unibz

Fakultät für Wirtschaftswissenschaften
Facoltà di Economia
Faculty of Economics and Management

BEMPS –
Bozen Economics & Management
Paper Series

NO 80/ 2021

Technology vs information to promote conservation: Evidence from water audits

Erik Ansink, Carmine Ornaghi, Mirco Tonin

Technology vs information to promote conservation: Evidence from water audits

Erik Ansink* Carmine Ornaghi** and Mirco Tonin***

February 2021

Abstract

We study the impact of audits on water conservation, distinguishing between the information and technological components. We observe water consumption for up to 18 months for 10,000 households in the South East of England who received the visit of a so-called Green Doctor. We find that water-saving devices decrease water consumption by 2-4%, with an effect that is persistent over 18 months. Devices reducing water pressure are particularly effective, while shower timers are ineffective. The information component of the water audit has a large initial impact, but this gradually fades to a drop in consumption of 2% after 12 months. Technology appears to be more cost-effective than information provision and this can help in the design of policy interventions.

Keywords: Water audits; Green Doctors; conservation, information; technology.

JEL codes: D12; H42; L95; Q25

* Erik Ansink, School of Business and Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam (the Netherlands). erik.ansink@vu.nl

** Carmine Ornaghi, School of Economics, Social and Political Sciences, University of Southampton, University Road, SO17 1BJ, Southampton (U.K.). C.Ornaghi@soton.ac.uk

*** Mirco Tonin, Faculty of Economics and Management, Free University of Bozen-Bolzano, piazza Università, 1, 39100, Bolzano (Italy); CESifo; IZA. mirco.tonin@unibz.it

We gratefully acknowledge financial support by the Economic and Social Research Council, UK [ES/K01210X/1]. We would like to thank participants at various conferences and seminars for helpful comments. We also thank Carlos Pineda for excellent research assistance, Ben Coles from Groundwork UK for valuable insights into the Green Doctor programme and Southern Water for providing the data.

1. INTRODUCTION

Non-pecuniary strategies to induce behavioural changes, including nudging and social comparisons, have received wide interest in recent years since they are believed to be cost-effective and relatively uncontroversial as they do not impose a price on "bad" behaviour. A notable disadvantage of non-pecuniary strategies is that their effects may be short-lived as compared with the effect of market-based policies (d'Adda et al, 2017).

In this paper we use data from a large water audit programme in the South East of England to assess the effectiveness and persistence of the following two specific nonpecuniary components of the programme: (1) information provision on current water use, potential water use savings, and comparison of water use to comparable households and (2) retrofitting existing devices including the issuing or installation of water-saving devices such as low-flow shower heads. We will refer to these two components as, respectively, information and technology (cf. Ferraro and Price, 2013).

Water audit programmes are quite common in parts of the US and Australia, and have become more popular in several European countries, including England, over the past two decades. The water audit programme that we evaluate, the Green Doctor (GD) programme, was implemented in the South East of England in 2010-2015 and was targeted at households with above-average water use.

There is ample reason to believe that the effects of both information and technology are not persistent. Starting with information, studies on the effectiveness of information provision via Home Energy Reports and Home Water Reports suggests that such reports have an instant effect that subsequently wanes over time (cf. Allcott & Rogers, 2014; Brent et al, 2015). For water, the effect of such reports is limited, inducing conservation of up to 5% on household water use (cf. Bernedo et al, 2020; Jessoe et al, 2020; Kažukauskas et al, 2020; West et al, 2020). It is not clear whether this result carries over to an audit program where, instead of receiving information by mail or email, the information and social comparison is conveyed in a face-to-face meeting. Like information, technology effects may also not be persistent. Álpizar et al (2020) summarize a list of examples from recent programs, including bednets, fluorescent light bulbs, and cook stoves, where technologies that generate positive externalities (as well as internalities) are subject to dis-adoption, even after substantial experience with the good in question. Anecdotally, after former US President Donald Trump expressed his unhappiness with the shower flow due to the dire implications for his hair, the U.S. Energy Department eased standards on shower heads.[1] To this effect we can add the possibility of rebound effects, which occur when the adoption of a more efficient technology – such as low-flow shower heads – leads to increased use, offsetting its potential benefits (Campbell et al, 2004; Olmstead & Stavins, 2009; Millock and Nauges, 2010). As a result, both information and technology effects may have limited persistence.

