Future options for managing customer demand for water

White paper prepared for Yorkshire Water by London Economics

January 2018

About London Economics

London Economics is one of Europe's leading specialist economics and policy consultancies. Based in London and with offices and associate offices in five other European capitals, we advise an international client base throughout Europe and beyond on economic and financial analysis, litigation support, policy development and evaluation, business strategy, and regulatory and competition policy.

Our consultants are highly-qualified economists who apply a wide range of analytical tools to tackle complex problems across the business and policy spheres. Our approach combines the use of economic theory and sophisticated quantitative methods, including the latest insights from behavioural economics, with practical know-how ranging from commonly used market research tools to advanced experimental methods at the frontier of applied social science.

We are committed to providing customer service to world-class standards and take pride in our clients' success. For more information, please visit www.londoneconomics.co.uk.

Head Office: Somerset House, New Wing, Strand, London, WC2R 1LA, United Kingdom.

w: londoneconomics.co.uk e: info@londoneconomics.co.uk : @LondonEconomics
t: +44 (0)20 3701 7700 f: +44 (0)20 3701 7701

Acknowledgements

We would like to acknowledge the useful guidance and feedback provided by the Steering Committee throughout this research. The Committee also provided us with much of the data that was employed in our analysis. We would like to thank the following stakeholders that the project team had met with over the course of the project for the valuable advice and information they have provided… We would also like to thank all of the stakeholders surveyed for their informative questionnaire responses. The research also benefited at a number of stages from suggestions by …. A.N. Other also provided helpful feedback on some of the methodological issues at an early stage in the research. Responsibility for the contents of this report remains with London Economics.

Authors

Dr Annette Cerulli-Harms

James Suter

Ryan Perkins

Table of Contents

1 Introduction

Scientific evidence suggests that due to climate change severe weather events like the heatwave of summer 2018 are likely to become more frequent and intense, **increasing the risk of droughts**. Climate change models predict that in the UK there will be an increased likelihood of hotter, drier summers and wetter, milder winters (Environment Agency, 2017). While water supplies are now more resilient to periods of dry weather than in the past, a larger population and higher water use per capita places increased strain on the system. Hence, the Environment Agency (2017) noted that authorities and the sector **must be more careful to manage demand especially in summer months** to ensure water restrictions do not need to be applied too frequently.

Managing demand is important for customers and water companies alike. For companies, under the regulatory framework the performance measures on which they are judged include reducing water consumption, increasing resilience, and overall customer satisfaction. For example, at the last price review Yorkshire Water committed to ensuring that customers always have enough water, with associated performance commitments to reduce water use and leakage and ensure the stability and reliability of the water network.[1] From the customer's perspective water restrictions during periods of drought have the potential to cause disruption and lead to misunderstanding and anger.

Temporary use bans (TUBs) are currently the default measure to be put in place by water companies when water supplies fall below certain levels. However, the body of available evidence on the effectiveness of TUBs has not yet, to our knowledge, been collated and reviewed.

Objectives of this white paper

Therefore, the objectives of this white paper are to:

- assess the effectiveness of TUBs from existing evidence by means of summarising and critically appraising this evidence (see section 2);
- assess what alternatives to TUBs are available for water companies and other authorities to manage water demand (see section 3) with lessons to be drawn particularly from behavioural economics; and
- provide an overview of methods for piloting possible interventions to contribute to the evidence-base and find out what works best in given circumstances (see section 4).

These objectives have been addressed via a rapid evidence assessment (REA) as described in further detail in the following subsection.

