Beyond Single-Resource Nudges: Experimental Evidence on Integrated Consumption Feedback
- Pre-Analysis Plan -

Jacopo Bonan[2,3,4], Giovanna d'Adda[2,3,4], Cristina Cattaneo[3,4], Jacopo Lunghi[3,4], and Massimo Tavoni[1,3,4]

[1]Politecnico di Milano
[2]University of Milan
[3]RFF-CMCC European Institute on Economics and the Environment (EIEE)
[4]Euro-Mediterranean Centre on Climate Change

1 Introduction

This pre-analysis plan describes a field experiment implemented in collaboration with an Italian multi-utility to evaluate the effects of a Consolidated Home Report (CHR). Eligible residential customers already receive separate digital Home Reports for at least two resources among water, gas, and electricity. Customers assigned to treatment continue to receive these resource-specific reports and additionally receive the CHR, which presents an Ecoscore summarizing their relative environmental performance across the resources in their bundle. Customers assigned to control continue to receive only the standard resource-specific reports.

Households consume several resources jointly, and the use of water, gas, and electricity may be connected through both technologies and household behavior. Nevertheless, behavioral interventions and consumption reports are generally designed and evaluated one resource at a time. This fragmented approach may obscure relationships across resource use and may be difficult for customers to process when several similar reports are received simultaneously. The CHR adds a consolidated summary of overall environmental performance to the existing resource-specific reports and is intended to make multi-resource information easier to interpret.

The study addresses five empirical research questions. Research Question 1 evaluates effects on pooled and resource-specific consumption. Research Question 2 examines contract deactivation. Research Questions 3, 4 and 5 use a customer survey to explore, respectively, self-reported recent opening of the Home Reports, satisfaction with the reporting service and the behavioral channels associated with the CHR.
At the time of writing this pre-analysis plan, the randomized controlled trial had already been implemented and treatment assignment had been completed. However, the multi-utility had not yet shared any post-treatment administrative or survey outcome data with the research team. Consequently, all empirical specifications, outcome definitions, hypothesis tests, and robustness analyses described in this document were pre-specified before access to post-treatment data. The only data available to the researchers at the time of writing were the pre-treatment administrative data required to define the study sample, construct the randomization strata, and implement treatment assignment.

2 Related literature

The paper contributes to four strands of the literature.
Social information and household resource conservation. First, the project contributes to the large literature on social information, feedback, and non-price interventions for household resource conservation. Social information programs are widely used because they can change behavior without relying on prices, taxes, or direct regulation. In the energy domain, evidence shows that home energy reports using social comparison reduce residential electricity consumption (Allcott, 2011; Ayres et al., 2013; Allcott and Rogers, 2014; Bonan et al., 2020, 2021). Meta-analyses show that information-based interventions can reduce electricity, gas, and water consumption, although effect sizes vary across contexts, designs, and populations (Delmas et al., 2013; Nemati and Penn, 2020).
A related strand evaluates home water reports showing that norm-based messages and social comparisons can reduce residential water use (Ferraro et al., 2011; Ferraro and Price, 2013; Ferraro and Miranda, 2013; Bernedo et al., 2014; Brent et al., 2015; Jessoe et al., 2021; Jaime Torres and Carlsson, 2018; Miranda et al., 2020; Bonan et al., 2024). Recent evidence also emphasizes that the level of detail at which feedback is provided can materially affect its effectiveness. In a field experiment with smart-meter customers, **?** compare standard aggregate electricity feedback with feedback that additionally disaggregates consumption at the appliance level. They find that appliance-level information reduces electricity consumption by approximately 5 percent relative to aggregate feedback alone, while also improving consumers' knowledge of appliance energy use. Their results suggest that summary information and disaggregated feedback need not be substitutes: aggregate indicators can facilitate overall evaluation, whereas more granular information may help customers identify the behaviors responsible for their consumption.

Most of this literature studies feedback aimed at a single resource or at specific uses within one resource. The present project instead evaluates an integrated report that summarizes the household's environmental performance across water, gas, and electricity. At the same time, the CHR preserves resource-specific information alongside the aggregate Ecoscore. It therefore combines two levels of feedback: a simplified overall indicator intended to facilitate cross-resource evaluation and disaggregated resource-level information intended to preserve diagnostic and actionable detail.

Cross-resource spillovers and integrated environmental behavior. Second, the project also contributes to the literature on spillovers across pro-environmental behaviors. A behavioral intervention targeting one domain may produce positive spillovers if it activates environmental identity, consistency motives, or awareness of complementarities across behaviors. It may produce negative spillovers if it generates moral licensing, fatigue, or substitution across conservation efforts. However, evidence on pro-environmental spillovers is mixed and depends on the type of intervention and outcome (Truelove et al., 2014; Dolan and Galizzi, 2015; Nilsson et al., 2017; Galizzi and Whitmarsh, 2019; Maki et al., 2019). Several recent studies estimate spillovers using administrative resource-use data, with mixed results, (Jessoe et al., 2020; Carlsson et al., 2020; Goetz et al., 2022; Bonan et al., 2024; Goette and Lim, 2026; Boomsma et al., 2025). Field evidence also shows that cross-resource effects can be negative or asymmetric. Water-use feedback has been found to reduce water consumption while increasing electricity use, consistent with moral licensing (Tiefenbeck et al., 2013). More recent evidence finds that electricity-focused social comparisons can reduce use in several energy domains, while analogous water comparisons generate limited direct and spillover responses (Kažukauskas et al., 2025).
The present experiment is designed around spillovers rather than treating them only as secondary outcomes. The CHR explicitly highlights resource interdependence through an overall ecoscore, resource-specific bars, CO_2 translation, and tips that emphasize behavioral synergies. This allows us to test whether making the relationship among resources more salient generates broader conservation effects.

Multiple interventions, limited attention, and information burden. Third, the project contributes to the literature on interactions among multiple behavioral interventions. As behavioral policies become more common, the same individual may receive several nudges at once. The combined effect of these interventions is theoretically ambiguous. Multiple nudges may reinforce each other if they activate related motives or address complementary behavioral frictions. They may also crowd each other out if attention, cognitive effort, or willingness to process information is limited. Evidence on nudge interactions is still limited (Brandon et al., 2019; Bonan et al., 2020; Fang et al., 2023; Ling et al., 2026). Bonan et al. (2024) show that the water nudge was effective mainly among customers who were not already receiving other resource-specific reports. This suggests that multiple separate nudges may strain users' limited attention and reduce the effectiveness of additional reports. The CHR experiment tests a design response to this problem: rather than adding another independent single-resource report, the intervention integrates information from multiple resources into a unified environmental-performance metric. More broadly, the CHR shifts attention from the effectiveness of individual behavioral messages to the design of the communication architecture through which multiple messages are delivered. Existing interventions typically add resource-specific reports independently, leaving customers to integrate the information themselves. By contrast, the CHR explicitly reorganizes multi-resource information into a single environmental-performance metric while preserving the underlying resource-specific details. The experiment therefore contributes to the emerging literature on how the organization and presentation of information influence behavioral responses, beyond the content of the messages themselves.

