Urgency and engagement: a field experiment of the impact of message content on consumption awareness

Giovanna d'Adda[1], Arianna Galliera[2] and Massimo Tavoni[3]

May 10th, 2019

Abstract

We study how to foster engagement in the energy sector, a market where signals about consumption are opaque and infrequent. We investigate the impact of an intervention aimed at fostering natural gas self-reading, which allows utilities to bill customers on the basis of their real and not of their estimated consumption. Since utilities all over the world are relying heavily on users to submit their meter readings to achieve billing accuracy, understanding how to engage consumers is an important policy issue. In our study, messages that induce a sense of urgency are two times more effective than the generic messages in encouraging self-readings, consistent with previous research on the urgency effect. Our findings suggest that the increased sense of urgency moves to action customer with both high and low levels of prior engagement, but that the effect on the former is stronger.

1. Introduction

Utilities all over the world rely heavily on users to submit their energy readings to achieve billing accuracy.[4] This is especially in the residential natural gas market, for which smart meters - that transfer consumption data automatically and make meter reading by the customer or by the utility unnecessary - are lacking. In Europe, for instance, just five countries, among which Italy, have decided to roll out natural gas smart meters. Even in these countries, penetration of gas smart meters is low: in Italy, for instance, less than 50 per cent of residential customers have smart meters.[5] Identifying strategies to effectively encourage customers to submit meter readings and thus improve the accuracy of energy bills is thus a widely relevant policy and marketing issue.

Increased billing accuracy is likely to improve both consumer welfare and administrative efficiency through a number of channels. First, billing accuracy reduces the likelihood of bill

---

[1] University of Milan, Department of Economics, Management and Quantitative Methods
[2] Politecnico di Milano, School of Management
[3] Politecnico di Milano, School of Management
[4] Indeed, an online search of "submit meter readings" shows a long list of guidelines and persuasive messages issued by energy companies from all over the world (New Zealand, UK, Australia, Canada, Hong Kong, Singapore, to name a few). We thank a referee for pointing this out.
[5] Source: ARERA, https://www.arera.it/it/operatori/smartmetering.htm.

shocks. Bill shocks are one of the main causes of delayed and incomplete bill payments, and of complaints customers make to utilities' customer care services.[6] Second, bills based on customers' actual consumption, rather than on their estimated consumption, send users a more precise signal of their energy usage and increase the fairness of the billing system. The opacity of the billing system is a serious issue: for instance, in our sample of Italian gas users, 45 per cent of customers had received bills based on estimated consumption for more than a year. This figure is even more striking, if one thinks that gas and electricity are responsible on average for 7 per cent of household expenditures among Italian consumers. Third, existing evidence suggests that making information on consumption clearer, more salient or more frequent, results in expenditure reductions (Gilbert and Zivin 2014; Allcott 2011; Ferraro and Price 2013; Costa and Kahn 2013; van Houwelingen and van Raaij 1989). More specifically, feedback on consumption to residential customers, who read their own meters, has been identified a promising method for motivating them to achieve savings, when smart meters are not available (Darby 1999). Understanding how to motivate individuals to attend to their gas consumption is therefore both challenging and policy relevant.

We study an intervention aimed at making gas customers more attentive to their consumption. Specifically, we evaluate a communication campaign encouraging customers to submit self-readings, so as to be billed on the basis of their real and not of their estimated consumption. One version of the campaign message imposes a sense of urgency on customers, by stressing the immediate impact of submitting a meter read on the subsequent bill. In particular, this version of the campaign message informs customers that they are entering a special time window to submit the self-reading in order to be billed exclusively on the submitted read, i.e., on real consumption. The other version of the message more generally links self-readings to having one's bill based on real consumption.

We examine the relative impact of the two versions of the campaign message on self-readings, and compare it to self-readings rates among customers not receiving the campaign. We use propensity score matching to select a sample of treated and untreated customers with similar characteristics. Knowing that submitting a reading would result in a bill based on real consumption, increases self-reading after the campaign by 12 percentage points relative to the control group of customers excluded from the campaign; being given a deadline, by which to submit a reading in order to have the following bill based exclusively on real consumption, leads to an additional increase in submitted reads of 16 percentage points. We explore the heterogeneous effect of the campaign, depending on customers' prior level of engagement with the utility, and show that, while effective also on previously unengaged customers, both versions of the campaign messages raise self-reading rates more among active customers, i.e. customers who were active on the utility's portal or submitted self-readings before the campaign.

Our results are consistent with evidence showing how agents, when made aware of time restrictions on a task, tend to prioritize it even if secondary with respect to others (Zhu, Yang,

[6] We control for the correlation between those customers late on bill and a bill good shock after submitting a self-reading; we find a negative and significant correlation ($\beta = -0.044$, $se = 0.009$).

and Hsee 2018; Zhu, Bagchi, and Hock 2019). Since in our setting the presence of a time window does not prevent customers from submitting a self-reading and benefitting from the improved informational content of the resulting bill to (Zhu, Yang, and Hsee 2018), [7] likely explanations for the urgency effect that we observe point to the urge of completing a certain goal (Kivetz, Urminsky, and Zheng 2006), or to individual's tendency to evaluate immediate benefits as larger than future ones (Frederick, Loewenstein, and O'Donoghue 2002; McClure et al. 2004). Urgency draws attention (Botti et al. 2008; Cialdini 2007; Pribram and McGuinness 1975; Zhu, Yang, and Hsee 2018; Zhu, Bagchi, and Hock 2019), creating enough tension to trigger action, even among those customers who are generally less reactive to informational campaigns. We confirm the urgency effect by exploring different types of heterogeneity of the gas users; even among the customers displaying the lowest level of attention towards the service, urgency significantly and marginally increases the response rate.

Our results confirm that urgency affects significantly individuals who are more tuned into the time dimension (Zhu, Yang, and Hsee 2018), such as customers regularly paying bills or submitting self-readings. To the best of our knowledge, the current work is novel in its focus on the heterogeneity of the urgency effect, namely in investigating whether urgency has the same effect on not-engaged customers. We find a significant reaction even from customers who displayed low attention to their own consumption prior to the campaign, even if not as strong as that of engaged customers.

Our results also contribute to the growing literature that puts behavioural insights to work and demonstrates their effectiveness in policy relevant domains in framed field experiments (Andor and Fels 2018; Brandon et al. 2019; Harrison and List 2004; List and Price 2016). In the energy sector, existing studies have overwhelmingly focused on fostering efficiency in electricity consumption and used social information as behavioural lever (Allcott 2011; Allcott and Rogers 2014; Allcott and Mullainathan 2010; Mark Andreas Andor et al. 2017; Sudarshan 2017). However, existing evidence shows the potential of enriching the behavioural policy toolbox (Abrahamse and Steg 2009; Fischer 2008) and how even interventions not directly aimed at reducing consumption can improve efficiency, thanks to their impacts on increased awareness and attention to resource usage (Ayres, Raseman, and Shih 2013; Darby 2010; Hargreaves, Nye, and Burgess 2013; Jessoe and Rapson 2014; Wichman 2017).

The paper proceeds as follows: Section 2 briefly describes the setting of our study, Section 3 presents the design of the study, Section 4 the empirical results, and Section 5 concludes.

2. Gas self-reading

[7] Urgency effect is thus distinguished from scarcity effect (Lynn 1989; Shah, Shafir, and Mullainathan 2015; Verhallen and Robben 1994), which is strictly related to the limited availability of a commodity.

Gas is used by about 95 per cent of Italian households for heating, as well as being the main energy source for hot water and cooking, and makes up about 65 per cent of energy consumption in the country (Istat, 2014).[8]

Gas consumption is tracked through meters (see Appendix Figure A1). Traditional gas meters do not send consumption data automatically to utilities, but need to be read manually.[9] Energy distributors are obliged by law to send meter readers to collect consumption data at least once a year per household. However, distributors' reads can occur even less frequently, if users cannot be found at home. Alternatively, customers can submit self-readings.

In the absence of readings or self-readings, gas consumption is estimated by energy utilities through an algorithm, as a function of each household's type of use (heating, hot water, and/or cooking), geographical area, past expenditure and number of occupants. The estimate gets refined as data on the household's real consumption becomes available. Households get billed for their estimated consumption, if real consumption data is not available. Once a reading or self-reading is submitted, the difference between real and estimated consumption is computed. If positive, households have to pay the difference. If negative, they get discounts on subsequent bills up to the value of the difference.

The issue of engaging gas customers to submit readings is a timely and policy relevant one following the liberalization of European energy markets. Recently introduced laws prevent customers from being billed for estimated consumption going back more than 24 months. This has placed utilities under great pressure to encourage self-readings. Furthermore, energy utilities have private incentives to promote self-readings, as these typically reduce the likelihood of large shocks to bill amounts and consequently improve payment rates and timing. The intervention that we study in this paper is an example of a utility's efforts in this realm.

