature})$. If the average indoor temperature is constant across months, we can exploit monthly outdoor temperature variation to approximate b_i by the formula $\frac{\Delta \textit{montly cost}}{\Delta \textit{outdoor temperature}}$. The coefficient of variation of the slope parameter using data collected using our readers is

[23] Visual evidence of the linearity of the cost function is illustrated in Figure C.5. The figure shows the relationship between the outdoor and indoor temperature gap and cost using data from all readers. To estimate individual slope parameters, we run a fixed-effect regression with interaction effects on the parameter associated with the difference between indoor and outdoor temperature. Estimates are restricted to the period Feb 8 - April 5, 2019, when all readers were activated. To avoid extreme values, we restrict the difference between indoor and outdoor temperature to be between 20 and 60 degrees. This is 94% of the sample.



Notes: The figure shows the distribution of estimated individual-level cost parameters.

Figure 1: Heterogeneity in costs (slope parameters)

0.6481. The coefficient of variation of the slope parameter using administrative data is between 0.5544 and 0.7004 depending on the months of data used. This suggests that heterogeneity of costs captured using our readers is a good representation of the heterogeneity of costs in the entire sample.

We assess whether there is selection into the information treatments based on costs by testing if the distribution of slope parameters is different between those who had a working reader and the remainder of the sample. The p-value of the Kolgomorov-Smirnov test of equality of distribution ranges from 0.671 to 0.998 for several different models to estimate marginal costs. This lack of evidence of selection is consistent with the hypothesis that the average participant was not aware a priori of the relationship between marginal cost and indoor temperature.

5.2 Methodology

5.2.1 Estimating intent-to-treat effects

The two outcome variables we use are the difference in logs of total consumption during the period Feb 1 - April 5 for the year of the intervention and the previous year and the difference in logs of consumption in the year after the intervention and the year previous to the intervention.[24] This is equivalent to a difference-in-difference estimator and recommended by Gerber and Green (2012) for more precise estimates.

We do not use high-frequency data as our outcome variable for two reasons. First, we do not have these data for all treatment groups. Second, we only have consistent data for the study period. There is some data for a 1-2 week pre and post study period, but this is not uniform or for all households in the information groups. Thus, for our main results, we use aggregate consumption during the period because it is available for all groups and all periods. Also, we cannot look at pretrends because we do not have consistent high-frequency data pre study. We note the average Feb-March 2018 bill from administrative data is balanced across treatments (Table 1).

The two outcomes are used to identify average and median intent-to-treat (ITT) effects for each treatment group. Equation (2) serves as our baseline estimation equation, where $\Delta y^{\alpha}_{i,t}$ represents the mean or $\alpha-$quantile of the distribution of changes in gas consumption and Z_i represent treatment assignment.

$$\Delta y^{\alpha}_{i,t} = \beta_0 + \beta_1 Z_i + \varepsilon_{i,t} \tag{2}$$

Regressions are performed jointly on all treatments unless explicitly mentioned, and estimations include inverse probability weighting to account for attrition. Median regressions are reported, in addition to average, to assess distributional effects since these are possible due to endogeneous attention (see Section 2 for discussion).

[24]Because of different billing cycles, we calculate daily consumption for the cycle and multiple by the number of days in the cycle within the Feb 1 - April 5 period.

Quantile regressions are also used to estimate average treatment effects under the assumption of rank similarity (see discussion in Section D.3).

5.2.2 Estimating treatment-on-the-treated effects

We use the same two outcome variables described in the previous section and an indicator of compliance to estimate treatment on the treated effects. Equation (3) serves as our baseline estimating equation,

$$\Delta y_{i,t}^{\alpha} = \beta_0 + \beta_1 T_i + \varepsilon_{i,t} \tag{3}$$

where $\Delta y_{i,t}^{\alpha}$ represents the mean or the median of the distribution of changes in gas consumption and T_i represent a measure of compliance with the treatment. We use treatment assignment as an instrument. Regressions are performed jointly on all treatments unless explicitly mentioned, and estimations include inverse probability weighting.

5.3 Intent-to-treat effects on energy usage

The main outcome of interest is the change in gas usage during the study period in 2019 compared to usage during the same period in 2018. We first report the distribution of change in usage across the treatments, and associated non-parametric tests, and then present intent-to-treat regression estimates.

Figure 2 illustrates our main findings. The left panel shows the cumulative distribution of change in energy usage for each household from 2018 to 2019 for each of the four treatments. Data are pooled from Feb 1 - April 5.[25] The Control group decreased

[25] Pooling is done to simplify the analysis and maintain treatment comparisons across households to the same time frame. Table C.2 replicates Table 3 using only overlapping study time frames, i.e. Feb 9-28 and March 9-April 5. Tables C.7 and C.8 show results separately for each period. These results are noisier, and there are some differences in the responsiveness to treatment. However, this is not a challenge to our analysis since we are mainly interested in the directional response to information.

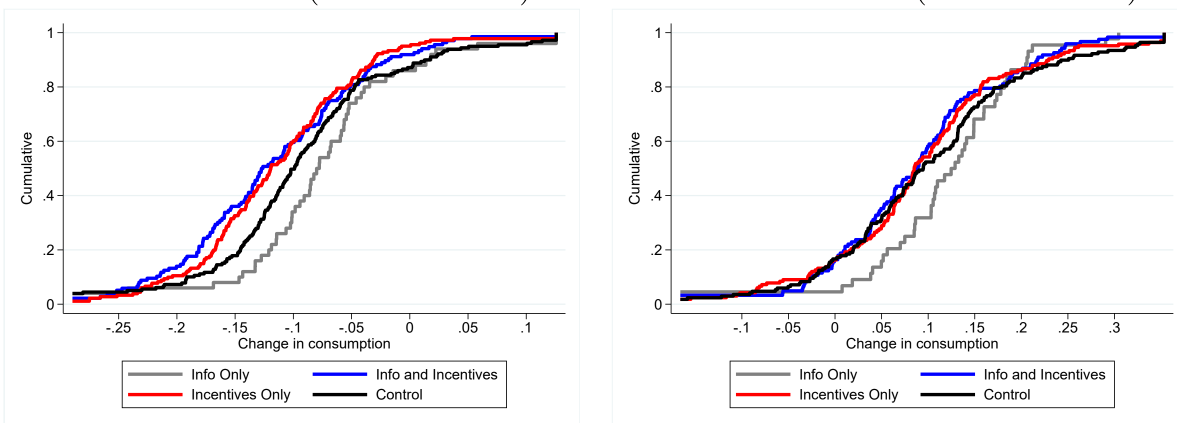
energy usage by a little over 10% on average. This group had no incentives or information, and the decline reflects the fact that 2019 was a warm winter in Alaska. The two groups with incentives (Incentive Only and Information and Incentive) decreased usage by about 15%, and the Information Only group decreased usage by about 8%. First-order stochastic dominance tests (Table 2, Panel A) show the distributions of energy use are not significantly different between Incentive Only and Information and Incentive, but both incentives groups are significantly different from the Control and the Information Only groups. The Information Only group uses significantly more energy than the Control group. These nonparametric tests provide clear support for Hypothesis 1 that incentives drive attention and energy usage.

The right panel illustrates a similar pattern is observed one year later, comparing energy usage after the experiment, in 2020, to usage the year prior to the experiment, in 2018. Higher energy usage by the Information Only group, compared to the other groups, is durable one year later. On average, the Control group increased consumption by 10%, the Information Only group by 11% and the Information and Incentive and the Incentive Only group by 9%. First-order stochastic dominance tests (Table 2, Panel B) show the increase in usage for the Information Only group is significantly larger than all other groups. These nonparametric tests provide clear support for Hypothesis 2 that learning under monetary incentives reverts to old patterns post-treatment, but learning without incentives persists.

Intent-to-treat estimates for the distributional changes for 2018 to 2019, based on Equation (2), are shown in Table 3, Columns 1-2, and lend further support for Hypothesis 1. The table reports estimates of the model at the mean (OLS) and median.[26] The overall pattern is those in Information Only increased usage relative

[26] OLS regressions exclude outliers, i.e. observations that are 3SD away from the mean. Winsorizing is suggested by Angrist and Krueger (1999) in the presence of outliers. Table C.3 reports estimates using robust regression as suggested by Han et al. (2021) and Coibion et al. (2019). Similar results obtain. Quantile regressions at each decile for intent to treat are shown in Figure C.2 and for treatment on the treated in Figure C.3.

Feb 1 - April 5 (2019 vs 2018) Feb 1 - April 5 (2020 vs 2018)



Notes: The figure uses winsorized changes (2.5%) for presentation purposes.

Figure 2: Change in energy usage by treatment

Table 2: First-order stochastic dominance tests

Panel A: Tests for 2019 v 2018

	p-values for $(H_0 : F_1 \succ_{FOSD} F_2, H_0 : F_2 \succ_{FOSD} F_1)$		
$F_1 \backslash F_2$	Info and Incentive	Incentive Only	Control
Information Only	(0.9640,0.0003)	(0.9158,0.0003)	(0.9077,0.0379)
Info and Incentive	-	(0.2252,0.5187)	(0.0025,0.9069)
Incentive Only	-	-	(0.0130,0.7810)

Panel B: Tests for 2020 v 2018

	p-values for $(H_0 : F_1 \succ_{FOSD} F_2, H_0 : F_2 \succ_{FOSD} F_1)$		
$F_1 \backslash F_2$	Info and Incentive	Incentive Only	Control
Information Only	(0.6866,0.0071)	(0.6403,0.0087)	(0.4718,0.0280)
Info and Incentive	-	(0.4155,0.6447)	(0.0523,0.8231)
Incentive Only	-	-	(0.1529,0.6225)

Notes: This is Barrett and Donald (2003)'s test for first-order stochastic dominance. Estimation uses changes in log of gas consumption between February 1 and April 5, 2018 and February 1 and April 5, 2019. p-values are calculated using 10,000 bootstrap replications. The first ordinate shows the p-value associated with the hypothesis that distribution F_1 first-order stochastically dominates distribution F_2. The second ordinate shows the p-value associated with the hypothesis that distribution F_2 first-order stochastically dominates distribution F_1. 546 observations, one per household.

to all groups (by 2.1 - 2.4 percentage points at the mean and median compared to the Control), and those in the incentive groups decreased usage (by 1.9 - 2.5 percentage points at the mean and median compared to the Control). All results in the table are robust to permutation tests (Table C.4), recommended by Young (2019), confirming that our results in the short and longer-run are not due to chance or small samples.

The bottom panel of the table reports equality of coefficient tests. These show that, conditional on receiving incentives to conserve energy, information on usage does not have a significant effect on consumption. However, conditional on receiving information on usage, incentives to conserve reduce consumption by 4.6 - 4.9 percentage points. This interaction effect also holds in a long main effects regression (Muralidharan et al., 2019) (Table C.5).

Columns 3-4 in Table 3 report the intent-to-treat estimates for changes in consumption from 2018 to 2020, one year after the intervention, and lend further support for Hypothesis 2. The results confirm those shown in the right panel of Figure 2. Participants in the Information Only treatment continue to consume significantly more than the Control group (2.7 - 3.3 percentage points) and the Information and Incentive group (3.6 - 3.7 percentage points). We reject the hypothesis that the treatment effect of Information Only is null in both the short and long run jointly.[27]

If having access to high-frequency information on energy usage allows a household to learn that costs are lower than anticipated, then we would expect those in both information groups to have learned the same thing. That is not what we observe. Behavior a year later reflects different learning across the two information groups. This finding supports Hypothesis 3 that attention cannot be modeled as unidimensional, but rather is multidimensional.

[27] To test the joint hypothesis, we run a stacked regression for changes in energy usage from 2018 to 2019 and 2018 to 2020, with clustered standard errors by household, and test the hypothesis that the treatment effect of Information Only in the short and long run are jointly zero. The F-test for the OLS regression is 4.60 (p-value=0.0105).

Table 3: Intention to treat

	2019 v. 2018		2020 v. 2018	
	OLS	Median	OLS	Median
	(1)	(2)	(3)	(4)
Information Only (T1)	0.021	0.024	0.033	0.027
s.e.	(0.010)	(0.010)	(0.011)	(0.014)
p-val	[0.040]	[0.018]	[0.004]	[0.051]
Information and Incentive (T2)	-0.025	-0.025	-0.003	-0.010
s.e.	(0.009)	(0.011)	(0.010)	(0.012)
p-val.	[0.006]	[0.026]	[0.799]	[0.424]
Incentive Only (T3)	-0.025	-0.019	-0.006	-0.005
s.e.	(0.008)	(0.009)	(0.010)	(0.012)
p-val.	[0.001]	[0.030]	[0.538]	[0.678]
Constant	-0.103	-0.105	0.088	0.091
s.e.	(0.006)	(0.006)	(0.007)	(0.010)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]
T2-T3	0.000	-0.006	0.003	-0.005
H0:T2-T3=0, p-val.	[0.976]	[0.630]	[0.731]	[0.651]
T2-T1	-0.046	-0.049	-0.036	-0.037
H0:T2-T1=0, p-val.	[0.000]	[0.000]	[0.002]	[0.004]
F-test/Chi2-test	8.858	20.944	5.719	9.257
p-val.	[0.000]	[0.000]	[0.004]	[0.026]

Notes: Dependent variable is change in log of total consumption for period Feb 1 - April 5 between years listed in column heading. 546 obs for Columns 1-2, 500 obs for Columns 3-4, one observation per household. Estimates include inverse probability weighting controlling for treatment and survey completion payment (either \$40 or \$60) assignment. Standard errors in parentheses, p-values in square brackets. OLS estimates exclude observations three standard deviations away from the mean. Results using robust regression methods are shown in Table C.3. Table C.4 reproduces this table and includes permutation p-values.

5.4 Treatment-on-the-treated effects on energy usage

Compliance with treatment assignment was not complete. In our setting, noncompliance is one-sided, since participants in the Control group could not access treatment, thus our estimates are the effects of treatment on the treated.

Compliance with treatment is defined as follows. In Information Only, compliers are those who had a working reader. In Information and Incentive, compliers are those who had a working reader or completed the endline survey (since that was a requirement to receive the incentive payment). In Incentive Only, compliers are those who completed the endline survey.[28]

The average and median effects of treatments for compliers is examined using Equation (3) and estimated with instrumental variable OLS and median regressions. Assignment to treatment is used to predict compliance. We confirm our first three hypotheses and the significant patterns found with the intent-to-treat estimates (see Table C.6).[29] Those in the incentive groups consume less than the other groups, but they are no different than the Control group the year after. Those in the Information Only group consume more than the other groups, and this persists a year later.

Estimates using Chernozhukov and Hansen (2005) IVQR and Abadie et al. (2002) LQTE (Tables C.10 and C.11) yield similar results. These are quantile regressions and are needed to estimate implied average treatment effects. This is an important validation because Wuthrich (2020) shows there is a direct relationship between these two estimates in the sense that IVQR uses treatment effect on compliers to extrapolate treatment effects to the populations of never and always takers. The similarity of estimates of both methods is viewed as support of the underlying assumptions of IVQR (see Chernozhukov and Hansen, 2004).

The implied average treatment effects can be recovered from treatment on the

[28] Compliance is not correlated with proxies for environmental consciousness (Table C.9).

[29] The treatment effects on the treated are estimated using Stata sivqr which implements the procedures introduced by Kaplan and Sun (2017).

treated quantile regressions assuming rank similarity holds (see Section D.3 for discussion and test). The average treatment effect, relative to the Control group, on the change in energy usage from 2018 to 2019 is an increase of 2.9 percentage points in Information Only, a decrease of 2.7 percentage points in Information and Incentive and a decrease of 2.3 percentage points in Incentive Only (Table C.12). The corresponding effects for the change in energy usage from 2018 to 2020 is an increase of 3.3 percentage points, a decrease of 1.4 percentage points and a decrease of 0.9 percentage points respectively.

In sum, those in Information Only behave, in the short and long-run, as if they learned about comfort, and those in Information and Incentive seem to focus on securing their bonuses, thus not exploring the option of comfort. This is consistent with the main hypotheses derived from the model in Section 2.