Our working hypothesis is that the technology component of the water audits is more likely to persist over time because the impact of information on water use requires behavioural change that is costly in terms of effort, whereas most water-saving devices, such as tap aerators and save-a-flush bags, save water mechanically without requiring constant attention or particular effort. Note that shower timers represent an exception, for they just convey information and their effectiveness relies on users' willingness to act upon that information. An interesting study by Tiefenbeck et al (2018) finds that such timers can have a substantial effect on both water and energy use. However, their study looks only at the short-term effect over the following 2 months. Our data allows us to analyse the effectiveness of shower timers, as well as several other types of water-saving devices over a period of up to 18 months. Furthermore, by comparing the relative effectiveness of information vs technology, our findings can help to design more effective policy interventions.

There are only few studies on the effectiveness of water audits, mostly covering US programmes. For instance, Nelson (1992) finds a 5% reduction in water use as a result of a 90-minutes audit amongst customers of the North Marin Water District (California, USA) in 1988. This audit included both the information and the technology component: installation of water-saving devices, identification of the most effective lawn irrigation schedule, and customized recommendations to save water. Similar audit programmes have been assessed by a.o. Bruvold and Mitchell (1993), Sarac et al (2003), Keen et al (2010), and Tsai et al (2011). Compared with these early programmes, the GD programme is much larger. Also, we control for unobserved variables that could affect water use using fixed

[1] https://www.reuters.com/article/us-usa-energy-efficiency-idUSKBN28P2XZ

effects models. In addition, since we have information on the number and type of water-saving devices installed or issued, we can separate the impact of the audits' information component from its technology component, i.e. the impact of the water-saving devices. Finally, while most water audit programs are offered continuously to all customers of a particular water utility, the GD program specifically targets a sub-sample of households with above-average water use. Since conservation programmes tend to have the largest effects among high-use households (Ferraro and Miranda, 2013; Wichman et al, 2016; Brent et al, 2020), we may expect sizable effects of the GD programme.[2]

Our results show that the effect of information is initially high at 39-46 litres/day but this effect wanes to a stable 10 litres/day after one year. This effect size is well within the range of earlier studies that assessed the effectiveness of Home Water Reports. The effect of technology persists at 10-20 litres/day, based on an average of 2 devices per household. Relative to baseline household water use of almost 500 litres per day, the information component results in an initial 8-10% water use decrease that wanes to 2% after one year, while the technology component accounts for a stable 2-4% decrease. The stability of the technology effect points to absence of dis-adoption or rebound effects over time. Depending on the number of devices issued or installed per household, technology is arguably more effective than information in this water audit program targeting households with above-average water use. One caveat is that this result does not account for spillover effects on other domains. Recent studies find that water conservation brings about important spillovers to energy conservation, partly via mechanical complementarities, but mostly via behavioral change (Carlsson et al, 2020; Jessoe et al, 2020). A back-of-the-envelope calculation (see Section 3 for details) suggests that the information component is twice as expensive as the technological component in reducing water use.

The paper proceeds as follows. In Section 2 we introduce the Green Doctors programme and our empirical model. In Section 3 we present our results. Conclusions are presented in Section 4.

[2] The only recent assessments of water audits' effectiveness in the economics literature are by-catch in studies by Brent et al (2015) and Browne et al (2021). Both use very specific samples. Brent et al (2015) due to self-selection into the programme and Browne et al (2021) due to the water audit being offered as replacing a fine for first-time perpetrators of outdoor water use restrictions.

2. DATA and EMPIRICAL MODEL

From the Autumn 2010 until Spring 2015 Southern Water installed more than 400,000 meters across its supply area in the South East of England, as part of the first large scale Universal Metering Programme (UMP) in the UK. The Green Doctor (GD) programme was carried out in parallel with the installation of meters and it consisted in the performance of water and energy audits by trained advisors – GDs hired by the charity Groundwork[3], including the provision of water-saving devices and the offer of advice on how to be efficient with water and cut household bills. The GD programme was targeted at households with above-average water use, who could therefore see large increases in their water bill due to metering. According to industry sources, when offered, households were generally well inclined towards receiving a visit. Indeed, once the major hurdle of getting in contact with someone in the household is cleared, in the vast majority of cases a visit is booked. As a result of this initiative, the company carried out more than 50,000 home visits and more than 165,000 water-saving devices - such as water-efficient showerheads and tap aerators - were provided into some 46,000 properties (Ofwat, 2015, page 23). Although GD visits were prevalently geared towards households that had a meter installed for the first time as part of the UMP, around 15% of the total visits referred to households that had already a meter installed. As explained below, to improve comparability our empirical analysis is limited to UMP customers only.

We have information about the number and type of water-saving devices issued or installed for around 24,000 households that are offered a water audit during the period Autumn 2010 - Summer 2014 (i.e., one year before the end of the UMP programme). Hereafter, we refer to these households as GDH, mnemonics for Green Doctor Households.

