Energy Research & Social Science 88 (2022) 102496

Contents lists available at ScienceDirect

Energy Research & Social Science

journal homepage: www.elsevier.com/locate/erss

Double trouble: Concurrently targeting water and electricity using normative messages in the Middle East

Ukasha Ramli [a,*], Kate Laffan [b]

[a] *Department of Psychological and Behavioural Science, London School of Economics, UK*
[b] *School of Economics & Geary Institute, University College Dublin, UK*

ARTICLE INFO

Keywords:
Social norms
Pro-environmental
Eco-feedback
Water usage
Energy usage

ABSTRACT

Personalised normative messages have been shown to be effective at encouraging both electricity and separately water savings. As use of this approach to promote resource savings becomes increasingly widespread, an important question is whether providing such feedback on consumption of the two resources together can yield reductions in both areas. In a field experiment with over 200,000 households in the Middle East, we send households personalised normative messages regarding both their water and electricity consumption on a monthly basis. This intervention saw a statistically significant reduction of around 1.2% for electricity but not for water consumption. Furthermore, we test different ways of concurrently presenting normative messages of both water and energy, including presenting it as a combined eco score. Local treatment effects of these were around 1.2% reduction. Our findings contribute towards nexus thinking around how (not) to concurrently achieve energy and water savings using normative feedback.

1. Introduction

The past decade has seen the widespread diffusion of technologies that collect fine-grained, and in some cases real-time, data on consumption of two critical resources: energy and water [1,2]. Given the environmental significance of these resources, the question of how best to communicate this information to end-users, including households and businesses, in order to encourage them to reduce their consumption is important. A wealth of existing research has examined the effectiveness of feedback. Overall, this literature suggests that feedback can bring about energy and water savings but that its impact varies according to features of the feedback, such as comparison type, frequency and delivery mode [3,4].

Among the most well-evidenced forms of feedback in the environmental domain is personalised normative feedback, including both descriptive and injunctive social norms [5–11]. Descriptive norms provide target populations with information about their consumption relative to relevant others, like neighbours, with the aim of encouraging them to conserve. Injunctive social norms provide messages that communicate the perceived levels of approval or disapproval of relevant others.

A seminal paper combines these two forms of messages to target reductions in energy consumption [5]. The results indicate that while descriptive norms messages are effective at reducing consumption of energy among high consumers, they give rise to a boomerang effect among people with low energy consumption at baseline. The paper also demonstrates that the boomerang effect can be undone by adding an injunctive message which signals approval of the performance of those low energy using households. This work serves to highlight the potential of normative messages to promote conservation behaviours, as well as their potentially heterogeneous effects across populations.

Off the back of the results of initial studies in this area, there has been a proliferation of utility companies and other organisations (e.g., OPower in the US) targeting energy or water consumption using interventions involving personalised normative feedback [12,13]. The companies typically send home resource reports which include both normative feedback as well as conservation tips and other information about energy or water use. Where these efforts have been robustly evaluated, they have tended to provide further evidence of the effectiveness of such reports at encouraging resource conservation [6,9,10]. Other work has explored the effectiveness of personalised normative messages in other environmentally significant domains including

* Corresponding author at: Department of Psychological and Behavioural Science, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK.

E-mail address: u.ramli@lse.ac.uk (U. Ramli).

https://doi.org/10.1016/j.erss.2022.102496
Received 7 July 2021; Received in revised form 21 October 2021; Accepted 3 January 2022
Available online 2 February 2022

recycling [14], the uptake of green technologies [15,16] and support for carbon capture technology [11].

Existing research indicates personalised normative interventions targeting resources in one area of consumption can have knock on effects in other environmentally significant consumption areas [17,18]. For example, Jessoe et al. [17] find that sending households home water reports including normative messages induces a 1.3 to 2.2% reduction in electricity use in summer months. Further analysis indicates that just over a quarter of these reductions were attributable to the indirect reductions in electricity consumption via water savings, suggesting that the messages encouraged electricity saving behaviour as well. Even larger positive spillovers from norms based messages targeting water into energy consumption were documented by Carlsson et al. [18]. Such spillovers have important implications for the cost-effectiveness and attractiveness of home resource report based interventions [19].

As the literature on this topic has developed, researchers have begun to ask questions about how to optimise the delivery of personalised normative feedback. For example, examining whether coupling it with other interventions like commitments [20] or incentives [21,22] makes them more effective, as well as examining different delivery modes [22–24]. The research has also provided further insights into the differential effects of personalised normative messages across different groups and contexts. For example, descriptive norms based messaging is found to be far less impactful on the consumption of political conservatives compared to liberals [25] and on residential energy consumption in Germany compared to the US [26].

Both the academic and policy discourses surrounding household resource consumption are placing increasing emphasis on 'nexus thinking' [27], i.e., accounting for the linkages between environmentally significant consumption across multiple domains including water, energy, and food. Given the now widespread prevalence of personalised norm interventions in the environmental space and the evidence of spillovers between resource areas, an important question is how to combine normative messages relating to the consumption of different resources in order to yield the greatest levels of conservation across domains and maximise the messages' environmental benefits. This question is of relevance both in cases where utility companies have the opportunity to collaborate and align their normative based feedback and in situations where utility companies provide services relating to more than one resource, e.g., water and energy or energy and waste.

Against this backdrop, in the current study we carry out a large-scale field experiment with over 200,000 households to examine whether providing both energy and water based personalised normative feedback via email can achieve savings in both domains. The study was carried out in a Middle Eastern metropolis in conjunction with a state-run utility company who supplies both electricity and water to its customers. The metropolis is characterised by high levels of both electricity and water consumption making the site an interesting test case for potentially achieving substantial environmental benefits through personalised norm interventions. Looking at the effectiveness of combining personalised messages on electricity and water provides insights into the feasibility of concurrently encouraging pro-environmental behaviours using this intervention strategy.

In addition to investigating the overall effectiveness of presenting personalised norms based feedback on consumption of both electricity and water, we also examine three different ways of presenting the information: two frames which present the norms based message for each area of consumption separately either sequentially side-by side or with a shared x-axis forming a wing style format or in a combined eco-score which is computed based on households' relative performance in both areas of consumption (see Figs. 1–3). Examining the different presentations of the personalised norms based information speaks directly to the literature around how to optimise the delivery of normative messages. Existing research into the relative effectiveness of different level goals indicates that high-level goals are less effective than low-level goals in promoting energy-saving and other pro-environmental behaviours [28]. At the same time, results in the spillovers literature suggest that thinking about pro-environmental goals in concrete rather than abstract terms results in less behavioural consistency [29]. As a result, it is unclear whether normative messages in the form of a combined eco-score may be less effective than information on both resources delivered separately.

Overall, the results indicate that personalised normative feedback does yield electricity, but not water, savings when compared to a control group that did not receive this information. When we analyse the electricity consumption of those people who opened the email with the personalised normative messages, we see an annual overall reduction of 1.21%[1] over the course of a 12-month period. This is lower than reductions documented in other high consumption contexts such as the US [6], which may be attributable to context specific features or to features of the treatment frames. That there is no significant impact on water indicates that, at least in the current study context, delivering normative messages on both water and electricity does not deliver reductions in both areas of consumption. Among the potential explanations for these findings include that the combined message focuses attention on electricity at the expense of water, that the complexity of the message may undermine its effectiveness, or other contextual factors relating to water make reductions harder to achieve in this domain. Further work which compares the relative impact on normative feedback on each domain in isolation to that of the combined feedback is required to explore these potential reasons.