1.1 Approach

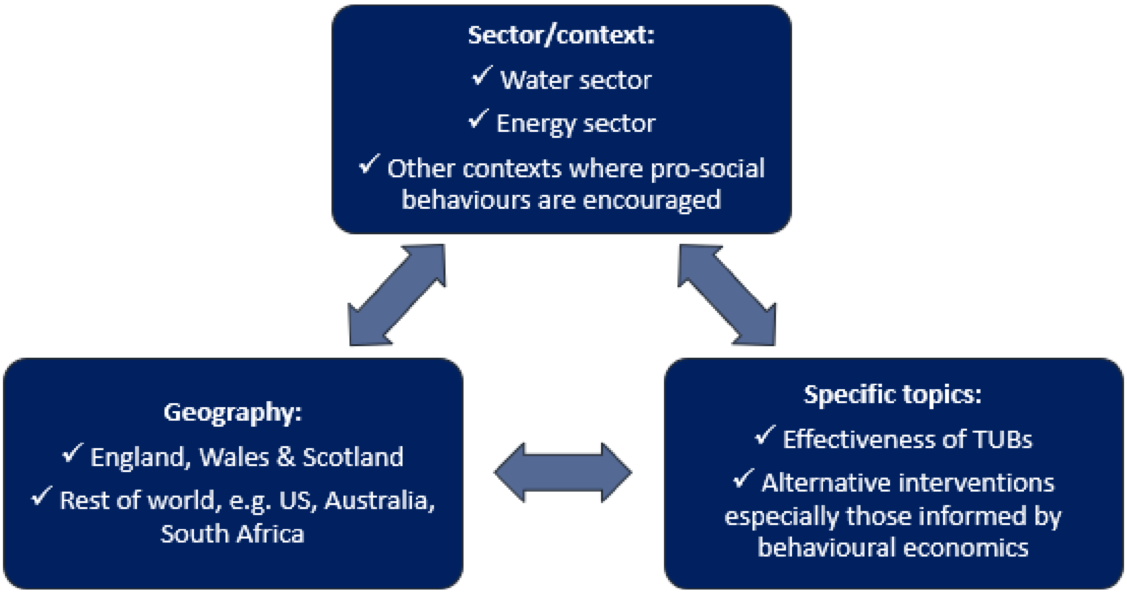
The evidence presented in this white paper was collected using a Rapid Evidence Assessment (REA). Whilst conducting the assessment, we looked at a variety of sources such as academic papers, government agency reports and research conducted by utility providers. This was to evaluate both the effectiveness of TUBs and the existence of alternative interventions used to encourage resource

[1] See Ofwat's Final price control determination notice company-specific appendix for Yorkshire Water, December 2014.

conservation by consumers in the water sector. Alongside literature related to the water sector we also looked at interventions utilised within the energy sector.

The research put a focus on sources from the UK, but we also expanded the range of the research geographically to capture the effect of water restrictions and alternative interventions used in other countries such as Australia, the USA, Israel, South Africa and multiple European countries. When examining the alternative interventions we evaluated multiple types of interventions such as social norm messaging, pricing strategies and feedback provision to consumers.

Figure 1 Dimensions of the Rapid Evidence Assessment presented in this white paper



Source: London Economics

The dimensions of **context**, **geography**, **source type** and **intervention type** describe the focus of our research strategy. Information on the reviewed papers was catalogued in a database and for each source we recorded the source type, a description of the intervention, evidence on the effectiveness of the intervention, and the methodology and robustness of the results. This Excel-based database made it easy to see an overview of the available evidence as well as to identify potential gaps in the literature. This helped to target our research throughout the process to ensure evidence collected was sufficient to support our conclusions.

2 Temporary use bans

Temporary use bans (TUBs) are currently the default measure to be put in place by water companies when water supplies fall below certain levels. While restricting the ways in which water can be used seems like a logical way of regulating water consumption, especially during droughts, there has been relatively little research into how effective TUBs actually are when they are put in place.[2] Therefore, this section seeks to collate the existing evidence on the effectiveness of TUBs that have been implemented in the UK and internationally, as well as evidence on consumer reactions to TUBs.

[2] The overall lack of evidence is also due to the fact that TUBs are not frequently implemented.

Box 1 Key findings

- **General evidence on effectiveness of TUBs**:
  - **Evidence** on the effectiveness of temporary use bans (TUBs) **is relatively scarce**. There have been two recent instances of TUBs being implemented in the UK, in 2006 and 2012, which have been the subject of analysis.
  - The **quality and robustness** of some available assessments are **questionable**, or at least highly circumstantial.
  - The most reliable sources come from the Environment Agency (2013) and UKWIR (2007), who analysed the impacts of TUBs implemented in 2006 and 2012 and found overall reductions **ranging from 1% (i.e. negligible) to 9%**:
    - The Environment Agency (2013) found that **the 2012 TUBs had no measurable impact on water consumption**. This was likely because the weather during the TUB was very wet and therefore required low compliance.
    - The **2006 TUBs significantly impacted consumption levels** according to both sources. According to UKWIR (2007) in the absence of wet weather, as seen in 2012, TUBs would be expected to reduce summer water use by 5-9%. Yet the authors caveat that these effects were likely circumstantial and might not translate to other TUBs.
  - International evidence from the USA and Australia shows that **voluntary restrictions are unlikely to be effective**. Mandatory water restrictions reduced consumption by 4-23% and were most effective when combined with financial incentives. It is however unclear how much these effects translate to the UK as the environments are very different.
- **Consumer reactions and attitudes towards TUBs:**
  - **Consumers widely seem to misunderstand TUBs**.
  - Most consumers are **unable to pinpoint which types of water uses are restricted** exactly.
  - Qualitative evidence suggests that **consumers would broadly accept TUBs especially in severe weather conditions**. However, this **acceptance falls drastically** when consumers feel that water levels have fallen to critical levels because of **mismanagement**, or other possibly avoidable causes such as leakages.
  - Multiple sources, from the UK and internationally, suggest that **consumers prefer TUBs to seeing their bills increase**. This is because they feel that restrictions imposed by TUBs are often not unsurmountable, or regard non-necessary types of usage only.
  - Academic research has shown that **the strongest driver of compliance with water restrictions is an individual's perception of their capacity and ability to comply**.