The GDH we use for the empirical analysis consists of *UMP households* that have a meter installed at least one month before the Green Doctor visit and for whom we can observe *monthly* water consumption for *at least* 12 months after the visit.[4] While half-yearly data

[3] https://www.groundwork.org.uk/projects/green-doctor/

[4] Households that have already a metered installed (i.e., no-UMP households), are not part of our sample, even if they received a GD visit, because the vast majority of these households have an old meter which, differently from the new meters installed during the UMP, do not automatically record consumption at the end of the month. Accordingly, monthly data for no-UMP customers that receive a GD visits are available

corresponding to the typical billing cycle are available for all the customers of SW, water consumption at the higher *monthly* frequency requires the construction of balanced panel starting from raw data that are very unbalanced and have a lot of missing observations. More information about the construction of this balanced *monthly* panel can be found in a companion paper by two of the authors (Ornaghi and Tonin, 2019), which investigates the impact of metering on water consumption and its implications for efficiency and equality.

Here we note the following two things. First, the use of *monthly* data (instead of the half-yearly data) is necessary *(a)* to identify the effects of Green Doctor visits separately from the reduction in water consumption due to meter installation and the ensuing change from unmetered to metered tariff and *(b)* to evaluate the dynamics triggered by the GD in the months following the visit. Second, while GDH are not randomly selected among all the newly metered customers, the set of customers for which we can observe higher-frequency *monthly* data is completely orthogonal to the customers' characteristics or consumption dynamics (see Ornaghi and Tonin, 2019). Furthermore, we restrict the attention to the set of households for whom we have at least twelve data points after a visit because one of the aims of our analysis is to investigate whether the effects of GD visits have persistent effects on water consumption. Out of the initial 24,000 GDH in our records, the final sample for whom we can observe *monthly* data for at least twelve months after the visit consists of 9,496 households.[5]

The median (average, resp.) duration of the water audits for these households is 40 minutes (41 minutes) and more than 90% of the audits last between 30 and 60 minutes. Although we do not have any information on the time necessary to arrange a visit or to reach the customers at their properties, we can safely assume that average auditors' time for each visit is well above one hour. As said, we also observe whether the GD issued or installed one or more of the following water-saving devices: "save-a-flush" bags, tap aerators, shower heads, shower regulators and shower timers. Table 1 below reports

only for a small number of customers (those that had a new meter installed because the old one stopped working).

[5] Comparing descriptive statistics for all available GDHs to the statistics for the sample of GDHs used in the regression analysis reported in Table 2, we find a similar average number of occupants (2.76 vs 2.88) and rateable values of the house (171 vs 172), but a lower daily water consumption at baseline (250 vs 472).

descriptive statistics for these water-saving devices. Given that tap aerators, shower heads and shower regulators are all used to manage the pressure and flow of water, we put them together in Table 1 under the label "Water flow and pressure".

The median and average number of devices issued or installed are respectively 2 and 1.95, with on average 0.61 "save-a-flush" bags and water flow and pressure devices per household and 0.74 shower timers. More specifically, around 16 percent of the households did not receive any device, 26 percent received one device, 29 percent received two devices, 15 percent received three devices, 9 percent received four devices while 5 percent received between five and eight devices.

Table 1. Statistics for Water-saving Devices.

Devices	Mean	S.D.	Median	Min	Max
All types	1.95	1.53	2	0	8
"Save-a-Flush" Bag	0.61	0.78	0	0	4
Water Flow and Pressure	0.61	0.94	0	0	4
Shower Timer	0.74	0.65	1	0	3

We use this information to estimate the effects of information on reducing water consumption vis-a-vis the water-saving effect of technological devices. Recall from the introduction that our working hypothesis is that the effects of devices are more likely to persist over time since their effect is in most cases mechanical, while information requires behavioral change.

GDH are subject to two different treatments: first the installation of a meter and then the visit of a Green Doctor. To identify the effects of the latter net of the former, we compare the dynamics of consumption of GDH (treated group) to those households that also have received a meter but not a GD visit (control group). Our identification strategy rests then on the assumption that this control group can mimic what the consumption of GDH would have been in the absence of the visit. The selection of the control group proceeds as follows. For each GDH, we first choose all never-treated UMP households (i.e., households that never receive a visit by a Green Doctor) with identical number of occupants and decile of house rateable values (RV)[6]. Then, among all these counterfactual households, we select the one with the closest level of water consumption in the pre-switch period, i.e., the very first observation available in the dataset (notice that this is always before the GD visit).[7] Matching on the same decile of RV ensures that the households had similar water bills before meter installation, since the unmetered tariff consists of a standing charge, fixed for all properties, and a rateable value charge, based on the RV of the house. Matching on the same number of occupants (and, as a second step, on water consumption in the pre-switch period) ensures that the treated and control group should have a similar level of water usage. However, we expect Green Doctor households to use somewhat more water at baseline for they are targeted because of their (known or assumed) higher level of consumption.