The project also relates to research on salience, simplification, and the architecture of environmental information. Information can affect behavior not only by changing knowledge, but also by changing which attributes are salient and how easily consumers can evaluate performance (Tiefenbeck et al., 2016, 2019; Asensio and Delmas, 2015, 2016; Steinhorst et al., 2015; Johnson et al., 2012; Sejas-Portillo et al., 2025). Our experiment sheds light on possible mechanisms explaining the potential treatment effect, through a customer survey. The survey allows us to investigate whether the intervention works by increasing environmental salience, perceived usefulness, and understanding of cross-resource relationships, or whether it instead increases perceived repetition and report overload.

More generally, the CHR contributes to the literature on choice architecture by evaluating whether reorganizing existing information into a more integrated presentation can improve decision making without changing either prices or the underlying information available to consumers (Johnson et al., 2012; Loewenstein and Chater, 2017).

Customer engagement and retention. Finally, the project contributes to the literature on customer engagement and retention in liberalized utility markets. In gas and electricity markets, customers can switch suppliers, and switching behavior depends on both price and non-price factors (Hortaçsu et al., 2017; Shin and Managi, 2017; Fontana et al., 2019; Schleich et al., 2019). Additional information need not generate additional engagement. Information demand depends on salience and the anticipated valence of the information received (Golman et al., 2022), while repeated behavioral communications may increase the intended behavior but also induce avoidance, such as unsubscribing from future messages (Damgaard and Gravert, 2018). These findings motivate examining report opening, digital engagement, and churn alongside consumption. Behavioral reports may affect not only resource use but also customer engagement and loyalty. Bonan et al. (2024) find that receiving the water report reduced deactivation of gas and electricity contracts, suggesting that green nudges may improve the customer experience. The CHR experiment extends this question by testing whether an integrated environmental-performance report increases retention, or whether the addition of an extra report/page generates fatigue.

2.1 Conceptual framework and hypotheses

The CHR adds a consolidated synthesis of multi-resource environmental performance to the separate electricity, gas, and water reports that customers continue to receive. It therefore changes the communication architecture through which existing resource-specific information is presented, while also increasing the total volume of information delivered to treated customers.

The CHR may affect behavior through several channels. Its summary Ecoscore may attract attention and reduce the effort required to combine information across resource domains. The cross-resource presentation may improve customers' understanding of the relationships among water, gas, and electricity use, while the environmental framing may increase the salience of conservation motives. At the same time, adding another page may increase perceived repetition or information overload. The net effects on attention and information burden are therefore theoretically ambiguous.

These channels may affect two main classes of administrative outcomes. First, changes in attention, comprehension, and environmental salience may alter resource consumption. Second, the perceived value or burden of the communication may affect customers' propensity to deactivate contracts. The customer survey is used to measure report opening, satisfaction, and the intermediate perceptions targeted by the intervention.

The empirical analysis distinguishes between one primary confirmatory hypothesis, a limited set of secondary confirmatory hypotheses, and exploratory analyses.

Primary confirmatory hypothesis.

H1. Assignment to receive the CHR affects pooled household resource consumption relative to the standard resource-specific reports.

The pooled standardized resource-usage outcome provides the main test of the intervention. Because the intervention is intended to promote conservation, a negative estimated effect would be consistent with its intended direction. However, the confirmatory statistical test is two-sided.

Secondary confirmatory hypotheses. Conditional on the hierarchy and multiple-testing procedure described in Section 5, we test the following secondary hypotheses:

H2a. Assignment to receive the CHR affects water consumption.

H2b. Assignment to receive the CHR affects gas consumption.

H2c. Assignment to receive the CHR affects electricity consumption.

H3. Treatment effects on pooled resource consumption differ across customers' resource bundles.

H4. Treatment effects on pooled resource consumption differ according to customers' baseline Ecoscore.

H5. Assignment to receive the CHR affects contract deactivation in the liberalized gas and electricity markets.

The resource-bundle-specific marginal effects, monthly event-study coefficients, recent report opening, satisfaction measures, and individual survey mechanism outcomes are used to decompose, interpret, or explore the confirmatory findings. They are not treated as additional confirmatory hypotheses.

3 Treatment, Sample, and Randomization

3.1 Setting

The project is implemented in collaboration with a large Italian multi-utility providing water, electricity, and gas services to more than four million customers. The intervention targets residential customers who have received digital Home Reports for at least two resources for no less than 12 months.

3.2 Experimental design

Customers in the control group receive the standard, "as-is" service: separate Home Reports for each resource they hold among electricity, gas, and water. These single-resource reports provide information on the customer's own consumption, social-comparison feedback relative to similar customers, dynamic comparisons with previous consumption, and resource-saving tips. Figures 1, 2, and 3 show examples of the standard electricity, gas, and water reports. Customers receive all reports on the same date, at the frequency of the least frequently billed resource in the bundle, which ranges from monthly to bimonthly.

Customers in the treatment group receive the same standard single-resource reports, one for each resource, plus an additional integrated report, the CHR. The CHR summarizes the customer's overall environmental performance across the resources included in the bundle using the *Ecoscore*. It also reports resource-specific positioning within the integrated ranking, translates consumption changes into environmental impact, and provides personalized tips emphasizing cross-resource synergies. Figure 4 shows the CHR received by treated customers.