2.1 Effectiveness

The rapid evidence assessment conducted for this white paper found that the **evidence on the effectiveness of TUBs is relatively scarce**. The evidence presented here is based on only six publications.[3] Moreover, the quality and robustness of some of the available assessments are

[3] The sources of these papers include UK Water Industry Research (UKWIR), the Environment Agency (EA); and a handful of academic papers.

questionable, or at least highly circumstantial. Some sources could not be assessed, as they were inaccessible or improperly cited.[4]

The most reliable estimates of the impact of TUBs in the UK come from the Environment Agency's (EA) 2013 paper discussed below. This study found that the **TUBs in 2012 had an impact of just 1-2% which is not statistically significant** (it is within the confidence intervals of the study's model), which may be explained by the fact that there was very **wet weather during the period of these TUBs**. The **2006 TUBs, in contrast, did have a statistically significant effect on consumption**, but the EA report does not provide any percentage figures for the size of this impact.

UKWIR (2007) used a similar econometric approach to the EA (2013) to analyse the impact of the 2006 TUBs and found results that are in line with those of the EA. Based on the reductions observed in 2006, UKWIR found that, **in the absence of wet weather as seen in 2012, TUBs would be expected to reduce summer water use by 5-9%**.

The studies from overseas, specifically the USA and Australia, found larger reductions in water usage due to use restrictions.[5] **Demand reductions range from 4% to 23% in Australia and 4% to 20% in the USA**. However, differences in the context of the restrictions in other countries and differences in the restrictions themselves compared to the UK mean that the applicability of these results to the UK is limited, especially compared to the UK-specific analyses mentioned above.

Below we discuss the relevant studies in more detail starting with studies relating to TUBs applied in the UK, followed by studies concerning water use restrictions in other countries.

Studies relating to the impact of TUBs in the UK in 2006 and 2012

The EA (2013) used an econometric approach to estimate the impacts of TUBs imposed in 2006 and 2012, using data from all six companies that implemented restrictions during 2012. The researchers first constructed a forecasting model to estimate[6] *expected water usage without any restrictions* based on meteorological factors (e.g. temperature). They then compared this forecast to actual recorded usage during the period of the TUBs. The model was informed by a prior feasibility study and a short review of possible approaches drawing on examples from the UK and overseas. The study is thorough and robust, especially given the limitations imposed by using historic datasets as opposed to experimental or trial data.[7]

While there was a decrease in all regions **during the 2012 TUB, the estimated impact of the water use restrictions was not statistically significant**[8] with an average weekly reduction in use over the TUB period of around 1-2% overall. The report notes that the lack of a statistically significant impact in most areas may be due to the short period of time that TUBs were in place in 2012 as a result of exceptionally wet weather that summer.

[4] For example, the 2017 Thames Water drought plan attributed water reduction figures to an earlier report which could not be traced. The report is referenced as 'Thames Water/R&D report/RST010-LTOADrought2012/Version 0.8, L. Kiernan, 27-03-2012'

[5] See Mini et al. (2014), Castledine et al. (2014) and Haque et al. (2014), all discussed below.

[6] Using Generalised Least Squares, since the variances of the observations were expected to be unequal (i.e. heteroscedastic) and the explanatory variables are likely to be correlated (i.e. multicollinear).

[7] A potential limitation (which is the case for any econometric analysis) is that there may be some drivers that are not included in the model leading to a bias in the forecast values (omitted variable bias). For more detail on possible methods for trialling different types of water management interventions, see section 4.

[8] Except for two water resource zones, Sutton and Lincoln.

In contrast, the EA found that **the reduction in water use during the 2006 restrictions was statistically significant for all regions**, although the results for the 2006 TUBs are not translated into percentage figures.[9] This is in line with the results of an earlier analysis by UKWIR (2007) of the impact of the 2006 hosepipe bans, which used a similar econometric approach to the EA's analysis with a similar set of explanatory variables.[10] UKWIR's analysis suggests that in the absence of wet weather (as experienced in 2012) **TUBs would be expected to reduce summer water use by around 5-9%**, depending on the area in question.

Specifically, UKWIR (2007) analysed the effects of TUBs implemented in six water resource zones in 2006, with separate analysis for different types of restrictions ranging from sprinkler bans to non-essential use bans. UKWIR found that unattended hosepipe and sprinkler bans[11] resulted in average reductions between 4.4% and 4.9%, full hosepipe bans resulted in average reductions between 4.5% and 9.3%[12] and a non-essential use ban resulted in an estimated reduction of 18.6%.