Disaggregating the impact across the different information frames, we find that all three frames, side-by-side, wings and the combined eco-score, yielded significant impacts on electricity but not on water. Although the greatest reductions were found in the case of wings, there were no significant differences between the treatment groups when looking at the average treatment effect. When looking at the local average treatment effect of those who opened the email, however, the wings framing did have a significantly larger impact to the other framing on electricity consumption. Taken together the results indicate that while personalised normative messages can bring about reductions in electricity consumption in the study context, targeting both electricity and water does not yield reductions in both and that the framing of the normative feedback influences its effectiveness.

In what follows, Section 2 will present the study context and data, Section 3 will present the estimation and results, and Section 4 will discuss the findings and avenues for future research.

2. Study context and data

The study was carried out in a Middle Eastern metropolis in collaboration with its nationally owned utility company that provides both energy and water to the population. To be eligible for the study, households had to meet the following criteria, 1) the customer account had to have an active account, 2) the household needed to have at least one month's worth of consumption data, 3) the household had their cooling provided by the same the utility company, 4) the customer did not have multiple accounts, 5) the household did not consume more than 40,000 kW h per day (kWh/day) of electricity and/or 200,000 imperial gallons of water. A sample of 218,737 households that met the criteria was selected for the study. The metropolis has a population of around 3 million, with an average household occupancy of 4.2, which means our sample represents around 30% of the population. The average bill amount based on the average consumption of the sample here is around USD$217.27, with electricity costing USD$136.84 and water costing USD$ 80.43. The selected households were then randomised into one out of the four experimental groups, three treatment

[1] This is based on taking the local average treatment effect of all three treatments combined (Table A2) and dividing it by the electricity consumption of the control group at baseline (Table A1).

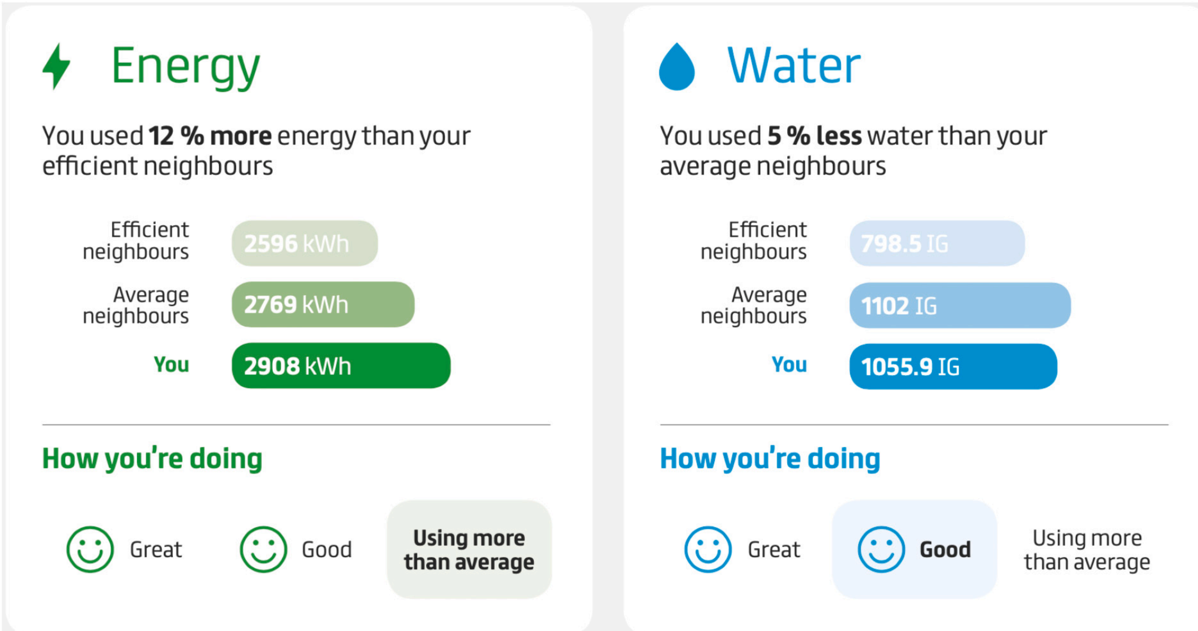

Fig. 1. Example of side-by-side treatment group message.

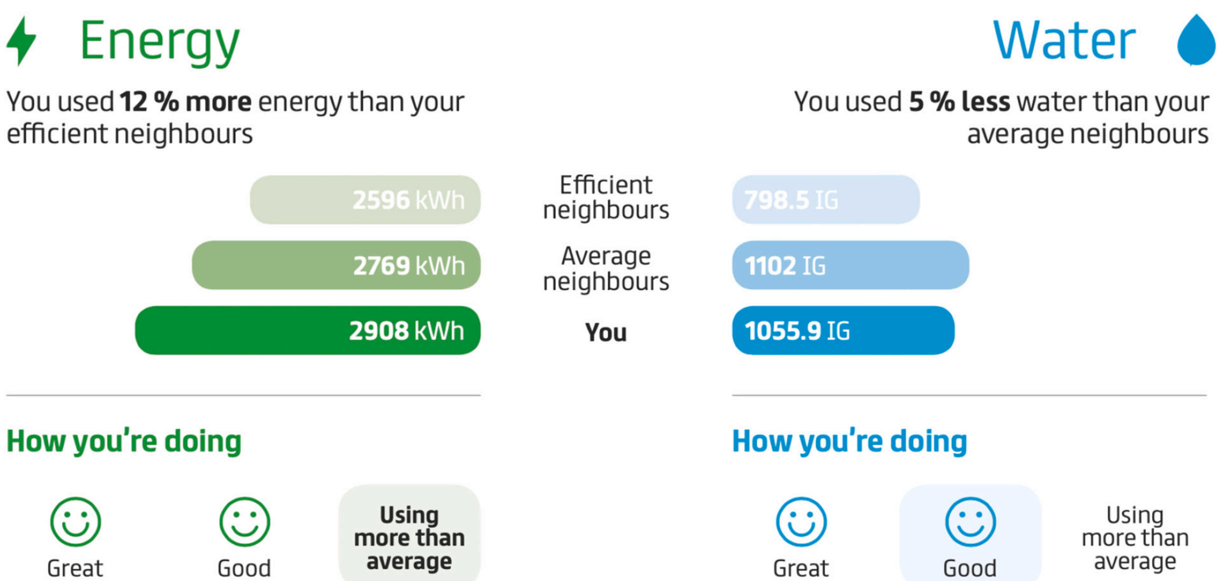

Fig. 2. Example of wings treatment group message.

groups and one control.