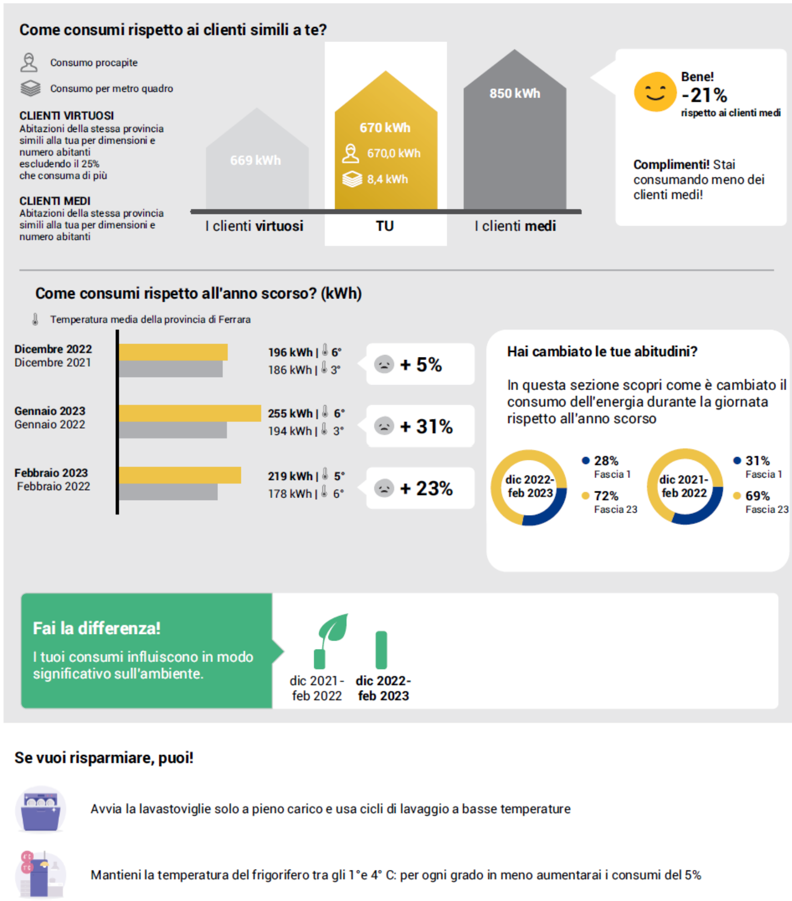

Figure 1: Standard electricity report

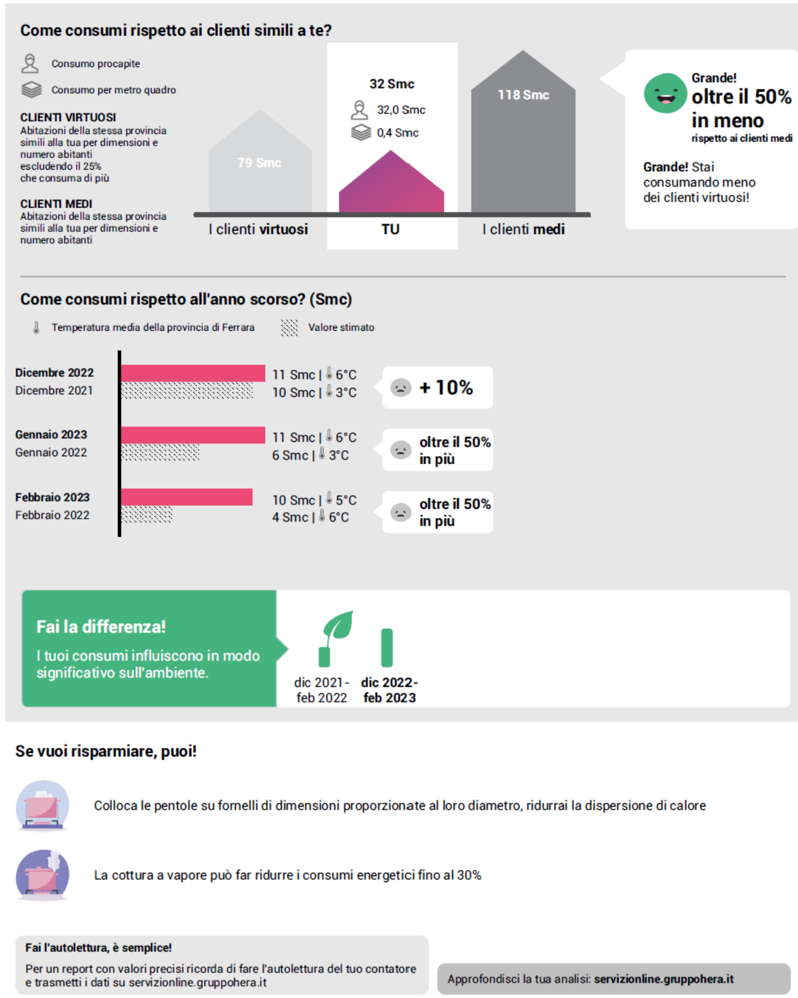

Figure 2: Standard gas report

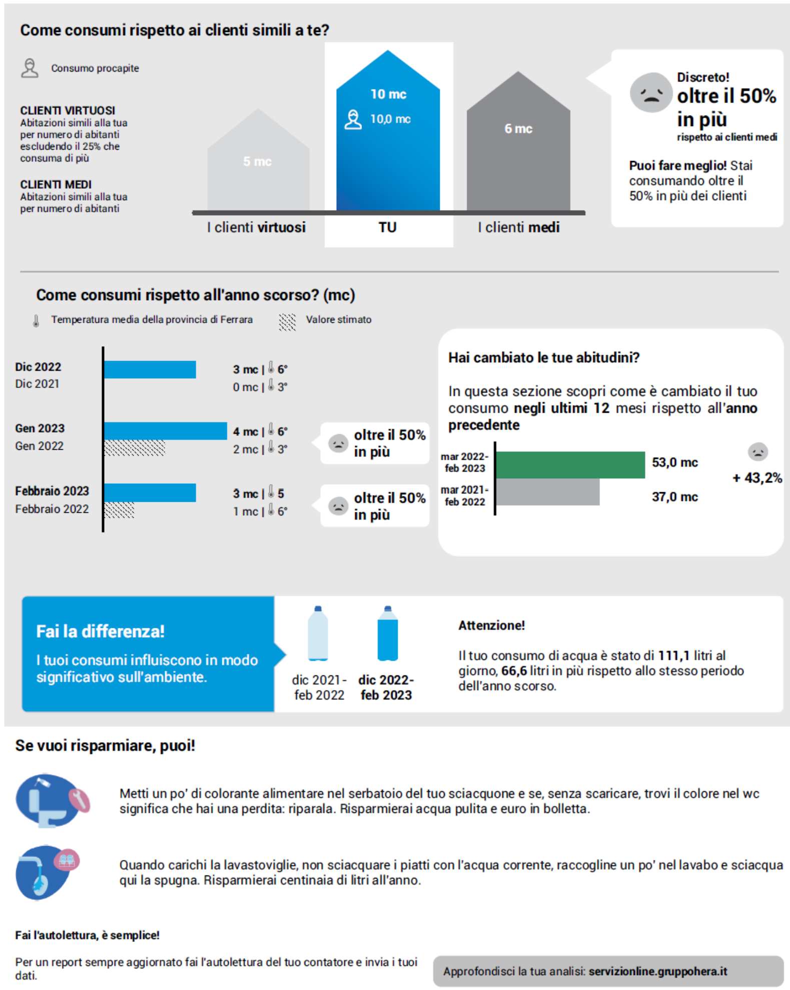

Figure 3: Standard water report

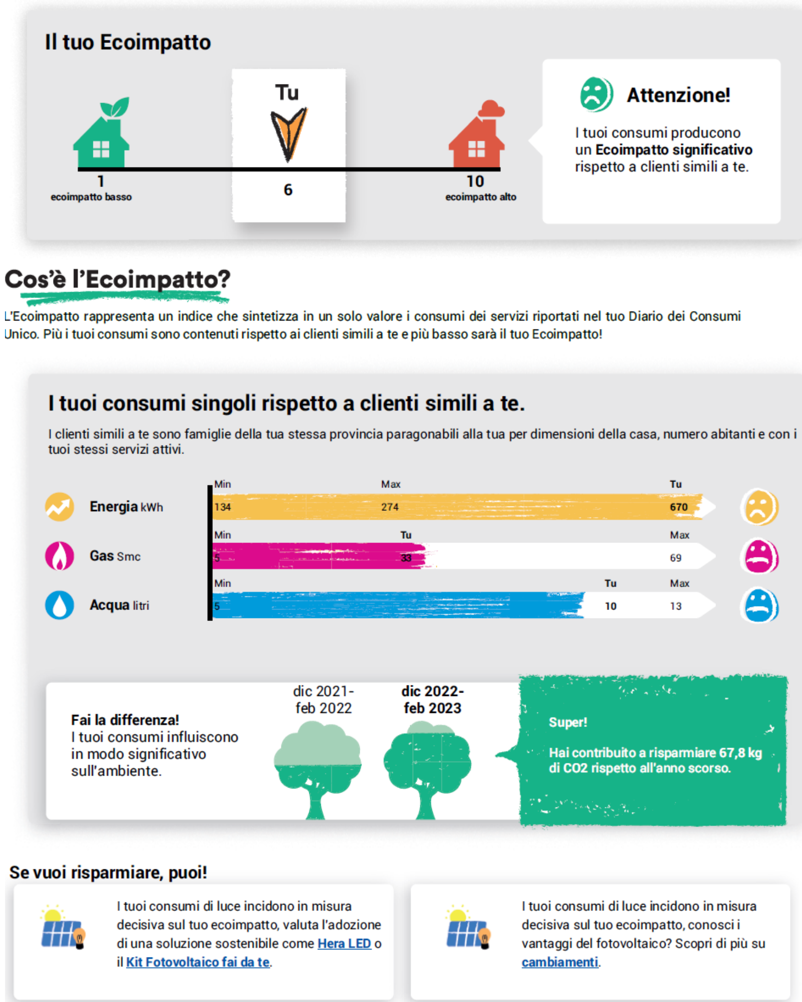

Figure 4: Consolidated Home report

The CHR is constructed from the customer's resource-specific consumption information for the resources included in the customer's existing Home Reports bundle. Similar customers are defined using province of supply, number of occupants, household square meters, and the same bundle of resources included in the CHR. The Ecoscore is calculated as the rounded average of the deciles of consumption for the individual resources included in the report. It ranges from 1 to 10, with lower values indicating more virtuous performance. For example, a customer in decile 1 for gas, decile 6 for electricity, and decile 3 for water would have an average score of 3.33 and an Ecoscore equal to 3.

The report includes four main elements:

(i). Overall Ecoscore: a single metric summarizing the customer's relative environmental performance across the resources included in the report.

(ii). Resource-specific details: bars for each resource showing the customer's consumption in the relevant unit of measurement and their position relative to similar customers.

(iii). Environmental-impact section: a "Make the difference" section translating changes in consumption into CO_2 implications.

(iv). Personalized tips: tips targeting the worst-performing resource, high-impact behaviors, and synergies across resources.

3.3 Study sample

The study sample consists of customers who:

- have a valid email address;
- receive the Home Report for every resource for which they hold a contract.
- already receive more than one Home Report;
- have been customers for at least 12 months;
- receive all relevant Home Reports on the same day and in the same PDF file;
- are included in strata (described below) with at least 20 customers.

The sample contains 190,638 customers. The distribution by Home Reports bundles is shown in Table 1.

Table 1: Sample distribution by Home Reports bundles

Home Reports bundle	Frequency	Percent	Cumulative percent
Gas and Electricity	123,565	64.82	64.82
Water and Electricity	12,711	6.67	71.48
Water and Gas	16,789	8.81	80.29
Water, Gas and Electricity	37,573	19.71	100.00
Total	190,638	100.00	

3.4 Data

The analysis uses administrative and survey data provided by the utility after anonymization.
Administrative data include:

- resources' usage at monthly or bimonthly level. Monthly observations for resources billed at lower frequencies are obtained by allocating cumulative consumption between two consecutive meter readings to the corresponding calendar months under a constant daily-consumption assumption. This procedure produces a monthly customer–resource panel that is comparable across resources while preserving total observed consumption over each billing interval;