5.5 Attention and treatment effects

We examine if treatment effects are larger among those paying more attention (Hypothesis 4)

While we do not have direct measures of attention, we do have proxies from the process data of web portal usage. The portal was designed to record logins and clicks on the participant's personalized usage and cost page. To avoid data mining, the proxy we use is a coarse measure of attention: the total number of times a participant logs in or clicks on panels over the course of the experiment.[30] For ease of exposition, we call this clicks henceforth, but it includes logins as well.

First, we validate that our measure of attention is consistent with economic rationality. Rationality implies that attention should be (weakly) larger in Information and Incentive than in Information Only. Consistent with that implication, we reject

[30] The data are too thin to decompose the analysis by clicks on different panels. Figure C.9 shows the distribution of interactions with web portal panels. Table C.15 shows who pays attention is weakly correlated with covariates. A detailed analysis of portal usage is in Appendix D.1.

Table 4: EFFECT OF ATTENTION ON CONSUMPTION - IVQR - 2019 v. 2018 AND 2020 v. 2018

	2019	2020
(Clicks) Information Only	0.012	0.017
s.e.	(0.005)	(0.009)
p-val	[0.022]	[0.068]
(Clicks) Information and Incentive	-0.011	-0.005
s.e.	(0.005)	(0.005)
p-val.	[0.025]	[0.347]
Constant	-0.105	0.091
s.e.	(0.006)	(0.009)
p-val.	[0.000]	[0.000]

Note: Dependent variable is change in log of total consumption for period Feb 1 - April 5 from 2018 to year listed in column heading. 365 (excluding Incentives Only treatment) obs in Column 1, 334 obs in Column 2, one observation per household. Attention is measured by the inverse hyperbolic sine of the number of clicks on the portal. To estimate the effect of attention for each information treatment simultaneously, this measure is interacted with the indicator of treatment assignment. Treatment assignments are used as instrumental variables. Estimates use Chernozhukov and Hansen (2005) instrumental variable quantile regression (IVQR), and the median regression is reported. Standard errors in parentheses, and p-values in square brackets.

the hypothesis that the number of clicks in the Information Only treatment is larger than in Information and Incentive (p-value = 0.0534 for levels, p-value = 0.04 for logs).[31] That treatment assignment has a monotone effect on the demand for information is important to interpret the results. Angrist and Imbens (1995) show that if treatment status (attention) is monotone in treatment assignment, we can interpret instrumental variable estimates as causal, i.e. they estimate the effect of those who change behavior due to the treatment.

Next, we examine the effect of information usage on treatment effects. We use treatment assignment as an instrument for number of clicks and the inverse hyperbolic sine transformation of this number to account for the unusually large dispersion.

[31] Figure C.9, bottom right panel, shows the distributions of clicks in Information Only and Information and Incentive treatments.

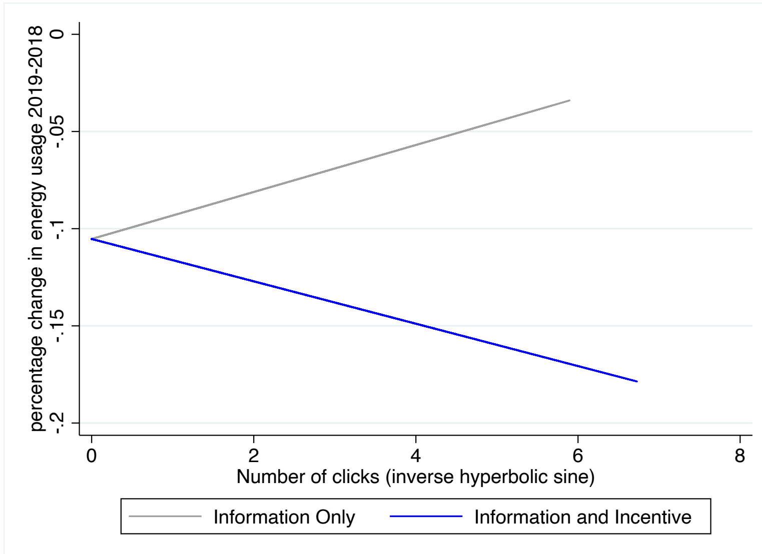


Figure 3: Attention (clicks) and treatment effects on changes in energy usage

Essentially, we reproduce the treatment on the treated analysis (Table C.10), using clicks as the measure of treatment intensity. We exclude Incentive Only from the analysis because we do not have a similar measure for that treatment.

Table 4 shows the treatment effect of attention, proxied by clicks, on changes in energy usage (see Figure 3 for a graphical representation of the results). Column 1 shows that a one SD increase in the number of clicks (∼2 clicks in inverse hyperbolic sine units) reduces median consumption from 2018 to 2019 by 2.2 percentage points in the Information and Incentive treatment and increases consumption by 2.4 percentage points in the Information Only treatment.[32] Since the maximum number of clicks is 6 in the Information Only treatment, the gap in the median change in consumption across treatments for those with the maximum number of clicks is over 13.8 percentage points. Column 2 shows that a similar, significant effect is present a year after the study concluded for those in the Information Only treatment.[33] There is no significant long-run effect for those in the Information and Incentive treatment. For households who used the portal intensely in the second month, the median treatment effect sizes

[32] Calculation for Information Only is (approximately) $0.012 * 2 = 2.4$.

[33] Table C.13 reports full quantile regressions.

in Information Only are three times larger (Table C.14).

The findings show that intensity of portal usage has a causal effect on energy consumption. The observed gap between the two information treatments is further exacerbated as participants pay more attention to their energy usage. This finding supports Hypothesis 4 that increased attention widens the gap in energy usage across the two information treatments. That intensity of usage has long-run effects for Information Only, but not for Information and Incentive, provides further support for Hypothesis 2.

6 Examining household behavior

Additional factors, outside of our model predictions, are examined to unpack how households might update and maintain their beliefs about preferences and costs for warmth.

6.1 Habit formation

We examine if treatment effects could be a result of habit formation. To do so, we examine the durability of treatment effects on the treated during the month after the experiment ended using ENSTAR data (Figure C.4). The consumption of those in the incentive groups was no different than the Control group, but those in the Information Only group consumed more than other groups. At least in this immediate post-experiment period, those in Information Only appear to have habituated to the new consumption level, however those in the incentive treatments abandoned all savings practices. If habits were formed, it was not uniform across treatments.[34]

[34] Byrne et al. (2022) document habits forming quickly and decaying slowly in a field experiment on water usage with variable duration of usage feedback. Larcom et al. (2017) document that commuters forced to experiment with alternative routes because of a transit strike change routes and reduce commute times. This is attributed to information frictions. Fowlie et al. (2021) show significant behavioral changes for households randomly defaulted to an electricity pricing program in California.

An alternative hypothesis is that those in the incentive treatments felt compelled to save due to moral obligation. Our field experiment was not designed to directly test whether those in the Information and Incentive treatment did learn it was affordable to consume more but felt morally compelled to save energy, as moral suasion would suggest (see Ito et al., 2018). However, if moral suasion is a strong motivation, we would expect to observe continued energy savings right after the experiment, especially for those in the Incentive Only treatment since they had no information to learn it was affordable to consume more. We do not find evidence for habit formation in the incentive treatments.

6.2 Response to weather conditions

Even though the marginal cost of an extra degree of warmth is constant and small (see Section 5.1), participants across the two information treatments may have understood this differently. Research by Ito (2014) suggests that individuals confound marginal and average costs. If individuals do, and do so differently across the two treatments, this might explain the differential treatment effects we observe.

With a constant marginal cost for an extra degree of warmth, participants should disregard day to day weather variations since the costs due to weather changes are akin to a sunk cost. Individuals unaware that marginal costs are fixed might assume that an extra degree of warmth is more expensive on a cold day and adjust their behavior. More importantly, it is possible that using average cost per degree rather than marginal cost of an extra degree underestimates the cost of warmth.[35] Thus, a possible reason for differences across information treatments is a difference in knowledge.

We test if participants in the Information Only and Information and Incentive

[35] To see this, note that the cost of keeping a house at X degrees when the outside temperature is Y degrees equals $a + b(X - Y)$. The average cost of X is $\frac{a+b(X-Y)}{X}$ which is smaller than b whenever a is small and the outside temperature is above 0.

groups differed in the way they responded to changes in outside temperature. The reaction is the same across treatments (Figure C.6). Both groups reduce inside temperature when it is colder outside, suggesting they are unaware that marginal costs are fixed. However, among those that used the portal more intensely, there is no response to outside temperature (Figure C.7). This is consistent with heavy-usage participants learning the true marginal cost, or simply adopting time invariant strategies, and that the response to weather is not mechanical. Also, we find no evidence that participants have different models of energy costs across treatments.[36]

Taken together, these findings illustrate that participants reacted similarly to weather conditions across treatments and have similar models of energy costs. The differences in energy usage between the Information Only and Information and Incentive treatments cannot be attributed to responses to weather conditions.

6.3 Temperature exploration

Most learning models (e.g., Erev and Roth, 1998; Camerer and Ho, 1999) require an individual to experiment with different options to be able to update propensities or priors. We examine whether exploration of inside temperature settings differ across treatments. If so, this would be consistent with differential learning and produce the different effects we observe across information treatments.

Figure 4 shows the evolution of average inside temperature for each week over the course of the study for the Information Only and Information and Incentive participants. The figure is restricted to those participants for whom we have at least two days of data prior to the start of the study.[37] There is a difference of

[36]There are no significant treatment effects on the relationship between portal usage and reaction to outside temperature. This is tested by regressing aggregate hourly data of inside temperature on outside temperature, interacted with portal use by treatment and controlling for fully interacted dummies for hour of the day, participant and a weekend dummy. Table C.17 reports the full estimates, and Figure C.8 displays estimates.

[37]Appendix D.2 presents detailed information on the strategies used by participants in the Information Only and Information and Incentive treatments. Appendix D.4 uses data from the endline

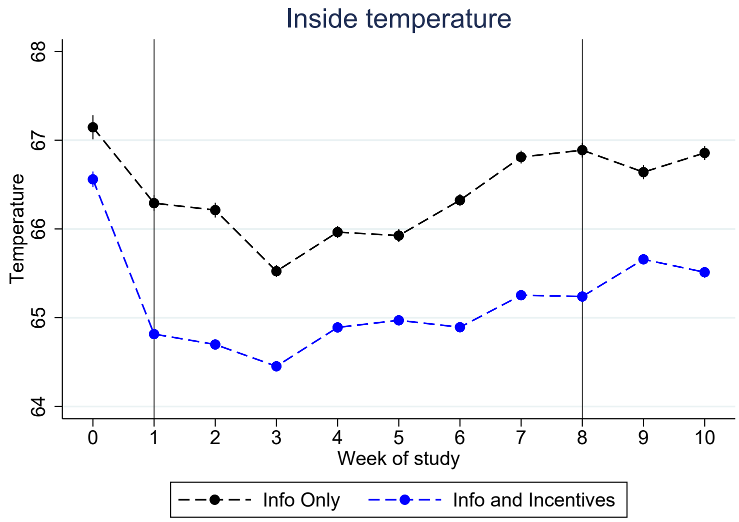


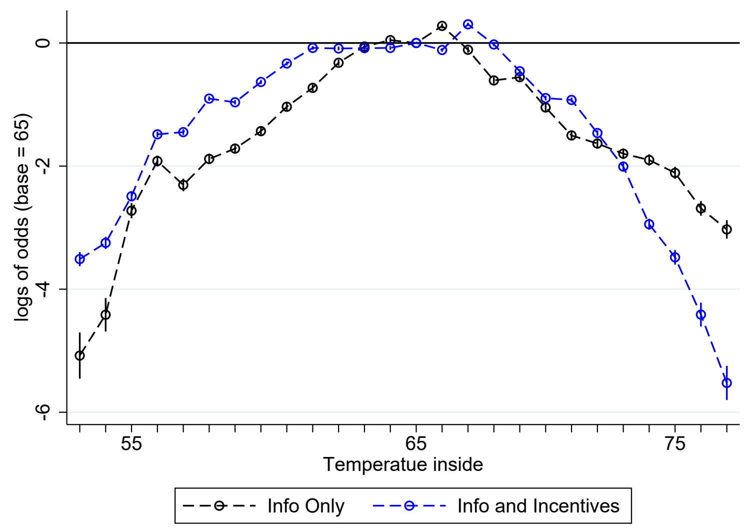
Figure 4: Inside household temperature by study week

0.4 degrees at the start of the study between treatments, and this gap increases as the study progresses to 1.7 degrees at 8 weeks. Participants in the Information and Incentive treatment decrease the temperature more than those in Information Only (by 1 degree) during the first week of the study and slowly increase it as time passes. The gap between the two treatments remains two weeks after the experiment concluded.[38]

Households explored different temperature settings across treatments. Figure 5 uses hourly data to show the log odds relative to 65F of choosing other temperatures. Two patterns are noticeable in the data. First, there is a gap in temperature settings at low temperatures. Those in the Information and Incentive treatment have larger odds of setting their thermostat below 65F in comparison to those in the Information Only treatment. Second, there is a gap at high temperatures. Those in the Information Only treatment have larger odds of setting the thermostat at temperatures 72F and above than those in the Information and Incentive treatment.

survey and documents strategies used by participants across all treatments. For example, participants in the two incentive treatments were more likely to adopt behavioral changes that sacrificed comfort (Table D.3).

[38]Readers were returned two weeks after the experiment ended.



Notes: The sample is restricted to participants who have at least two days of data prior to the start of the study. Hourly temperature is calculated as the average temperature setting during that hour. Standard error bars are calculated using 1,000 bootstrap replications.

Figure 5: Distribution of hourly inside household temperature

If learning requires experience, as predicted by the behavioral model in Section 2, those in the Information and Incentive treatment had less experience with a warmer house. This would affect learning about their preferences for warmth. By contrast, those in Information Only tried out a warmer house much more frequently, so they had a better chance to learn their preferences for warmth.

It is possible that those in the Information and Incentive treatment learned the true cost of a warmer house and decided to keep their houses cold anyway. To test if a failure to learn occurred, we explore if those who tried higher-than-usual temperatures in the Information and Incentive treatment consumed more the year after the experiment. If they did, this would suggest that incentives distracted most participants in that treatment from learning, i.e. too few people tried a warmer house, so that average effects the year after are null. We run OLS regressions on gas consumption in 2020 relative to 2018 as a function of the distance between the 90-quantile observed temperature for the household during the experiment and the average household temperature reported in the confirmation survey (Table C.16). Households that tried higher-than-usual temperatures during the experiment consumed more a year later. The magnitude is considerable and concentrates during weekends when participants are more likely to be home. A one standard deviation increase in attempted temperatures leads to a 1.7 percent increase in consumption. This result is consistent with the hypothesis of a failure to learn rather than obeisance to a conservation request.

In sum, further examination of household behavior shows that a potential influence on the differential effects we observe across the two information treatments is different temperature exploration. Those in the Information Only treatment were more likely to set their thermostat at warmer temperatures. They find they liked it, and they continued to consume more energy immediately following the end of the study and a year later. Those in the Information and Incentive treatment kept the house colder on average during the study and failed to learn the cost of a warmer house. They consumed similarly to the Control group immediately post-experiment and one year later. Neither group appears to have learned that the marginal cost of energy usage is constant, although those who used the portal more heavily might have. We conclude that incentives drive different learning about preferences for warmth but not necessarily learning about cost.

7 Conclusion

Our field experiment shows that learning from the same information technology depends strongly on the available incentives and how these incentives direct learning. Households in our study who had monetary incentives to reduce gas consumption decreased usage by 2.5 percentage points compared to the control group during the study but reverted to previous consumption patterns once the incentives were removed. Those who had no monetary incentives, just real-time information on energy usage, increased consumption by 2.1 percentage points compared to the control group during the study and this persisted one year later (3.3 percentage points increase). In principle, in-home information technologies can help consumers better respond to energy prices and manage usage. However, energy conservation may increase, decrease or have no effect, depending on how information is delivered or with what it is coupled.