Table 2. Statistics for Treated and Control Group.

Sample	# Households	Occupants	RV	Litres at Baseline
Control - 12 months	9496	2.886	171.813	472.353
Treated - 12 months	9496	2.877	170.748	496.160
Mean Diff (p-values)		0.610	0.163	<0.001
Control - 18 months	4972	2.898	170.849	465.815
Treated - 18 months	4972	2.889	169.468	483.729
Mean Diff (p-values)		0.708	0.187	0.008

Table 2 shows that our matching procedure performs well as the number of households' occupants and the RV are not statistically different between treated and control groups. As

[6] The rateable value was an indicator of the rental value of the house as of 31 March 1990.

[7] We obtain very similar results with an alternative matching procedure, when for each GDH we choose first never-treated UMP households by matching number of occupants and decile of house rateable values (RV) as well as postcode and calendar quarter of meter installation. Then, if multiple counterfactual households are found, we select the one with the highest number of observations. With this procedure, for around 5% of observations for which there are no matches at all, we move to a second round of matching where we match on semester of installation instead of quarter.

expected, the baseline consumption of treated households (recorded in the first month of meter installation, when they have not received the visit of a Green Doctor yet), is higher.

The empirical model we use to quantify the impact of GD visits consists in a matched (treated-control) DiD of the following form:

$$y_{it} = \sum_{k=1}^{N} \beta_k^{GD} GD_{i,t-k} + \sum_{k=1}^{N} \beta_k^{WD} WD_{i,t-k} + \sum_{j=2}^{36} \gamma_j^I I_{i,t-j} + \mu_i + \tau_t + u_{it} \quad (1)$$

where y_{it} is water consumption of households i at time t; GD_{it} is a one/zero indicator for GD visits that took place at time t-k, i.e., k months in the past relative to time t; WD_{it} is a variable taking values between 0 and 8 depending on the number of water-saving devices that were installed or issued to the households at the time of the visit (8 being the maximum number observed). $I_{i,t-j}$ is an indicator for meter installation j months in the past relative to time t that captures the dynamics of consumption after a meter is installed. Finally, μ_i are households' fixed effects and τ_t monthly fixed effects.

The coefficients of interests are the β_k^{GD} and β_k^{WD}: they will measure the reduction in water usage for the GDH from month 1 to month N after the visit, with N=12 or N=18. If the effects of the visits and devices are persistent over time, we would expect the βs not only to be negative and significant, but also to be stable over time. Two clarifications about equation (1) are in order. First, as we study the effects of water audits over a period of one year and a longer period of 18 months, our specification includes either twelve or eighteen *GD* and *WD* indicators. Although our data allows us to track a longer period, the number of households for which we have such observations drops sharply after 18 months. Second, as we observe households for a maximum of three years from the moment a meter is installed, equation (1) comprises thirty-five *I*-indicators, not including the very first month after the meter is installed, which is our baseline consumption.