- Home Report combination and report-delivery information, including date and report customized contents;
- treatment assignment and randomization strata;
- customer tenure and contract characteristics;
- contract activation and deactivation dates. We construct a dummy variable for deactivation in the post-treatment period

Survey data come from the short customer survey administered to treated and control customers participating in the experiment. The survey was sent to 10,000 randomly selected customers on October 10, 2025, hence after the treatment implementation, and a reminder was sent on October 21.

3.5 Randomization

Based on the information available at the time of randomization, strata were constructed from the interaction of the customer's Home Report bundles (four categories) and categorical indicators of baseline resource consumption.
For each resource, baseline consumption was classified as:

(i). below or equal to the median among customers with available consumption data;

(ii). above the median among customers with available consumption data;

(iii). missing.

In particular, baseline water consumption was not available at the time of randomization and was therefore coded as missing for all customers. For electricity and gas, consumption was coded as missing when the customer did not hold the corresponding contract or when baseline consumption data were otherwise unavailable.

Strata with fewer than 20 eligible customers were excluded before treatment assignment. Within each retained stratum, eligible customers were randomly assigned to treatment or control with equal probability using a computer-generated random number. Treatment assignment was performed once and remained fixed throughout the study.

4 Research Questions and Empirical Strategy

This study combines administrative and survey data to evaluate the effects of the CHR on resource consumption, customer engagement, contract retention, and customers' responses to the reporting service. The study hypotheses are classified as primary confirmatory, secondary confirmatory, or exploratory.

4.1 Research Question 1 (H1–H4): What is the effect of the CHR on resource usage?

The first research question asks whether receiving the CHR changes customers' consumption of water, gas, and electricity. Because customers are randomly assigned to treatment, identification of the causal effect relies on the experimental design. The panel difference-in- differences specification is adopted to exploit repeated observations, improve statistical precision, and account for permanent differences across customer–resource pairs through fixed effects.

The average effect on pooled standardized resource consumption is the sole primary confirmatory test. The resource-specific effects and the two pre-specified heterogeneity tests are secondary confirmatory analyses and are included in the common secondary multiple-testing family defined in Section 5. The event-study analysis and the reported subgroup-specific marginal effects are descriptive decompositions of these confirmatory estimates.