The divergent patterns of energy usage in the presence of information are consistent with consumers having multidimensional goals and attention being selective. Information theories with multidimensional attention predict that treatment effects will be larger for those paying more attention. Indeed, we find households who use the web portal more to check energy usage have stronger treatment effects. The gap in the median change in consumption across information treatments for those who pay the most attention is over 13.8 percentage points.

By focusing on reducing gas consumption to get monetary incentives, participants were distracted and did not try out warmer household temperature settings on average. They did not experience a warmer house nor learn the cost to provide this (estimated to be about $10 extra per month, see Section E for discussion). By contrast, those participants without incentives did try out warmer temperatures. They discovered they preferred it and it was not so expensive to keep the house warmer. Indeed, this lesson was also learned by the households that received incentives to conserve energy but nonetheless tried out a warmer house. One year later, they ended up consuming as much as those receiving information only.

Given the perception that information on energy usage should decrease consumption, the reader might find our results anomalous. The results replicate those in Battalio et al. (1979). These authors also found that information alone (delivered weekly) increased energy consumption and information and incentives decreased it.

Our 2×2 design informs the additive and interaction effects of information and incentives, and by looking at energy usage a year after the study, it tests the durability of these effects. The process data we collect allow us to peek into the black box of decision-making to understand that divergent learning took place through different temperature exploration across our treatment groups. Also, our sample is large enough to be consider a replication of the information only result in a strong sense (see Simonsohn, 2015).

Replication of treatments is important to discern credibility and offer guidance on how to update priors with new evidence from study replications (Maniadis et al., 2014). If we consider a prior probability the hypothesis is true of 0.10, a low power of 0.20, and assume there are a total of ten comparable studies, our findings lead to an updated probability the hypothesis is true of 0.45.[39] While still below 50 percent, this is a sizeable change in beliefs that information alone increases consumption and information and incentives decrease it. The fact that the effect of information on consumption is durable one year later advocates for more research on the interaction of information and incentives on energy consumption and their long-term effects.

[39] See Maniadis et al. (2014), Table 4, for guidance on this calculation. To our knowledge, there are no RCTs that have all the treatments in Battalio et al. (1979), so this is a conservative analysis. Many studies combine information with tips, feedback and/or social comparisons or norms, with only a handful of studies that test the effect of access to granular usage information only or combine that with incentives.

References

Abadie, A., Angrist, J., Imbens, G., 2002. Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. Econometrica 70, 91–117.

Alaska Climate Research Center, 2022. Data on heating degree days. https://akclimate.org/data/. Data retrieved from Temperature Normals, https://akclimate.org/data/,.

Allcott, H., 2011. Consumers' perceptions and misperceptions of energy costs. American Economic Review 101, 98–104.

Allcott, H., Rogers, T., 2014. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. American Economic Review 104, 3003–3037.

Angrist, J.D., Imbens, G.W., 1995. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. Journal of the American statistical Association , 431–442.

Angrist, J.D., Krueger, A.B., 1999. Chapter 23 empirical strategies in labor economics. Handbook of Labor Economics 3, 1277 – 1366.

ASHRAE, 2017. Handbook - Fundamentals. American Society of Heating Refrigeration and Air Conditioning Engineers.

Asmare, F., Jaraite, J., Kazukauskas, A., 2021. The effect of descriptive information provision on electricity consumption: Experimental evidence from Lithuania. Energy Economics 104.

Attari, S.Z., DeKay, M.L., Davidson, C.I., Bruine de Bruin, W., 2010. Public perceptions of energy consumption and savings. Proceedings of the National Academy of sciences 107, 16054–16059.

Barrett, G., Donald, S., 2003. Consistent tests for stochastic dominance. Econometrica 71, 71–104.

Battalio, R., Kagel, J., Winkler, R., Winett, R., 1979. Residential electricity demand: An experimental study. Review of Economics and Statistics 61, 180–189.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013. Salience and consumer choice. Journal of Political Economy 121, 803–843.

Brandon, A., Clapp, C., List, J., Metcalf, R., Price, M., 2021. Smart Tech, Dumb Humans: The Perils of Scaling Household Technologies. Technical Report. Working Paper.

Buchanan, K., Russo, R., Anderson, B., 2015. The question of energy reduction: The problem(s) with feedback. Energy Policy 77, 89–96.

Burkhardt, J., Gillingham, K., Kopalle, P.K., 2019. Experimental Evidence on the Effect of Information and Pricing on Residential Electricity Consumption. Technical Report. NBER Working Paper 25576.

Byrne, D., Goette, L., Martin, L., Jones, A., Miles, A., Schob, S., Staaked, T., Tiefenbeck, V., 2022. The habit-forming effects of feedback: evidence from a large-scale field experiment. Technical Report. Working Paper.

Byrne, D.P., La Nauze, A., Martin, L.A., 2018. Tell me something i don't already know: Informedness and the impact of information programs. The Review of Economics and Statistics 100, 510–527.

Calonico, S., Cattaneo, M.D., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica 82, 2295–2326.

Camerer, C., Ho, T., 1999. Experience-weighted attraction learning in normal form games. Econometrica 67, 827–874.

Caplin, A., 2016. Measuring and modeling attention, in: Aghion, P., Rey, H. (Eds.), Annual Review of Economics. volume 8 of *Annual Review of Economics*, pp. 379–403.

Chernozhukov, V., Hansen, C., 2004. The effects of 401(k) participation on the wealth distribution: An instrumental quantile regression analysis. Review of Economics and Statistics 86, 735–751.

Chernozhukov, V., Hansen, C., 2005. An IV model of quantile treatment effects. Econometrica 73, 245–261.

Chernozhukov, V., Hansen, C., 2006. Instrumental quantile regression inference for structural and treatment effect models. Journal of Econometrics 132, 491–525.

Chetty, R., Looney, A., Kroft, K., 2009. Salience and Taxation: Theory and Evidence. American Economic Review 99, 1145–1177.

Coibion, O., Gorodnichenko, Y., Hong, G.H., 2019. The cyclicality of sales, regular and effective prices: Business cycle and policy implications: Reply. American Economic Review 109, 314–324.

Costa, D.L., Kahn, M.E., 2013. Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. Journal of the European Economic Association 11, 680–702.

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

U.S. Department of Energy, 2022. Energy data facts. https://rpsc.energy.gov/energy-data-facts.

U.S. Environmental Protection Agency, 2022. Data on heating and cooling degree days. https://www.epa.gov/climate-indicators/climate-change-indicators-heating-and-cooling-degree-days.

Erev, I., Roth, A., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. American Economic Review 88, 848–881.

Fang, X., Goette, L., Rockenbach, B., Sutter, M., Tiefenbeck, V., Schoeb, S., Staake, T., 2023. Complementarities in behavioral interventions: Evidence from a field experiment on resource conservation. Journal of Public Economics 228, 105028.

Farhi, E., Gabaix, X., 2020. Optimal Taxation with Behavioral Agents. American Economic Review 110, 298–336.

Ferraro, P., Price, M., 2013. Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. Review of Economics and Statistics 95, 64–73.

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

Frandsen, B., Lefgren, L., 2018. Testing Rank Similarity. Review of Economics and Statistics 100, 86–91.

Gabaix, X., 2014. A Sparsity-Based Model of Bounded Rationality. Quarterly Journal of Economics 129, 1661–1710.

Gabaix, X., 2019. Behavioral inattention, in: Bernheim, B., DellaVigna, S., Laibson, D. (Eds.), Handbook of Behavioral Economics - Foundations and Applications 2. Handbooks in Economics, pp. 261–343.

Gans, W., Alberini, A., Longo, A., 2013. Smart meter devices and the effect of feedback on residential electricity consumption: Evidence from a natural experiment in northern ireland. Energy Economics 36, 729–743.

Gerber, A., Green, D., 2012. Field Experiments Design, Analysis, and Interpretation. Norton.

Han, J.S., Houde, J.F., Van Benthem, A.A., Abito, J.M., 2021. When does regulation distort costs? lessons from fuel procurement in us electricity generation: Comment. American Economic Review 111, 1356–1372.

Handel, B., Schwartzstein, J., 2018. Frictions or mental gaps: What's behind the information we (don't) use and when do we care? Journal of Economic Perspectives 32, 155–178.

Hanna, R., Mullainathan, S., Schwartzstein, J., 2014. Learning through noticing: Theory and evidence from a field experiment. Quarterly Journal of Economics 129, 1311–1353.

Harding, M., Lamarche, C., 2016. Empowering consumers through data and smart technology: Experimental evidence on the consequences of time-of-use electricity pricing policies. Journal of Policy Analysis and Management 35, 904–931.

Harrison, G., List, J., 2004. Field experiments. Journal of Economic Literature 42, 1009–1055.

Hastings, J.S., Shapiro, J.M., 2013. Fungibility and consumer choice: Evidence from commodity price shocks. Quarterly Journal of Economics 128, 1449–1498.

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

Ito, K., 2014. Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. American Economic Review 104, 537–563.

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

Jessoe, K., Rapson, D., 2014. Knowledge is (less) power: Experimental evidence from residential energy use. American Economic Review 104, 1417–38.

Kaplan, D.M., Sun, Y., 2017. Smoothed estimating equations for instrumental variables quantile regression. Econometric Theory 33, 105–157.

Larcom, S., Ferdinand, R., Willems, T., 2017. The benefits of forced experimentation: Striking evidence from the london underground network. Quarterly Journal of Economics 132, 2019–2055.

Leland, H., 1968. Saving and uncertainty - Precautionary demand for saving. Quarterly Journal of Economics 82, 465–473.

List, J., 2020. Non Est Disputandum De Generalizability? A Glimmpse Into the External Validity Trial. Technical Report. NBER Working Paper 27535.

List, J.A., Sadoff, S., Wagner, M., 2011. So you want to run an experiment, now what some simple rules of thumb for optimal experimental design. Experimental Economics 14, 439–457.

Mackowiak, B., Matejka, F., Wiederholt, M., forthcoming. Rational Inattention: A Review. Journal of Economic Literature .

Maniadis, Z., Tufano, F., List, J.A., 2014. One swallow doesn't make a summer: New evidence on anchoring effects. American Economic Review 104, 277–290.

Martin, S., Rivers, N., 2018. Household electricity consumption: Evidence from a large-scale field deployment. Journal of the Association of Environmental and Resource Economists 5, 207–231.

Muralidharan, K., Romero, M., Wuthrich, K., 2019. Factorial designs, model selection, and (incorrect) inference in randomized experiments. Technical Report. NBER Working Paper 26562.

Nemati, M., Penn, J., 2020. The impact of information-based interventions on conservation behavior: a meta-analysis. Resource and Energy Economics 62.

Rees-Jones, A., Taubinsky, D., 2020. Measuring "Schmeduling". Review of Economic Studies 87, 2399–2438.

Sandmo, A., 1971. Theory of competitive firm under price uncertainty. American Economic Review 61, 65–73.

Saylor, B., Haley, S., Szymoniak, N., 2008. Estimated Household Costs for Home Energy Use. Technical Report. Institute of Social and Economic Research University of Alaska Anchorage.

Schultz, P., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., Griskevicius, V., 2007. The constructive, destructive, and reconstructive power of social norms. Psychological Science 18, 429–434.

Simonsohn, U., 2015. Small Telescopes: Detectability and the Evaluation of Replication Results. Psychological Science 25, 559–569.

Stigler, G., 1961. The Economics of Information. Journal of Political Economy 69, 213–225.

Taubinsky, D., Rees-Jones, A., 2018. Attention variation and welfare: theory and evidence from a tax salience experiment. The Review of Economic Studies 85, 2462–2496.

Wuthrich, K., 2020. A comparison of two quantile models with endogeneity. Journal of Business & Economic Statistics 38, 443–456.

Young, A., 2019. Channeling fisher: Randomization test and the statistical insignificance of seemingly significant experimental results. Quarterly Journal of Economics 134, 557–598.

APPENDICES INTENDED FOR ONLINE PUBLICATION

A Decision framework model

We use a simple model of endogenous attention and energy consumption. The model follows the approach proposed by Farhi and Gabaix (2020) who analyze welfare implications of behavioral biases. The model captures testable implications that models without endogenous attention do not predict.

We distinguish between perceived utility u_s and real utility u. Let the real utility of temperature τ given income y be equal to $u(\tau, c) = \frac{\theta}{1-1/e}\left(\frac{\tau}{\theta}\right)^{1-1/e} + y - p\tau$ where p is the marginal cost of an extra degree of warmth. c is consumption and substituted out with the budget constraint, $y = c + p\tau$. Parameter θ captures heterogeneous preferences for warmth and parameter e represents the price elasticity. We assume a constant elasticity representation for clarity.

Consumers may perceive their preferences for warmth and the marginal cost of warmth as different to what they are, i.e. $\theta_s \neq \theta$, and $p_s \neq p$. For given parameters θ_s and p_s, the perceived utility is defined by $u_s(\tau, c) = \frac{\theta_s}{1-1/e}\left(\frac{\tau}{\theta_s}\right)^{1-1/e} + y - p_s\tau$. We represent utility this way because we do not know ex-ante if consumers in the experiment will learn their true preferences or the true cost of their actions.

To allow for endogenous attention and learning, we define a_θ as the amount of attention a consumer devotes to learning her preferences and define a_p as the amount of attention a consumer devotes to learning about prices. We assume that $a_\theta, a_p \in [0, 1]$. We also assume that attention is costly and equal to $\kappa(a_\theta + a_p)$. We make this assumption for three reasons. First, as shown by Gabaix (2014), this function allows for a corner solution in which an individual pays no attention. Second, as show in the empirical analysis, we do not have separate proxies for these two types of attention. Last, the reveal preference intuition we present is independent of this latter assumption.

We assume that perceived parameters are a linear combination of real parameters and default values and that attention is multidimensional. In particular, perceived preference is $\theta_s(a_\theta) = \theta_d + a_\theta(\theta - \theta_d)$ where a_θ is the attention placed on preferences, θ_d is a default preference and θ is the real preference parameter. The perceived price is $p_s(a_p) = p_d + a_p(p - p_d)$ where a_p is the attention placed on prices, p_d is a default price and p is the real price. If the consumer is fully attentive in both domains, i.e. $a_\theta, a_p = 1$, they respond to real preferences (θ) and constraints (p). We let $\tau_s(\theta_s(a_\theta), p_s(a_p)) = \theta_s p_s^{-e}$ be the optimal temperature choice given perceived preferences and prices.

In the case of new information or changes in the cost of attention, consumers allocate attention to preferences and prices to maximize real utility knowing that their own behavior will follow perceived parameters rather than real parameters. The intuition is that consumers will develop attention-constrained optimal choices. Given an amount of attention (a_θ, a_p), a behavioral-constrained individual will choose according to $(\theta_s(a_\theta), p_s(a_p))$, the perceived parameter, not the real ones. So, an omniscient ego will maximize the real utility in anticipation of constrained decisions. In particular, attention will be allocated to solve:

$$max_{a_\theta, a_p} \theta^{1/e} \frac{1}{1 - 1/e} \tau_s(\theta_s(a_\theta), p_s(a_p))^{1-1/e} - p\tau_s(\theta_s(a_\theta), p_s(a_p)) - \kappa(a_\theta + a_p) \tag{4}$$

Equation (4) shows that the consumer chooses the amount of attention to place on learning preferences and price by maximizing real utility under the assumption that decisions made later are based on perceived parameters not real parameters, i.e. according to $\tau_s(\theta_s(a_\theta), p_s(a_p))$. We assume that default parameters θ_d and p_d are the result of previous attention allocation decisions. We do not assume that these default parameters are correct, rather the individual's model of the world evolves as circumstance change.