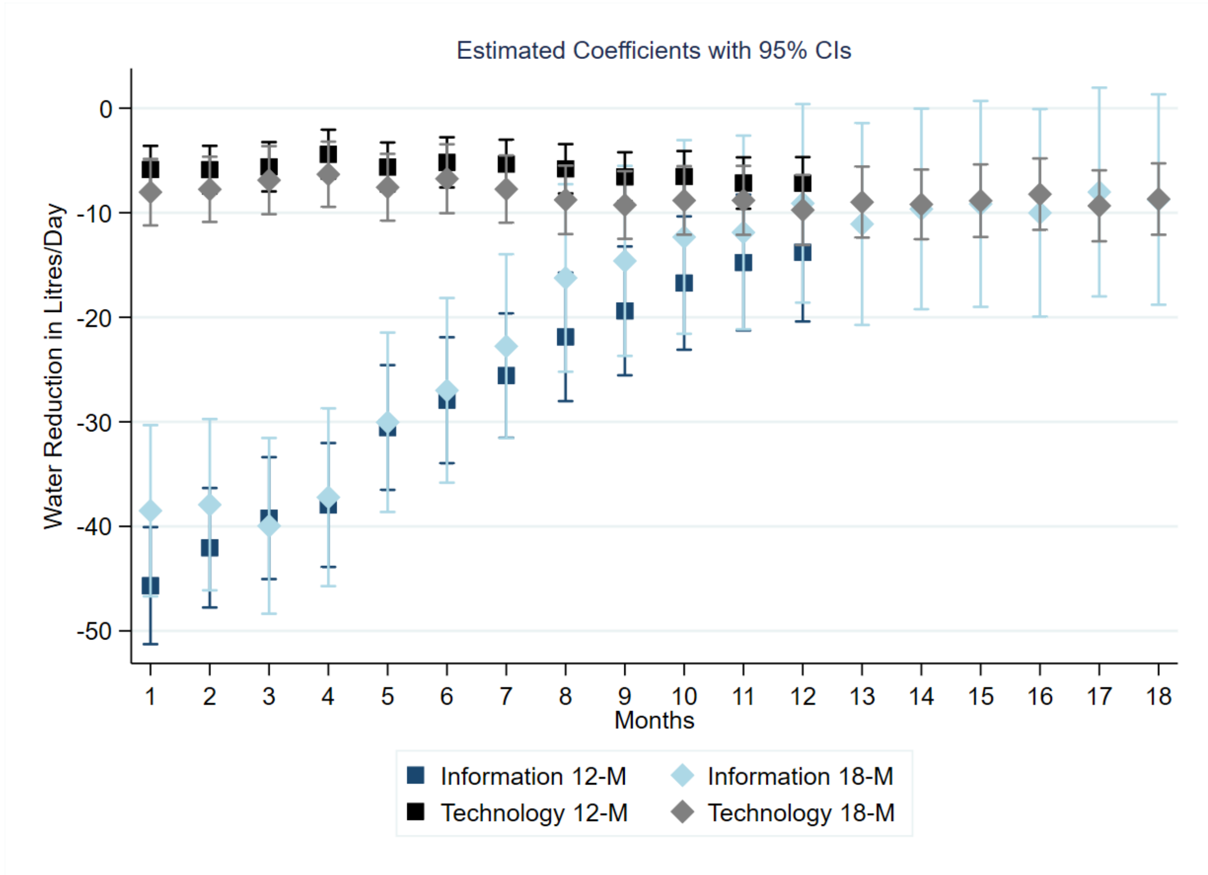
Before concluding this section, we note that our identification strategy rests solely on differences in water usage over time between treated and never-treated units. An alternative identification strategy would have been to use GDH with later visits as control for GDH with early visits. The recent paper by Goodman-Bacon (2018), however, discusses the identification problems that arise when using the timing of treatment to identify treatment effects. In particular, the author shows that the DiD estimator with staggered treatment timing is a weighted average of many different treatment effects, with groups treated in the middle of the panel receiving higher weights than those treated earlier and later. By matching each treated household with a similar but never treated subject, our identification strategy does not suffer from problems associated with DiD models with staggered timing.

3. RESULTS

Figure 1 shows the values of the β coefficients from equation (1), distinguishing between the impact of the technology and information component of the GD program. The technology impact is measured via the effect of water-saving devices (β_k^{WD} in the equation, black and grey lines in the figure) and the information impact is measured by the remaining effect of visits by Green Doctors (β_k^{GD} in the equation, dark blue and light blue lines in the figure), thus including the impact of information about current water use, potential water use savings, and comparison of water use to comparable households. The 12- and 18-months samples behave in a very similar way in the period in which they overlap. So, all in all, we can consider the dynamics displayed by the longer sample as applying overall.

Two patterns emerge clearly from Figure 1. First, we find that one additional device is associated with a reduction in consumption of 5-10 litres/day. Considering that on average households had 2 devices installed, this translates into an overall reduction due to devices of 10-20 litres per day (equivalent to 2-4%, given a baseline consumption from Table 2 that is almost 500 litre per day). The fact that this reduction remains rather stable over the 18-month time-window suggests that there is no rebound effect nor a rejection of technology over time, due for instance to the dis-adoption of shower heads or regulators. Interestingly, when we re-estimate the model distinguishing between devices installed and just issued, we obtain point estimates that are very similar for both the 12M and 18M time-window. This suggests that issued devices are actually installed.