However, **UKWIR stress a number of caveats to these results**, in particular:

- these effects need to be seen in their specific contexts, as each drought is likely to be associated with a unique pattern of events and these effects would not necessarily be the same in future droughts;
- these effects are water resource zone (WRZ) specific and responses to these restrictions are likely to differ between WRZs; and
- the estimated effects of the restrictions include the effects of other associated factors and activities such as media coverage, and thus the observed impact of a restrictions needs to be seen as a function of all activity preceding it and cannot be wholly attributed to the restriction itself.

In a later study UKWIR (2013) also analysed the impacts of TUBs implemented by three English companies in 2012.[13] For two companies (Thames Water and Anglian Water) the study examines the number of 'hosepipe events'[14] between 3am and 4am (to measure sprinkler use) and 6pm and 11pm (to measure evening gardening watering), while for the third company (South East Water) the paper examines monthly averages for household consumption (since this is the only data available). UKWIR's approach is to compare water use during the TUB period to water use during the summers before and after the restrictions, without controlling for other factors that differ between these periods, meaning a proper counterfactual is not used. Hence, although some quite substantial

[9] The report does present estimated coefficients from the econometric model which indicate the average impact on usage in megalitres per day by water resource zone, with values ranging from 1.77ML/day for Kent Thanet to 83.3ML/day for London. Although, of course these WRZs vary greatly in size and usual daily consumption. Note that, since values these are model coefficients, this is not the same as saying that Thames Water actually saved 83.3ML/day in the London WRZ as a result of the 2006 restrictions.

[10] The UKWIR (2007) analysis used multiple linear regression to model the impact of restrictions and included temperature in their model (like the EA) but (unlike the EA) not potential evapotranspiration and soil moisture deficit.

[11] Unattended hosepipe bans are less stringent TUBs which essentially refer to sprinkler use and 'automatic' hosepipes which are usually used to water gardens. An 'unattended hosepipe' is often used as a synonym for sprinkler in this context.

[12] The estimated average summer restriction effects for the four water resource zones analysed (i.e. the zones in the sample which such as ban was put in place) were 5.0%, 7.7%, 9.3% and 4.5%.

[13] In addition to making their own analysis UKWIR (2013) also report another headline figure of a 10% decrease in demand. Note that this is taken from the EA (2013) study (discussed above). This 10% figure in fact refers to the *maximum* reduction in water usage in a single week compared to the forecasted value for one water resource zone, specifically the WRZ that saw the largest reduction, East Surry. The overall average reduction over the TUB period for that WRZ was 2.8%, whereas the overall average across all WRZs analyse was around 1-2% (as noted above).

[14] Households in these areas are recorded as using a hosepipe if its usage exhibits four periods of 15 minutes where flows are greater than 480 litres per hour (high flow periods).

differences in water use are found between the TUB and non-TUB periods these cannot necessarily be attributed to the TUBs, with the report noting that during the TUB period there were periods of unusually heavy rain which may have led to a suppression in the demand for hosepipe use.

For Thames Water and Anglian Water, the difference in the frequency of hosepipe events between the TUB period and the comparison period (the summer before or after) ranged from approximately -70% to +20%[15] (including 0% in some cases), meaning the results are inconclusive overall. The results varied depending on the company in question, time of day observed (3-4am of 6-11pm) and whether the house was metered or unmetered. For South East Water, average domestic usage during the TUB was 32.6% lower than the following summer, while for non-domestic properties there was an average reduction of 1%.[16] However, UKWIR note potential limitations with these results, in particular that there were periods of heavy rainfall over the TUB period which may have suppressed demand in any event.

In addition, UKWIR (2013) also report figures based on a survey of customers which asked about the frequency that they watered their lawns (domestic customers) or used water for outdoor activities (non-domestic customers) in the summers of 2012 and 2013. By applying (assumed) figures for water use rates for these activities, the study calculates that the reduction in water use for these activities during the TUB compared to 2013 was 22% for domestic consumers and ranged from 11.4% (Thames) to 17.7% (South East) for non-domestic consumers.

Evidence on the impact of water use restrictions from overseas

Various studies provide evidence from overseas. Mini et al. (2014) analysed data on water billing and lot size in Los Angeles from the start of 2000 to the end of 2010.[17] Between 2008 and 2010, local water agencies introduced restrictions in LA in an attempt to manage demand in a response to warm and dry conditions experienced in California. In 2008, only voluntary restrictions were put in place. Mandatory restrictions began the next year including no irrigation[18] between 9am and 4pm, limits on the frequency and duration for some irrigation techniques, and limits on some water practices (no washing vehicles with a hose, using a hose to wash paved surfaces, irrigation during rain or watering with excess water flow). In the final year the restrictions became more stringent including only two-days watering allowed per week, no washing of vehicles in streets or filling of residential pools and spas with potable water, and further reductions in watering times and frequency for all types of irrigation.[19] In addition to these restrictions, a change in the price structure was introduced, namely a 15% reduction in the consumption threshold before a single-family household moved into the (higher) Tier 2 price of a block pricing system, combined with a 44% increase in the Tier 2 rate. However, this analysis does not separate the effect of the price structure change from the effect of water restrictions.