More generally, policy makers are increasingly concerned with educating consumers and raising their awareness in the use of energy. The first step towards this goal consists in drawing consumers' attention towards their energy and gas use, which are typically poorly known and understood, despite representing one of the major expenditure items of European households. The intervention we examine in the next section offers interesting insights on the behavioural mechanisms that can be leveraged to this purpose.

3. Data and design

The study is based on an intervention conducted in collaboration with a large energy company with operations in Italy and worldwide. As part of such collaboration, the research team provided inputs on the content of campaign messages and on the need to randomize the

---

[8] https://www.istat.it/it/files//2014/12/StatReport_Consumi_energetici.pdf

[9] As mentioned above, traditional gas meters are currently being replaced with new generation, smart meters, which will send automatically gas consumption data to the provider monthly. However, at the time of the interventions described here, the coverage of smart meters was negligible.

different message variants across customers, but the utility ultimately retained full control both on the design and on the implementation of the campaigns.

The campaign we study was anticipated by a test of different messages on self-reading conducted by the same utility on almost 3000 households. The versions of the messages piloted in this test were all broadly inspired by behavioural principles.[10] The test suggested that the desire to keep, or gain, control over one's gas expenditures was a strong motivator of customers' engagement. Consistent with this result, both versions of the campaign message make explicit the link between self-reading and receiving a bill with information on one's real consumption.

The utility launched the campaign in May 2017, in order to get consumption data from customers at the end of the heating season. More than 1 million customers received the campaign message by email over the course of one week: they received on May 5th a first message, and on May 8th a reminder. We have data on the 1,083,369 customers targeted by the campaign, as well as on 1,637,002 customers that did not receive any message, who constitute our control group. Data include customers' characteristics, such as their location, use of gas, activity on the utility's portal, type of bill received (paper or mail) and payment type (e.g direct debit); as well as information on their billing and self-reading history from January 2016. Table A1 provides a description of the variables in our dataset, while Table A2 shows summary statistics for the samples of customers targeted and not targeted by the campaign.

The utility introduced two versions of the campaign message. Customers in the Deferred Feedback (DF) treatment received a message similar to the most effective message proposed in the pilot test, emphasizing how submitting a self-reading resulted in a bill in line with one's real gas consumption. The Immediate Feedback (IF) treatment augmented this message with information on a special time window to submit the self-reading. Namely, the IF message told customers by what date they should submit a self-reading for it to fall within this special time window. Figure A2 shows the two campaign messages.

The eligibility criteria are the same for both groups: active residential gas customers without any severe billing problem due to repeated missing payments, with no meter-reading the month before the campaign. DF treatment included customers not receiving a bill on May

[10] One version stressed the link between self-reading and keeping one's bill under control, and was inspired by evidence in economics and behavioural sciences shows that individuals tend to be risk averse and dislike risky event (Arrow 1971; Pratt 1964). A second version emphasized the possibility of paying more than one's due that came with an estimated bill, and was based on the idea that individuals weigh losses more heavily than gains (Thaler and Benartzi 2004; Tversky and Kahneman 1992; Kahneman, Knetsch, and Thaler 1986). The third version took inspiration from a recent literature showing how an extension of social norms (Allcott and Mullainathan 2010; Sudarshan 2017 for different applications), which renders them dynamic, can promote behaviours that are not yet consolidated (Loschelder et al. 2019; Sparkman and Walton 2017). The resulting messages appeal to the desire to avoid surprises in the bill, to prevent negative bill shocks and to conform to the behaviour of a growing number of customers in order to foster self-reading. Since we have no data on customers' participating in this test, thus we cannot check or improve the quality of randomization to treatment, and the test messages differed along various dimensions, beside the behavioural levers, we do not report results for it in the main text, but a description of the test and basic results are available in Appendix B.

2017 and only with a registered valid email, while customers involved IF treatment received a bill on May 2017 and included both customers with and without a valid email registered.

The IF treatment exploits the fact that, if a self-reading is submitted between 7 and 4 days from the last day of a billing cycle, the resulting bill will be based exclusively on the submitted read, i.e. on real consumption. For ease of exposition, we will hereafter refer to the 7 to 4 days' window before the end of the bill cycle as the IF window. Since most customers receive their bill bi-monthly or quarterly, submitting a read in the IF window will generate a clear signal of one's real consumption in the very short term, i.e. in the bill that will be received one week after the self-reading. On the contrary, submitting a read outside of the IF window will result in feedback on one's real consumption that is both more diluted and less timely, thus less informative. Self-readings submitted outside the IF window still count, but the resulting bill will include a component based on estimated consumption, for the days between the IF window and the last day of the billing cycle. In an extreme example, imagine a customer who submits a read after the IF window, but before the end of the bill cycle, and has not submitted reads in the months before nor will submit reads in the following ones: this customer will receive a bill completely based on estimated consumption a week later, and a bill based on real consumption two months after the self-reading, which would include two months of estimated consumption.

Bill cycles differ for each customer, and mainly depend on the date in which the customer's contract was activated. Since campaign messages were sent over the same one-week period (May 3 – May 8), they could reach a customer within or outside her IF window. The treatments exploit this variation: customers whose IF window occurred during the campaign week received the IF treatment message (24,748 customers), while other customers received the standard message (1,058,621 users).

While we may expect the IF treatment to be exogenous to customers' characteristics, since it depends on the contract activation date, and thus the IF and DF groups to be balanced in terms of their characteristics, this is not the case. Similarly, control group customers significantly differ from treated ones across several dimensions. Appendix Table A4 presents summary statistics of the three groups, and balance tests. The three samples of customers are significantly different across all the dimensions we have data for.

We thus use propensity score matching (PSM) techniques to select the sample for our empirical analysis. We proceed in two stages: first, we apply PSM to construct two balanced treated samples; second, we apply PSM to identify a suitable control group for our pooled sample of treated customers. We match customers on the basis of household characteristics and gas consumption habits. In terms of consumption, we consider average monthly bill amount, average consumption between January 2016 and April 2017and the type of gas use (heating, cooking, hot water). Since our outcome variable is engagement, in the form of self-reading, we include in the matching routine customers' characteristics that are potentially related with their engagement level: total self-readings submitted between January 2016 and April 2017, past activities on the utility's website, any delays in bill payment, registered email or phone number, electronic bill, direct debit, age,[11] gender and location (north-center-south of Italy) of the contract holder. Finally, we include some extra controls related to their account: the type of contract they subscribed to and dual contracts (electricity and gas).

In the first stage, we find the common support in the propensity score, defined as the probability to be included in the IF treatment, given customers characteristics, computed through a logit specification. Within the common support, we elicit the nearest neighbour (NN) matching estimator lying within the caliper (0.001), with no replacement option. This procedure produces a sub-sample of 47,535 customers.

We follow the same protocol to find a suitable control group for the pooled sample of treated customers among the clients excluded from the intervention, resulting in a control sample of 29,628 customers.[12]

Table 1 shows summary statistics of the three groups: the control, DF and IF. The PSM has reduced both the magnitude of the differences between the groups along all dimensions, and the number of traits where we observe significant differences between the two groups. Statistically significant differences, at the 5 per cent level or below, remain in terms of baseline consumption and bill amount, gas use and availability of email contact. However, the magnitude of these differences is small, for instance at most 6 per cent for consumption and 4.5 per cent for bills. We address this issue by including controls for all unbalanced covariates in the regression analysis.

[11] We expect age to play an important role in gas-consumption, thus we control for age squared as well.

[12] Among the 47,535 customers participating either to DF or IF treatment, 17,970 fall out from the common support in the second stage of the Propensity Score Matching, where we find a suitable control for the treated group. We finally end up with 14,743 customers in the DF treatment group and 14,822 in the IF one.)