Let i index customers, r index resources, and t index months. Resources are $r \in \{\text{water}, \text{gas}, \text{electricity}\}$, depending on the customer's Home Report combination. Let $Treat_i$ equal one for customers assigned to receive the CHR and zero otherwise. Let $Post_t$ equal one from February 2025 through June 2026 and zero from January 2024 through January 2025.

The primary estimand is the intention-to-treat (ITT) effect of being assigned to receive the CHR. Accordingly, treatment status is defined by random assignment rather than by actual exposure to the report, such as opening the email or reading its contents, for which no administrative data are available.

The unit of observation is the resource–customer–month. Because the three resources are measured in different physical units, the primary pooled outcome is standardized separately for each resource using the mean and standard deviation of the control group during the pre-treatment period. Specifically,

$$y_{irt} = \frac{Consumption_{irt} - \overline{Consumption}_{r,\text{Control,Pre}}}{\sigma_{r,\text{Control,Pre}}},$$

where $\overline{Consumption}_{r,\text{Control,Pre}}$ and $\sigma_{r,\text{Control,Pre}}$ denote the mean and standard deviation of consumption for resource r among customers assigned to the control group during the pre-treatment period.

This standardization places all three resources on a common scale while relying on a reference distribution that is unaffected by the intervention. Consequently, the estimated treatment effects can be interpreted as changes in resource consumption measured in pre-treatment control-group standard deviation units.

4.1.1 Average treatment effect

We first estimate the average treatment effect using the following weighted difference-in-differences specification.

$$y_{irt} = \beta\,(Treat_i \times Post_t) + \mu_{ir} + \tau_{rt} + \varepsilon_{irt}, \tag{1}$$

where $Treat_i$ is an indicator for assignment to the CHR; $Post_t$ equals one from February 2025 onward, μ_{ir} are customer–resource fixed effects; τ_{rt} are resource-by-calendar-month fixed effects. The model is estimated by weighted least squares using observation weights

$$w_i = \frac{1}{N_i},$$

where $N_i \in \{2, 3\}$ denotes the number of resources included in customer i's Home Report bundle. Consequently, customers receiving reports for two resources receive a weight of $1/2$, while those receiving reports for three resources receive a weight of $1/3$. This weighting ensures that each customer contributes equally to the pooled treatment effect, rather than giving greater influence to customers with reports covering more resources.

Customer–resource fixed effects absorb all time-invariant differences in baseline consumption across customer–resource pairs. Resource-by-calendar-month fixed effects account for resource-specific seasonality and aggregate shocks, allowing the common time path of water, gas, and electricity consumption to differ flexibly.

Because treatment is assigned at the customer level, each customer contributes repeated observations over time and across multiple resources. Standard errors will therefore be clustered at the customer level, allowing for arbitrary correlation of the error term across both months and resources within the same customer.

This specification tests H1. The corresponding primary confirmatory null is

$$H_0^P : \beta = 0,$$

against the two-sided alternative $H_A^P : \beta \neq 0$. This hypothesis is tested at the 5% level without adjustment because pooled standardized resource use is the sole primary outcome.

Unless otherwise stated, all pooled resource-use specifications include customer–resource fixed effects and resource-by-calendar-month fixed effects, are estimated using inverse-resource weights, and cluster standard errors at the customer level.

4.1.2 Dynamic treatment effects

To characterize the dynamics of the intervention and verify the balance generated by the randomized assignment, we estimate an event-study specification in which treatment assignment is interacted with indicators for event time (months relative to the first delivery of the CHR).

Specifically,

$$y_{irt} = \sum_{k \neq -1} \beta_k \left(Treat_i \times \mathbf{1}\{EventTime_t = k\} \right) + \mu_{ir} + \tau_{rt} + \varepsilon_{irt}, \tag{2}$$

where $EventTime_t$ denotes the number of months relative to the first delivery of the CHR, with $k = 0$ corresponding to February 2025. The omitted category is $k = -1$, corresponding to January 2025, the month immediately preceding the intervention.

The coefficients β_k measure treatment effects k months relative to the introduction of the CHR.

Under successful randomization, treatment and control customers should exhibit similar consumption dynamics before the intervention. Consequently, the pre-treatment coefficients, for $k < 0$, provide a graphical assessment of baseline balance. These coefficients should be close to zero and should not display systematic trends before treatment.

The coefficients for $k \geq 0$ characterize the dynamics of the intervention, allowing us to assess whether the effects emerge immediately after customers receive the first CHR, strengthen with repeated exposure, or attenuate over time. The estimated coefficients will be presented graphically together with their 95% confidence intervals.

The individual event-study coefficients are not treated as separate confirmatory hypotheses. We will report a joint test of the null hypothesis that all pre-treatment interaction coefficients are equal to zero:

$$H_{0,\text{pre}} : \beta_k = 0 \quad \text{for all } k < 0.$$

To summarize the timing and persistence of the intervention, we will also report average treatment effects over three pre-specified post-treatment windows:

(i). the initial treatment period, from February to June 2025;

(ii). the subsequent treatment period, from July to December 2025;

(iii). the longer-run treatment period, from January to June 2026.

These estimates will be obtained by replacing the single post-treatment indicator with indicators for the three pre-specified periods. They are descriptive decompositions of the primary average treatment effect and are not included as additional confirmatory hypotheses.

4.1.3 Resource-specific effects

The baseline specification estimates the average treatment effect across all resources. To investigate whether the effect of the CHR differs across water, gas, and electricity consumption, we estimate the following heterogeneous-effects specification:

$$\begin{aligned} y_{irt} = {} & \beta\,(Treat_i \times Post_t) \\ & + \delta_G\,(Treat_i \times Post_t \times Gas_r) \\ & + \delta_E\,(Treat_i \times Post_t \times Electricity_r) + \mu_{ir} + \tau_{rt} + \varepsilon_{irt}, \end{aligned} \tag{3}$$

where Gas_r and $Electricity_r$ are indicators for the corresponding resource, with water serving as the omitted reference category.

The coefficient β measures the treatment effect on water consumption, while δ_G and δ_E capture the differential treatment effects for gas and electricity relative to water. Consequently, the treatment effects for gas and electricity are given by $\beta + \delta_G$ and $\beta + \delta_E$, respectively.

These specifications test H2a–H2c through the following secondary confirmatory two-sided null hypotheses:

$$H_{0,W} : \beta_W = 0, \qquad H_{0,G} : \beta_G = 0, \qquad H_{0,E} : \beta_E = 0$$

where

$$\beta_W = \beta, \qquad \beta_G = \beta + \delta_G, \qquad \beta_E = \beta + \delta_E.$$

Inference for these three hypotheses is included in the common secondary confirmatory family. The differential coefficients δ_G and δ_E will help interpret differences across resources but are not additional confirmatory hypotheses.