Mini et al. (2014) used an approach similar to that of the EA (2013), estimating water consumption via a linear regression (using variables such as rainfall, temperature, unemployment and seasons) and comparing predicted consumption without the restrictions to the actual observed values to

[15] These approximate figures are based on interpolations and calculations from charts in the UKWIR (2013) paper.

[16] Data was available for 44,536 non-domestic properties. Water use by these customers was generally lower post-drought than pre-drought', except for those in sanitary and construction services who had an increase in usage by 4% and 11% percent respectively.

[17] This database, which contained data on approximately 480,000 customers, was used to calculate water consumption.

[18] Referring to domestic irrigation, such as watering a garden.

[19] Los Angeles Department of Water and Power (2010), '*2010 Urban Water Management Plan*'

assess the effects of these restrictions. **The voluntary restrictions were found to result in a reduction of less than 2m³ per household, an almost negligible effect**. **The first phase of mandatory restrictions resulted in savings ranging from 4% to 15%, the second phase achieved reductions of 19% to 23%** over the spring and summer periods of the year they were introduced. Furthermore, the paper also found that the **mandatory restrictions had the greatest effect on larger households with higher incomes**.[20]

Water restrictions have also been used in other parts of the USA, including northern Nevada. The Truckee Meadows Water Authority implemented water restrictions as early as 1992 (making them permanent in 1996), preventing consumers from using sprinklers except in mornings and evenings of two assigned watering days (increased to three days during the 2010 watering season).[21] Castledine et al. (2014) tested the effectiveness of this watering schedule over the summers of 2008 and 2010 by examining the water use of those following the schedule precisely, those who followed the schedule with some additional watering on scheduled days, and those who did not follow an assigned schedule.[22] The study found that, **those following a watering schedule used more water overall with significantly higher usage than those with no schedule**. This 'rigidity penalty' represented a usage that was 20-25% higher in weekly consumption than those with no schedule.

Haque et al. (2014) analysed the **impact of water use restrictions in place between 2003 and 2009 in the Blue Mountains region of Australia**. These restrictions took three escalating levels between 2003 and 2005, before changes were made to the restrictions in 2008 and 2009. The level 1 restrictions (October 2003) comprised of no hosing of hard surfaces and vehicles and no use of sprinklers or other watering systems; the level 2 restrictions (June 2004) additionally prohibited hosing of lawns and gardens except hand-held hosing at certain hours/days, and no filling of pools over 10,000 litres (except with a permit); and the level 3 restrictions (June 2005) further limited the days on which lawns could be watered by hand (to just two days per week), prohibited hoses or taps from being left running unattended, and ruled that fire hoses only be used for firefighting (not cleaning). The analysis by Haque et al. (2014) used a new estimation technique referred to as Yearly Base Difference Method.[23] The paper found that the **water savings achieved by the level 1, 2 and 3 restrictions were, respectively, approximately 9%, 18% and 20% for single-dwelling households, and 4%, 8% and 9% for multiple dwelling households** (such as apartment blocks).

2.2 Consumer reactions and attitudes towards TUBs

This section presents evidence on consumers' reactions and attitudes towards TUBs. The reactions allow us to understand consumers' attitudes regarding TUBs such as:

- What are reasons, drivers and barriers for compliance?
- What affects the acceptability of TUBs among consumers?

[20] More specifically, those in the 75th quantile of households in terms of income.

[21] These watering days were allocated using a resident's address.

[22] Household watering statistics were gathered by driving along a route with remote sensing devices for a total of 126 days, across both years.

[23] In this study, the base period was defined as 1997-2002 and the restriction period as 2003- 2009. In this analysis the use over a month in the drought period was subtracted from the average monthly use in the base period. The water savings due to the restrictions were then estimated using the Yearly Base Difference Method and linear regression. This paper evaluated four different models and found YBDM to be the most robust and accurate model of the four.

We present this evidence since it helps to inform the debate about the role of TUBs and why they may or may not be effective. The evidence on consumer reaction is mostly sourced from qualitative research (e.g. surveys, interviews, focus groups) commissioned by water companies, industry bodies and academic research.

In summary, **domestic consumers, seem to lack understanding of what TUBs entail**. Additionally, there is evidence that suggests that **consumers would rather accept water restrictions over an increase to their billing costs**. Non-domestic consumers generally claim to be greatly affected by TUBs.