After some algebra, the corresponding FOCs are:

$$[\theta^{1/e}\tau_s^{-1/e} - p]\frac{\partial \tau_s}{\partial \theta_s}\frac{\partial \theta_s}{\partial a_\theta} - \kappa \leq 0 \tag{5}$$

$$[\theta^{1/e}\tau_s^{-1/e} - p]\frac{\partial \tau_s}{\partial p_s}\frac{\partial p_s}{\partial a_p} - \kappa \leq 0 \tag{6}$$

where the term $\theta^{1/e}\tau_s^{-1/e} - p$ is the marginal gain from an extra degree of warmth under the true parameters.

Equations (5)-(6) provide useful insights. Suppose there is a solution in which an individual overconsumes, i.e., $\theta^{1/e}\tau_s^{-1/e} - p < 0$. From (6), we have that attention will be put on prices ($a_p > 0$) only if $p - p_d > 0$. This follows from the fact that $\frac{\partial \tau_s}{\partial p_s} < 0$ and $\frac{\partial p_s}{\partial a_p} = p - p_d$ under our assumptions. Attention will be focused on prices if the price increases, i.e. p is equal to $p + tax$ as in our experiment.[40]

Under the assumption of overconsumption, from (5), we can also conclude that attention will be put on preferences for warmth $a_\theta > 0$ only if $\theta - \theta_d < 0$ since $\frac{\partial \tau_s}{\partial \theta_s} > 0$. That is, a consumer will pay attention to her preferences if by doing so she learns that she cares less about warmth than at the default. This knowledge would help reduce over-consumption. The model therefore allows situations in which subsidies to reduce consumption completely crowd out attention to one's own preferences, as we observed in our experiment. This occurs when the omniscient ego expects over-consumption.

The model supports other solutions. Suppose a consumer underconsumes, i.e., $\theta^{1/e}\tau_s^{-1/e} - p > 0$. For a consumer to put attention on preferences for warmth, $a_\theta > 0$, it requires that $\theta - \theta_d > 0$. For a consumer to put attention on prices, $a_p > 0$, it requires $p - p_d < 0$. This is equivalent to a consumer being cautious and over-

[40]In a model in which attention is only devoted to price, it can be shown that attention is increasing in the absolute distance between p and p_d.

estimating the cost of warmth. Note also that it is possible that all attention is devoted to preferences for warmth if $p - p_d > 0$. In this case, learning that prices are over-estimated will help reduce under-consumption. The reasoning is analogous to the previous case. If a consumer underconsumes on average, there is less incentive to update information on prices that would further lead to even lower consumption.

The model produces stark predictions, in that attention can be devoted to either preferences or prices, but not both. It highlights the potential differential effect of access to information when attention is multidimensional. The strategic attention and avoidance stem from the omniscient ego trying to reduce sub-optimal decisions. Note that equation (4) implies that the welfare effect of a change in the cost of attention κ is increasing in the optimal level of attention $a_\theta^* + a_p^*$. We expect treatments effects on welfare to be increasing in attention.

We would like to determine the relationship between parameters of the utility function and the demand for information. Intuitively, we expect that those who would benefit from information the most will be the ones who demand it more. For instance, suppose that, in equilibrium, an individual underconsumes and only pays attention to comfort (a_θ). Inspection of equation (5) shows that equilibrium attention to comfort increases in θ and decreases in e.[41] Similarly, suppose that, in equilibrium, an individual overconsumes and only pays attention to price (a_p). Inspection of equation (6) shows that equilibrium attention to price decreases in θ and increases in e.[42] This suggests that it is possible that in one treatment subjects sort according to their preferences for comfort (θ) and in another they sort according to their price responsiveness (e). Importantly, we expect that those demanding information in a particular dimension are those more likely to respond in that dimension.[43] Finally, we observe that regardless of the values of parameter θ and e, those with lower costs

[41] We obtain this results by taking derivatives of equation (5) and making use of the held assumption of the corner solution of only paying attention to comfort.

[42] The second result requires e to be sufficiently small.

[43] Given the definition of τ_s, it is increasing in a_θ if $\theta > \theta_d$ and a_p if $p > p_d$.

to attention κ will demand more information.

The main prediction from the model is that attention is driven by treatment incentives. In the extreme, attention could be solely devoted to learning about one's own preferences for warmth or learning about price instead. Since the model predicts that attention should be a function of the incentive to learn in that dimension, it also predicts that more attention resources should be devoted to learning about price in the treatment that provides monetary incentives to reduce gas consumption.

There are other testable implications. For instance, if attention is unidimensional and consumers only care about the price of energy usage, we should observe that giving consumers access to real-time information on energy usage has the same directional effect, whether or not there are also monetary incentives to reduce energy usage.[44] This case is consistent with Jessoe and Rapson (2014) whereby real-time information on energy usage increases price responsiveness. However, when attention is multidimensional, it is possible that information on energy usage by itself, without any monetary incentives, produces reactions of the opposite sign. This occurs because, in the case of incentives to reduce energy usage, no attention is placed on one's own preferences for warmth and then little is learned about how to improve comfort. In other words, if consumers only pay attention to price and price is found to be lower (or less variable) than anticipated, this information should be learned in all treatments and behavior adjusted accordingly. If, however, attention is multidimensional, then the same information about price may not be learned across treatments. This would be observed when monetary incentives are removed. Behavior learned with incentives reverts to old patterns, but behavior learned in the absence of incentives persists.

Our experimental evidence is consistent with participants not being fully aware of their own preferences and thus a decision framework of multidimensional attention on preference and price. If instead the model is modified to allow for differential at-

[44] Gabaix (2014); Farhi and Gabaix (2020) shows that price responsiveness will increase with reductions in attention costs.

tention to price and taxes/subsidies, but not preferences, those two types of attention are complementary. A focus of attention on tax/subsidy also implies learning about price. If price is lower than anticipated, those in the Information and Incentive treatment would partially learn this as well as those in the Information Only treatment. Participants in the Information Only treatment increase energy consumption during and after the study (Figure C.4, Figure C.3), suggesting they learned the price was lower than anticipated. But, we do not see similar behavior from those in Information and Incentive. These patterns of behavior across the two treatments are more consistent with a model of multidimensional attention on preference and price, rather than a model of unidimensional attention on prices.

B Field experiment material

B.1 Dashboard

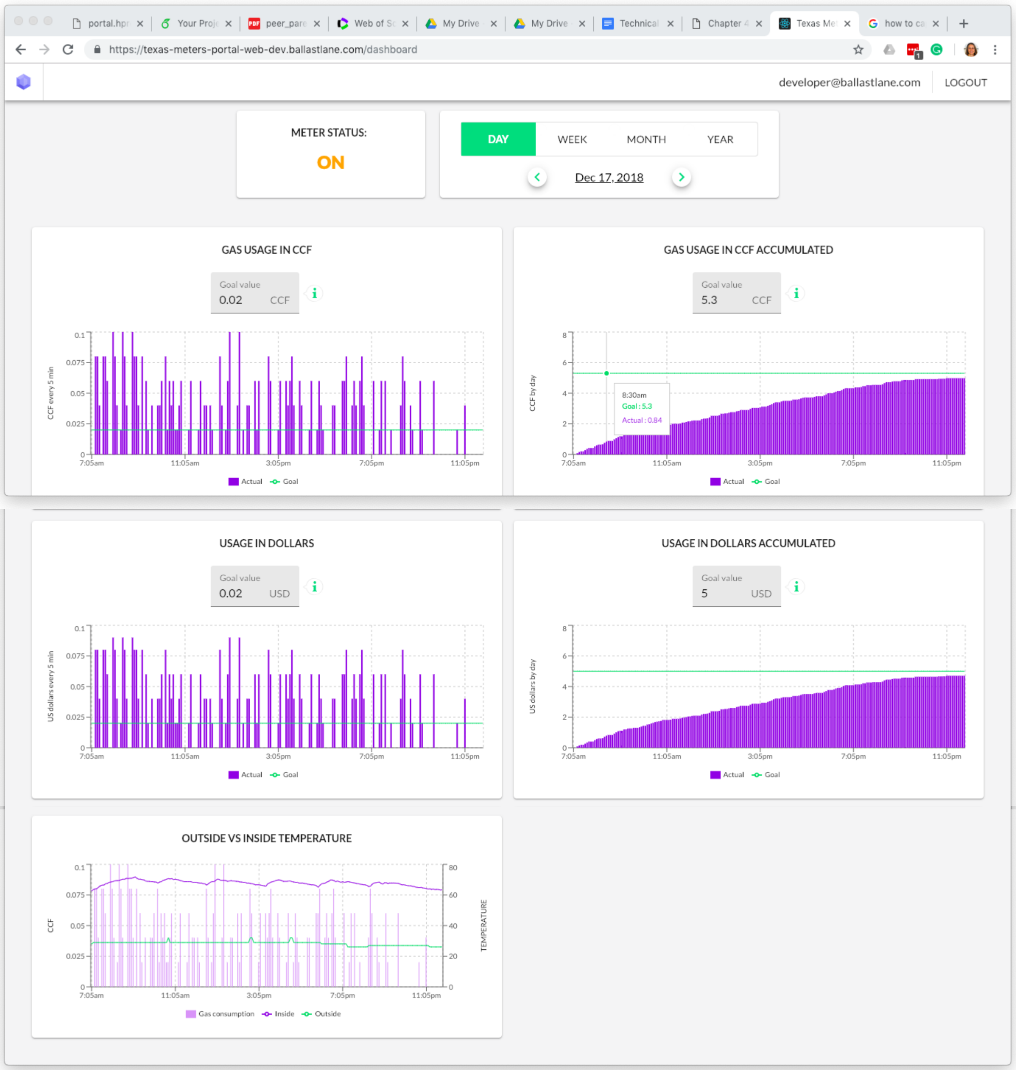


Figure B.1: Dashboard - Temperature and consumption

B.2 Invitation email from ENSTAR

[Subject line:] Enroll in AHFC/Texas A&M's Energy Usage in Anchorage Study

[Text in email:]

Introduction: Alaska Housing Finance Corporation (AHFC) and Texas A&M University, in collaboration with ENSTAR Natural Gas Company, are conducting a study, funded by the National Science Foundation, on energy usage in Anchorage.

AHFC and Texas A&M University are recruiting households who are willing to participate in this study. The aim of the study is to measure natural gas usage on a frequent basis and examine ways to help reduce gas expenditures. For this purpose, AHFC has developed high-frequency readers that can register gas usage at regular intervals, such as every 5 minutes.

If you are interested in participating, we ask that you complete an online survey to enroll. The survey can be accessed by clicking this link which will direct you to our website. The enrollment period ends January 22 at 12pm.

If you participate in and complete the study, you will receive a payment of $40 [$60]. This payment is offered by Texas A&M University and will be paid and distributed by Texas A&M University. As a condition of participation, households are asked to respond to two short surveys, one to enroll in the study and one upon study completion, and allow researchers to analyze information on your household energy usage. All information will be kept confidential.

Participants must have broadband home internet service, have a dedicated gas meter and reside in a single-family home. Households will be selected at random to be part of the study. Some will receive a high frequency reader at no financial cost to measure gas and/or additional monetary incentives to save energy.

Participation in the study is voluntary. You are not compelled to enroll or participate.

What do you stand to gain by participating? Usage of a high-frequency gas reader. Better understanding of your gas usage and costs. Compensation for study completion. Additional monetary incentives to save energy

Information Privacy: We will keep your personal information private. Information obtained will be used to implement and analyze the effectiveness of alternative energy conservation policies and no response will be individually traced to you or your household.

Study Description: The study will provide some participants technology to make more informed energy usage decisions and/or incentives to achieve them. This might be, but not limited to, visibility of usage and additional monetary incentives to reduce energy consumption. All incentives are offered by Texas A&M University and will be paid and distributed by Texas A&M University.

Study Period, Rewards: We plan to conduct the study from January 2019 through the end of March 2019.

All participants will receive a payment upon completion of the study. Some participants will receive a high-frequency gas reader and/or additional monetary incentives to reduce energy usage. All payments and incentives are offered by Texas A&M University and will be paid and distributed by Texas A&M University

For inquiries about the study, please contact Professor XXX (Texas A&M University) at xxx@tamu.edu or XXX (AHFC) xxx@ahfc.org.

If you are interested in volunteering to participate in the study, please select this link.

B.3 Assignment to treatment email

Dear X,

Thank you for your interest in the AHFC/Texas A&M Anchorage Energy Usage Study and completing the online signup survey.

We received a large response to our request for participants, and your household has been randomly chosen among those eligible to be part of the study.

We request that you confirm or decline participation in this study by clicking this personalized link, to complete a short confirmation survey. The deadline to do so is

Sunday (Jan 27) by 5pm.

If you do not confirm participation by this deadline, we will assume you are no longer interested. We hope the tight deadline does not cause undue burden.

If you participate,

[*Only included for those in Information Only or Information and Incentives*]

1. You will receive a high-frequency gas reader by mail to self-install in your house. The reader will arrive next week, with detailed instructions for easy installation. You can also contact us if you have any questions about installation. Once installed, you will be able to see your gas usage and costs via an online dashboard designed for this project. An email providing a link to the dashboard and login information for your household will be sent to you next week (this will come from "Anchorage Energy Usage Study"). If you do not receive an email providing this information by the time you receive the reader, please contact us at xxx@tamu.edu.

[*Only included for those in Incentive Only or Information and Incentives*]

2. You will receive a monetary incentive to reduce gas usage during the month of February. For every 1% reduction in average daily gas usage in February 2019 compared to average daily gas usage in February 2018, you will receive $10. This incentive is capped at a 10% reduction in usage (i.e. capped at $100).

This means that if your average daily usage was 10 CCFs (hundred cubic feet of fuel) in Feb 2018 and you reduce your usage to 9.80 CCFs in Feb 2019, you would receive $20. If you reduce it to 9.40 CCFs, you would receive $60. If you reduce it to 9.32 CCFs, you would receive $68.

Average daily usage will be determined by prorating and averaging your usage for the month of February in 2018 and 2019 using gas usage data collected by ENSTAR and the billing cycles that occur during February.

[*Included for all groups*]

3. Upon completion of the study, you will receive a payment of [$X]. Completion

of the study includes us collecting data from ENSTAR on your energy usage and you completing a second survey in April 2019.

This will be paid to you via PayPal, at no cost to you, and you can transfer the payment from PayPal to your bank account at no charge. If you do not have a PayPal account, you can open one for free. We prepared a step-by-step guide to open a PayPal account, and you can see the guide here.

We are using PayPal to facilitate prompt and efficient payment. If you require another form of payment, indicate this in the confirmation survey and/or contact the research team.

After you confirm participation [*Information Only and Information and Incentives* and complete installation of your high-frequency reader], we will send $10 of your completion payment via PayPal. The remainder will be paid when the study concludes.