Table 1. Study 2: summary statistics

Variable	Control Obs	Control Mean	Control Std. Dev.	DF Obs	DF Mean	DF Std. Dev.	IF Obs	IF Mean	IF Std. Dev.	Prob > F
Baseline self-reading	29,628	2.235	2.733	14,822	2.298	2.609	14,743	2.254	2.642	0.064
Baseline consumption	29,628	89.339	154.363	14,822	88.594	101.668	14,743	94.226	78.575	0.000
Baseline bill	29,628	74.794	102.122	14,822	78.398	337.563	14,743	76.626	58.845	0.038
North	29,628	0.407	0.491	14,822	0.407	0.491	14,743	0.410	0.492	0.733
Center	29,628	0.172	0.378	14,822	0.169	0.374	14,743	0.168	0.374	0.433
South	29,628	0.373	0.483	14,822	0.376	0.485	14,743	0.374	0.484	0.720
Gas use: heating	29,628	0.861	0.346	14,822	0.866	0.340	14,743	0.881	0.323	0.000
Gas use: hot water	29,628	0.956	0.205	14,822	0.961	0.194	14,743	0.968	0.175	0.000
Gas use: cooking	29,628	0.972	0.166	14,822	0.971	0.169	14,743	0.969	0.173	0.403
Gas use: other_use	29,628	0.000	0.000	14,822	0.000	0.000	14,743	0.000	0.000	.
Age	29,628	53.154	15.647	14,822	53.068	15.614	14,743	53.071	15.513	0.806
Female	29,628	0.414	0.493	14,822	0.413	0.492	14,743	0.409	0.492	0.586
Bill pay: direct debit	29,628	0.182	0.386	14,822	0.187	0.390	14,743	0.178	0.383	0.148
Late on bill	29,628	0.100	0.301	14,822	0.105	0.306	14,743	0.102	0.302	0.403
Contact: mobile	29,628	0.901	0.298	14,822	0.899	0.302	14,743	0.902	0.297	0.549
Contact: email	29,628	0.983	0.130	14,822	0.978	0.148	14,743	0.989	0.104	0.000
Electronic bill	29,628	0.225	0.418	14,822	0.230	0.421	14,743	0.231	0.422	0.279
Active on website	29,628	0.420	0.494	14,822	0.424	0.494	14,743	0.422	0.494	0.698
Contract type: free mkt 1	29,628	0.012	0.110	14,822	0.013	0.112	14,743	0.013	0.115	0.544
Contract type: free mkt 2	29,628	0.209	0.406	14,822	0.216	0.412	14,743	0.213	0.409	0.188
Contract type: free mkt 3	29,628	0.251	0.433	14,822	0.254	0.435	14,743	0.248	0.432	0.557
Contract type: regulated mkt	29,628	0.036	0.187	14,822	0.037	0.189	14,743	0.035	0.183	0.567
Dual contracts	29,628	0.092	0.289	14,822	0.093	0.291	14,743	0.093	0.291	0.830

Note: summary statistics for Control, Deferred Feedback (DF) and Immediate Feedback (IF) groups. To test balance among treatments, we pool all observations across the three groups; for each variable we estimate $y_i = \beta_0 + \beta_1 \times DF_i + \beta_2 \times IF_i + \epsilon$ and report the p-value of the joint test of the following null hypothesis $H_0\colon \beta_1 = \beta_2 = 0$. Variable definition: baseline self-reading, consumption and bill are computed between January 2016 and April 2017; late on bill is a dummy denoting customers who have been late in their bill payment but with no stop in issuing invoices; dual contract is a dummy indicating whether a customer has both a gas and an electricity contract.

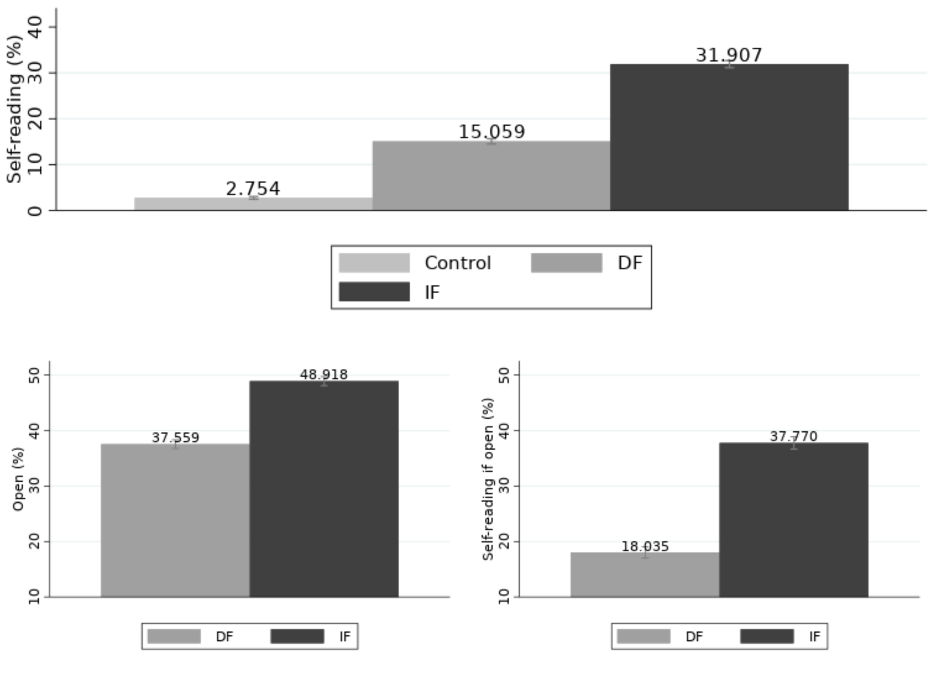
4. Results

We now turn to the analysis of the campaign's impact on self-reading. All subjects in the treatment sample received the campaign email, and 12,779 of them opened the email, equivalent to a 43.29% per cent average open rate. Of them, 29.17 per cent submitted self-readings in the 15 days following the campaign.

Figure 2 shows the share of customers submitting self-readings within 15 days from the message by treatment and control groups (top panel), the share of customers opening the campaign email by treatment group among treated customers (bottom left panel), and the share of self-readings by treatment group among treated customers who opened the email.

Both campaign messages had a strong, positive and statistically significant effect on self-reading. With respect to customers in the control group, customers receiving the DF treatment are 12.5 percentage points more likely to submit a self-reading. The IF treatment leads to an even more striking increase: customers included in IF treatment were 16 percentage points more likely to submit a self-reading than customers in the DF treatment. Similarly, among treated customers, those in the IF treatment were 12 percentage points more likely to open the email, and 20 percentage points more likely to submit a reading once they opened the email. These effects are large, as self-reading rates more than double under the IF treatment, while open rates increase by about 33 per cent with respect to the DF treatment values.

Figure 2. Self-reading, opened email and self-reading if opened.



Notes: bars indicate means, whiskers indicate 95 per cent confidence intervals.

Since a key goal of self-reading campaigns of the type analysed here is that of engaging customers who had not submitted self-readings before, we next ask whether treatment effects vary depending on customers' baseline level of engagement.

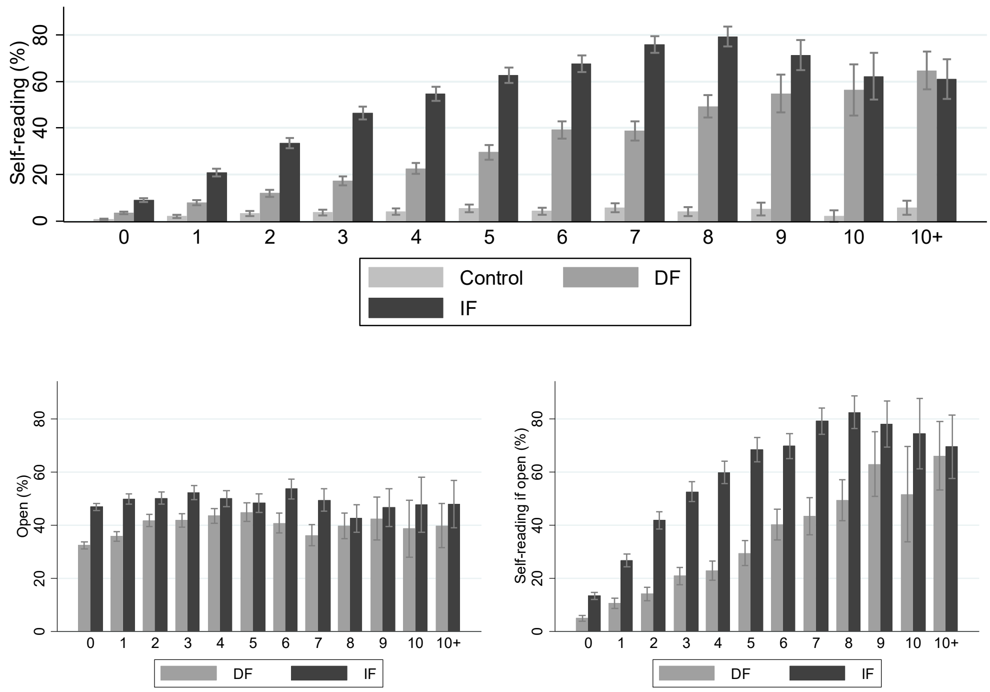
A first indicator of prior engagement with one's own gas consumption, available in our data, is the number of self-readings submitted by a customer between January 2016 and April 2017, just before the campaign. Since customers are billed typically bi or tri-monthly, and do not receive any additional information by submitting more than one self-reading within each billing cycle, we would not expect customers, who submit self-readings to have their bills based on real consumption, to submit more than 8 self-readings over the 16 months' period that we observe before the campaign.[13] Indeed, Figure A3 shows that the share of customers submitting more than 8 self-readings before the campaign is negligible. Consistent with the need to engage customers to attend to their gas consumption, almost 40 per cent of customers did not submit any self-readings in the 16 months prior to the campaign.

Figure 2 reports the share of customers submitting self-readings after the campaign and opening the campaign email, by treatment and number of self-readings submitted at baseline. The fact that DF and IF treatments both significantly increase the share of self-readings submitted, and that the IF treatment is about twice as effective as the DF one are confirmed at all levels of prior engagement. Similar results obtain when we focus on treated subjects and compare open rates and self-reading rates among customers who opened the email.