The main reasons for consumer acceptability of these bans involve the perceived seriousness of the situation and how well they believe their water provider is managing water supply (e.g. leakages). If low water supply can be attributed to factors such as severe drought rather than mismanagement, they are more likely to be accepting of a restriction.

2.2.1 Evidence from UK water sector

According to a survey conducted by UKWIR (2013), 68% and 75% of domestic and non-domestic respondents respectively, were aware there was a hosepipe ban in their local area. Less than 1% of domestic respondents could correctly identify all the restricted activities (from a list of 17).[24] These responses are evidence that those affected by the hosepipe ban failed to fully understand what the ban entailed and it seems likely that the impact of this TUB was limited by this lack of understanding.

Yorkshire Water (2018a) also examined consumer reaction by commissioning research with the aim of looking into the acceptability of hosepipe bans to customers.[25] Quotes from the research emphasise that the acceptability of a hosepipe ban is linked to a customer's perception of the provider's current water management. If the water company is perceived as responsible and the reason for implementing a ban is a result of a serious shortage customers are more likely to accept a ban.

Further research commissioned by Yorkshire Water (2018b) also attempted to gather consumer reaction and attitudes towards hosepipe bans.[26] It emerged that **for consumers the main areas of concern are garden watering, car washing and social relationships**:

- "The driveway is a gateway to my house; people view that, and they view you as a presentable person."
- "My son likes to water the plants with me and help his dad wash his bike; it's lovely to do things together."
- "The lawn being burned away is literally your money being burned away."

[24] The mean of correct answers was 6.45 out of 17. Only 38% of non-domestic responses were correct when asked to identify the restricted activities.

[25] This survey was completed by 775 respondents; These respondents were selected to ensure representativeness based on gender, age, SEG & location; the data was weighted on unmetered and metered status for a 50/50 split. Focus groups were also conducted. These focus groups had 8-9 participants each, these participants had a variety of attitudes towards the environment (and were also a mix of ethnicities, ages and genders).

[26] This research took the form of 6 qualitative workshops, designed to spark debate, comprising of 9 participants: 3 whom were supporters of hosepipe bans, 3 participants whom were against hosepipe bans and 3 participants whom were 'on the fence'. In addition to these workshops, 5 'at home in-depth interviews' with selected customers were also conducted.

This report presents over thirty consumer quotes representing reaction to TUBs; these quotes mostly fall into the category of concerned homeowners who are worried that their garden, cars and social relationships will be damaged or diminished by a ban. However, these problems could be solved with handheld watering (which would not be banned under a hosepipe ban), implying that these consumers do not fully understand the ban, which is consistent with previous findings.

During the interviews, consumers were consulted on their views before and after being given information on 'water use and the impact of hosepipe bans.' As seen in Table 1, providing consumers with evidence on general water management almost always weakened support for TUBs. The two main concerns found in these interviews were that bans did not appear to save great amounts of water, and the amount of leakage.

Table 1 The effect that water usage statistics had on household views regarding TUBs

View prior to receiving info	View after having considered provided info	Associated Quote
Pro-ban	Pro-ban	"I'd still support a ban despite what we've heard, because every little helps"
Pro-ban	'On the fence'	"I just assumed hosepipes accounted for so much more water use"
Pro ban	Anti-ban	"When you see what a ban achieves, it just feels ridiculous"
'On the fence'	Anti-ban	"Even though I use my hosepipe a lot, the info today doesn't suggest me not using would make a great deal of difference"
Anti-ban	Anti-ban	"I'm even more anti that I was before given how much water they lose a day"

Source: Yorkshire Water (2018b), Reducing Demand for Water Research

Some non-domestic consumer reactions can be found in a code of practice and guidelines on water use restrictions for water companies, published by UKWIR (2014). This included interviews with non-domestic public and private associations. Their responses suggest that water use restrictions would have a range of negative effects on their activities and businesses, including both financial and social consequences such as job losses and safety concerns. The table below presents some views of respondents in different sectors.

Table 2 Views of non-domestic water users on the impacts of TUBs on their organisations

Organisation	How TUBs affect their activities/business
British Swimming Pool Federation	'Demand in 2012 fell to 1000-1500 [swimming pools purchased] compared to 2500-3000 at time of [this] report (pre-2008 were 5-6000.'
Car Wash Association	'If car washes could not operate during droughts, dedicated car wash businesses would be closed.'
England Golf	'During 2012, there was a talk of a six month ban on sports in Thames Water area. Had this actually happened, 30,000 jobs would have been affected.'
Horticultural Trades Association	'If landscapers cannot use water efficiently to establish plants during the first phase of TUBs, then they will not be hired to establish or replant gardens. There is an immediate knock on effect on the supply chain (garden centres and growers).'
Kent Cricket Board	'Watering between weekend fixtures is vital for pitch safety.'