4. You can find energy saving tips in the AHFC pamphlet here

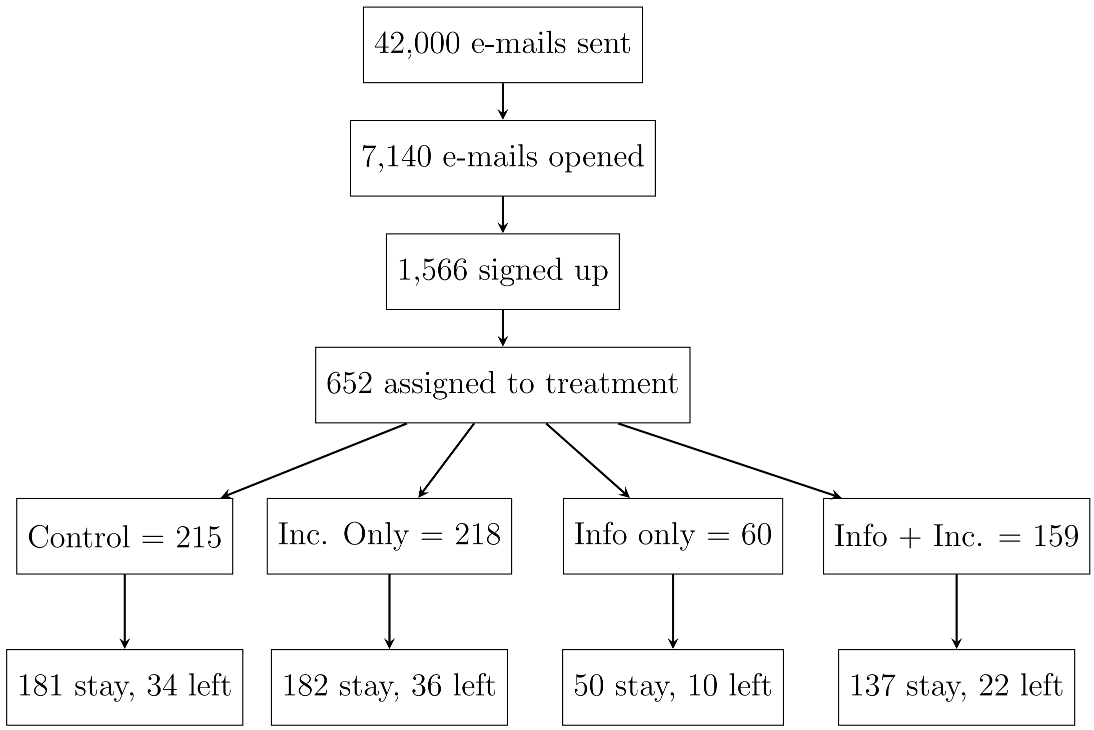
It is important that you confirm or decline participation in the study by clicking your personalized link and completing the confirmation survey. We ask that you complete the survey in either case so that, should you choose to not participate, we can invite another household to take the available slot in the study.

We appreciate your willingness to help us learn more about energy usage in Anchorage. If you have any questions about the study, you can contact XXX, xxx@tamu.edu, or XXX, xxx@ahfc.us.

Sincerely,

The Research Team

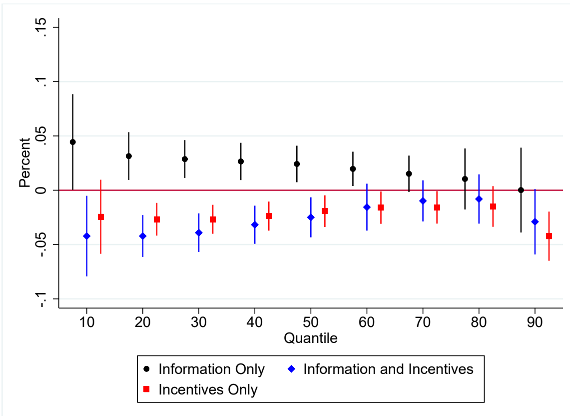
C Supplemental figures and tables



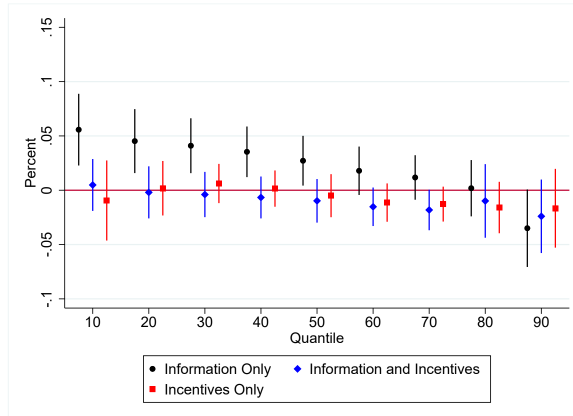
Notes: The figure shows the construction of our final estimation sample. The final row reports the number of participants assigned to treatment who remained in the study after treatment assignment and the number who left the study. The estimation sample is the 550 participants who stayed.

Figure C.1: Sample construction

Feb 1 - April 5 (2019 vs 2018)



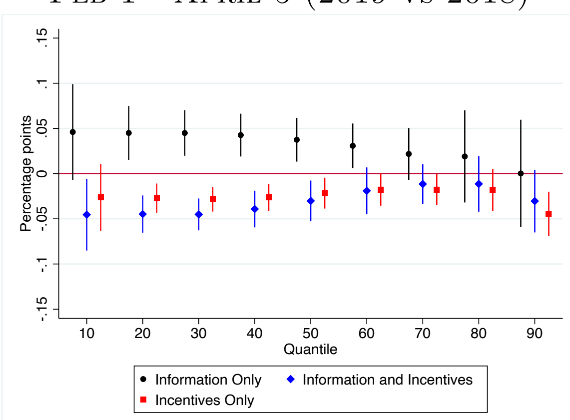
Feb 1 - April 5 (2020 vs 2018)



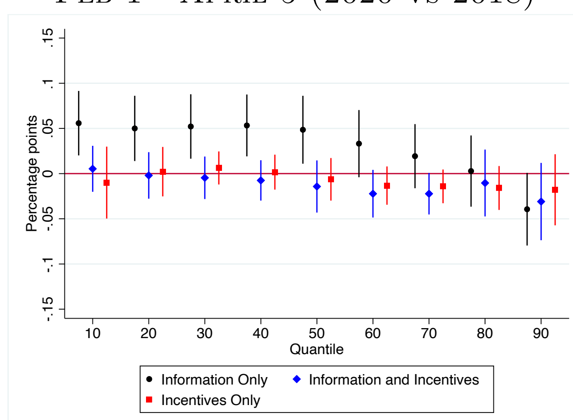
Note: All parameters are estimated jointly using the data from the treatments and the control group. Regression coefficients and robust standard error bars are shown by decile.

Figure C.2: Intent-to-treat effects by change in usage quantiles

Feb 1 - April 5 (2019 vs 2018)

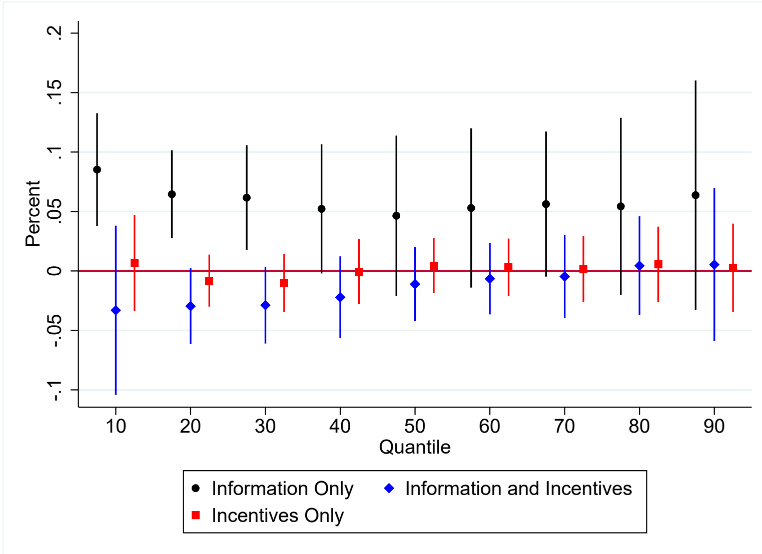


Feb 1 - April 5 (2020 vs 2018)



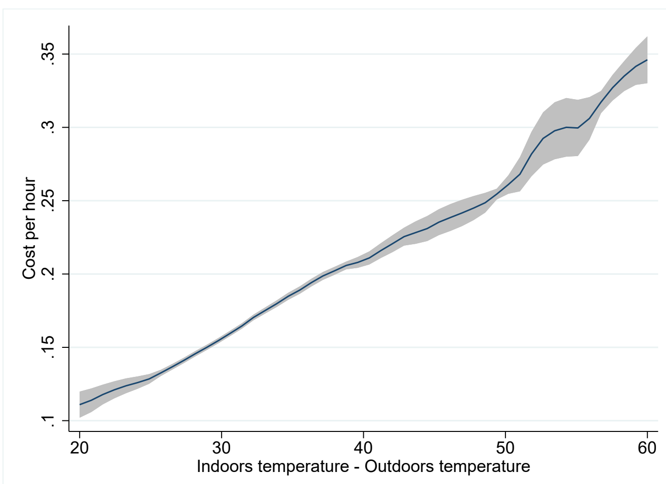
Note: The measure of compliers in Information Only is an indicator of an installed reader. In Information and Incentive, it is an indicator of having an installed reader or completion of the endline survey (endline survey completion was a condition to receive incentive payments). In Incentive Only, it is completion of the endline survey. To estimate the TOT for each treatment simultaneously, we create three variables that equal 1 if the treatment is adopted and zero if not. Treatment assignments are used as instrumental variables. Estimates use Chernozhukov and Hansen (2005) IVQR. Regression coefficients and standard error bars are shown by decile.

Figure C.3: Treatment on the treated effects by change in usage quantiles



Notes: To make the figure comparable to previous results, the estimation uses the change in gas consumption relative to consumption in the same period in the year prior to the experiment (2018). Regression coefficients and standard error bars are shown by decile.

Figure C.4: Treatment on the treated effects during the month after the experiment ended



Notes: Figure reports estimates with 95-percent confidence interval.

Figure C.5: Energy cost per hour by difference in indoor and outdoor temperature

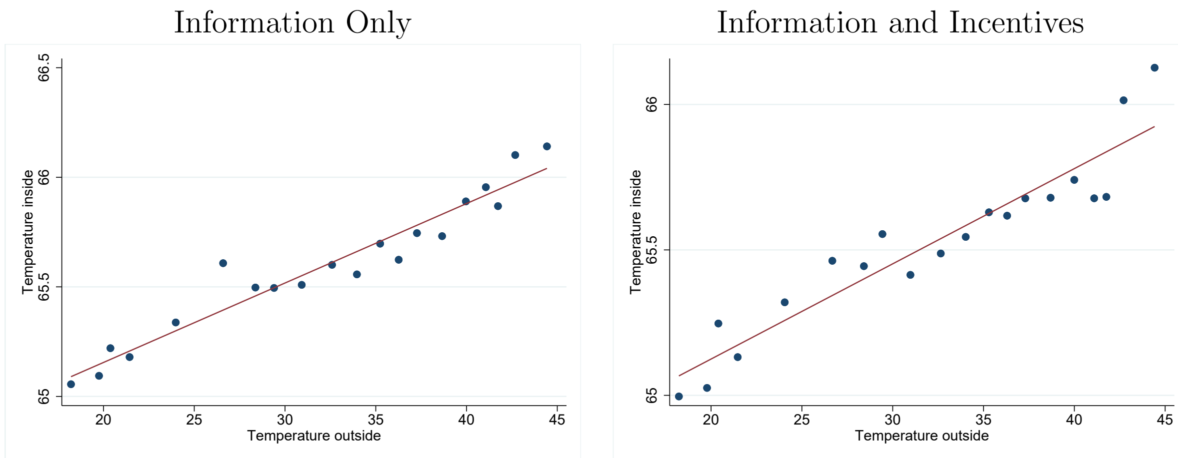


Figure C.6: Response to outside temperature by treatment condition

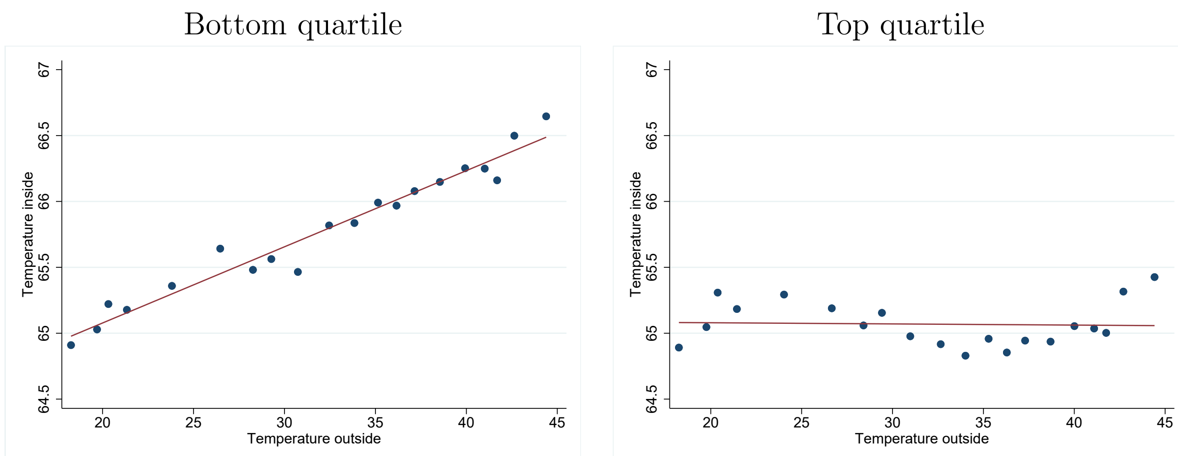


Figure C.7: Response to outside temperature by portal use intensity

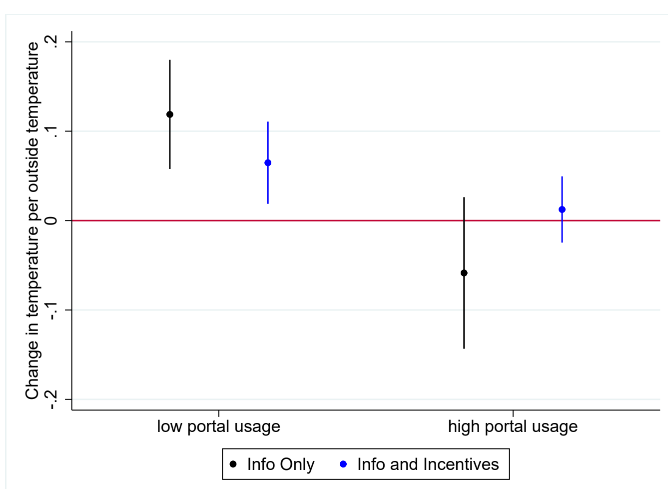


Figure C.8: Treatment effects on response to outside temperature by portal use intensity