The passive customers, no self-readings submitted at baseline (10,315 customers, 5,033 in DF group and 5,282 in IF group) in IF treatment react (almost) three times more than the customers receiving DF treatment ($M_{DF} = 3.537$ vs. $M_{IF} = 8.898$). The difference between the two groups becomes larger as self-readings submitted increase; the maximum difference between IF (546 customers) and DF (534 customers) groups is reached for seven self-readings are baseline ($M_{DF} = 38.645$ vs. $M_{IF} = 75.843$). By contrast, active customers submitting at least a self-reading every two-month, react similarly to the two treatments, showing that urgency has smaller effect on those customers paying a lotto of attention to the service.

Figure 2. Self-reading, opened email and self-reading if open by past self-reading frequencies.

[13] A customer may need to send frequent self-readings in case of change of contract, tariff or meter, or problems with the bill.

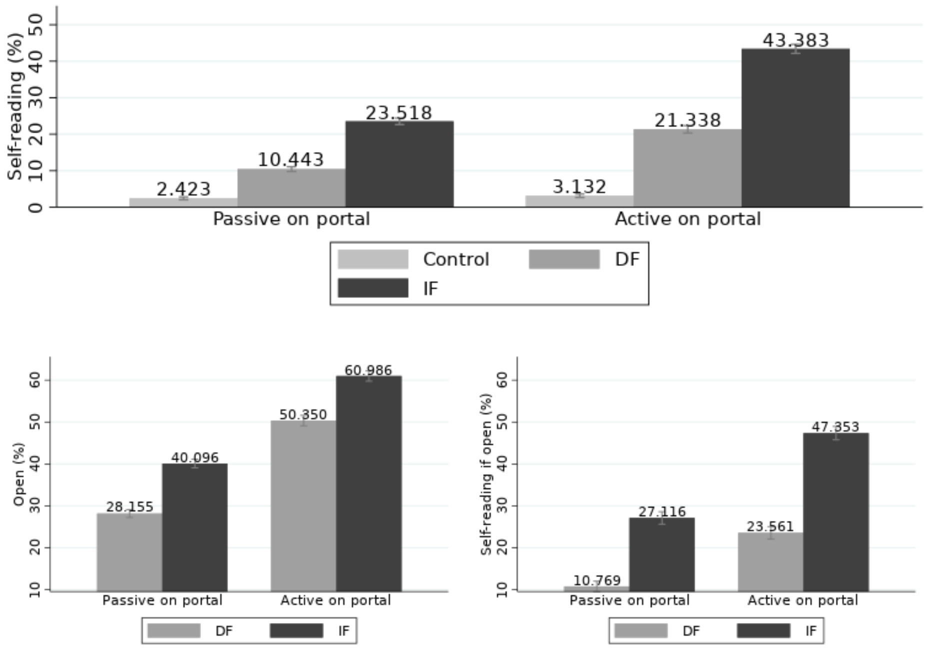


Notes: bars indicate means, whiskers indicate 95 per cent confidence intervals.

Next, we consider activity on the utility's website as another indicator of customer engagement. On the utility's website, each customer, if registered, can access a private area with information on past bills, tariffs, an area to submit a self-reading, etc. Being active on the website is not only a signal of engagement, but also a proxy for higher digital literacy. We cannot disentangle these two dimensions, but we expect them both to impact self-reading and email opening rates for a number of reasons. First, customers who are more digitalized are able to submit a reading more easily through the utility's app or website and could, for this reason, respond more to the treatment. Theory and evidence from behavioural economics shows how small practical barriers may have large impact on behaviour, due to status quo bias, inertia and procrastination (Carroll et al. 2009). Second, customers who have actively looked at their personal page on the utility's web portal demonstrate to be more digitalized and engaged.

Figure 3 shows that, indeed, levels of engagement are higher among customers active on the web portal, both in terms of self-readings and of email open rates. As for treatment effects, the share of passive customers submitting self-readings under the OF treatment is twice as large as that under the DF treatment and 10 times larger than that of the control group. The results are qualitatively similar, but the magnitude of the differences somewhat larger, among active customers. When comparing open rates and self-reading rates among customers who opened the email, treatment effects remain large and similar across active and inactive customers.

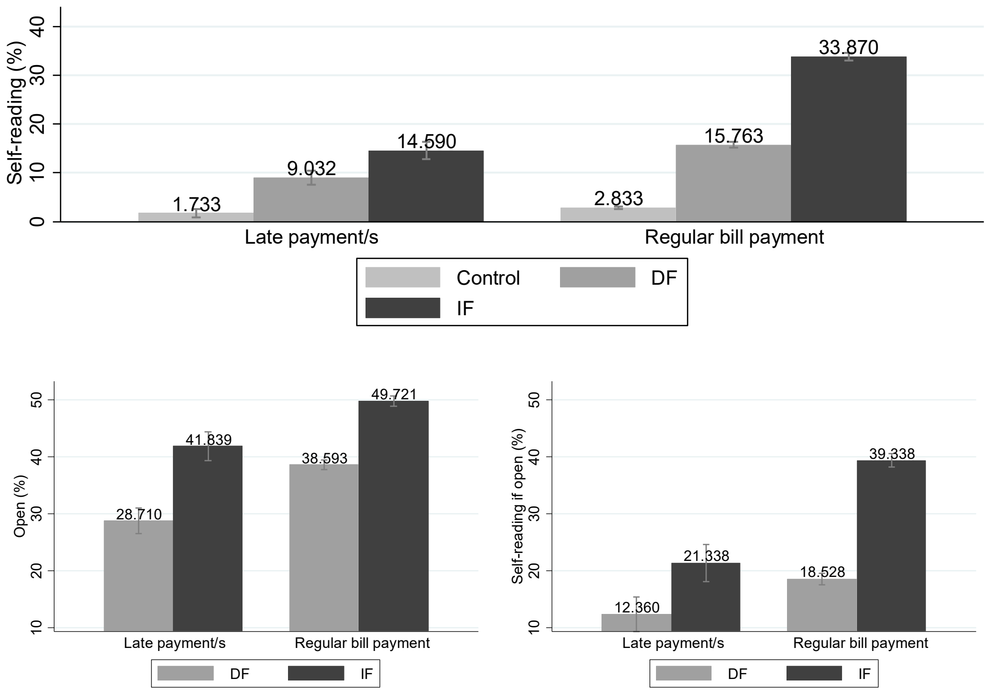
Figure 3. Self-reading, opened email and self-reading if opened by customers active on website.



Notes: bars indicate means, whiskers indicate 95 per cent confidence intervals.

Finally, we consider a third indicator of engagement with one's gas use: timely bill payment. Although late bill payment may also signal lack of financial resources, customers who pay little attention to their bills may forget to pay them on time. We use an indicator variable, provided by the utility, of whether a customer experienced delays in paying her bills. Figure 4 shows how engagement levels, proxied by self-reading rates and email opening rates, are higher among customers who regularly pay their bills. Even on late-paying customers, the campaign appears to be effective in raising self-reading rates with respect to the control group, and the two treatments show qualitatively similar and strong results among both regular and late bill payers. When focusing on treated customers, the IF treatment leads to significantly higher email open rates and self-reading rates among customers who opened the email than the DF treatment, and this holds for both regular and late bill payers.

Figure 4. Self-reading, opened email and self-reading if opened by customers late on bill payment.



Notes: bars indicate means, whiskers indicate 95 per cent confidence intervals.

We test the statistical significance of these results, and in particular of differences between treatments and over each source of heterogeneity, using regression analysis. Table 2 shows results of linear regressions of self-reading on treatment indicators and their interaction with the three dimensions of engagement considered in the graphical analysis: number of baseline self-reading, being active on the web portal, and being on time with bill payment. We use the regular bill payment indicator in the regression, so that all our measures of heterogeneity are increasing in engagement. All regressions control for unbalanced covariates across treatment groups, namely baseline self-reading, baseline bill amount, baseline consumption, dummies for whether gas is used for heating or hot water, and a dummy if the utility has the customer's email address. Columns 1 to 3 consider each dimension in isolation, while Column 4 pools them in a single regression. To correct for the false discovery rate (FDR) due to multiple testing, in Column 4 we include FDR-adjusted q-values in square brackets (M. L. Anderson 2008).

The regression results confirm the basic patterns we observed in the graphical analysis. First, the likelihood of submitting a self-reading after the campaign is significantly and positively correlated with the number of self-readings submitted in the period before the campaign. However, once we control for baseline self-reading and the other covariates, neither being active on the portal nor paying bills on time is significantly or consistently correlated with self-reading after the campaign.[14]

[14] The 3 proxies of engagement are correlated with each other: correlation coefficients are 0.259 between baseline self-reading and being active on portal; 0.143 between baseline self-reading and paying bills on time;

Second, both treatments significantly increase self-reading among both unengaged and engaged customers, with the IF treatment having the strongest effect. Among unengaged customers, the DF treatment increases self-reading between 1.5 and 15.5 percentage points, depending on the dimension of engagement considered, while the effect of the IF treatment ranges between 12.1 and 33.6 percentage points. When we pool all dimensions of engagement together in one regression, the DF treatment does not significantly affect self-reading among unengaged customers, while the effect of the IF treatment remains statistically significant. All differences between treatments are statistically significant, according to Wald tests of the equality of the IF and DF coefficients (all p's = 0.000). Among engaged customers, the IF treatment is also always significantly more effective than the DF one (Wald tests, all p's = 0.000). When we pool all dimensions of engagement together, the DF treatment does not significantly increase self-reading among regular bill payers.