Racecourse Association	'If the going (firmness of the ground) cannot be managed within industry guidelines, trainers may pull horses from races to avoid injury.' This will lead to lower attendance levels which has a negative economic impact on the rest of the industry.'
Turf Grass Growers Association	'If [...] customers cannot establish turf there will be: adverse environmental impact, adverse social impact, job losses, loss of income and wasted crops [for turf growers].'
Wandsworth Borough Council	'The impacts of not being able to water planted bedding plants and street trees is financial (tens of thousands of pounds).'

The business plans of some UK water companies also include some additional insight on consumer attitudes towards TUBs, as presented in Table 3. Mixed consumer attitudes on TUBs emerge from these plans. Most providers stated that their consumers would want to either avoid TUBs entirely, or to keep the frequency to as low as 1 in every 10 years. However, some of these consumer responses imply that there may be some support for the use of TUBs and did not see a reduction of their frequency as a priority. These perceptions are also consistent with previous willingness to pay research, which found that consumers would prefer to experience water restrictions than see an increase in their bills.

Table 3 Consumer Reaction towards TUBs from other UK water companies

Company	Consumer Perception
Anglian Water	'Customers are satisfied with the current levels of service for temporary use bans, at not more than 1 in 10 years, and non-essential use bans, at not more than 1 in 40 years. Customers do not see reducing the frequency of these restrictions as a priority area for investment.'
Northumbrian Water	'1. Very few customers (none for NW) have experienced restrictions, although some have experienced long-term interruptions 2. Customers are more accepting of reduced pressure and hosepipe bans than other types of restrictions 3. They may see 1 in 20-year risk as acceptable'
South West Water	'Our conversations with customers show us that customers want to avoid water restrictions where possible. Water restrictions are currently very rare in the region and customers want us to be incentivised to maintain this position. A notable proportion of PR19 customer research has taken place during the 2018 heatwave and has confirmed that customers have a strong preference to avoid these.'
United Utilities	'Customers place a high value on continuous water supply and avoiding the need for severe restrictions on use is a priority.'
Affinity Water	'Most people felt Temporary Use Bans should apply to all customers equally; be imposed no more than once in every 10 years and would rather experience these restrictions than see their water bill increase - Blue Marble, June 2016 Pre-SDS Consultation Online Survey' 'Respondents supported temporary use bans (56%) but wanted all other options to be exhausted first such as better water management and reducing leakage. 72% of them did not think that we should spend more to reduce the likelihood of

	temporary use bans and would rather experience them than see their water bill increase. Those who did suggested investment in infrastructure and education of customers'
Bristol Water	'We also asked customers their opinions on hosepipe bans and TUBs. Hosepipe bans are consistently given as a low priority across the customer research, and due to a perceived low risk of drought given few recall having experienced one, customers do not value a reduction in the instances of TUBs.'

Source: Business Plans, 2020-2025

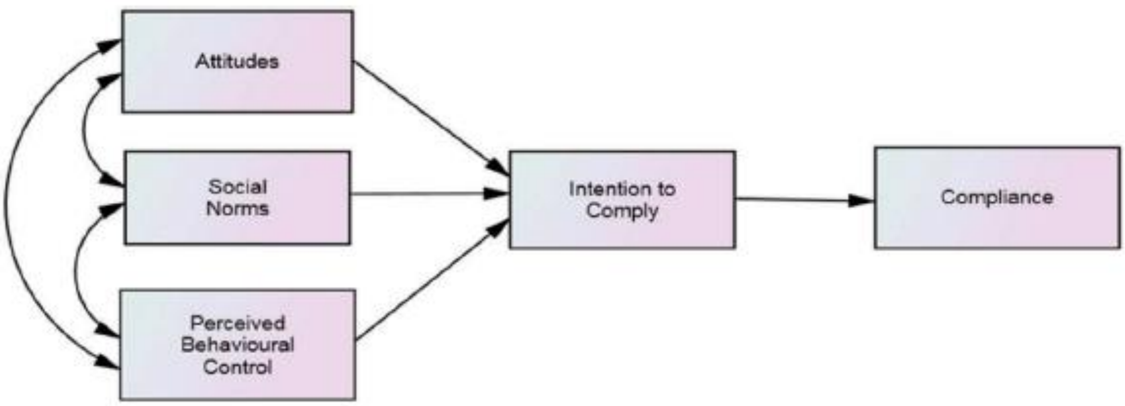
2.2.2 Academic literature on consumer reactions to water restrictions

In reaction to the restrictions widely implemented in Australia, Cooper (2016) conducted research to investigate the motivations behind consumer compliance with these restrictions, hypothesising four drivers of compliance:

1) **Attitudes:** 'An evaluative predisposition towards the behaviour as a function of its determinant personal consequences.'
2) **Social Norm:** 'The perceived social acceptability of behaviour.'
3) **Perceived Behavioural Control:** 'Related to an individual's perception of their capacity to achieve a behaviour.'
4) **Intention:** 'The individual's intention towards the restriction (whether they intend to comply).'