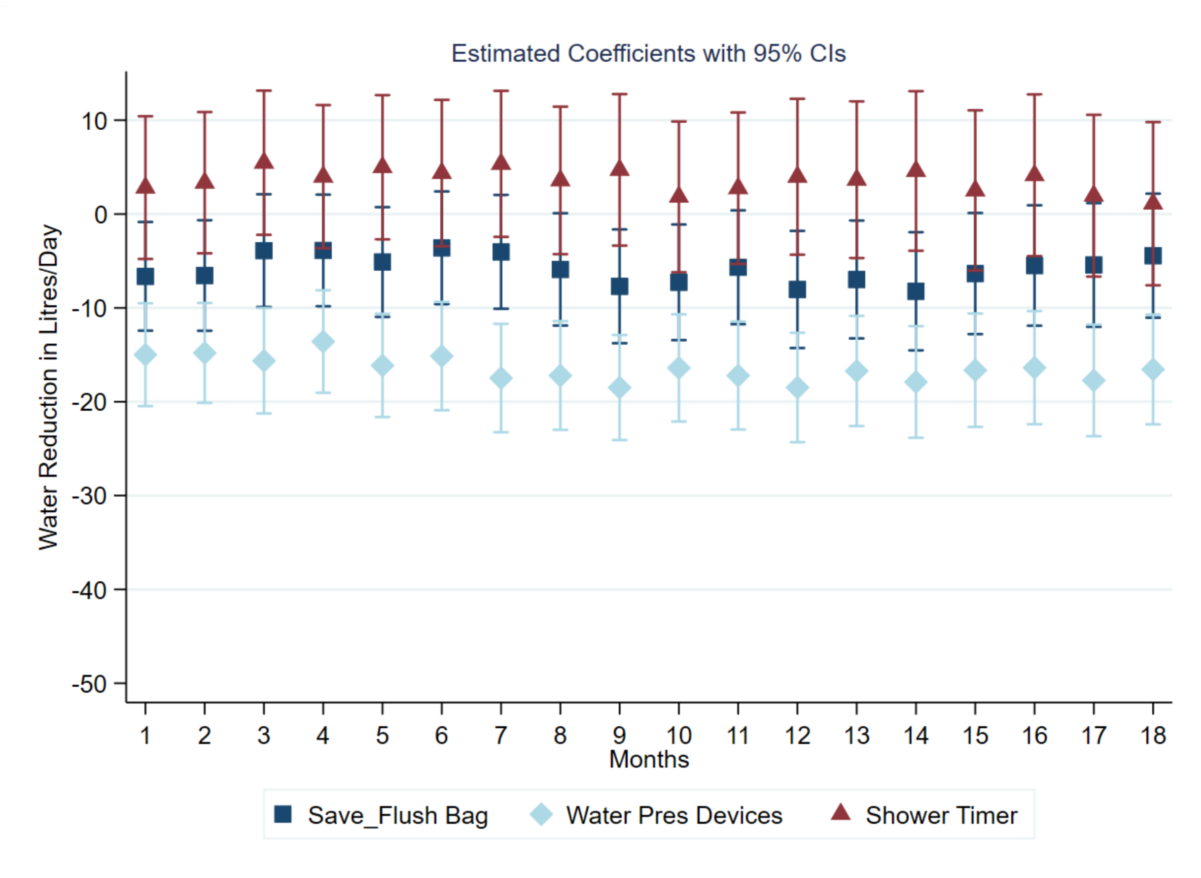
Figure 1. Impact of information (dark and light blue) vs impact of technology (black and grey)



Second, the impact of the information component of the GD program shows instead a fading path, starting with a drop of around 39-46 litres/day (8-10%) that stabilizes to a level of around 10 litres/day (2%) after one year. Thus, the information component weakens its impact over time. The large standard errors around point estimates suggests that there is a large heterogeneity in the way households react to the information component of water audits. Indeed, Figure 1 shows that, starting from month 12, the 95% confidence intervals around the point estimates often include zero (i.e. we cannot reject the null hypothesis that GD has no effect on water consumption in several months starting from the 12th).

Although we do not find evidence of rebound effects or dis-adoption of the water-saving devices, note that we lump the three types of devices together in our analysis, which may conceal differentiated effects of each. In order to have a better understanding of the relative effectiveness and persistence of the three types of devices, we extend our empirical model to assess them separately and estimate equation (1) with three different categories of devices. Figure 2 shows that devices that regulate the flow of water (i.e. shower heads, shower regulators and tap aerators) are the most effective in reducing usage, with an average reduction of around 18 litres/day per device, followed by "save-a-flush" bags that are responsible for a reduction of around 8 litres/day per bag. Interestingly, shower timers are associated with an increase in water usage, although none of the coefficients is statistically significant. Shower timers, differently from the other devices, only provide information about water usage but they require attention and effort on the side of the customers to act upon such information. In this respect, shower timers are more similar to water audits to the extent that they require a costly change in habits in order to reduce water consumption. It is then not surprising that there is no effect of shower timers on water usage and, if anything, they are associated with a small increase in consumption. Instead, devices that save water mechanically, that is, by the simple fact of being installed without requiring further involvement by people, are effective and persistent.