Third, engaged customers react more strongly to the campaign. The coefficients on the interaction terms between the treatment dummies and the engagement indicators are always positive and statistically significant. When we test for the difference in coefficients between each treatment dummy in isolation and interacted with the engagement proxy, we confirm this result across all the engagement proxies that we consider, but one. Namely, both treatments are significantly more effective among engaged customers (Wald tests, all p's = 0.000), with one exception: the IF treatment works equally well among customers who are active and passive on the portal (p = 0.761).

---

and 0.101 between being active on portal and paying bills on time. All correlations are significant at the 1 per cent level (all p's = 0.000).

Table 2. Treatment effect on self-reading: heterogeneity by customers' engagement level

Dependent variable	Self-reading after the campaign			
	(1)	(2)	(3)	(4)
DF	0.017***	0.091***	0.155***	0.003
	(0.00)	(0.00)	(0.00)	(0.00)
IF	0.121***	0.222***	0.336***	0.097***
	(0.00)	(0.00)	(0.00)	(0.01)
Baseline self-reading	0.004***	0.044***	0.045***	0.004***
	(0.00)	(0.00)	(0.00)	(0.00)
Active on portal		-0.072***		-0.002
		(0.00)		(0.00)
Regular bill payment			-0.060***	0.004
			(0.01)	(0.00)
DF x Baseline self-reading	0.048***			0.046***
	(0.00)			(0.00)
				[0.001]
IF x Baseline self-reading	0.078***			0.072***
	(0.00)			(0.00)
				[0.001]
DF x Active on portal		0.127***		0.048***
		(0.01)		(0.01)
				[0.001]
IF x Active on portal		0.218***		0.107***
		(0.01)		(0.01)
				[0.001]
DF x Regular bill payment			0.084***	0.009
			(0.01)	(0.01)
				[0.123]
IF x Regular bill payment			0.196***	0.080***
			(0.01)	(0.01)
				[0.001]
Constant	-0.017	-0.084***	-0.123***	-0.008
	(0.01)	(0.01)	(0.01)	(0.01)
Number of Obs	40786000	40786000	40786000	40786000
R-Squared	0.260	0.223	0.211	0.270

Notes: OLS regressions, robust standard errors in parentheses, FDR-adjusted q-values in square brackets (Anderson 2008). All regressions control for: baseline self-reading, baseline bill amount, baseline consumption, dummies for whether gas is used for heating or hot water, and a dummy if the utility has the customer's email address. * p<0.1, ** p<0.05, *** p<0.01.

Taken together, the results of the empirical analysis suggest that the campaign, with its message emphasizing the link between self-reading and control over one's real gas expenditures, was broadly effective in fostering self-reading, both among customers with and low levels of baseline engagement. However, adding urgency to this basic message resulted in much larger effects, among all customer segments.

Our results seem the confirm that urgency best works for those people who are more tuned into the time dimension (engaged customers) thus chronically paying more attention to task expiration time (Zhu, Yang, and Hsee 2018).

5. Discussion

We study the impact of messages encouraging customers of a large utility to submit gas meter readings. Treatments vary the content of the messages, in particular making salient a near deadline for submitting self-readings, which would make the upcoming bill completely based on real consumption, as opposed to estimated one. Imposing a sense of urgency is a strong motivator of engagement, especially for customers who already pay some attention to their gas consumption.

This study has implications both for business and policy. Following market liberalization and technological advancement, the energy industry has transformed from pure energy producer and distributor to provider of energy services. Engaging with customers by becoming advisors is now an integral part of energy companies' strategies. Transparency and accountability have also become key indicators in a market, which has become more complex over time. Companies seek to encourage customer feedback because of regulatory requirements, because they aim to build a reputation, as well as to ensure that payment is timely and complaints are limited –something which reducing bill fluctuations, thanks to more frequent readings, can provide.

Policymakers have pushed customer awareness and empowerment as key levers to reduce energy waste and promote a sustainable energy transition. The so called 'Energy Paradox' (Jaffe and Stavins 1994; Allcott and Wozny 2014; Attari et al. 2010) has highlighted how consumers mis-optimize their energy investments. This has a double negative impact: it affects households welfare directly and indirectly by increasing negative externalities, such as air pollution and climate change. To help consumers make better choices, a great number of countries have legislated policies aimed at reducing the information gap (Abrahamse and Steg 2009; Sudarshan 2017; S. T. Anderson and Newell 2004). But providing information is ineffective if people do not attend to it. The advent of smart metering and real time measurements will generate a large quantity of new data, but it remains to be seen whether this will actually lead to better energy decisions by consumers (Lynham et al. 2016; Lurie 2004). The insights of this paper suggest that behavioural insights can be successfully leveraged to increase customers' engagement, even among those who are least attentive.

References

Abrahamse, Wokje, and Linda Steg. 2009. “How Do Socio-Demographic and Psychological Factors Relate to Households’ Direct and Indirect Energy Use and Savings?” *Journal of Economic Psychology* 30 (5): 711–20. https://doi.org/10.1016/j.joep.2009.05.006.

Abrahamse, Wokje, Linda Steg, Charles Vlek, and Talib Rothengatter. 2005. “A Review of Intervention Studies Aimed at Household Energy Conservation.” *Journal of Environmental Psychology* 25 (3): 273–291.

Allcott, Hunt. 2011. “Social Norms and Energy Conservation.” *Journal of Public Economics* 95 (9–10): 1082–1095.

Allcott, Hunt, and Sendhil Mullainathan. 2010. “Behavior and Energy Policy.” *Science* 327 (5970): 1204–1205.

Allcott, Hunt, and Todd Rogers. 2014. “The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation.” *American Economic Review* 104 (10): 3003–37.

Allcott, Hunt, and Nathan Wozny. 2014. “Gasoline Prices, Fuel Economy, and the Energy Paradox.” *Review of Economics and Statistics* 96 (5): 779–795.

Anderson, Michael 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.

Anderson, Soren T, and Richard G Newell. 2004. “Information Programs for Technology Adoption: The Case of Energy-Efficiency Audits.” *Resource and Energy Economics* 26 (1): 27–50.

Andor, Mark A, and Katja M Fels. 2018. “Behavioral Economics and Energy Conservation–a Systematic Review of Non-Price Interventions and Their Causal Effects.” *Ecological Economics* 148: 178–210.

Andor, Mark Andreas, Andreas Gerster, Jörg Peters, and Christoph M. Schmidt. 2017. “Social Norms and Energy Conservation beyond the US.” 714. Ruhr Economic Papers. RWI - Leibniz-Institut für Wirtschaftsforschung, Ruhr-University Bochum, TU Dortmund University, University of Duisburg-Essen. https://ideas.repec.org/p/zbw/rwirep/714.html.

Arrow, Kenneth J. 1971. “The Theory of Risk Aversion.” *Essays in the Theory of Risk-Bearing*, 90–120.

Attari, Shahzeen Z, Michael L DeKay, Cliff I Davidson, and Wändi Bruine De Bruin. 2010. “Public Perceptions of Energy Consumption and Savings.” *Proceedings of the National Academy of Sciences* 107 (37): 16054–16059.

Ayres, Ian, Sophie Raseman, and Alice Shih. 2013. “Evidence from Two Large Field Experiments That Peer Comparison Feedback Can Reduce Residential Energy Usage.” *The Journal of Law, Economics, and Organization* 29 (5): 992–1022.

Botti, Simona, Susan Broniarczyk, Gerald Häubl, Ron Hill, Yanliu Huang, Barbara Kahn, Praveen Kopalle, Donald Lehmann, Joe Urbany, and Brian Wansink. 2008. “Choice under Restrictions.” *Marketing Letters* 19 (3–4): 183–199.

Brounen, Dirk, Nils Kok, and John M Quigley. 2013. “Energy Literacy, Awareness, and Conservation Behavior of Residential Households.” *Energy Economics* 38: 42–50.

Byrne, David P, Andrea La Nauze, and Leslie A Martin. 2017. “Tell Me Something I Don’t Already Know: Informedness and the Impact of Information Programs.” *Review of Economics and Statistics*, no. 0.

Carroll, Gabriel D, James J Choi, David Laibson, Brigitte C Madrian, and Andrew Metrick. 2009. “Optimal Defaults and Active Decisions.” *The Quarterly Journal of Economics* 124 (4): 1639–1674.

Cialdini, Robert B. 2007. *Influence: The Psychology of Persuasion*. Collins New York.