4.1.4 Heterogeneity by resource bundle

We first investigate whether the treatment effect differs according to the combination of resources included in the customer's Home Report bundle. We estimate

$$\begin{aligned} y_{irt} = {} & \beta\,(Treat_i \times Post_t) \\ & + \delta_{WE}\,(Treat_i \times Post_t \times WaterElectricity_i) \\ & + \delta_{WG}\,(Treat_i \times Post_t \times WaterGas_i) \\ & + \delta_{WGE}\,(Treat_i \times Post_t \times WaterGasElectricity_i) \\ & + \mu_{ir} + \tau_{rt} + \varepsilon_{irt}, \end{aligned} \tag{4}$$

where the omitted category is the *gas + electricity* bundle, which represents the largest group in the experimental sample. The indicators $WaterElectricity_i$, $WaterGas_i$, and $WaterGasElectricity_i$ identify customers receiving the corresponding combination of Home Reports.

The coefficient β measures the average treatment effect for customers in the *gas + electricity* bundle. The treatment effects for the other resource combinations are given by: $\beta + \delta_{WE}$, $\beta + \delta_{WG}$, and $\beta + \delta_{WGE}$, respectively.

H3 is evaluated through the omnibus null:

$$H_{0,B} : \delta_{WE} = \delta_{WG} = \delta_{WGE} = 0.$$

This tests whether treatment effects are equal across all four resource bundles. The bundle-specific treatment effects will be reported as marginal effects to interpret the omnibus result, but they are not treated as separate confirmatory hypotheses.

4.1.5 Heterogeneity by baseline Ecoscore

We also investigate whether the treatment effect varies with the customer's baseline environmental performance, as summarized by the Ecoscore associated with the first CHR received.

For treated customers, the baseline Ecoscore is the score displayed in the first CHR. For control customers, an equivalent baseline Ecoscore will be constructed using the same algorithm, the same reference period, and the same resource-specific consumption information that would have been used to generate the CHR.

In particular, control customers will be assigned resource-specific consumption deciles relative to similar customers, and these deciles will be averaged and rounded using the same procedure applied to treated customers.

Because the baseline Ecoscore is constructed using consumption information referring to the period preceding the first delivery of the CHR, it is not affected by treatment assignment.

The primary heterogeneity variable is an indicator for having a baseline Ecoscore above the median within the customer's resource bundle:

$$HighEcoscore_i = \mathbf{1}\left\{Ecoscore_{i0} > Median\left(Ecoscore_0 \mid Bundle_i\right)\right\},$$

where $Bundle_i$ identifies the combination of resources included in customer i's Home Reports.

A higher Ecoscore indicates worse relative environmental performance. Defining the threshold within resource bundle ensures that customers are compared with others receiving reports for the same set of resources and avoids confounding baseline performance with systematic differences in the distribution of the Ecoscore across two-resource and three-resource bundles.

We estimate

$$\begin{aligned} y_{irt} = {} & \beta\left(Treat_i \times Post_t\right) \\ & + \gamma\left(Post_t \times HighEcoscore_i\right) \\ & + \delta\left(Treat_i \times Post_t \times HighEcoscore_i\right) \\ & + \mu_{ir} + \tau_t + \varepsilon_{irt}. \end{aligned} \tag{5}$$

The coefficient β measures the treatment effect among customers with a baseline Ecoscore at or below the median. The coefficient δ captures the differential treatment effect among customers with an above-median baseline Ecoscore.

The secondary confirmatory hypothesis H4 is tested using the following two-sided hypothesis:

$$H_{0,S} : \delta = 0,$$

As a secondary specification and robustness check, we will replace the binary indicator with the continuous baseline Ecoscore.

4.2 Research Question 2 (H5): What is the effect of the CHR on customer retention?

The second research question contains a confirmatory hypothesis concerning contract deactivation.

Customer churn is analyzed separately at the customer–resource level, restricting the sample to gas and electricity contracts, since water services are not subject to customer switching (the market is not liberalized). Each customer therefore contributes either one or two observations, depending on whether they hold a contract for one or both market-based resources.

We estimate the intention-to-treat effect using the following ordinary least squares linear probability model:

$$Churn_{ir} = \alpha + \beta_C Treat_i + \phi_r + \lambda_s + \varepsilon_{ir}. \tag{6}$$

where $r \in \{\text{gas}, \text{electricity}\}$ and $Churn_{ir}$ equals one if customer i deactivates the contract for resource r during the post-treatment period and zero otherwise. The model includes resource fixed effects, ϕ_r, to account for differences in baseline churn rates between gas and electricity contracts, and strata fixed effects, λ_s. The model is estimated using inverse-resource weights: customers holding only one gas or electricity contract therefore receive a weight of one, while customers holding both contracts receive a weight of $1/2$. Standard errors are clustered at the customer level to allow for arbitrary correlation in churn outcomes across contracts belonging to the same customer.

H5 is tested using the following two-sided hypothesis:

$$H_{0,C} : \beta_C = 0,$$

4.3 Research Question 3 [Exploratory]: Does the CHR increase customers' engagement with the reports?

The third research question investigates whether the CHR increases customers' attention to the information provided by the utility. While administrative data allow us to study engagement with all customers' digital channels, the customer survey provides a direct measure of engagement with the reports themselves on a sub-sample of them.

The first survey outcome is based on a question asking respondents when they last opened a report. Responses range from "Never opened" to "Less than two months ago." Since the survey is administered during the experimental period about eight months after its launch, a more recent opening is interpreted as a proxy for a higher report opening rate and greater attention to the communication.

The main specification is

$$RecentOpen_i = \theta Treat_i + \lambda_s + X'_{i0}\Gamma + u_i,$$

where $Treat_i$ indicates assignment to the CHR, λ_s are randomization-stratum fixed effects, and X_{i0} denotes pre-treatment controls listed in Section 6. The outcome is a binary indicator equal to one if the respondent reports having opened a Home Report within the previous six months and zero otherwise, including respondents who report never having opened one.

The hypothesis is tested using the following two-sided hypothesis:

$$H_{0,O} : \theta_O = 0.$$

The model is estimated as a linear probability model with heteroskedasticity-robust standard errors.

Because recent report opening is observed only for survey respondents, the treatment-control comparison may be affected by differential survey participation. We therefore classify this analysis as exploratory and report the effect of treatment assignment on survey response before presenting the opening results.

4.4 Research Question 4 [Exploratory]: Does the CHR affect customer satisfaction?

The fourth research question evaluates whether the CHR changes customers' overall satisfaction with the utility's reporting service.

This analysis is exploratory and is not included in the confirmatory multiple-testing family. This classification reflects both the secondary role of satisfaction in the study and the fact that the satisfaction question is administered only to survey respondents who report having opened at least one Home Report.

The primary outcome is a question which asks respondents to rate their overall satisfaction with the reports on a scale from 1 to 10. Customers in the treatment group evaluate the reporting service, including the CHR, whereas customers in the control group evaluate the traditional set of single-resource reports.