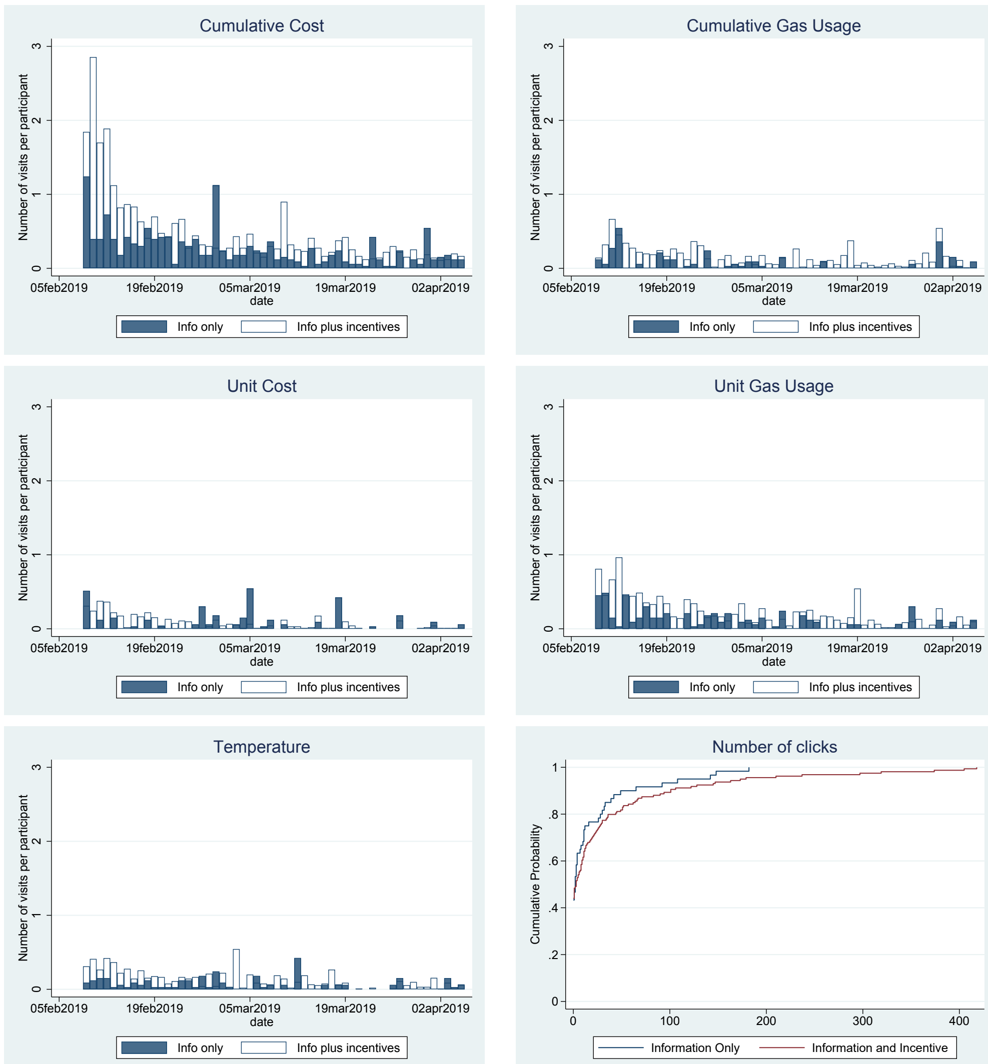


Figure C.9: Portal usage by panel and number of clicks

Table C.1: Balance across samples - Signed-up and Assigned to Treatment

	Signed-up	Assigned to treatment	Control	Treatments: Information Only	Treatments: Information and Incentive	Treatments: Incentive Only	F-test (p-value)
Household size	2.9	3.0	3.0	3.0	3.1	3.0	0.894
# Children	0.9	0.9	0.9	0.9	0.9	0.9	0.994
Yrs in residence	5.4	6.3	5.7	6.7	6.6	6.5	0.025
Household Income	90,516	97,185	97,442	94,167	99,811	95,833	0.582
Have prog thermostat	72.4	78.4	74.4	78.3	82.4	79.4	0.306
HERP	22.1	35.3	31.6	38.3	37.7	36.2	0.572
Avg temp (M-F, 8am-5pm)	n/a	66.3	66.5	66.6	66.2	66.1	0.761
Avg temp (M-F, 5-10pm)	n/a	68.3	68.6	68.4	68.2	68.1	0.236
Avg temp (M-F, after 10pm)	n/a	66.0	66.2	66.2	66.1	65.8	0.653
Avg temp (Sat-Sun, day)	n/a	68.0	68.2	67.9	67.9	67.9	0.692
Avg temp (Sat-Sun, night)	n/a	66.2	66.3	66.8	66.3	65.8	0.372
Year built	1981	1981	1982	1978	1981	1980	0.146
Home value	345,427	351,273	355,316	358,741	352,631	344,239	0.806
Avg Jan bill (reported)	177.3	186.8	184.0	195.8	187.6	186.3	0.854
Feb-March 2018 bill (admin)	254.8	378.1	385.8	389.6	377.2	368.1	0.610
Obs	1,566	652	215	60	159	218	

Notes: Signed-up is the sample of participants that completed the sign-up survey. Assigned to treatment is the sample of participants that were eligible for the study. Treatments include all participants assigned to treatment (n=652). Household size (adults and children), # children, years in residence, household income and having a programmable thermostat are from the signup survey. Participation in the Alaska Home Energy Rebate Program (HERP) and reported average January gas bill are from the confirmation survey. Year built and home value are from tax records. Feb-March 2018 gas bill is from ENSTAR administrative records. Average temperature in the house (Avg temp) are reported temperature settings from the confirmation survey. The F-test is the joint test that the treatments are jointly zero.

Table C.2: Intention to treat - Feb 9-Feb 28 and March 9-April5, i.e. excluding 2 weeks of data

	2019 OLS	2019 Median	2020 OLS	2020 Median
Information Only (T1)	0.016	0.025	0.025	0.031
s.e.	(0.011)	(0.011)	(0.015)	(0.015)
p-val	[0.148]	[0.025]	[0.100]	[0.034]
Information and Incentive (T2)	-0.027	-0.021	-0.010	-0.013
s.e.	(0.009)	(0.011)	(0.012)	(0.015)
p-val.	[0.004]	[0.048]	[0.387]	[0.372]
Incentive Only (T3)	-0.027	-0.020	-0.005	-0.007
s.e.	(0.008)	(0.008)	(0.011)	(0.012)
p-val.	[0.001]	[0.014]	[0.650]	[0.600]
Constant	-0.102	-0.107	0.085	0.084
s.e.	(0.006)	(0.005)	(0.008)	(0.010)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]
T2-T3	-0.000	-0.001	-0.005	-0.006
H0:T2-T3=0, p-val.	[0.965]	[0.920]	[0.656]	[0.605]
T2-T1	-0.043	-0.047	-0.035	-0.044
H0:T2-T1=0, p-val.	[0.000]	[0.001]	[0.023]	[0.002]
F-test/Chi2-test	7.565	19.406	2.594	11.745
p-val.	[0.001]	[0.000]	[0.076]	[0.008]

Notes: Dependent variable is change in log of total consumption for period Feb 1 - April 5 from 2018 to year listed in column heading. 546 obs for Columns 1-2, 500 obs for Columns 3-4, one observation per household.. Estimates include inverse probability weighting controlling for treatment and survey completion payment (either \$40 or \$60) assignment. Standard errors in parentheses, p-values in square brackets. OLS estimates exclude observations three standard deviations away from the mean. Results using robust regression methods are shown in Table C.3. Table C.4 reproduces this table and includes permutation p-values.

Table C.3: Intention to treat - Robust regressions

	2019 M	2019 S	2019 MM	2020 M	2020 S	2020 MM
Information Only (T1)	0.023	0.024	0.024	0.024	0.031	0.031
s.e.	(0.011)	(0.010)	(0.010)	(0.012)	(0.017)	(0.012)
p-val	[0.043]	[0.023]	[0.015]	[0.049]	[0.060]	[0.009]
Information and Incentive (T2)	-0.025	-0.025	-0.025	-0.009	-0.017	-0.006
s.e.	(0.010)	(0.025)	(0.010)	(0.011)	(0.016)	(0.011)
p-val.	[0.012]	[0.316]	[0.015]	[0.408]	[0.302]	[0.557]
Incentive Only (T3)	-0.021	-0.017	-0.022	-0.005	-0.003	-0.003
s.e.	(0.008)	(0.012)	(0.008)	(0.010)	(0.014)	(0.010)
p-val.	[0.012]	[0.170]	[0.007]	[0.613]	[0.854]	[0.737]
Constant	-0.103	-0.108	-0.103	0.090	0.090	0.089
s.e.	(0.006)	(0.006)	(0.006)	(0.007)	(0.012)	(0.007)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
T2-T3	-0.004	-0.008	-0.003	-0.004	-0.014	-0.003
H0:T2-T3=0, p-val.	[0.683]	[0.751]	[0.783]	[0.711]	[0.253]	[0.770]
T2-T1	-0.048	-0.048	-0.049	-0.033	-0.048	-0.037
H0:T2-T1=0, p-val.	[0.000]	[0.059]	[0.000]	[0.009]	[0.002]	[0.002]
F-test/Chi2-test	7.841	3.441	4.714	3.560	4.970	3.066
p-val.	[0.000]	[0.033]	[0.001]	[0.029]	[0.007]	[0.016]

Notes: Dependent variable is change in log of total consumption for period Feb 1 - April 5 from 2018 to year listed in column heading. 546 obs for Columns 1-3, 500 obs for Columns 4-6, one observation per household.. Robust estimates, as suggested by Han et al. (2021) and Coibion et al. (2019), include M-estimator (or Huber), S-estimator (or least-trimmed regression) and MM-estimator for 2019 v. 2018 and 2020 v. 2018. Estimates include inverse probability weighting controlling for treatment and survey completion payment (either \$40 or \$60) assignment. Robust standard errors in parentheses, p-values from the regression in square brackets.

Table C.4: Intention to treat including permutation tests

	2019 OLS	2019 Median	2020 OLS	2020 Median
Information Only (T1)	0.021	0.024	0.033	0.027
s.e.	(0.010)	(0.010)	(0.011)	(0.014)
p-val	[0.040]	[0.018]	[0.004]	[0.051]
perm. p-val	<0.084>	<0.084>	<0.022>	<0.088>
Information and Incentive (T2)	-0.025	-0.025	-0.003	-0.010
s.e.	(0.009)	(0.011)	(0.010)	(0.012)
p-val.	[0.006]	[0.026]	[0.799]	[0.424]
perm. p-val	<0.004>	<0.012>	<0.769>	<0.367>
Incentive Only (T3)	-0.025	-0.019	-0.006	-0.005
s.e.	(0.008)	(0.009)	(0.010)	(0.012)
p-val.	[0.001]	[0.030]	[0.538]	[0.678]
perm. p-val	<0.001>	<0.044>	<0.502>	<0.581>
Constant	-0.103	-0.105	0.088	0.091
s.e.	(0.006)	(0.006)	(0.007)	(0.010)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]
perm. p-val	<0.006>	<0.151>	<0.984>	<0.759>
T2-T3	0.000	-0.006	0.003	-0.005
H0:T2-T3=0, p-val.	[0.976]	[0.630]	[0.731]	[0.651]
T2-T1	-0.046	-0.049	-0.036	-0.037
H0:T2-T1=0, p-val.	[0.000]	[0.000]	[0.002]	[0.004]
F-test/Chi2-test	8.858	20.944	5.719	9.257
p-val.	[0.000]	[0.000]	[0.004]	[0.026]

Notes: Dependent variable is change in log of total consumption for period Feb 1 - April 5 from 2018 to year listed in column heading. 546 obs for Columns 1-2, 500 obs for Columns 3-4, one observation per household. Estimates include inverse probability weighting controlling for treatment and survey completion payment (either $40 or $60) assignment. Standard errors in parentheses, p-values from the regression in square brackets and p-values from a permutation test in angle brackets. OLS estimates excludes observations three standard deviations away from the mean.

Table C.5: 2019 v. 2018: Intention to treat main effects Muralidharan et al. (2019) long regression

	2019 OLS	2019 Median	2020 OLS	2020 Median
Information (T1)	0.019	0.022	0.033	0.022
s.e.	(0.010)	(0.009)	(0.011)	(0.009)
p-val	[0.074]	[0.019]	[0.004]	[0.019]
Incentive (T2)	-0.028	-0.017	-0.006	-0.017
s.e.	(0.008)	(0.008)	(0.010)	(0.008)
p-val.	[0.001]	[0.029]	[0.538]	[0.029]
Information and Incentive (T3)	-0.019	-0.027	-0.033	-0.027
s.e.	(0.014)	(0.014)	(0.015)	(0.014)
p-val.	[0.181]	[0.046]	[0.033]	[0.046]
Constant	-0.100	-0.100	0.088	-0.100
s.e.	(0.006)	(0.005)	(0.007)	(0.005)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]
T2-T3	0.037	0.049	0.066	0.049
H0:T2-T3=0, p-val.	[0.103]	[0.020]	[0.009]	[0.020]
T2-T1	-0.009	0.010	0.027	0.010
H0:T2-T1=0, p-val.	[0.635]	[0.606]	[0.242]	[0.606]

Notes: Dependent variable is change in log of total consumption for period Feb 1 - April 5 from 2018 to year listed in column heading. 546 obs for Columns 1-2, 500 obs for Columns 3-4, one observation per household. Estimates include inverse probability weighting controlling for treatment and survey completion payment (either \$40 or \$60) assignment. Standard errors in parentheses, p-values from the regression in square brackets and p-values from a permutation test in angle brackets.

Table C.6: Treatment on the treated

	2019 OLS	2019 IVQR (Median)	2020 OLS	2020 IVQR (Median)
Information Only (T1)	0.033	0.037	0.050	0.049
s.e.	(0.016)	(0.014)	(0.018)	(0.024)
p-val	[0.047]	[0.010]	[0.007]	[0.043]
Information and Incentive (T2)	-0.030	-0.030	-0.003	-0.014
s.e.	(0.011)	(0.014)	(0.012)	(0.017)
p-val.	[0.006]	[0.027]	[0.798]	[0.404]
Incentive Only (T3)	-0.028	-0.022	-0.007	-0.006
s.e.	(0.009)	(0.010)	(0.011)	(0.015)
p-val.	[0.001]	[0.028]	[0.535]	[0.666]
Constant	-0.103	-0.105	0.088	0.091
s.e.	(0.006)	(0.006)	(0.007)	(0.010)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]
T2-T3	-0.002	-0.009	0.003	-0.008
H0:T2-T3=0, p-val.	[0.850]	[0.543]	[0.759]	[0.553]
T2-T1	-0.062	-0.068	-0.053	-0.063
H0:T2-T1=0, p-val.	[0.000]	[0.000]	[0.002]	[0.003]

Note: Dependent variable is change in log of total consumption for period Feb 1 - April 5 from 2018 to year listed in column heading. 546 obs for Columns 1-2, 500 obs for Columns 3-4, one observation per household. The measure of compliers in Information Only is an indicator of an installed reader. In Information and Incentive, it is an indicator of having an installed reader or completion of the endline survey (endline survey completion was a condition to receive incentive payments). In Incentive Only, it is completion of the endline survey. To estimate the TOT for each treatment simultaneously, we create three variables that equal 1 if the treatment is adopted and zero if not. Treatment assignments are used as instrumental variables. Estimates use Chernozhukov and Hansen (2005) IVQR.

Table C.7: Intention to treat - first month

	2019 OLS	2019 Median	2020 OLS	2020 Median
Information Only (T1)	0.029	0.032	0.026	0.025
s.e.	(0.015)	(0.012)	(0.013)	(0.015)
p-val	[0.058]	[0.008]	[0.050]	[0.103]
Information and Incentive (T2)	-0.015	-0.019	-0.004	-0.002
s.e.	(0.010)	(0.012)	(0.010)	(0.010)
p-val.	[0.137]	[0.135]	[0.668]	[0.806]
Incentive Only (T3)	-0.015	-0.011	-0.005	0.005
s.e.	(0.009)	(0.011)	(0.010)	(0.009)
p-val.	[0.115]	[0.303]	[0.628]	[0.623]
Constant	-0.089	-0.087	0.115	0.107
s.e.	(0.007)	(0.007)	(0.007)	(0.007)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]
T2-T3	-0.001	-0.008	0.001	-0.007
H0:T2-T3=0, p-val.	[0.956]	[0.560]	[0.951]	[0.440]
T2-T1	-0.044	-0.051	-0.030	-0.027
H0:T2-T1=0, p-val.	[0.005]	[0.000]	[0.022]	[0.070]
F-test/Chi2-test	4.110	14.564	2.746	3.515
p-val.	[0.017]	[0.002]	[0.065]	[0.319]

Notes: Dependent variable is change in log of total consumption for first month from 2018 to year listed in column heading. 546 obs for Columns 1-2, 500 obs for Columns 3-4, one observation per household. Estimates include inverse probability weighting controlling for treatment and survey completion payment (either \$40 or \$60) assignment. Standard errors in parentheses, p-values from the regression in square brackets and p-values from a permutation test in angle brackets.

Table C.8: Intention to treat - second month

	2019 OLS	2019 Median	2020 OLS	2020 Median
Information Only (T1)	0.017	0.023	0.019	0.023
s.e.	(0.015)	(0.014)	(0.014)	(0.015)
p-val	[0.254]	[0.102]	[0.187]	[0.121]
Information and Incentive (T2)	-0.034	-0.022	-0.019	-0.010
s.e.	(0.012)	(0.014)	(0.013)	(0.016)
p-val.	[0.003]	[0.126]	[0.151]	[0.525]
Incentive Only (T3)	-0.030	-0.025	-0.013	-0.008
s.e.	(0.011)	(0.011)	(0.013)	(0.015)
p-val.	[0.005]	[0.020]	[0.310]	[0.593]
Constant	-0.134	-0.141	0.104	0.096
s.e.	(0.008)	(0.008)	(0.009)	(0.009)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]
T2-T3	-0.004	0.004	-0.007	-0.002
H0:T2-T3=0, p-val.	[0.755]	[0.794]	[0.614]	[0.901]
T2-T1	-0.051	-0.045	-0.038	-0.034
H0:T2-T1=0, p-val.	[0.001]	[0.007]	[0.010]	[0.060]
F-test/Chi2-test	7.071	15.668	3.355	4.637
p-val.	[0.001]	[0.001]	[0.036]	[0.200]

Notes: Dependent variable is change in log of total consumption for second month from 2018 to year listed in column heading. 546 obs for Columns 1-2, 500 obs for Columns 3-4, one observation per household. Estimates include inverse probability weighting controlling for treatment and survey completion payment (either \$40 or \$60) assignment. Standard errors in parentheses, p-values from the regression in square brackets and p-values from a permutation test in angle brackets.