Costa, Dora L, and Matthew E Kahn. 2013. "Energy Conservation 'Nudges' and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment." *Journal of the European Economic Association* 11 (3): 680–702.
Darby, Sarah. 1999. "Energy Advice - What Is It Worth?" *Proceedings, European Council for an Energy-Efficient Economy Summer Study, Paper III* 5 (3–5): 13.
———. 2010. "Smart Metering: What Potential for Householder Engagement?" *Building Research & Information* 38 (5): 442–57. https://doi.org/10.1080/09613218.2010.492660.
Ferraro, Paul J, and Michael K Price. 2013. "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment." *Review of Economics and Statistics* 95 (1): 64–73.
Fischer, Corinna. 2008. "Feedback on Household Electricity Consumption: A Tool for Saving Energy?" *Energy Efficiency* 1 (1): 79–104. https://doi.org/10.1007/s12053-008-9009-7.
Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature* 40 (2): 351–401. https://doi.org/10.1257/002205102320161311.
Gilbert, Ben, and Joshua Graff Zivin. 2014. "Dynamic Salience with Intermittent Billing: Evidence from Smart Electricity Meters." *Journal of Economic Behavior & Organization* 107: 176–190.
Hargreaves, Tom, Michael Nye, and Jacquelin Burgess. 2013. "Keeping Energy Visible? Exploring How Householders Interact with Feedback from Smart Energy Monitors in the Longer Term." *Energy Policy* 52: 126–134.
Harrison, Glenn W, and John A List. 2004. "Field Experiments." *Journal of Economic Literature*, 96.
Houwelingen, Jeannet H. van, and W. Fred van Raaij. 1989. "The Effect of Goal-Setting and Daily Electronic Feedback on In-Home Energy Use." *Journal of Consumer Research* 16 (1): 98. https://doi.org/10.1086/209197.
Jaffe, Adam B, and Robert N Stavins. 1994. "The Energy-Efficiency Gap What Does It Mean?" *Energy Policy* 22 (10): 804–810.
Jessoe, Katrina, and David Rapson. 2014. "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review* 104 (4): 1417–38.
Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler. 1986. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *The American Economic Review* 76 (4): 728–41.
Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng. 2006. "The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention." *Journal of Marketing Research* 43 (1): 39–58.
List, John A., and Michael K. Price. 2016. "The Use of Field Experiments in Environmental and Resource Economics." *Review of Environmental Economics and Policy* 10 (2): 206–25. https://doi.org/10.1093/reep/rew008.
Loschelder, David D., Henrik Siepelmeyer, Daniel Fischer, and Julian A. Rubel. 2019. "Dynamic Norms Drive Sustainable Consumption: Norm-Based Nudging Helps Café Customers to Avoid Disposable to-Go-Cups." *Journal of Economic Psychology*, February. https://doi.org/10.1016/j.joep.2019.02.002.
Lurie, Nicholas H. 2004. "Decision Making in Information-Rich Environments: The Role of Information Structure." *Journal of Consumer Research* 30 (4): 473–86. https://doi.org/10.1086/380283.
Lynham, John, Kohei Nitta, Tatsuyoshi Saijo, and Nori Tarui. 2016. "Why Does Real-Time Information Reduce Energy Consumption?" *Energy Economics* 54: 173–181.

Lynn, Michael. 1989. “Scarcity Effects on Desirability: Mediated by Assumed Expensiveness?” *Journal of Economic Psychology* 10 (2): 257–274.
McClure, Samuel M., David I. Laibson, George Loewenstein, and Jonathan D. Cohen. 2004. “Separate Neural Systems Value Immediate and Delayed Monetary Rewards.” *Science* 306 (5695): 503–7. https://doi.org/10.1126/science.1100907.
Pratt, John W. 1964. “Robustness of Some Procedures for the Two-Sample Location Problem.” *Journal of the American Statistical Association* 59 (307): 665–80. https://www.jstor.org/stable/2283092.
Pribram, Karl H, and Diane McGuinness. 1975. “Arousal, Activation, and Effort in the Control of Attention.” *Psychological Review* 82 (2): 116.
Ramos, Ana, Alberto Gago, Xavier Labandeira, and Pedro Linares. 2015. “The Role of Information for Energy Efficiency in the Residential Sector.” *Energy Economics* 52: S17–S29.
Shah, Anuj K, Eldar Shafir, and Sendhil Mullainathan. 2015. “Scarcity Frames Value.” *Psychological Science* 26 (4): 402–412.
Sparkman, Gregg, and Gregory M. Walton. 2017. “Dynamic Norms Promote Sustainable Behavior, Even If It Is Counternormative.” *Psychological Science* 28 (11): 1663–74. https://doi.org/10.1177/0956797617719950.
Sudarshan, Anant. 2017. “Nudges in the Marketplace: The Response of Household Electricity Consumption to Information and Monetary Incentives.” *Journal of Economic Behavior & Organization* 134: 320–335.
Thaler, Richard H., and Shlomo Benartzi. 2004. “Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving.” *Journal of Political Economy* 112 (S1): S164–87. https://doi.org/10.1086/380085.
Tversky, Amos, and Daniel Kahneman. 1992. “Advances in Prospect Theory: Cumulative Representation of Uncertainty.” *Journal of Risk and Uncertainty* 5 (4): 297–323. https://doi.org/10.1007/BF00122574.
Verhallen, Theo MM, and Henry SJ Robben. 1994. “Scarcity and Preference: An Experiment on Unavailability and Product Evaluation.” *Journal of Economic Psychology* 15 (2): 315–331.
Wichman, Casey J. 2017. “Information Provision and Consumer Behavior: A Natural Experiment in Billing Frequency.” *Journal of Public Economics* 152: 13–33.
Zhu, Meng, Rajesh Bagchi, and Stefan J. Hock. 2019. “The Mere Deadline Effect: Why More Time Might Sabotage Goal Pursuit.” *Journal of Consumer Research* 45 (5): 1068–84. https://doi.org/10.1093/jcr/ucy030.
Zhu, Meng, Yang Yang, and Christopher K. Hsee. 2018. “The Mere Urgency Effect.” *Journal of Consumer Research* 45 (3): 673–90. https://doi.org/10.1093/jcr/ucy008.