We estimate via OLS:

$$Satisfaction_i = \theta_S Treat_i + \lambda_s + X'_{i0}\Gamma + u_i.$$

where notation follows the previous specification.

4.5 Research Question 5 [Exploratory]: Which behavioral mechanisms are associated with the CHR?

The fifth research question explores whether assignment to the CHR is associated with the behavioral channels targeted by the intervention. The survey measures cross-resource understanding, environmental salience, perceived usefulness, cognitive burden, repetition, and information overload.

These questions are administered only to respondents who report having opened at least one Home Report. Estimates from this analysis will therefore be interpreted as conditional treatment-control differences among report openers and not as population intention-to-treat effects or formal causal mediation estimates. These analyses serve as exploratory manipulation checks.

For each mechanism outcome m, we estimate via OLS:

$$Mechanism_{im} = \theta_m Treat_i + \lambda_s + X'_{i0}\Gamma + u_{im}.$$

The coefficients θ_m will be reported with heteroskedasticity-robust standard errors.

The first manipulation-check outcome is agreement with the statement

> "The information in the report helps me understand the relationship between the use of different resources."

This question directly measures whether the CHR succeeded in increasing awareness of cross-resource relationships, which represents one of the central objectives of the intervention.

Secondary mechanism outcomes include:

- environmental salience: "The information in the Home Report reminds me of the importance of acting in respect of the environment;"
- perceived usefulness: "The information in the Home Report helps me save resources;"
- perceived difficulty: "The Home Reports are difficult to read and understand";
- perceived repetition: "The information in the Home Reports is repetitive";
- perceived information overload: "I receive too many Home Reports."

Each question is coded with a 1-10 index.

These outcomes are observed only among respondents who report having opened at least one Home Report. The estimates are therefore interpreted as conditional treatment-control differences among report openers rather than population ITT effects or formal causal mediation estimates.

5 Multiple-Hypothesis Testing

The empirical analysis distinguishes between one primary confirmatory hypothesis, a set of secondary confirmatory hypotheses, and exploratory analyses.

The primary confirmatory hypothesis concerns the average treatment effect of the CHR on pooled standardized resource consumption (H1). Because this is the sole primary outcome of the study, it will be tested at the two-sided 5% significance level without multiple-testing adjustment.

All remaining confirmatory hypotheses constitute a single secondary confirmatory family. Specifically, this family consists of:

(i). H2a: treatment effect on water consumption;

(ii). H2b: treatment effect on gas consumption;

(iii). H2c: treatment effect on electricity consumption;

(iv). H3: omnibus test of treatment-effect heterogeneity across resource bundles;

(v). H4: treatment-effect heterogeneity by baseline Ecoscore;

(vi). H5: treatment effect on contract deactivation (churn).

Inference for these six hypotheses will control the familywise error rate at the 5% level using the Romano–Wolf stepdown procedure, following the experimental-economics literature on multiple outcomes and heterogeneous treatment effects (Anderson, 2008; List et al., 2019). The resampling procedure will preserve the original customer-level randomization and stratification. For each confirmatory hypothesis, we will report both the conventional and the Romano–Wolf adjusted p-value.

The following analyses are not treated as additional confirmatory hypotheses:

- individual monthly event-study coefficients;
- bundle-specific treatment effects reported to interpret the omnibus resource-bundle test;
- the continuous baseline-Ecoscore interaction;
- recent report opening
- customer satisfaction;
- individual survey mechanism outcomes and exploratory mechanism indices;
- any additional subgroup analyses not explicitly pre-specified.

These analyses are reported to characterize the intervention, interpret the confirmatory findings, or generate hypotheses for future work. Where appropriate, exploratory families of outcomes (e.g., the survey mechanism questions) will additionally be accompanied by Benjamini–Hochberg false-discovery-rate adjusted q-values.

6 Sample Balance at Baseline

For each baseline variable available before treatment assignment, we estimate:

$$X_{i0} = \pi Treat_i + \lambda_s + e_i,$$

where X_{i0} is a baseline characteristic and λ_s are randomization-stratum fixed effects. We report the control mean, treatment-control difference, standard error, and p-value.
Balance variables include:

- Home Report bundle;
- baseline standardized consumption for the resources in the customer's bundle;
- reconstructed baseline Ecoscore;
- customer tenure;

A joint orthogonality test will also be reported by regressing treatment assignment on the full set of baseline variables and strata fixed effects.

7 Power Calculations

For the administrative consumption analysis, a simple cross-sectional calculation with 190,638 customers equally allocated to treatment and control implies an MDE of approximately 0.013 pre-treatment control-group standard deviations for the pooled standardized outcome. The actual panel difference-in-differences model with repeated monthly observations and fixed effects should generally have higher precision, depending on serial correlation and within-customer persistence.
For survey outcomes, power depends on the response rate among the 10,000 customers invited to participate in the survey. Assuming equal allocation between treatment and control, a two-sided significance level of 5%, power of 80%, and a standard deviation of 2.5 for outcomes measured on a 1–10 scale, the approximate minimum detectable effects are reported in Table 2.

8 Robustness and Sensitivity Analyses

The following robustness checks are pre-specified:

(i). use inverse-hyperbolic-sine consumption outcomes;

Table 2: Approximate minimum detectable effects for survey outcomes

Response rate	Respondents	MDE: 1–10 outcome	MDE: binary outcome
5%	500	0.63	0.125
10%	1,000	0.44	0.089
20%	2,000	0.31	0.063

Notes: Calculations assume equal treatment allocation, a two-sided significance level of 5%, 80% power, a standard deviation of 2.5 for outcomes measured on a 1–10 scale, and a control-group probability of 0.5 for binary outcomes. Binary-outcome MDEs are expressed as probability changes.

(ii). winsorize 0.5% of extreme consumption observations

(iii). assess sensitivity to differential post-treatment outcome availability using Lee Bounds;

(iv). estimate heterogeneity by baseline Ecoscore using the continuous score rather than the above-median indicator;

References

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

Allcott, H. and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10):3003–3037.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Asensio, O. I. and Delmas, M. A. (2015). Nonprice incentives and energy conservation. *Proceedings of the National Academy of Sciences*, 112(6):E510–E515.

Asensio, O. I. and Delmas, M. A. (2016). The dynamics of behavior change: Evidence from energy conservation. *Journal of Economic Behavior & Organization*, 126:196–212.

Ayres, I., Raseman, S., and Shih, A. (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization*, 29(5):992–1022.

Bernedo, M., Ferraro, P. J., and Price, M. (2014). The persistent impacts of norm-based messaging and their implications for water conservation. *Journal of Consumer Policy*, 37(3):437–452.