The experimental design closely follows the same procedure as previous similar experiments [e.g., 6]. Each household in the treatment group was sent an email on a monthly basis that included a link to a portal that contained a horizontal bar graph that compared each household's water and energy consumption against the consumption of their neighbours. Neighbours are defined as households within a geographical area with similar occupancy and house type. Email was chosen as the medium for delivering the treatment by the utility company to avoid the high cost of paper, printing and postage, as well as to reduce the environmental impact of the programme. In each treatment group, households saw a different design of the neighbour comparison bar graph, as shown in Figs. 1–3. The first treatment group saw two sets of bar graphs side-by-side with electricity on the left and water on the right. The second treatment group – wings – was a similar design but with the graphs sharing an x-axis that is adjacent to each other. The third treatment group was a single bar graph that displayed a consumption score that combined both electricity and water consumption. The score is a standardisation of the two metrics combined. Aside from the neighbour comparison bar graph to deliver the descriptive social norm, a set of smiley faces were also included to highlight an injunctive norm. This serves the purpose of preventing a 'boomerang' effect where high performing households end up licensing themselves to increase their consumption [5]. Finally, a set of water and electricity conservation tips were also included in the reports. These tips were refreshed every month with new content.

In order to assess the impact of the treatments, consumption data for both water and electricity was recorded on a monthly basis over a 12 month baseline period (March 2017 to March 2018), and for a subsequent 12 months (March 2018 to March 2019) when households in the

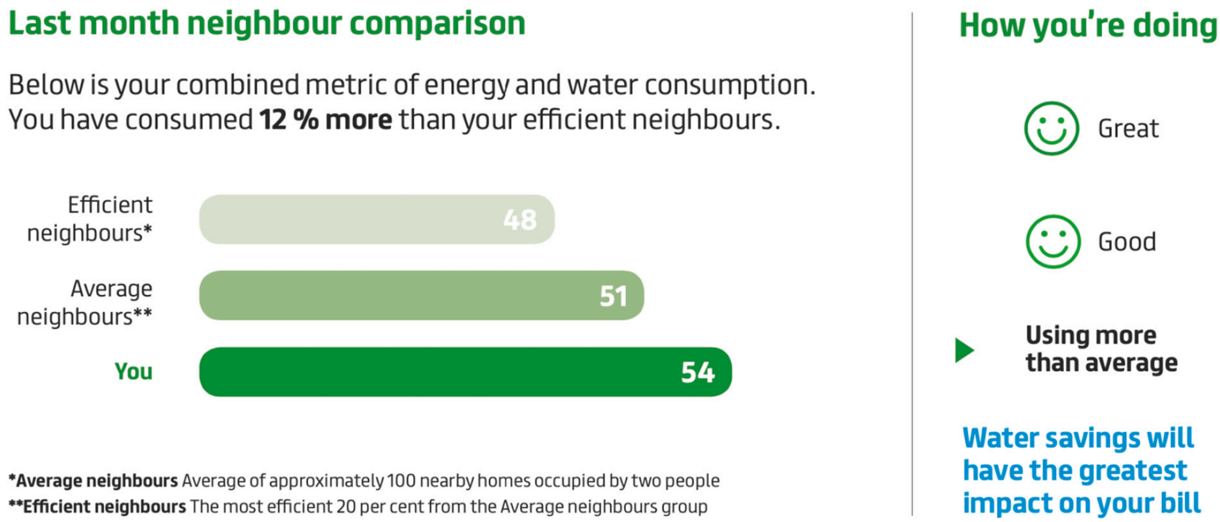

Fig. 3. Example of combined eco-score treatment group message.

treatment group received monthly emails containing the treatment. The dependent variables, water and electricity consumption are analysed as imperial gallons per day and kilowatt/hour per day respectively. 3279 of those meter readings had negative values, suggesting an error in the read, and were therefore dropped from the dataset. Out of the initial sample of 218,737 households, 50,738 households did not have the full 12 months of baseline data available. Missing baseline data was imputed by taking the closest adjacent month's consumption data. For example, if a household had data for January but not for February, then February's baseline was imputed using January's data. This method was preferred over the use of the mean, median or multiple imputation because it better preserved the seasonal differences in consumption. Missing baseline data was balanced between all groups. Furthermore, 9928 households were missing endline data and were dropped from the dataset. The final dataset consisted of 208,809 households all with positive reads and 12 months' worth of baseline data, though not necessarily all 12 reads from the study duration. In total, these households provided 2,163,393 observations across all four groups in combination. Both electricity and water consumption at baseline were balanced across all groups, and can be observed in Fig. 4 for water and Appendix Fig. A2 for water. Descriptive statistics can be seen in Appendix Table A1.

3. Estimation & results

3.1. Average treatment effect

We begin by examining the impact of being assigned to any treatment versus being in the control group – the average treatment effect (ATE). We model electricity, and separately water consumption, conditional on being treated T_i. We first estimate the simple relationship between treatment and consumption. See Model I and III in Table A2. We then include a vector of controls including house type and whether the household is local or foreign C_i, as well as month fixed effects Mi and baseline consumption B_i. See Model II and IV in Table A2).

$$Y_i = \beta_1 T_i + \beta_2 C_i + \beta_3 M_i + \beta_4 B_i + \varepsilon_i$$

We estimate these models using Ordinary Least Squares Regression (OLS) with standard errors clustered at the household level. The results indicate that assignment to treatment leads to reductions in people's electricity consumption by 0.33%[2] over the 12 month period but does not significantly impact water consumption.

We then go on to model the impact of the three different treatment groups. The model is specified as above except that T_i now indicates which of the three treatment groups or control the participants were assigned to: side-by-side, wings or eco-score. See Models I and II of Tables 1 and 2. Here we see that of the three treatments, only the wings treatment significantly reduced electricity consumption on average over the 12 month period. None of the treatments impact water consumption.

Looking longitudinally (Fig. 4), we see that the impact of the wings treatment was greatest in the initial period which coincided with the Spring/Summer months, when consumption is expectedly high due to the increased use of air conditioning, with the treatment remaining significant but decreasing later in the year. For a graph showing the longitudinal estimates for water please see Appendix Fig. A1.

3.2. Quantile regression analysis

To understand the heterogeneity of the treatment effect, we run unconditional quantile regressions for each treatment group separately. As can be observed in Fig. 5, the treatment effect on electricity consumption is heterogeneous for all three treatments, in a largely similar way. Those below the 50th percentile are not moving at all, representing a precise zero effect. Between the 50th and 90th percentiles, there is a steady increase in electricity consumption. The average treatment effect seems to be carried by those moving above the 90th percentile, although this is difficult to discern due to the noise. Overall, these quantile distributions indicate that all three treatments have very similar impacts. This suggests that the significant ATE found in wings does not necessarily reflect systematic differences between the different treatments but could simply be a statistical artifact of the sample. This explanation is further supported by pairwise *t*-tests presented in Table 1. This is an example of quantile regressions being able to better interpret ATEs as the distributional impact of the wings treatment group looks almost identical to the others. This uniformity can also be observed for water consumption, although with mostly a precise zero effect (Fig. A2).

[2] This was calculated by taking the −0.191 kWh ATE from the OLS of all three treatments combined (Table A2) and dividing it by 57.30 kWh, the average daily electricity consumption of the control group at baseline (Table A1).