Appendix Tables and Figures

Figure A1. Gas meter

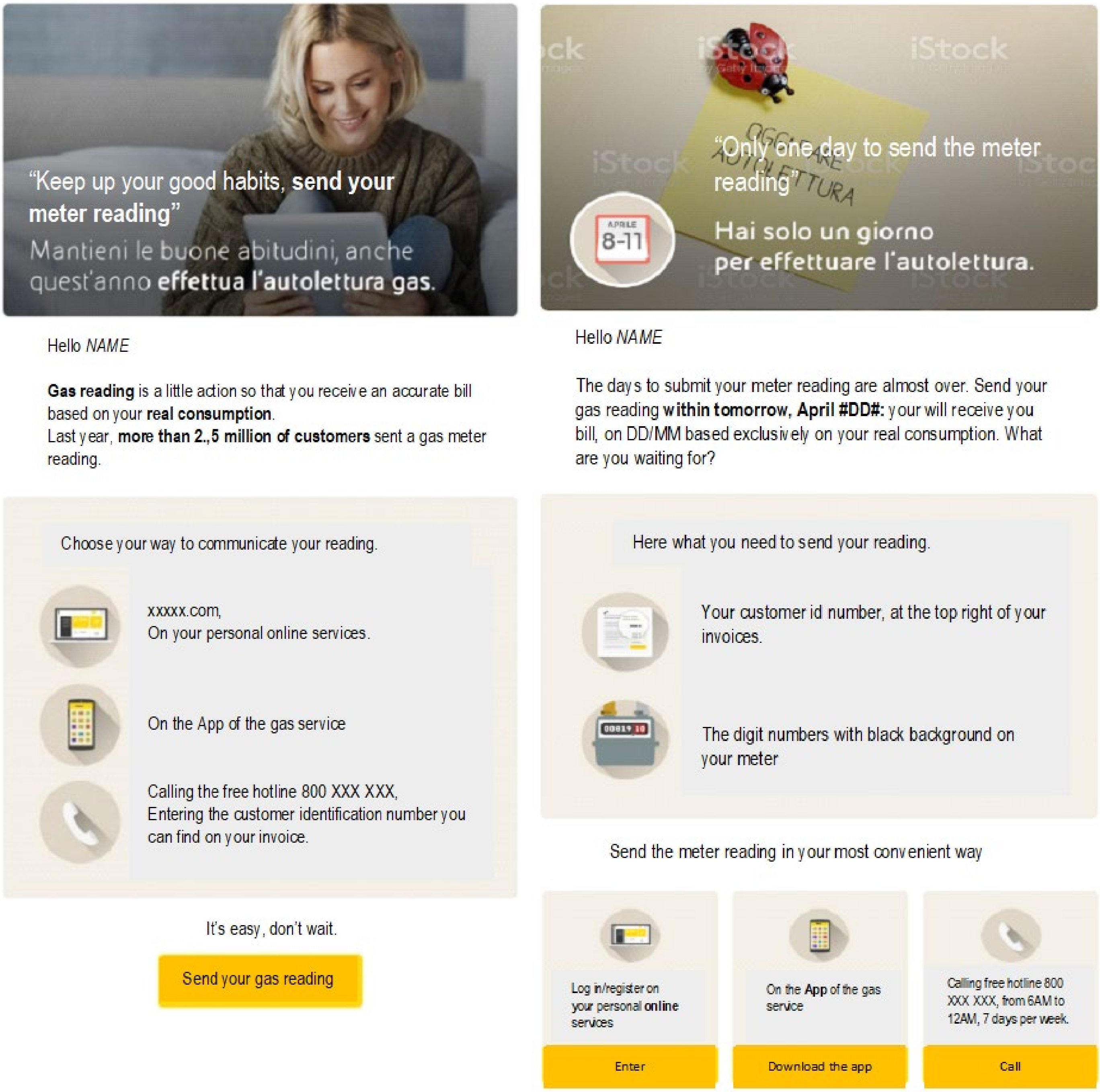
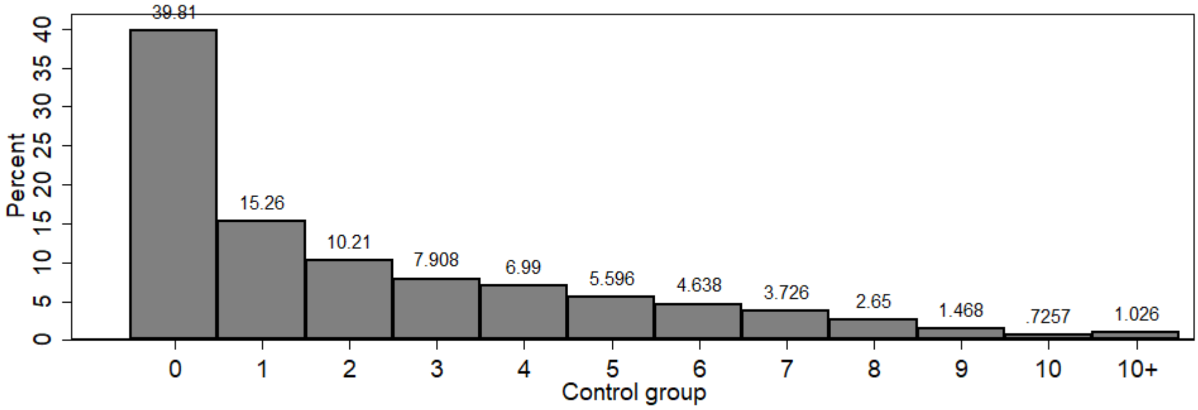
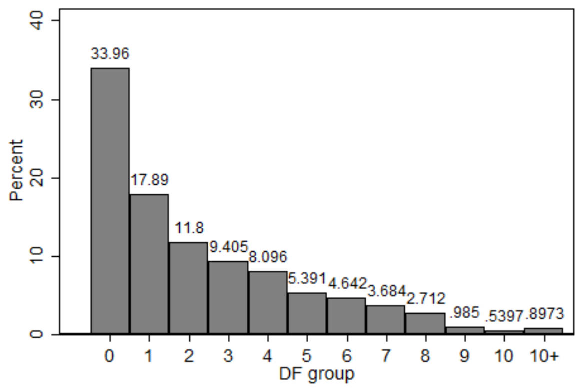


Figure A2. Structure of campaign messages, from May 2017 self-reading campaign (Study 2)

Figure A3: Self-reading before the campaign





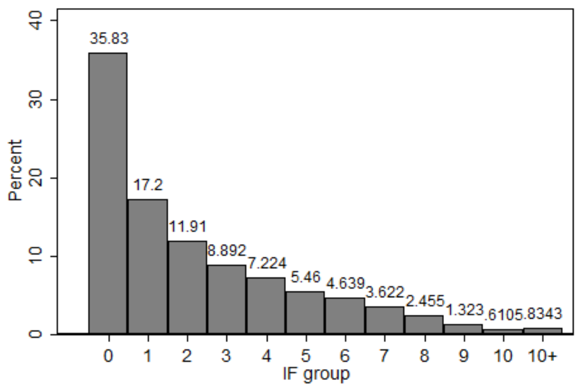


Table A1: List of variables

Variable name	Description
Self-reading after the campaign	Customer submitted a self-reading after the campaign (15 days)
Baseline self-reading	Monthly self-reading between January 2016-April 2017
Baseline consumption	Average monthly consumption (m^3) between January 2016-April 2017
Baseline bill	Average monthly billing (Euro) between January 2016-April 2017
North	Contract address location: North
Center	Contract address location: Center
South	Contract address location: South
Gas use: heating	Domestic gas use: heating
Gas use: hot water	Domestic gas use: hot water
Gas use: cooking	Domestic gas use: cooking
Gas use: other_use	Domestic gas use: other use
Age	Age of contract holder
Female	Gender of contract holder
Bill pay: direct debit	Bill payment type: direct debit
Late on bill	At least one bill paid with delay or not paid
Contact: mobile	Phone number available
Contact: email	Email available
Electronic bill	Bill received by email
Active on website	At least one operation made on the utility's web portal in the prior 12 months
Contract type: free mkt 1	Type of contract: free market type 1
Contract type: free mkt 2	Type of contract: free market type 2
Contract type: free mkt 3	Type of contract: free market type 3
Contract type: regulated mkt	Type of contract: regulated market
Dual contracts	Has both a gas and electricity contract with the utility

Table A2: summary statistics before PSM Balance test

	Control			DF			IF			
	Obs	Av	Std. Dev.	Obs	Av	Std. Dev.	Obs	Av	Std. Dev.	Prob > F
Baseline self-reading	1,637,002	2.148	2.559	1,058,621	1.769	2.284	24,748	2.245	2.691	0.000
Baseline consumption	438,571	97.479	88.160	1,056,097	86.732	91.260	24,504	101.575	88.104	0.000
Baseline bill	438,571	105.802	68.114	1,056,097	70.076	1062.021	24,504	83.260	67.034	0.000
North	1,636,977	0.410	0.480	1,058,601	0.374	0.484	24,748	0.413	0.492	0.000
Center	1,636,977	0.149	0.388	1,058,601	0.189	0.391	24,748	0.147	0.354	0.000
South	1,636,977	0.386	0.490	1,058,601	0.364	0.481	24,748	0.384	0.486	0.000
Gas use: heating	1,636,325	0.919	0.382	1,058,582	0.796	0.403	24,746	0.925	0.263	0.000
Gas use: hot water	1,636,325	0.978	0.294	1,058,582	0.892	0.311	24,746	0.980	0.138	0.000
Gas use: cooking	1,636,325	0.975	0.181	1,058,582	0.973	0.161	24,746	0.975	0.157	0.000
Gas use: other_use	1,637,002	0.000	0.042	1,058,621	0.000	0.011	24,748	0.000	0.000	0.000
Age	1,591,404	53.332	16.983	1,043,370	56.013	16.094	24,717	53.316	15.467	0.000
Female	1,583,277	0.420	0.495	1,036,383	0.405	0.491	24,636	0.418	0.493	0.000
Bill pay: direct debit	1,637,002	0.114	0.483	1,058,621	0.437	0.496	24,748	0.110	0.313	0.000
Late on bill	1,637,002	0.087	0.178	1,058,621	0.035	0.184	24,748	0.087	0.281	0.000
Contact: mobile	1,637,002	0.890	0.500	1,058,621	0.739	0.439	24,748	0.890	0.313	0.000
Contact: email	1,637,002	0.990	0.369	1,058,621	0.891	0.312	24,748	0.993	0.081	0.000
Electronic bill	1,637,002	0.184	0.215	1,058,621	0.318	0.466	24,748	0.187	0.390	0.000
Active on website	1,637,002	0.430	0.269	1,058,621	0.401	0.490	24,748	0.433	0.496	0.000
Contract type: free mkt 1	1,636,681	0.009	0.032	1,058,416	0.003	0.052	24,643	0.009	0.094	0.000
Contract type: free mkt 2	1,636,681	0.143	0.185	1,058,416	0.041	0.198	24,643	0.140	0.347	0.000
Contract type: free mkt 3	1,636,681	0.297	0.185	1,058,416	0.032	0.177	24,643	0.301	0.459	0.000
Contract type: regulated mkt	1,636,681	0.022	0.340	1,058,416	0.866	0.341	24,643	0.021	0.144	0.000
Dual contracts	1,637,002	0.096	0.135	1,058,621	0.030	0.171	24,748	0.094	0.292	0.000

Note: summary statistics for Control group, Deferred Feedback (DF) and Immediate Feedback (IF) . To test balance among treatments, we pool all observations across the three groups; for each variable we estimate $y_i = \beta_0 + \beta_1 \times DF_i + \beta_2 \times IF_i + \epsilon$ we report the p-value of the joint test of the following null hypothesis $H_0: \beta_1 = \beta_2 = 0$. Variable definition: baseline self-reading, consumption and bill are computed between January 2016 and April 2017; late on bill is a dummy denoting customers who are late (with at least) one of their bill payment; dual contract is a dummy indicating whether a customer has both a gas and an electricity contract.