Table C.9: Characteristics of compliers

	Complied with any treatment				Installed gas reader			
	Compliers (C)	Never-takers (NT)	$C < NT$	$C > NT$	Compliers (C)	Never-takers (NT)	$C < NT$	$C > NT$
Gas Bill (log)	5.84	5.94	0.93	0.07	5.85	5.92	0.85	0.15
Age of house (log)	3.55	3.55	0.51	0.49	3.58	3.46	0.08	0.92
Building size (log)	7.54	7.62	0.93	0.07	7.54	7.59	0.78	0.22
Property value (log)	12.71	12.74	0.71	0.28	12.71	12.73	0.61	0.39
HH size	3.03	2.93	0.31	0.69	2.94	3.19	0.83	0.17
Children	0.97	0.84	0.20	0.81	0.94	0.97	0.59	0.41
Tenure in house	4.95	5.75	1.00	0.00	4.82	5.58	0.99	0.01
Participated in HERP	0.24	0.23	0.40	0.59	0.24	0.23	0.51	0.49
Has nest thermostat	0.14	0.09	0.13	0.87	0.15	0.09	0.13	0.86
Has prog thermostat	0.79	0.77	0.37	0.63	0.80	0.77	0.36	0.64
Has gas stove	0.62	0.63	0.57	0.43	0.65	0.61	0.32	0.68
Has gas dryer	0.34	0.46	0.94	0.06	0.34	0.39	0.70	0.30

Notes: "Complied with any treatment" columns use data from all four treatments. The dependent variable in the (C) and (NT) columns is a dummy variable that equals one if the household is in any treatment and was a complier. "Installed a gas reader" drops the Incentive Only treatment. The dependent variable in the (C) and (NT) columns is a dummy variable that equals one if the household installed a reader. $C < NT$ is the p-value testing the null hypothesis that the Complier group average is less than the Never-Taker group average for that row. $C > NT$ is the p-value testing the null hypothesis that the Complier group average is greater than the Never-Taker group average for that row.

Table C.10: TREATMENT ON THE TREATED - LQTE COMPARED TO IVQR - 2019 v. 2018

	q(10)	q(25)	q(50)	q(75)	q(90)
			IVQR		
Information Only (T1)	0.045	0.046	0.037	0.020	0.001
s.e.	(0.031)	(0.018)	(0.016)	(0.024)	(0.037)
p-val	[0.146]	[0.011]	[0.020]	[0.402]	[0.981]
Information and Incentive (T2)	-0.046	-0.046	-0.030	-0.011	-0.031
s.e.	(0.023)	(0.012)	(0.014)	(0.015)	(0.023)
p-val.	[0.046]	[0.000]	[0.030]	[0.452]	[0.167]
Incentive Only (T3)	-0.025	-0.028	-0.022	-0.016	-0.044
s.e.	(0.022)	(0.009)	(0.010)	(0.011)	(0.015)
p-val.	[0.250]	[0.002]	[0.029]	[0.145]	[0.003]
			LQTE		
Information Only (T1)	0.054	0.047	0.046	0.020	0.000
s.e.	(0.032)	(0.020)	(0.021)	(0.028)	(0.051)
p-val	[0.096]	[0.020]	[0.032]	[0.471]	[1.000]
Information and Incentive (T2)	-0.034	-0.047	-0.034	-0.013	-0.032
s.e.	(0.026)	(0.014)	(0.014)	(0.016)	(0.025)
p-val.	[0.192]	[0.001]	[0.015]	[0.400]	[0.208]
Incentive Only (T3)	-0.024	-0.030	-0.021	-0.021	-0.046
s.e.	(0.024)	(0.011)	(0.011)	(0.012)	(0.019)
p-val.	[0.311]	[0.007]	[0.050]	[0.099]	[0.016]

Note: Dependent variable is change in log of total consumption for period Feb 1 - April 5. 546 observations, one observation per household. The measure of compliers in Information Only is an indicator of an installed reader. In Information and Incentive, it is an indicator of having an installed reader or completion of the endline survey (endline survey completion was a condition to receive incentive payments). In Incentive Only, it is completion of the endline survey. Estimates of TOT are obtained by comparing each treatment against the control separately. Estimates use Abadie et al. (2002) for LQTE and Chernozhukov and Hansen (2005) IVQR.

Table C.11: Treatment on the treated - LQTE compared to IVQR - 2020 v. 2018

	q(10)	q(25)	q(50)	q(75)	q(90)
			IVQR		
Information Only (T1)	0.056	0.052	0.045	0.012	-0.038
s.e.	(0.022)	(0.021)	(0.023)	(0.019)	(0.027)
p-val	[0.010]	[0.013]	[0.045]	[0.545]	[0.154]
Information and Incentive (T2)	0.005	-0.005	-0.016	-0.017	-0.031
s.e.	(0.017)	(0.014)	(0.017)	(0.016)	(0.026)
p-val.	[0.754]	[0.743]	[0.352]	[0.273]	[0.237]
Incentive Only (T3)	-0.012	0.005	-0.007	-0.015	-0.018
s.e.	(0.025)	(0.012)	(0.013)	(0.011)	(0.025)
p-val.	[0.651]	[0.675]	[0.574]	[0.192]	[0.460]
			LQTE		
Information Only (T1)	0.053	0.054	0.052	0.025	-0.041
s.e.	(0.029)	(0.024)	(0.029)	(0.045)	(0.042)
p-val	[0.069]	[0.023]	[0.067]	[0.581]	[0.327]
Information and Incentive (T2)	-0.000	-0.004	-0.007	-0.017	-0.031
s.e.	(0.021)	(0.016)	(0.018)	(0.018)	(0.029)
p-val.	[0.996]	[0.793]	[0.679]	[0.351]	[0.295]
Incentive Only (T3)	-0.010	0.002	-0.009	-0.011	-0.013
s.e.	(0.024)	(0.014)	(0.014)	(0.014)	(0.027)
p-val.	[0.666]	[0.904]	[0.518]	[0.427]	[0.625]

Note: Dependent variable is change in log of total consumption for period Feb 1 - April 5. 500 observations, one observation per household. The measure of compliers in Information Only is an indicator of an installed reader. In Information and Incentive, it is an indicator of having an installed reader or completion of the endline survey (endline survey completion was a condition to receive incentive payments). In Incentive Only, it is completion of the endline survey. Estimates of TOT are obtained by comparing each treatment against the control separately. Estimates use Abadie et al. (2002) for LQTE and Chernozhukov and Hansen (2005) IVQR.

Table C.12: Implied average treatment effects, assuming rank similarity (Chernozhukov and Hansen, 2005)

	2019 v 2018	2020 v 2018
Information Only	0.032	0.031
s.e.	(0.023)	(0.024)
Information and Incentive	-0.031	-0.012
s.e.	(0.012)	(0.013)
Incentive Only	-0.025	-0.007
s.e.	(0.010)	(0.012)

Notes: 546 obs for Columns 1, 500 obs for Columns 2, one observation per household. Implied average treatment effects are recovered from treatment on the treated quantile regressions under the assumption of rank similarity (Chernozhukov and Hansen, 2005). Section D.3 presents the test for rank similarity and shows it holds with the data.

Table C.13: Effect of attention on consumption - IVQR - 2019 v. 2018

	Q(10)	Q(25)	Q(50)	Q(75)	Q(90)
(Clicks) Information Only (T1)	0.016	0.015	0.012	0.007	0.000
s.e.	(0.010)	(0.005)	(0.005)	(0.007)	(0.011)
p-val	[0.133]	[0.005]	[0.022]	[0.355]	[0.998]
(Clicks) Info and Incentive (T2)	-0.017	-0.015	-0.011	-0.003	-0.010
s.e.	(0.012)	(0.006)	(0.005)	(0.005)	(0.006)
p-val.	[0.146]	[0.006]	[0.025]	[0.529]	[0.099]
Constant	-0.205	-0.149	-0.105	-0.058	0.014
s.e.	(0.017)	(0.006)	(0.006)	(0.007)	(0.012)
p-val.	[0.000]	[0.000]	[0.000]	[0.000]	[0.260]

Note: 365 observations (Incentives Only treatment excluded) observations, one observation per household. Attention is measured by the inverse hyperbolic sine of the number of clicks on the portal. To estimate the effect of attention for each information treatment simultaneously, this measure is interacted with the indicator of treatment assignment. Treatment assignments are used as instrumental variables. Estimates use Chernozhukov and Hansen (2005) instrumental variable quantile regression (IVQR). Standard errors in parentheses, and p-values in square brackets.

Table C.14: Effect of attention on consumption - IVQR - 2020 v. 2018 - using only clicks in 2nd month

	q(10)	q(25)	q(50)	q(75)	q(90)
(Clicks) Info only (T1)	0.027	0.035	0.059	0.033	-0.010
s.e.	(0.013)	(0.026)	(0.024)	(0.029)	(0.012)
p-val	[0.033]	[0.178]	[0.015]	[0.253]	[0.394]
(Clicks) Info and Incentive (T2)	0.004	-0.003	-0.007	-0.008	-0.016
s.e.	(0.011)	(0.009)	(0.008)	(0.008)	(0.010)
p-val.	[0.755]	[0.754]	[0.327]	[0.348]	[0.108]
Constant	-0.028	0.030	0.091	0.149	0.224
s.e.	(0.013)	(0.009)	(0.008)	(0.011)	(0.020)
p-val.	[0.031]	[0.001]	[0.000]	[0.000]	[0.000]

Note: 334 observations (Incentives Only treatment excluded), one observation per household. Attention is measured by the inverse hyperbolic sine of the number of clicks on the portal in the second month. To estimate the effect of attention for each information treatment simultaneously, this measure is interacted with the indicator of treatment assignment. Treatment assignments are used as instrumental variables. Estimates use Chernozhukov and Hansen (2005) instrumental variable quantile regression (IVQR). Standard errors in parentheses, and p-values in square brackets.

Table C.15: Determinants of attention (number of clicks)

	Number of visits to portal (ihs)
Expenditure in gas at baseline (2018, logs)	-0.99*
	(0.55)
Age of building (logs)	0.68*
	(0.37)
Building size (logs)	1.25*
	(0.60)
Family size	0.03
	(0.11)
Responder is female	-0.31
	(0.31)
Household income (logs)	-0.64*
	(0.37)
Info and Incentives	0.35
	(0.35)
Constant	3.40
	(5.40)
Observations	187
R2	0.070

Notes: Dependent variable is number of visits to the portal converted to inverse hyperbolic sin. Covariate data is from the confirmation survey and tax records for the household. * $p<0.10$, ** $p<0.05$, *** $p<0.010$.

Table C.16: Relationship between exploration and consumption 2020 v 2018

	(1)	(2)	(3)	(4)	(5)	(6)
90Q - mean reported inhouse temperature	0.004*					
	(0.002)					
	(0.089)					
90Q - reported temp. (weekday 8am-2pm)		0.003				
		(0.002)				
		(0.102)				
90Q - reported temp. (weekday 3pm-9pm)			0.002			
			(0.002)			
			(0.470)			
90Q - reported temp. (weekday 10pm-7am)				0.004*		
				(0.002)		
				(0.093)		
90Q - reported temp. (weekend 8am-9pm)					0.004*	
					(0.002)	
					(0.060)	
90Q - reported temp. (weekend 10pm-7am)						0.004**
						(0.002)
						(0.031)
Constant	0.074***	0.075***	0.080***	0.080***	0.072***	0.075***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs	85	84	85	85	83	83
R2	0.034	0.032	0.006	0.034	0.043	0.056

Notes: The regressions use data from readers to calculate the 90th quantile of the indoor temperature during different times of the week and hours of the day. The dependent variable is the difference in consumption from 2018 to 2020. The explanatory variable is the difference between the observed 90 quantile and the reported indoor temperature reported in the confirmation survey. * p<0.10, ** p<0.05, *** p<0.010.

Table C.17: Reaction to outside temperature by treatment

	Information Only	Information and Incentive
Temperature outside	0.116***	0.064**
	(0.036)	(0.027)
Rank of use × Temperature outside	-0.178**	-0.052
	(0.076)	(0.045)
Obs	32835	97738
R2	0.832	0.804

* p<0.10, ** p<0.05, *** p<0.010

D Additional analysis

D.1 Demand for information

Table D.1 presents detail information on the use of the portal. We aggregate information on an hourly basis. In particular, we create a variable that equals 1 if a panel of a portal is used in a particular hour and 0 if not. This allows us to aggregate information that is recorded on an event basis. The table shows linear regressions for six different binary measures of information acquisition by hour: 1. visited the site, 2. checked relative cost panel, 3. checked relative gas use panel, 4. checked cumulative cost panel, 5. checked cumulative gas use panel and 6. checked the panel with indoor temperature, outdoor temperature and cost. The analysis shows that the Information and Incentive treatment increased the use of the portal. There is a consistent decrease in use of the portal as time passed, but the announcement of an extension increased use. Consistent with the existence of costs to acquire/process information, we see that the use of the portal is less frequent during working hours.

D.2 Behavior changes in information treatments

The device used in the information treatments to read the gas meter also recorded the temperature inside the house. Thus, we observe at what temperature participants kept their house over the course of the study. Figure D.1 shows the indoor temperature for each hour of the day split by weekday and weekend for the Information Only and Information and Incentive treatments. The Information Only group kept their house warmer than the Information and Incentives group and did not differentiate between weekday and weekend. The Information and Incentive group kept the house temperature lower overall and lowered it during working hours on weekdays relative to weekends. Figure D.2 presents the average hourly cost per day per treatment using data from the readers. A gap in average costs emerges in the second week of

Table D.1: Portal usage by hour

	(1)	(2)	(3)	(4)	(5)	(6)
		Relative		Cumulative		
	Any	Cost	Gas	Cost	Gas	Temp.
Information and Incentive	0.0040	0.0042	0.0015	0.0005	0.0020	0.0008
	(0.0031)	(0.0029)	(0.0007)	(0.0003)	(0.0006)	(0.0006)
Weekday	-0.0037	-0.0009	0.0001	0.0001	0.0004	-0.0002
	(0.0008)	(0.0007)	(0.0003)	(0.0002)	(0.0004)	(0.0002)
Days since start of experiment	-0.0021	-0.0015	-0.0002	-0.0001	-0.0004	-0.0002
	(0.0002)	(0.0002)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
Days since start of experiment (sq)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Working hours (8 am - 5 pm)	-0.0070	-0.0070	-0.0001	-0.0005	-0.0019	-0.0007
	(0.0012)	(0.0012)	(0.0008)	(0.0002)	(0.0004)	(0.0003)
Announcement of extension	0.0042	0.0036	0.0006	0.0007	0.0014	0.0005
	(0.0016)	(0.0014)	(0.0003)	(0.0003)	(0.0006)	(0.0005)
Announcement of extension $\times$ Info and incentive	0.0009	0.0006	-0.0001	-0.0004	-0.0008	0.0004
	(0.0020)	(0.0018)	(0.0006)	(0.0004)	(0.0006)	(0.0006)
Constant	0.0441	0.0303	0.0037	0.0028	0.0078	0.0047
	(0.0044)	(0.0040)	(0.0008)	(0.0005)	(0.0010)	(0.0010)
Observations	165312	165312	165312	165312	165312	165312
Number of housholds	123	123	123	123	123	123

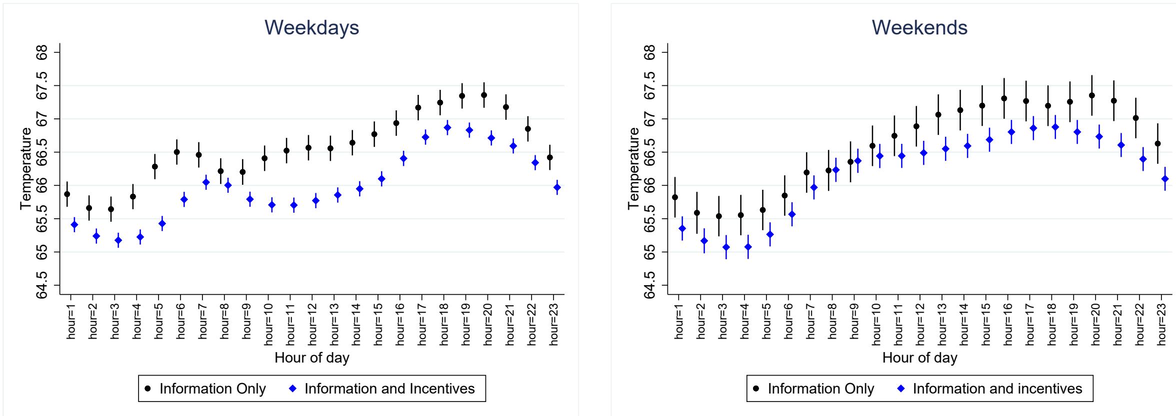
Notes: Linear regressions for six different measures of information acquisition. Observations are at the hour level for the two-month study period. All dependent variables are dummy variables that equal one if the outcome condition is met. In Column 1, the outcome is visiting the site, Column 2 is checking the relative cost panel, Column 3 is checking the relative gas use panel, Column 4 is checking the cumulative cost panel, Column 5 is checking the cumulative gas use panel and Column 6 is checking the panel with indoor temperature, outdoor temperature and cost. Standard errors are in parentheses and clustered by household.

the intervention.