Figure 2. Impact of Three Types of Devices



As a last step, we can combine the obtained results on water savings with estimated costs of the information and technology components of the GD programme to estimate the variable costs per unit of water saved. Starting with the information component, we find that the long-run reduction of 2% is obtained via a GD visit that lasts more than an hour when including the logistics of arranging appointments and travelling to different households, plus the time waste implied by last-minute cancellations.[8] Taking 1.5 hours at an estimated cost of £20/hr[9], the costs of the information component are roughly equal to £15.00 per 1 percentage-point water savings. The costs of the technology component largely consist of the costs of the devices, which is close to £9.00 per device.[10] With 2 devices per household, assuming £4.50 delivery costs, and an estimated water use reduction of 2-4% per household, this comes down to approximately £7.50 per 1 percentage-point water savings. This back-of-the-envelope calculation - ignoring possible interaction effects - suggests that the information component is twice as expensive compared to the technology component in achieving water use reduction. These are of course just approximate calculations, but can be informative about the relative cost-effectiveness of technology vs. information.

4. CONCLUSIONS

Our analysis shows that water audits offered via the GD programme are effective in reducing water consumption and this effect persists over the relatively long time period we consider. The technology component has a stable and persistent effect, and we find no evidence of a rebound effect or dis-adoption of water-saving devices. The information component is also instrumental in inducing households to save water, but in this case there is a clear fading of the impact over time, with a strong initial drop and a gradual convergence to a more modest reduction. The fact that this reduction is still present after 18 months and appears to stabilize suggests that information is successful in triggering

[8] These are estimated to be around 15% of visits (personal communication).
[9] This is consistent with an average hourly labor cost of around £17.5 per hour in the service sector in the period under consideration, plus some further costs related to transportation.
[10] This is based on current retail prices and is thus an upper bound to the per unit price associated with a bulk purchase.

some change in consumption habits. Comparing both components, we find that technology is both more persistent as well as more effective, both in terms of water saved per household and in terms of costs per unit of water saved.

Our results suggest that conservation programs such as water audits may benefit from redirecting their attention from information towards technological solutions. There are two caveats, however. One is that our results are obtained using a non-representative sample of households, i.e., those with above-average water use. This is an interesting population to study, in particular considering that audits can generally be targeted. In any case, earlier studies point out that water saving programs are most effective for high users (Ferraro and Price, 2013; Brent et al, 2015) and, as a result, we expect that our results would provide an upper bound of possible water conservation using audits directed toward the general population. The second caveat is that the GD programme combines information and technology so there may be interaction effects between those two components that we cannot extract from our data. Given the mechanical effect of the water-saving devices, however, we expect that such interaction effects are negligible. As a result, an alternative programme that would only distribute water-saving devices may achieve reductions in water consumption at lower cost. Programmes like Green Doctors are currently being implemented in other areas of England, such as Smarter Home Visits in London and in the Thames Valley region, and available on an ongoing basis to Southern Water customers. We expect that our results can help to inform the details of such programs and design more effective policy interventions in general.

References

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

Alpízar, F., Del Carpio, M. B., Ferraro, P. J., & Meiselman, B. S. (2020). Exposure enhanced goods and technology disadoption. Working paper.

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

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

Brent, D. A., Lott, C., Taylor, M., Cook, J., Rollins, K., & Stoddard, S. (2020). What causes heterogeneous responses to social comparison messages for water conservation?. *Environmental and Resource Economics*, *77*(3), 503-537.

Browne, O. R., Gazze, L., & Greenstone, M. (2021). Do conservation policies work? Evidence from residential water use. *Environmental and Energy Policy and the Economy, 2*, 190-225.

Bruvold, W. H., & Mitchell, P. R. (1993). Evaluating the effect of residential water audits. *Journal of the American Water Works Association*, *85*(8), 79-84.

Campbell, H. E., Johnson, R. M., & Larson, E. H. (2004). Prices, devices, people, or rules: The relative effectiveness of policy instruments in water conservation. *Review of Policy Research*, *21*(5), 637-662.

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

d'Adda, G., Capraro, V., & Tavoni, M. (2017). Push, don't nudge: Behavioral spillovers and policy instruments. *Economics Letters*, *154*, 92-95.

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

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

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. NBER Working Paper No. 25018.

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

Kažukauskas, A., Broberg, T., & Jaraite, J. (2020). Social comparisons in real time: A field experiment of residential electricity and water use. *Scandinavian Journal of Economics*, forthcoming.