Appendix B: Abtest

B.1. Design

The intervention took place at the end of April 2016 and consisted in email messages inviting customers to submit self-readings. Each message was composed by three main elements: the email title, which was visible to all message recipients before opening the email; a banner and the email text, both visible only after opening the email.[15] Since only email titles embodied clearly the framing manipulations, while all message bodies mention every behavioural mechanism, in what follows we only discuss the former (see Table B1)

The intervention tests three framings of the campaign message. The three framings are broadly inspired by evidence in economics and behavioural sciences. The first framing invites customers to submit a meter read, so as not to risk being surprised by their bill. Since it emphasised the control over one's gas bills that could be gained by submitting a reading, we refer to this framing as the Control treatment. By emphasizing the risk of surprises, i.e. the uncertainty over one's bill, associated with not submitting a self-reading, this framing leverages individuals' aversion to uncertain outcomes. Indeed, existing evidence shows people to be generally risk averse (Arrow 1971; Pratt 1964). This version of the message had been used by the utility in its previous self-reading campaigns.

In what we refer to as the Loss treatment, the message told customers that they should submit a read, if they didn't want to pay more than they owed. This framing emphasized the potential losses from not submitting a self-reading, rather than the potential gains from submitting one. This design choice was motivated by the assumption that the desire to avoid a loss would be a stronger motivator of behaviour than the wish of achieving a gain. Such assumption is based on existing evidence on loss aversion, defined as the tendency of individuals to weigh losses more heavily than gains of the same amount (Thaler and Benartzi 2004; Tversky and Kahneman 1992; Kahneman, Knetsch, and Thaler 1986).

Finally, the third framing gave prominence to the fact that more and more customers were submitting reads, namely 2 million over the previous year. By emphasizing the increasing frequency of a certain behaviour, this treatment leveraged the influence of dynamic norms (Sparkman and Walton 2017). Existing evidence suggests that providing static social information, i.e. information on the absolute frequency of a certain behaviour, is not effective in inducing behavioural change, or may even be counter-productive, when such frequency is low (Jessoe and Rapson 2014; Wichman 2017). In cases where a behaviour is not the norm, recent studies show how providing information on the increasing frequency with which it occurs, is more effective in motivating its adoption (Abrahamse et al. 2005; Mark A Andor and Fels 2018; Brounen, Kok, and Quigley 2013; Byrne, La Nauze, and Martin 2017; Darby 1999, 2010; Fischer 2008; Hargreaves, Nye, and Burgess 2013; Ramos et al. 2015; Lynham et al. 2016; van Houwelingen and van Raaij 1989). The dynamic norms treatment thus built on this

[15] A fourth element, consisting of instructions on how to submit a self-reading, is common to all framings. Self-readings can be submitted in three ways: through the customer's private area on the utility's website, which could be reached directly from the campaign email by clicking on it; by calling the utility's customer care; or by entering the self-reading on the utility's smartphone app.

evidence and complemented standard static social information – 2 million customers have submitted self-readings- with a dynamic component – the number of customers submitting self-readings is growing.

B.2. Implementation and data

The campaign was implemented by the utility, which makes the clean identification of treatment effects challenging for a number of reasons. First, as already mentioned, the text of email messages sent to customers do not cleanly capture the different treatments, but on the contrary combine all three behavioural mechanisms. For instance, all treatment messages mention the static social information on the number of customers who submitted self-readings over the previous year. We address this issue by exploiting the fact that the campaign was sent by email to customers. Customers thus saw the email title first, and only read the message body if they decided to open the email. The email titles, contrary to the message bodies, cleanly varied with treatment, each containing only one framing. We thus test whether the title of the campaign email influences open and self-reading rates.

A second issue relates to the randomization of messages to customers. Although our partner utility claims to have randomised the treatments, we have no way to test it, as we have no data available on customers' characteristics.

Third, the test did not involve a pure control group of customers not receiving any communication. Since the previous campaign conducted by the same utility to encourage self-readings was based on the Control framing, we consider this as our control treatment, and compare the two alternative framings, which we suggested, to it. However, to give a sense of the impact of the campaign, and to reassure readers that, overall, treated subjects did not submit fewer self-reading than untreated ones, we also compare the behaviour of customers, targeted by the campaign, with that of the remaining universe of customers, who did not receive any message at the time of the campaign

The available data show whether a customer opened the email, clicked on it, and whether she submitted a reading through any channel. The campaign targeted a sample of 3,491 customers, spread over the Italian territory. Only 2,791 customers, 79.98 per cent of the sample, actually received the email. Of these, 234 opened the email and 89 submitted a read. We attribute a self-reading to the receipt of the email if it was submitted by the customer within one week from the email date. We now turn to the analysis of differences in self-reading shares by treatment.

Table B1. Pilot test: treatment messages

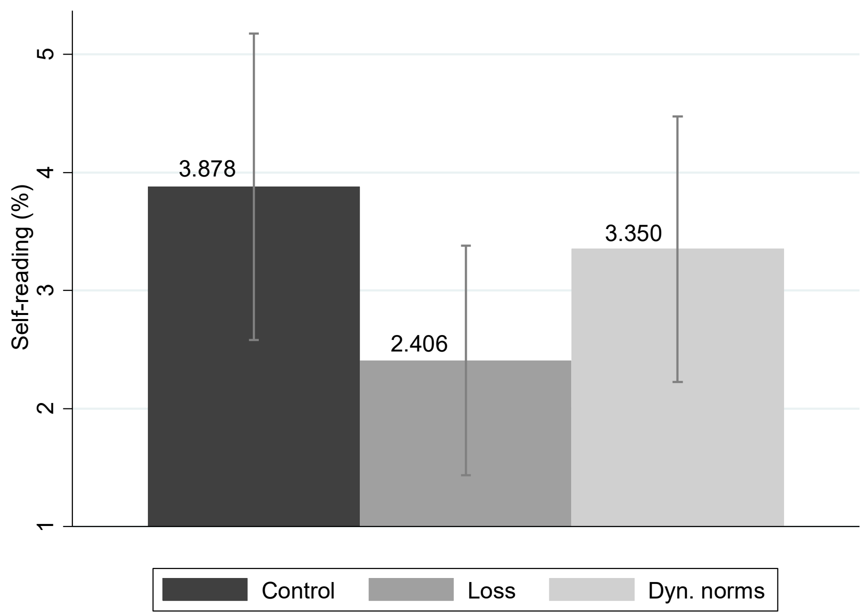
Treatment	Email title	Banner	Email Body
Control	***Do not risk bad surprises in the bill: send us you reading.***	The self-reading gas is convenient, do it now too	Hello XXX with the self-reading you will finally receive a bill without surprises, because in line with your real consumption. Over 2 million customers have already followed our advice. Communicate

			your gas consumption in a few steps.
Loss	***Gas self-reading: do not pay more than you need.***	Communicate the self-reading; to avoid paying more than you need in your bill	Hello XXX, did you know that gas reading allows you to finally have a bill based on your real consumption? Over 2 million customers have already followed this advice. It's quick and easy: what are you waiting for? Do it now too.
Dynamic norms	***Gas self-reading: more and more customers have already communicated their consumption. What are you waiting for?***	More than two million customers have already sent the self-reading gas, for a bill without surprises: do it now too!	Hello XXX, do you know that more and more customers (over the last year more than 2 million) already sent a gas self-reading? By communicating your consumption, you will no longer receive extra expenses in your bill. What are you waiting for? Do it now too.

B.3. Results

We look at the reduced-form effects of campaign treatments on self-reading, that is, at the effect of the treatment embodied in the email title on the likelihood that a customer submits a self-reading. Figure 1 shows the share of customers submitting self-readings by treatment. Overall, emphasizing risk, and how self-readings help customers keep their consumption under control, leads to a significantly higher share of self-readings than either messaging leveraging loss aversion or dynamic norms. Results from linear regressions, reported in Table A2, reveal that 3.9 per cent of customers receiving the Control message submit self-readings and that the Loss and Dynamic norms messages lower self-readings by 1.5 and .5 percentage points, respectively, a difference that is statistically significant at the 10 per cent for the Loss treatment. The campaign overall increased self-reading: average self-reading among (passive) registered customers not targeted by the campaign over the same week of April 2016 is equal to 1.58% (approximately 20,230 valid meter readings).

Figure B1. Self-reading if received email.



Notes: bars indicate means, whiskers indicate 95 per cent confidence intervals.

To summarise, the results from the pilot test suggest that avoiding the risk of surprises in one's bill is a more effective motivation to submit a self-reading than either social norms or loss aversion, significantly so in the latter case.