We can test if access to high-frequency data remains useful with the "surprise" extension of the experiment to a second period. The extension could generate renewed demand for information and behavioral changes. Figure D.3 presents the propensity to visit the portal by hour for each day of the intervention. The graph approximates use using polynomials for the period before and after the announcement of the extension of the experiment. The figure on the left shows the response of those in the Information Only treatment and figure on the right shows the response of those in the Information and Incentive treatment. There is a significant increase in visits to the portal in the Information and Incentive treatment but not in the Information Only.[45] This is consistent with information usage being instrumental and a validation of the number of logs into the portal as a valid proxy for attention.

[45]The p-value associated with the discontinuity test of Calonico et al. (2014) is 0.687 for Information Only and 0.007 for the Information and Incentive.

During the experiment



Last two weeks of the experiment

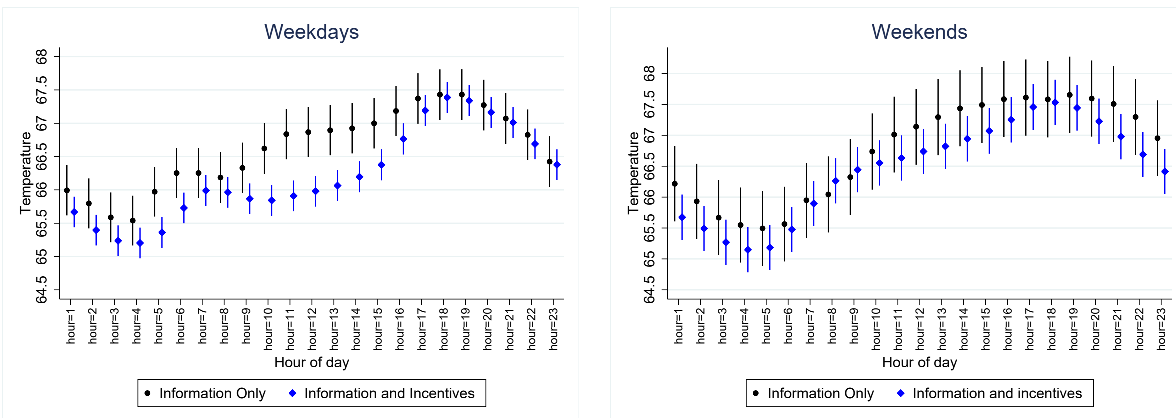


Figure D.1: Indoor temperature by hour of day

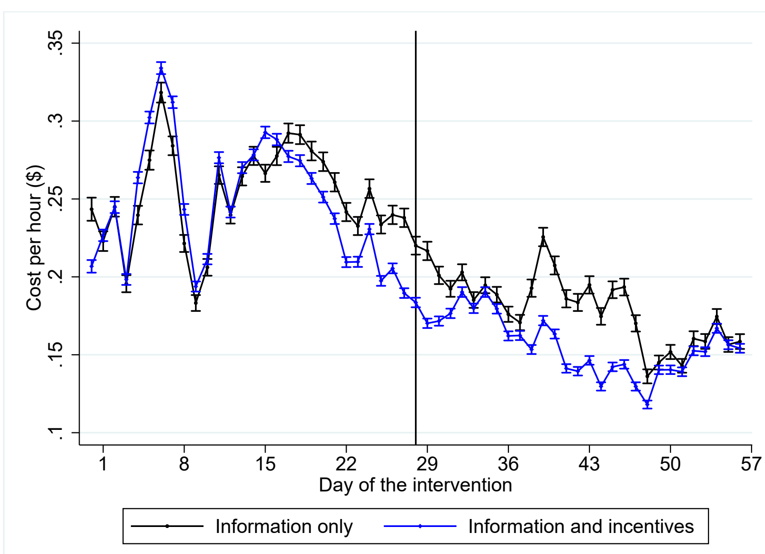
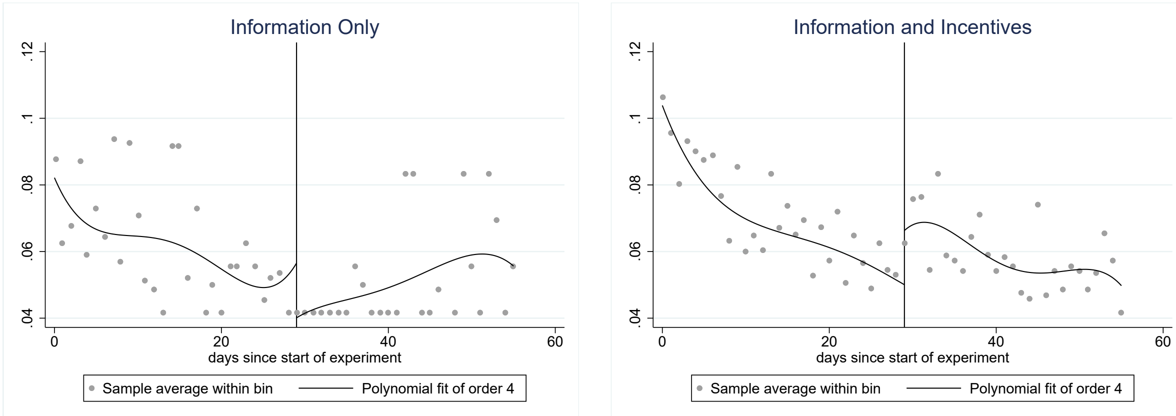


Figure D.2: Average hourly cost by day of intervention

Hourly use of portal (Clicked=1)



Note: Estimates use hourly measurements of hourly use of portal.

Figure D.3: Reaction to surprise extension of study

D.3 Rank similarity and the causal effect of information

We would like to know if real-time information on energy usage has a causal effect on consumption behavior in the population. To explore this, we appeal to the condition of rank similarity (Chernozhukov and Hansen, 2005; Frandsen and Lefgren, 2018). If rank similarity holds in our experiment, we can estimate the causal impact of portal usage on energy consumption using treatment assignment as an instrument.[46]

Rank similarity (Chernozhukov and Hansen, 2005) requires that the conditional distribution of ranks across counterfactual treatments to be identical in all treatment states. Frandsen and Lefgren (2018) develop a test of rank similarity based on the existence of variables that predict treatment ranks but not treatment assignment. We construct such a variable using administrative data on the yearly change in consumption during the month prior to the experiment (January 2019). We confirm that this variable is not correlated with treatment assignment or installation of the reader at conventional levels.

The rank similarity test examines if the rank in consumption pre-experiment pre-

[46]We test for rank similarity because we would like to estimate the causal effect of a continuous treatment, i.e. portal usage, on consumption. Following the approach in Chernozhukov and Hansen (2005), which requires rank similarity, provides such estimates.

dicts treatment outcome ranks but not treatment selection.[47] Table D.2 presents the results of the test. Columns 1 and 2 pertain to the Information Only treatment and Columns 3 and 4 to the Information and Incentive treatment. The Control group is included in all regressions. The dependent variable is the rank in either the treated or control condition. The rank of yearly change in consumption in January strongly predicts ranks. However, it does not have a differential impact on rank across treated and non-treated participants, as expected if rank similarity holds.

Table D.2: Test of rank similarity

	(1)	(2)	(3)	(4)
	Information Only		Information and Incentive	
	Median regression	OLS	Median regression	OLS
Reader installed	-0.1040	0.0222	-0.0463	0.0169
	(0.0986)	(0.0612)	(0.0701)	(0.0475)
Yearly change (%) in cons. in January 2019	1.6768	1.0905	1.7018	1.2201
	(0.2574)	(0.1599)	(0.2006)	(0.1361)
Reader installed×Yearly change (%) in cons. in Jan. 2019	-0.7495	0.0374	0.0199	0.2366
	(0.6968)	(0.4329)	(0.5880)	(0.3990)
Constant	0.6286	0.5948	0.6411	0.6032
	(0.0340)	(0.0211)	(0.0271)	(0.0184)
Observations	274	274	372	372
R2		0.168		0.207

Notes: Standard errors in parentheses. The test of rank similarity follows Frandsen and Lefgren (2018). Ranks are constructed using the Chernozhukov and Hansen (2006) procedure. Estimations are done separately for the Information Only and Information and Incentive treatments, hence the difference in observations.

D.4 Behavior changes by all participants

Table D.3 reports strategies and behavioral changes undertaken by participants across all treatments using data from the endline survey. The outcome variable is a dummy variable that equals one if the behavior is undertaken and zero otherwise. These Probit regressions incorporate inverse probability weight to adjust for attrition. We take advantage of the fact that participants were randomly assigned to different completion bonuses and these bonuses strongly predict endline survey completion. Since

[47] For our test, the dependent variable is defined as the rank of change in consumption during the study, Feb-April 2019 compared to Feb-April 2018. The independent variables are a dummy for whether the reader was installed, the rank of change in consumption in January 2019 compared to January 2018 and an interaction of the two variables. A set of regressions is run combining the Control and Information Only participants and another set with combining the Control and Information and Incentive participants and reported in Table D.2.

we have little evidence of selection into the experiment (see Section 3), we take these results to be representative of the population under study.

Columns 1-3 in the table show that participants in the Information and Incentive and Incentive Only treatments are more likely to report having made changes to their houses, their behavior and their future behavior. We also observe that those in Information Only report having changed their behavior. In Column 4, all participants report monitoring their energy consumption behavior, regardless of having access to a reader.

Columns 5-10 report specific changes in behavior. These questions were intended to capture potential welfare losses due to incentivized changes in behavior. Those in the incentive treatments report having adopted behavioral changes like taking shorter and colder showers. This points to a trade-off between lower consumption and comfort. There is no evidence that those in the Information Only treatment adopted these type of behavioral changes any differently than the Control group.

This analysis points to reasons why the treatment with monetary incentives to reduce consumption did not lead to permanent behavioral changes. First, participants might already be close to their minimum comfort level and extra reductions in indoor temperature might be too costly. Second, participants might not be able to optimize behavior as to minimize discomfort.

Table D.3: Responses in Endline survey

	(1) Changed Equip.	(2) Changed Behavior	(3) Plan Change	(4) Monitored Consum.	(5) Warmer Clothes	(6) Lower T. Night	(7) Lower T. Day	(8) More Blankets	(9) Shorter Showers	(10) Colder Showers
Info Only	0.0154	0.2160	0.0687	0.1265	0.0123	0.0123	0.0370	-0.0093	0.0864	0.0309
	(0.0707)	(0.0841)	(0.0942)	(0.0691)	(0.0785)	(0.0785)	(0.0785)	(0.0708)	(0.0524)	(0.0479)
Info and Incent.	0.1311	0.4160	0.1402	0.1776	0.1639	0.1639	0.2249	0.0609	0.1026	0.0662
	(0.0474)	(0.0564)	(0.0626)	(0.0463)	(0.0526)	(0.0526)	(0.0526)	(0.0475)	(0.0351)	(0.0321)
Incent. Only	0.0753	0.4450	0.1102	0.1670	0.1529	0.1529	0.2017	0.0506	0.1018	0.0898
	(0.0424)	(0.0504)	(0.0560)	(0.0414)	(0.0471)	(0.0471)	(0.0470)	(0.0424)	(0.0314)	(0.0287)
Constant	0.1235	0.2840	0.4313	0.0679	0.1543	0.1543	0.1296	0.1481	0.0247	0.0247
	(0.0301)	(0.0358)	(0.0394)	(0.0295)	(0.0335)	(0.0335)	(0.0335)	(0.0302)	(0.0223)	(0.0204)
Observations	474	474	456	474	474	474	474	474	474	474
R2	0.018	0.165	0.014	0.044	0.032	0.032	0.054	0.005	0.028	0.022

Notes: Probit regressions with dependent variable coded as one if behavior is reported, zero otherwise. Control group is the excluded category. Standard errors in parentheses. Correction for survey non-response uses inverse probability weighting. Regressions include household characteristics and randomly assigned incentive to complete survey. Outcome variables are: "Changed equip" is whether the household reported changing appliances, windows or thermostat, "Changed behavior" is whether the household reported changing activities, "Plan change" is whether the household reported they planned to change behavior in the future, "Monitored consum" is whether the household reported monitoring energy usage, "Warmer clothes" is whether the household reported putting on warmer clothes, "Lower T. Night" is whether the household reported lowering the temperature at night, "Lower T. Day" is whether the household temperature was lowered during the day, "More blankets" is whether more blankets were used while sleeping, "Shorter showers" is whether people in the household took shorter showers, "Colder showers" is whether people in the household took colder showers.

E Back of the envelope welfare calculations

Using the model derived in Section 2, we assess potential welfare effects from the intervention and distraction.

We start by assessing how much perceived gas costs would have to decrease to explain the results in the Information Only treatment. According to the U.S. Energy Information Administration (EIA) 2021 report, the estimated short-run price elasticity for residential gas is -0.08 while the long-run elasticity is -0.21 (30 year period).[48] Those using the reader in the Information Only treatment (i.e. compliers) consume about 5 percent more in gas than those in the Control group a year after the intervention. This is equivalent to a 2.9 percentage points increase in temperature from 68F, the average survey reported household temperature in the Control group, or a 36.8 percentage point reduction in the perceived cost of energy. For those in the Information and Incentive treatment, they would have to be distracted to have missed this perceived change in the cost of energy.

Alternatively, this result can be explained by participants miscalibrating their demand for a warm house by 2.9 percentage points and access to information reducing this error. Given that participants in the Control group spent \$398 on average during the experiment-equivalent period in 2020, we estimate the extra costs for compliers in the Information Only treatment to be about \$20 over the study period or about 10 extra dollars per month in winter months for them to achieve the desired house warmth. The population average treatment effect estimates in Section 5.4, while somewhat imprecise, suggest access to information would lead to roughly a 3 percentage point increase in gas consumption ($\sim$ \$6 per month). When accompanied with an incentive to reduce consumption, this leads to 2.7 percent reduction in gas consumption ($\sim$ \$5.4 per month). This represents a large welfare gap due to variation

[48] Estimates using our experimental data from the Incentive Only and Control conditions produce an even lower estimate of about -0.02.

in attention driven by incentives.