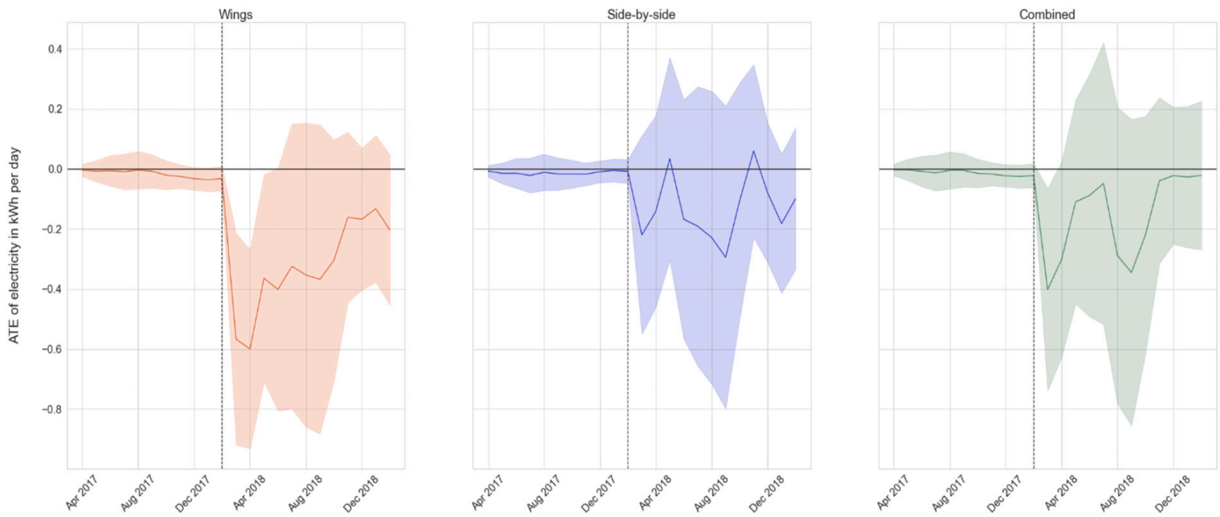

Fig. 4. Average treatment effect on electricity over time. The vertical dash line represents the start of the treatment period. The y-axis represents the coefficients are generated from an interaction between group assignment and every month in the pre and post intervention period on electricity consumption. Error bars are 95% confidence intervals.

Table 1
OLS and 2SLS analysis of electricity consumption.

	Model I	Model II	Model III
	ATE Electricity	ATE Electricity	2SLS Electricity
Side-by-side	−0.261	−0.129	−0.467***
	(0.450)	(0.104)	(0.194)
Wings	0.010	−0.302***	−1.099***
	(0.456)	(0.106)	(0.199)
Combined score	0.148	−0.146	−0.530***
	(0.458)	(0.105)	(0.197)
Side-by-side = combined	−0.409	0.011	0.103
	(0.458)	(0.105)	
Wings = combined	−0.138	−0.158	7.960***
	(0.464)	(0.107)	
Wings = side-by-side	0.271	−0.169	10.031***
	(0.456)	(0.106)	
Baseline	No	Yes	Yes
Time FE	No	Yes	Yes
House type	No	Yes	Yes
Foreign vs local	No	Yes	Yes
Number of observations	2,163,393	2,163,393	2,163,393
R2	0.000	0.904	0.904

Notes: For Model specification I and II, coefficients from pairwise *t*-test to test for equality of coefficients were included to show whether the coefficients of the treatment groups were statistically significantly different from each other. For Model specification III, coefficients from Wald test for equality of coefficients were also included.

Table 2
OLS and 2SLS analysis of water consumption.

	Model I	Model II	Model III
	ATE Water	ATE Water	2SLS Water
Side-by-side	1.063	0.641	2.3232
	(2.088)	(0.931)	(1.602)
Wings	0.0603	−0.577	−2.102
	(2.076)	(0.909)	(1.591)
Combined score	0.3528	0.069	−0.252
	(2.090)	(0.912)	(1.599)
Side-by-side = combined	0.710	0.710	2.503
	(2.126)	(0.943)	
Wings = combined	−0.413	−0.508	1.315
	(2.114)	(0.920)	
Wings = side-by-side	−1.123	−1.218	7.463***
	(2.11)	(0.939)	
Baseline	No	Yes	Yes
Month FE	No	Yes	Yes
House type	No	Yes	Yes
Foreign vs local	No	Yes	Yes
Number of observations	2,163,393	2,163,393	2,163,393
R2	0.000	0.680	0.680

Notes: For Model I and II, coefficients from pairwise *t*-test to test for equality of coefficients were included to show whether the coefficients of the treatment groups were statistically significantly different from each other. For Model specification III, coefficients from Wald test for equality of coefficients were also included.

3.3. Conditional average treatment effect

The treatment effects of these feedback interventions are rarely homogenous, and one of the main sources of heterogeneity is based on a household's baseline consumption [30,31]. We therefore look at the conditional average treatment effect (CATE), conditional on deciles of baseline consumption, to determine whether the treatment effect is greater for households with higher baseline consumption. We disaggregate average treatment effects by deciles based on a given household's mean consumption over the baseline period, and interact this with treatment group assignment within the same model specification used for determining the ATE. See Fig. 6.

Across all three CATE estimates the majority of the treatment effect occurs in the top decile with all other deciles having either very small or no effect at all. Households in the 8th and 9th deciles also seem to increase consumption, which may explain the overall small ATE for wings and the null effects for the other treatment groups. As a point of reference, the average daily consumption of households in the top decile is 251.9 kWh. The similarity in the pattern of distribution between these CATE figures and the quantile distribution is suggestive of rank invariance, which is to say that households are not swapping ranks over the distribution over time nor as a result of treatment. CATE of water consumption is available in Fig. A3.

3.4. Local average treatment effect

Importantly, being randomly allocated to be in a treatment group does not automatically ensure that individuals were exposed to the feedback on their consumption in a given month as they may not have opened the email. In fact, of the total emails sent, on average 27.51% of

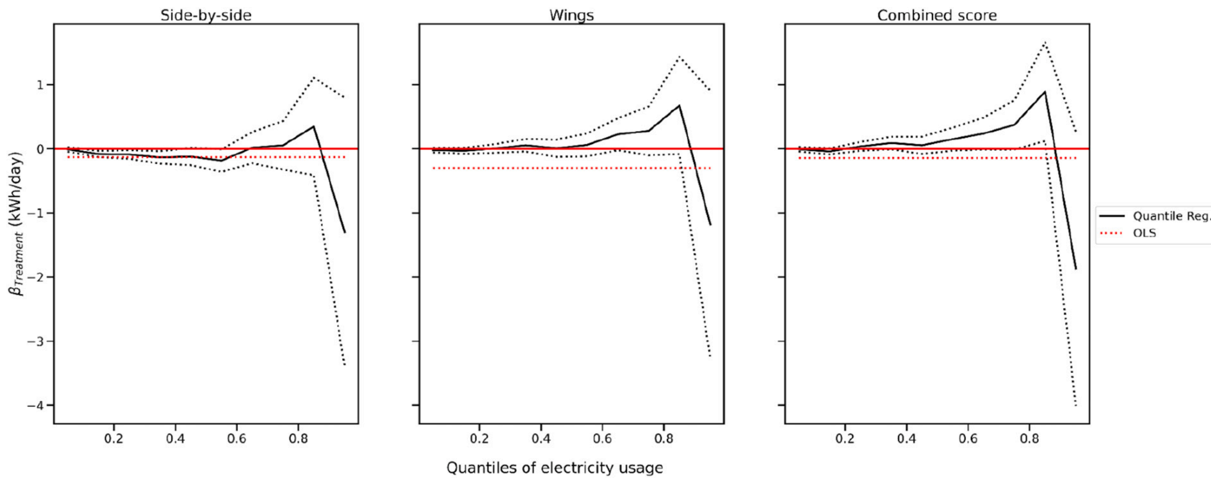

Fig. 5. Unconditional quantile regression of treatment effects on electricity. The red dashed lines represent the ATE based on Model I. The solid black line represents the treatment effect of each quantile. Therefore, no variables were included as control here. Confidence intervals are represented as black dashed lines. Higher deciles reflect higher baseline consumption. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