Bonan, J., Cattaneo, C., d'Adda, G., Galliera, A., and Tavoni, M. (2024). Widening the scope: The direct and spillover effects of nudging water efficiency in the presence of other behavioral interventions. *Journal of Environmental Economics and Management*, 127:103037.

Bonan, J., Cattaneo, C., d'Adda, G., and Tavoni, M. (2020). The interaction of descriptive and injunctive social norms in promoting energy conservation. *Nature Energy*, 5(11):900–909.

Bonan, J., Cattaneo, C., d'Adda, G., and Tavoni, M. (2021). Can social information programs be more effective? the role of environmental identity for energy conservation. *Journal of Environmental Economics and Management*, 108:102467.

Boomsma, M., Vringer, K., and Soest, D. v. (2025). The impact of real-time energy consumption feedback on residential gas and electricity usage. *Journal of Environmental Economics and Management*, 132:103163.

Brandon, A., List, J. A., Metcalfe, R. D., Price, M. K., and Rundhammer, F. (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences*, 116(12):5293–5298.

Brent, D. A., Cook, J. H., and Olsen, S. (2015). Social comparisons, household water use, and participation in utility conservation programs: Evidence from three randomized trials. *Journal of the Association of Environmental and Resource Economists*, 2(4):597–627.

Carlsson, F., Jaime, M., and Villegas, C. (2020). Behavioral spillover effects from a social information campaign. *Journal of Environmental Economics and Management*, 109:102325.

Damgaard, M. T. and Gravert, C. (2018). The hidden costs of nudging: Experimental evidence from reminders in fundraising. *Journal of Public Economics*, 157:15–26.

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

Dolan, P. and Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, 47:1–16.

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

Ferraro, P. J. and Miranda, J. J. (2013). Heterogeneous treatment effects and mechanisms in information-based environmental policies: Evidence from a large-scale field experiment. *Resource and Energy Economics*, 35(3):356–379.

Ferraro, P. J., Miranda, J. J., and Price, M. K. (2011). The persistence of treatment effects with norm-based policy instruments: Evidence from a randomized environmental policy experiment. *American Economic Review*, 101(3):318–322.

Ferraro, P. J. and Price, M. K. (2013). Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Review of Economics and Statistics*, 95(1):64–73.

Fontana, M., Iori, M., and Nava, C. R. (2019). Switching behavior in the italian electricity retail market: Logistic and mixed effect bayesian estimations of consumer choice. *Energy Policy*, 129:339–351.

Galizzi, M. M. and Whitmarsh, L. (2019). How to measure behavioral spillovers: A methodological review and checklist. *Frontiers in Psychology*, 10.

Goette, L. and Lim, Z. H. (2026). Testing for spillovers in resource conservation: Evidence from a natural field experiment. *Journal of Environmental Economics and Management*, 140:103401.

Goetz, A., Mayr, H., and Schubert, R. (2022). Beware of side effects? spillover evidence from a hot water intervention. Working paper.

Golman, R., Loewenstein, G., Molnár, A., and Saccardo, S. (2022). The demand for, and avoidance of, information. *Management Science*, 68(9):6454–6476.

Hortaçsu, A., Madanizadeh, S. A., and Puller, S. L. (2017). Power to choose? an analysis of consumer inertia in the residential electricity market. *American Economic Journal: Economic Policy*, 9(4):192–226.

Jaime Torres, M. M. and Carlsson, F. (2018). Direct and spillover effects of a social information campaign on residential water-savings. *Journal of Environmental Economics and Management*, 92:222–243.

Jessoe, K., Lade, G. E., Loge, F., and Spang, E. (2020). Spillovers from behavioral interventions: Experimental evidence from water and energy use. *Journal of the Association of Environmental and Resource Economists*, 8(2):315–346.

Jessoe, K., Lade, G. E., Loge, F., and Spang, E. (2021). Residential water conservation during drought: Experimental evidence from three behavioral interventions. *Journal of Environmental Economics and Management*, 110:102519.

Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C., Goldstein, D. G., Häubl, G., Larrick, R. P., Payne, J. W., Peters, E., Schkade, D., Wansink, B., and Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, 23(2):487–504.

Kažukauskas, A., Asmare, F., Broberg, T., and Jaraitė, J. (2025). The spillover effects of real-time social comparison information on water and energy use: Experimental evidence using in-home displays. *Environmental and Resource Economics*, 88(7):1879–1904.

Ling, M., Liu, Y., Xu, L., and He, J. (2026). Interactions of behavioral interventions across pro-environmental domains: A natural field experiment. *Journal of Environmental Economics and Management*, 138:103339.

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4):773–793.

Loewenstein, G. and Chater, N. (2017). Putting nudges in perspective. *Behavioural Public Policy*, 1(1):26–53.

Maki, A., Carrico, A. R., Raimi, K. T., Truelove, H. B., Araujo, B., and Yeung, K. L. (2019). Meta-analysis of pro-environmental behaviour spillover. *Nature Sustainability*, 2(4):307–315.

Miranda, J. J., Datta, S., and Zoratto, L. (2020). Saving water with a nudge or two: Evidence from costa rica on the effectiveness and limits of low-cost behavioral interventions on water use. *World Bank Economic Review*.

Nemati, M. and Penn, J. (2020). The impact of information-based interventions on conservation behavior: A meta-analysis. *Resource and Energy Economics*, 62:101201.

Nilsson, A., Bergquist, M., and Schultz, P. W. (2017). Spillover effects in environmental behaviors, across time and context: A review and research agenda. *Environmental Education Research*, 23(4):573–589.

Schleich, J., Faure, C., and Gassmann, X. (2019). Household internal and external electricity contract switching in eu countries. *Applied Economics*, 51(1):103–116.

Sejas-Portillo, R., Moro, M., and Stowasser, T. (2025). The simpler the better? threshold effects of energy labels on property prices and energy efficiency investments. *American Economic Journal: Applied Economics*, 17(1):41–89.

Shin, K. and Managi, S. (2017). Liberalization of a retail electricity market: Consumer satisfaction and household switching behavior in japan. *Energy Policy*, 110:675–685.

Steinhorst, J., Klöckner, C. A., and Matthies, E. (2015). Saving electricity — for the money or the environment? risks of limiting pro-environmental spillover when using monetary framing. *Journal of Environmental Psychology*, 43:125–135.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., and Staake, T. (2016). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management Science*, 64(3):1458–1476.

Tiefenbeck, V., Staake, T., Roth, K., and Sachs, O. (2013). For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy*, 57:160–171.

Tiefenbeck, V., Wörner, A., Schöb, S., Fleisch, E., and Staake, T. (2019). Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives. *Nature Energy*, 4(1):35–41.

Truelove, H. B., Carrico, A. R., Weber, E. U., Raimi, K. T., and Vandenbergh, M. P. (2014). Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental Change*, 29:127–138.