Keen, A. H., Keen, D., Francis, G. E., & Wolff, A. (2010). High-contact, hands-on outreach program changes customers' water use behavior. *Journal of the American Water Works Association*, *102*(2), 38-45.

Millock, K., & Nauges, C. (2010). Household adoption of water-efficient equipment: The role of socio-economic factors, environmental attitudes and policy. *Environmental and Resource Economics*, *46*(4), 539-565.

Nelson, J. O. (1992). Water audit encourages residents to reduce consumption. *Journal-American Water Works Association*, *84*(10), 59-64.

Ofwat (2015), Affordability and debt 2014-15, ISBN 978-1-910739-24-2, https://www.ofwat.gov.uk/publication/affordability-and-debt-2014-15/

Olmstead, S. M., & Stavins, R. N. (2009). Comparing price and nonprice approaches to urban water conservation. *Water Resources Research, 45*(4), W04301.

Ornaghi, C. and Tonin, M. (2019). The effects of the universal metering programme on water consumption, welfare and equity. Oxford Economic Papers, forthcoming. https://doi.org/10.1093/oep/gpz068

Southern Water (2015). Water Efficiency Plan - Saving customers water, energy and money https://www.southernwater.co.uk/Media/Default/PDFs/Water-Efficiency_Plan.pdf

Sarac, K., Day, D., & White, S. (2003). What are we saving anyway? The results of three water demand management programs in NSW, Australia. Water Supply, 3(3), 215-222.

Tsai, Y., Cohen, S., & Vogel, R. M. (2011). The impacts of water conservation strategies on water use: Four case studies. Journal of the American Water Resources Association, 47(4), 687-701.

West, J., Fairlie, R. W., Pratt, B., & Rose, L. (2020). Automated enforcement of irrigation regulations and social pressure for water conservation. Working paper.

Wichman, C. J., Taylor, L. O., & Von Haefen, R. H. (2016). Conservation policies: Who responds to price and who responds to prescription? *Journal of Environmental Economics and Management*, *79*, 114-134.

***Appendix* - Table A**. Coefficient estimates of Figure 1

	12-Month		18-Month	
Month	GD	Devices	GD	Devices
M1	-45.673***	-5.863***	-38.505***	-8.033***
	(2.87)	(1.16)	(4.20)	(1.63)
M2	-42.044***	-5.879***	-37.927***	-7.760***
	(2.93)	(1.17)	(4.20)	(1.60)
M3	-39.212***	-5.603***	-39.950***	-6.884***
	(2.99)	(1.21)	(4.31)	(1.67)
M4	-37.958***	-4.401***	-37.217***	-6.320***
	(3.04)	(1.20)	(4.36)	(1.60)
M5	-30.546***	-5.621***	-30.042***	-7.562***
	(3.06)	(1.20)	(4.40)	(1.64)
M6	-27.933***	-5.182***	-26.987***	-6.749***
	(3.09)	(1.23)	(4.53)	(1.69)
M7	-25.575***	-5.333***	-22.767***	-7.737***
	(3.05)	(1.19)	(4.51)	(1.65)
M8	-21.876***	-5.796***	-16.236***	-8.767***
	(3.15)	(1.21)	(4.60)	(1.68)
M9	-19.386***	-6.559***	-14.605***	-9.261***
	(3.16)	(1.20)	(4.66)	(1.66)
M10	-16.725***	-6.535***	-12.319***	-8.820***
	(3.27)	(1.25)	(4.75)	(1.68)
M11	-14.761***	-7.159***	-11.883**	-8.813***
	(3.33)	(1.26)	(4.75)	(1.69)
M12	-13.787***	-7.177***	-9.100*	-9.738***
	(3.39)	(1.28)	(4.87)	(1.72)
M13			-11.079**	-8.980***
			(4.95)	(1.74)
M14			-9.634*	-9.198***
			(4.92)	(1.71)
M15			-9.155*	-8.851***
			(5.05)	(1.78)
M16			-10.009**	-8.221***
			(5.09)	(1.75)
M17			-8.019	-9.335***
			(5.12)	(1.74)
M18			-8.735*	-8.687***
			(5.16)	(1.75)
R-squared	0.118		0.100	
N	367997		223115	

Note: The specification includes also household fixed effects, a complete set of monthly dummies as well as thirty-six dummies indicating how many months have passed since a meter has been installed (eg..*j-th* dummy will take a value of 1 if a metered was installed *j* months ago with respect to *t*, with *j*=1, …, 36) to capture the dynamics of consumption after a meter is installed.