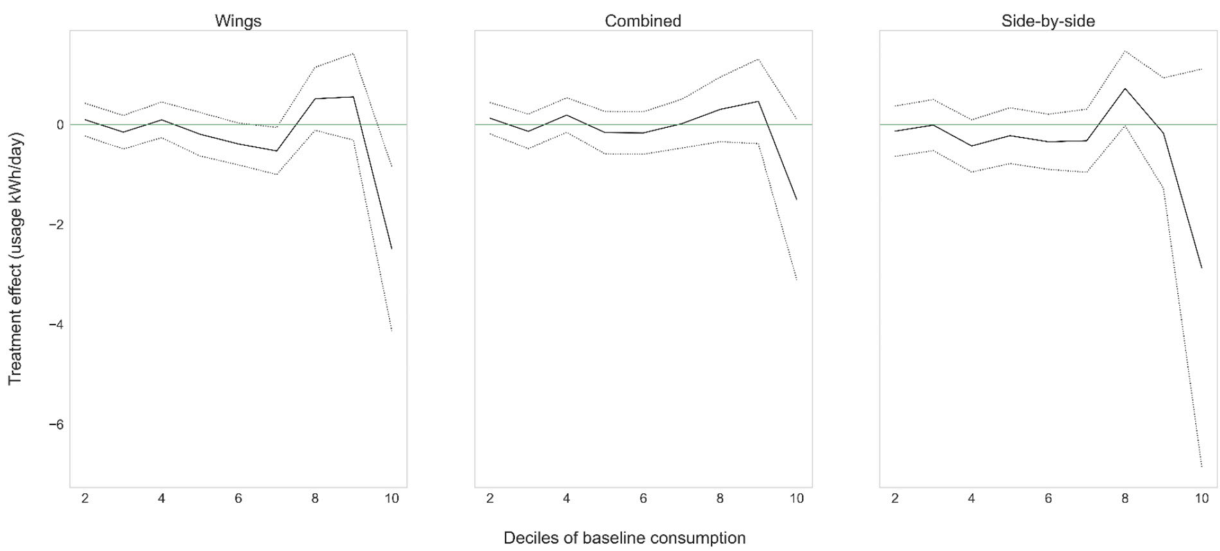

Fig. 6. Conditional average treatment effect (solid lines) based on an interaction between deciles of baseline electricity consumption and treatment group assignment. Decile 10 indicates households with the largest amount of electricity consumption during the baseline period. Decile 1 is omitted as the comparison decile. Error bars (dashed lines) are 95% confidence intervals.

treatment emails were opened per month over the course of the one year study.[3] We therefore follow up our ATE analysis by adopting a two stage least squares approach (2SLS), to estimate the local average treatment effect (LATE) of opening an email in each month [32].These models take the following form:

$$\textit{First Stage}: T_i = \rho_1 Z_i + \rho_2 C_i + \rho_3 M_i + \rho_3 M_i + \rho_4 B_i + \varepsilon_i$$

$$\textit{Second stage}: Y_i = \Gamma_1 T_i + \Gamma_2 C_i + \Gamma_3 M_i + \Gamma_4 B_i + \varepsilon_i$$

T_i is the endogenous variable, whether person i opened the email, and $\rho_1 Z_i$ is the instrumental variable of treatment group assignment. See Model III in Tables 1 and 2 for results relating to electricity and water consumption respectively. Examining the impact of the feedback on consumption for those who opened the email across the two resources, here again we see that although the treatments significantly impact electricity consumption, there is still no effect on water consumption.

Moving on to the disaggregated results, we interact treatment assignment with whether a household opened the email to generate coefficients for each treatment group. We see that while all three feedback frames reduce electricity, none have a significant impact on water (including when adjusting for testing multiple hypotheses using a Bonferroni correction). See Figs. A4 and A5 for coefficient plots. Of the three treatments the wings design yields the greatest magnitude reductions of −1.096 kWh/day (1.9% overall), followed by the combined eco-score with −0.529 kWh/day (0.911%) and side-by-side with −0.465 kWh/day (0.797%), with the difference between the impact of the wings treatment and the other two being statistically significant.

4. Discussion

In the current study, we present the first work to explore the effectiveness of concurrently presenting households with personalised normative feedback on two areas of environmentally significant

[3] Due to technical errors in the data storing of the email read receipts, 25% of the data of those read receipts were missing. To be conservative, we converted those missing values as ‘not open email’.

consumption: electricity and water. The intervention delivered a 1.9% reduction in electricity consumption over a 12-month period, when looking at the LATE using email open rate of the wing treatment, or 1.2% when looking at all treatments combined. Importantly, as participants were required to navigate to an online portal from the email to see their feedback and we are not able to monitor the frequency with which people did this, this LATE should be considered lower bounds for the impact of actually seeing the feedback and is likely an underestimate. The percentage change is similar to the treatment effects found in the US where personalised normative feedback delivered as part of paper home energy report letters achieved reductions in electricity consumption of around 2% [6,9].[4] The treatment effects found here are also expectedly higher than the 0.7% savings achieved with a similar intervention in a German context [26]. As suggested by Andor et al. [26] greater savings can be achieved in populations with higher levels of consumption. Although it is interesting to note, in our Middle Eastern context, electricity consumption is three times greater, at 90kWh/day, than that of the US, at around 30kWh/day. Furthermore, despite equivalent levels of water consumption in both the case study metropolis and the US (both of which are high by international standards), and evidence of the effectiveness of home water reports in the US [12], water consumption was unaffected by normative messages in this context.

Personalised normative messages, delivered through mail, and to a lesser extent online, have been shown to be effective at encouraging reductions in water consumption in other work [23]. In the current work, however, we find no evidence of an overall impact of the intervention on water consumption. Research on goal shielding suggests that when individuals have multiple goals they are prone to concentrate on only one goal. This is understood to be particularly likely to occur when the goals serve the same overarching purpose [33]. Although in the current design it is not possible to rule out that the null effect on water consumption is driven by contextual factors that are specific to water consumption, it may be that the intervention focused efforts on electricity at the expense of water.

Other work on the impact of dynamic pricing programs on electricity demand indicates that concurrently offering two forms of dynamic pricing is less effective than only offering one form in isolation, despite the increased incentive involved [34]. This result highlights that increasing the complexity of strategies aimed at encouraging demand reductions can backfire. Another interpretation of the findings in the current work, therefore, is that the complexity of the normative intervention targeting two resource areas may have diluted its impact. Future research should examine the differences in providing water feedback alone compared to water feedback paired with electricity to shed further light on these issues.

While existing research has documented electricity savings as knock-on effects from personalised norm interventions targeting water consumption [17,18], as far as we are aware no research to date examined the reverse, i.e., spillovers from interventions targeting electricity consumption on water consumption. While our null findings in relation to water in the current work cannot speak directly to this gap, they do raise the question of whether spillovers into electricity savings from interventions targeting water are more attainable than those arising in water from interventions targeting electricity. Given the increasing emphasis placed on the interdependencies between consumption of these resources in households – the water-energy-food nexus [27] – future work should explore the drivers of both direct and indirect reductions from personalised normative messages achieved across all of these domains when one area of consumption is targeted.

Looking at the different forms of combined feedback, we find that according to the LATE, looking at those that opened the email, the wings treatment has the largest coefficients, and is significantly different from that of the other two treatments. More specifically, the coefficient for the local average treatment effect on the wings treatment is 1.09 kWh/day, while that of the side by side treatment is 0.47 kWh/day and that on the combined score is 0.53 kWh/day. By way of comparison, the average daily reduction for wings is very close to that of the electricity an iron uses per hour at 1.08 kWh and for the side by side treatment and the combined score it is closer to half an hour.

Although existing research has found that adding embodied energy feedback to feedback on water consumption yielded significant reductions in water consumption [35], no research to date has examined the impact of presenting water and electricity personalised normative feedback together. Prior to the study, therefore, it was unclear whether combining the information or presenting it separately would best promote overall environmental performance. On the one hand, presenting the information in a combined eco-score makes salient the connection between the two areas of consumption and their relationship to environmental impact. Features which should theoretically limit negative behavioural spillovers and encourage positive ones [29]. On the other hand, presenting the feedback separately provides consumers feedback relating to concrete, rather than abstract, goals – something which should promote goal attainment [28]. That the wings score achieves larger reductions than the eco score suggests there may be benefits to separating out the feedback. However, as the eco score and the side by side treatments have equivalent impacts, the evidence is rather mixed and perhaps that these relative benefits counteract one another.

In regard to cost effectiveness of the intervention, if the average daily savings based on the ATE of the wings treatment is 0.302 kWh, which is around 9.19 kWh per month, and the cost of electricity in the metropolis is around USD$0.082 per kWh, then the savings per household per month is around USD$0.75. As the treatment itself costs around USD $0.21 per email per month, the intervention appears to be cost effective. Although true cost effectiveness is difficult to determine since households do not pay for these reports, and the benefit to the utility is difficult to calculate as it relates to the operational costs of delivering energy, as well as the subsidies provided to households.

The study design shares limitations with some previous work on normative messages targeting energy and water savings, highlighting directions for future work. First, the study does not examine the underlying behavioural processes causing resource reductions despite their relevance for both theory and practice. Future work could make use of graphical causal models and include surveying household energy- and water-saving measures in order to shed light on the behavioural pathways behind the impact on electricity consumption documented here [19,36]. A second issue is that the study focuses on water and electricity resource consumption as the outcomes of interest without providing insights into the overall consumer welfare effects. In contrast, Allcott and Kessler [37] elicit consumer willingness to pay for home energy reports in order to explore the welfare impacts on those receiving the intervention. Future work targeting both energy and water could adopt this approach to better understand the value of the intervention from the consumers' perspective.

Furthermore, in order to carry out the current study it was necessary to partner with a utility company. Such partnerships typically require flexibility and mutual benefits [38]. Compromises to the research design, in particular in relation to excluding treatment groups in which only single resources were targeted, were necessary in order to carry out the study. Including such groups would have allowed us to causally identify the spillover effect of just targeting one of the resources. This however, could not be done as it would have meant denying the service of the feedback to a portion of the population. Additionally, despite existing evidence to suggest that paper based messages are more effective than email [22,23], it was not possible in the current context to send paper letters and so we proceeded with emails. The environmental and financial cost of sending paper letters to a sample size of this size at that

[4] Importantly the ~2% documented by Allcott represents the average treatment effect. This is as a result of the study design being unable to identify whether recipients opened the home energy report letter they were mailed or not.

frequency was not seen as sustainable to the utility. Another suboptimal feature of the design was that recipients in the treatment group had to click on a link in the email in order to view the normative messages, creating further barriers to treatment. Despite these practical limitations, the intervention represented a cost-effective solution for the encouragement of electricity reduction, and they have now rolled out the intervention to their entire customer base. They have also subsequently embedded the normative messages into the body of the email that they send. This change will likely enhance the intervention's effectiveness. Future work should examine the relative effectiveness of targeting the two resources areas in isolation compared to together, and further explore the impact of different features of the delivery mode on the impact of normative messages.

Overall, the results of this study indicate that concurrently presenting personalised normative feedback on both electricity and water consumption yields reductions in the former but not the latter. As social norms based messages become more and more widespread there is an onus on utility companies to consider the effectiveness of sending normative messages targeting multiple domains. While existing studies demonstrate that personalised normative messages can be effective at encouraging water and separately electricity savings, the current results present evidence of a context in which concurrently providing feedback on electricity and water runs into some trouble: reducing consumption in one but not both domains.

Declaration of competing interest

Ukasha Ramli was employed by the tech company that delivered the intervention on behalf of the utility company. The utility company conducted the data collection of household water use through its meter vendors as part of its usual business operation. The authors worked alone in the analysis and interpretation of the data and the writing of this manuscript. The utility company and tech company opted to remain anonymous for the publication of the paper. While the two companies saw the results of the analysis, they did not see the final written work.

Kate Laffan has no conflict of interest in relation to this work.

Appendix A

Table A1
Descriptive statistics of each treatment group, subgroup in the pre and post period across water and electricity usage.

	Pre	Post	Pre	Post	n	n
	Water (Ig/day)	Water (Ig/day)	Electricity (kWh/day)	Electricity (kWh/day)	Households	Observations
Treatment group						
Control	243.01	241.75	58.09	57.23	52,986	548,174
	(400.19)	(400.45)	(91.46)	(90.48)		
Side-by-side	243.77	242.81	58.08	56.97	53,010	549,765
	(401.57)	(414.91)	(91.17)	(90.53)		
Wings	243.54	241.69	58.41	57.24	51,256	531,685
	(401.66)	(403.36)	(93.63)	(91.48)		
Combined score	243.30	242.10	58.39	57.38	51,557	533,769
	(403.05)	(409.84)	(93.46)	(92.23)		
House type						
Apartment	131.63	129.85	29.05	28.37	163,518	1,669,010
	(121.92)	(119.26)	(27.49)	(26.59)		
Villas	618.33	617.94	156.38	154.22	44,914	490,469
	(682.75)	(696.08)	(148.80)	(146.67)		
Other	922.86	1005.68	186.60	195.83	377	3914
	(996.94)	(1100.14)	(193.91)	(210.21)		
Nationality						
Local	746.72	748.12	171.66	170.58	24,621	278,554
	(800.74)	(820.32)	(178.24)	(176.06)		
Foreign	169.02	167.31	41.44	40.45	184,188	1,884,839
	(217.71)	(217.78)	(54.09)	(52.72)		

Note. Balance checks using OLS of baseline consumption and group assignment show no statistically significant differences between any of the control and treatment groups.

Table A2
OLS and 2SLS results of all treatment groups combined for water and energy.

	Model I OLS	Model II OLS	Model III OLS	Model IV OLS	Model V 2SLS	Model VI 2SLS
	Electricity (kWh/day)	Electricity (kWh/day)	Water (IG/day)	Water (IG/day)	Electricity (kWh/day)	Water (IG/day)
ATE of all treatments combined	−0.0372	−0.191	0.459	0.005	−0.695	−0.019
	(0.369)	(0.085)	(1.689)	(0.742)	(0.160)	(1.423)
Baseline	No	Yes	No	Yes	Yes	Yes
Time FE	No	Yes	No	Yes	Yes	Yes
House type	No	Yes	No	Yes	Yes	Yes
Number of observations	2,163,393	2,163,393	2,163,393	2,163,393	2,163,393	2,163,393
R2	0.000	0.904	0.000	0.680	0.904	0.680

Fig. A1. Average treatment effect on water over time. The vertical dash line represents the start of the treatment period. The y-axis represents the coefficients generated from an interaction between group assignment and every month in the pre and post intervention period on electricity consumption. Error bars are 95% confidence intervals.

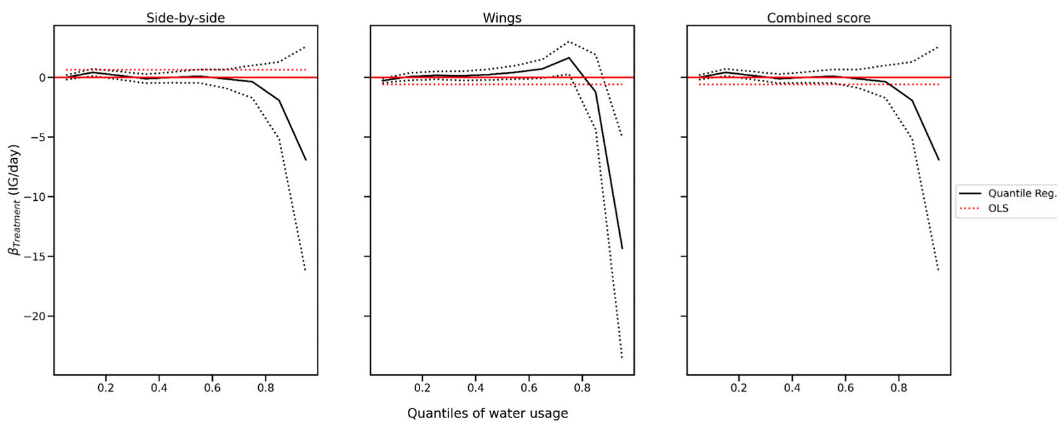

Fig. A2. Quantile treatment effects of water consumption across all three treatments. The red dashed lines represent the ATE based on Model I. The solid black line represents the treatment effect of each quantile. Confidence intervals are represented as black dashed lines. While there may seem like one of the upper quantiles is significant, the size is so small that the effect is likely to disappear once multiple corrections has been implemented.

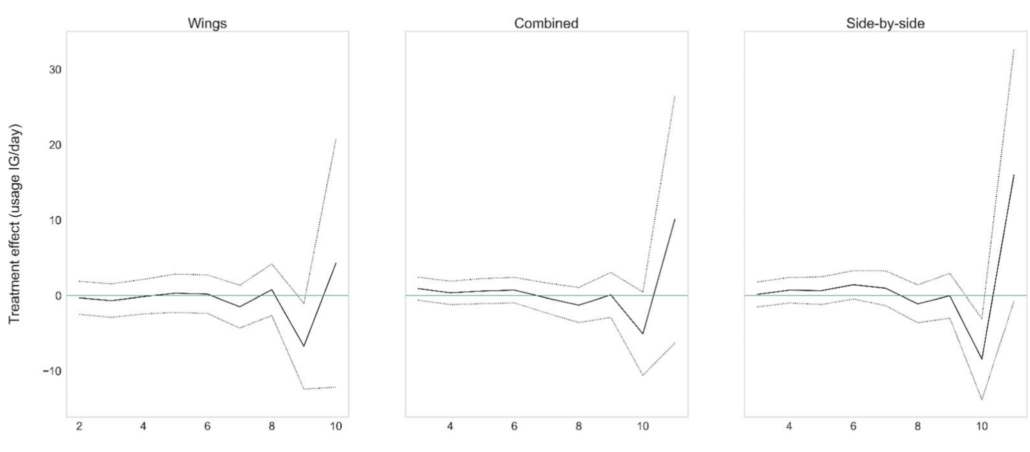

Fig. A3. Conditional average treatment effect (solid lines) based on an interaction between deciles of baseline water consumption and treatment group assignment. Decile 10 indicates households with the largest amount of electricity consumption during the baseline period. Decile 1 is omitted as the comparison decile. Error bars (dashed lines) are 95% confidence intervals.

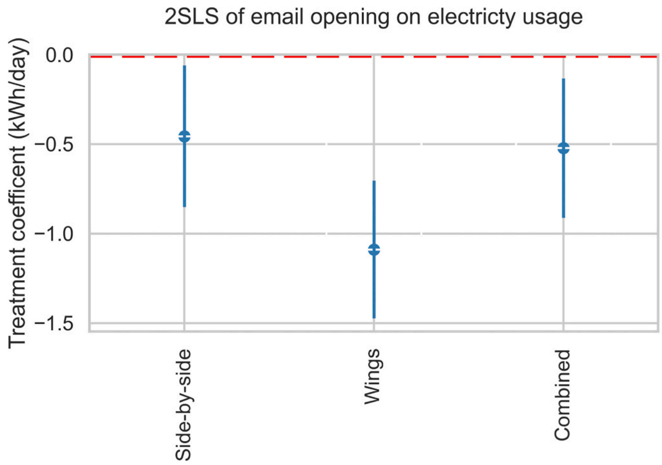

Fig. A4. Coefficient plot of electricity usage in kilowatt hours per day from 2SLS for each treatment group, using an interaction between treatment assignment as the IV and monthly email opening as the endogenous variable. Error bars are 95% confidence intervals.

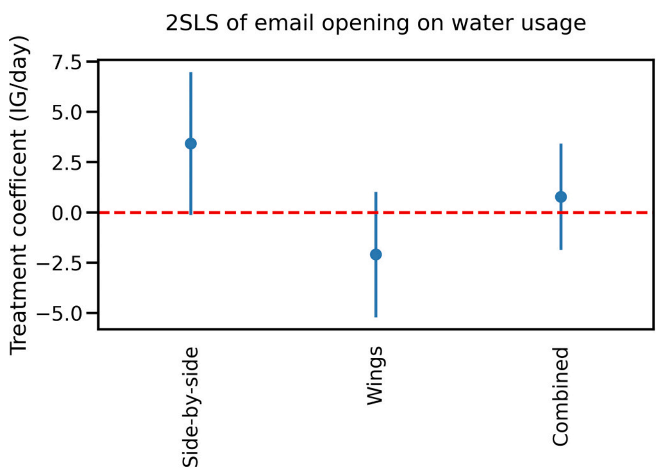

Fig. A5. Coefficient plot of electricity usage in imperial gallons per day from 2SLS for each treatment group, using an interaction of treatment assignment as the IV and monthly email opening as the endogenous variable. Error bars are 95% confidence intervals.

References

[1] D.B. Avancini, J.J. Rodrigues, S.G. Martins, R.A. Rabêlo, J. Al-Muhtadi, P. Solic, Energy meters evolution in smart grids: a review, J. Clean. Prod. 217 (2019) 702–715.

[2] M.J. Mudumbe, A.M. Abu-Mahfouz, Smart water meter system for user-centric consumption measurement, in: 2015 IEEE 13th International Conference on Industrial Informatics (INDIN), IEEE, 2015, pp. 993–998.

[3] B. Karlin, J.F. Zinger, R. Ford, The effects of feedback on energy conservation: a meta-analysis, Psychol. Bull. 141 (6) (2015) 1205.

[4] A. Liu, P. Mukheibir, Digital metering feedback and changes in water consumption–a review, Resour. Conserv. Recycl. 134 (2018) 136–148.

[5] P.W. Schultz, J.M. Nolan, R.B. Cialdini, N.J. Goldstein, V. Griskevicius, The constructive, destructive, and reconstructive power of social norms, Psychol. Sci. 18 (5) (2007) 429–434.

[6] H. Allcott, Social norms and energy conservation, J. Public Econ. 95 (9–10) (2011) 1082–1095.

[7] Ian Ayres, Sophie Raseman, Alice Shih, Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage, J. Law Econ. Organization 29 (5) (2013) 992–1022.

[8] P.J. Ferraro, M.K. Price, Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment, Rev. Econ. Stat. 95 (1) (2013) 64–73.

[9] H. Allcott, T. Rogers, The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation, Am. Econ. Rev. 104 (2014) 3003–3037.

[10] P.W. Schultz, S. Javey, A. Sorokina, Social comparison as a tool to promote residential water conservation, Front. Water 1 (2019) 1–9.

[11] T. Wang, B. Shen, C.H. Springer, J. Hou, What prevents us from taking low-carbon actions? A comprehensive review of influencing factors affecting low-carbon behaviors, Energy Res. Soc. Sci. 71 (2021), 101844.

[12] P.W. Schultz, J.M. Nolan, R.B. Cialdini, N.J. Goldstein, V. Griskevicius, The constructive, destructive, and reconstructive power of social norms: reprise, Perspect. Psychol. Sci. 13 (2) (2018) 249–254.

[13] K.S. Wolske, K.T. Gillingham, P.W. Schultz, Peer influence on household energy behaviours, Nat. Energy 5 (3) (2020) 202–212.

[14] P.W. Schultz, Changing behavior with normative feedback interventions: a field experiment on curbside recycling, Basic Appl. Soc. Psychol. 21 (1) (1999) 25–36.

[15] R. Hafner, D. Elmes, D. Read, M.P. White, Exploring the role of normative, financial and environmental information in promoting uptake of energy efficient technologies, J. Environ. Psychol. 63 (2019) 26–35.

[16] N. DellaValle, A. Zubaryeva, Can we hope for a collective shift in electric vehicle adoption? Testing salience and norm-based interventions in South Tyrol, Italy, Energy Res. Soc. Sci. 55 (2019) 46–61.

[17] Katrina Jessoe, et al., Spillovers from behavioral interventions: experimental evidence from water and energy use, J. Assoc. Environ. Resour. Econ. 8 (2) (2021) 315–346.

[18] F. Carlsson, M. Jaime, C. Villegas, Behavioral spillover effects from a social information campaign, J. Environ. Econ. Manag. 102325 (2020).

[19] A. Sanguinetti, C. McIlvennie, M. Pritoni, S. Schneider, Beyond Curtailment and Efficiency: Identifying Household Energy and Water Saving Measure Classes, 2020.

[20] C.M. Jaeger, P.W. Schultz, Coupling social norms and commitments: testing the underdetected nature of social influence, J. Environ. Psychol. 51 (2017) 199–208.

[21] J.A. Pellerano, M.K. Price, S.L. Puller, G.E. Sánchez, Do extrinsic incentives undermine social norms? Evidence from a field experiment in energy conservation, Environ. Resour. Econ. 67 (3) (2017) 413–428.

[22] P. Dolan, R. Metcalfe, Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation, Becker Friedman Institute for Research in Economics Working Paper, 2589269, 2015.

[23] P. Wesley Schultz, et al., Personalized normative feedback and the moderating role of personal norms: a field experiment to reduce residential water consumption, Environ. Behav. 48 (5) (2014) 686–710.

[24] R.J. Bator, K. Phelps, J. Tabanico, P.W. Schultz, M.L. Walton, When it is not about the money: social comparison and energy conservation among residents who do not pay for electricity, Energy Res. Soc. Sci. 56 (2019), 101198.

[25] D.L. Costa, M.E. Kahn, Energy conservation "nudges" and environmentalist ideology: evidence from a randomized residential electricity field experiment, J. Eur. Econ. Assoc. 11 (3) (2013) 680–702.

[26] M.A. Andor, A. Gerster, J. Peters, C.M. Schmidt, Social norms and energy conservation beyond the US, J. Environ. Econ. Manag. 103 (2020), 102351.

[27] M. Foden, A.L. Browne, D.M. Evans, L. Sharp, M. Watson, The water–energy–food nexus at home: new opportunities for policy interventions in household sustainability, Geogr. J. 185 (4) (2019) 406–418.

[28] L.S. Moussaoui, O. Desrichard, Act local but don't think too global: the impact of ecological goal level on behavior, J. Soc. Psychol. 156 (5) (2016) 536–552.

[29] H.B. Truelove, A.R. Carrico, E.U. Weber, K.T. Raimi, M.P. Vandenbergh, Positive and negative spillover of pro-environmental behavior: an integrative review and theoretical framework, Glob. Environ. Chang. 29 (2014) 127–138.

[30] D.A. Brent, C. Lott, M. Taylor, J. Cook, K. Rollins, S. Stoddard, What causes heterogenous response to social comparison messages for water conservation? Environ. Resour. Econ. 77 (2020) 503–537.

[31] P.J. Ferraro, J.J. Miranda, Heterogeneous treatment effects and mechanisms in information-based environmental policies: evidence from a large-scale field experiment, Resour. Energy Econ. 35 (2013) 356–379.

[32] J.D. Angrist, G.W. Imbens, D.B. Rubin, Identification of causal effects using instrumental variables, J. Am. Stat. Assoc. 91 (434) (1996) 444–455.

[33] J.Y. Shah, R. Friedman, A.W. Kruglanski, Forgetting all else: on the antecedents and consequences of goal shielding, J. Pers. Soc. Psychol. 83 (6) (2002) 1261.

[34] Jacobsen, G., & Stewart, J. I. (2020). How Do Consumers Respond to Price Complexity? Experimental Evidence from the Power Sector. *Experimental Evidence from the Power Sector (December 1, 2020)*.

[35] S.H. Jeong, R. Gulbinas, R.K. Jain, J.E. Taylor, The impact of combined water and energy consumption eco-feedback on conservation, Energy Build. 80 (2014) 114–119.

[36] N. Bhushan, L. Steg, C. Albers, Studying the effects of intervention programmes on household energy saving behaviours using graphical causal models, Energy Res. Soc. Sci. 45 (2018) 75–80.

[37] H. Allcott, J.B. Kessler, The welfare effects of nudges: a case study of energy use social comparisons, Am. Econ. J. Appl. Econ. 11 (1) (2019) 236–276.

[38] I.B. Müftüoglu, S. Knudsen, R.F. Dale, O. Eiken, D. Rajak, S. Lange, Rethinking access: key methodological challenges in studying energy companies, Energy Res. Soc. Sci. 45 (2018) 250–257.