Within this framework, drivers 1, 2 and 3 all influence the fourth driver, the 'Intention to Comply'. *Intention* is then specified to impact on compliance as illustrated in Figure 2.[27]

Figure 2 Conceptual framework of Cooper (2016)



Source: Cooper (2016), 'What drives compliance? An application of the theory of planned behaviour to urban water restrictions using structural equation modelling.'

[27] With the use of online surveys, these four drivers were measured. Participants in these surveys were given questions which they would answer using a five-point Likert scale anchored by 'totally disagree' and 'totally agree'. An example of a survey question is 'I think that is a good idea to comply with water restrictions.' In addition to this, some questions are also included to measure the level of compliance that respondents have with water restrictions, using questions such as 'How often would you say that you comply with water restrictions?' For these questions, respondents answered using five categories that ranged from 'less than 20% of the time' to '90% more of the time.'

Perceived Behavioural Control was found to have the strongest effect on *Intention*, with *Attitudes* being the second strongest and *Social Norm* being the weakest. *Intention* was also found to have a very strong effect on actual *Compliance*. Therefore, **the strongest driver of compliance with water restrictions is an individual's perception of their capacity and ability to comply**.

Hensher et al. (2006) assessed consumer willingness-to-pay (WTP) to avoid water restrictions. They examined WTP regarding restrictions which varied in frequency, duration, and stringency.[28] Their analysis found **a lack of WTP to avoid most types of drought-induced restrictions**, with respondents appearing unwilling to pay to avoid higher levels of restrictions that are not in place 'every day, and all year.'

Instead, when given the option to restrict their water usage to alternate days, customers would rather comply with these restrictions if this way they can avoid paying higher water bills. This may be because any inconvenience customers receive from restrictions may be offset by the novelty of 'feeling good' from using water responsibly, and low-level restrictions may not be particularly inconvenient if they are flexible enough and restrict only non-essential consumption.

Brennan et al. (2007) also completed research on the welfare costs of these Australian water restrictions. In this study, a model is constructed to examine how a ban on sprinklers affects consumer utility when they have the choice to substitute sprinkler use with the more labour-intensive options of hand-held watering from a watering can.[29] The estimated household welfare cost of a sprinkler restriction was found to be 'less than $100 per season when mild (two days per week) restrictions are in place.' This also results in an associated water savings of around 36 per cent of current consumption. More stringent restrictions (a complete sprinkler ban for all days of the week), may range between '$347 and $870 [welfare loss] per season when a complete sprinkler ban is in place' across all estimations regarding wage rates. This also only provides an associated water saving of 42%, an increase of 6% over mild restrictions of two days per week.

This research does have some limitations in the form of a lack of research on model parameters. The proposed model relies on a consumer's preference for 'greenness' of their lawn as well as the time cost associated with hand-held watering. This paper acknowledges that there is a lack of previous research on these two concepts that are core for this welfare analysis.

Dessai and Sims (2011) explored consumer attitudes towards drought in southeast England. From 102 questionnaires, "[in periods of drier summers and increased frequency of water shortages], 75% of respondents were willing to pay more to ensure there was enough water for everyone". In addition, a water company (Anglian Water) and the EA are noted as saying that "the application of non-essential water use bans will be perceived by customers to be a result of poor water management".

3 Alternative interventions

Given that TUBs seem to be unpopular and widely misunderstood among water customers and their effectiveness remains difficult to robustly assess, it is useful to look at alternative ways of managing

[28] The choice experiment was administered via a mail survey. The impacts of the various attributes of water restrictions on respondents' choices were modelled using maximum likelihood estimation, and WTP results were derived from the results.

[29] The welfare cost for these consumers is modelled using a consumer's preference for lawn greenness and the cost and time interaction between leisure/sprinkler use and hand watering.

water demand, during droughts and in general. This is done in this section, which proceeds as follows:

- Key findings are presented up front in Box 2.
- The remainder of this introductory section provides the broad motivation for alternative interventions to water restrictions.
- Section 3.1 provides a concise overview of these measures.
- Section 3.2 discusses each type of intervention in more detail, assesses its effectiveness and puts this into perspective with respect to the effectiveness of TUBs.

Box 2 Key findings

- Various types of alternative interventions to TUBs have been assessed, such as:
  1) Messaging interventions (see section 3.2.1)
  2) Feedback provision on consumption behaviour (see section 3.2.2)
  3) Changing the financial incentives of resource demand (see section 3.2.3)
  4) Education, information and awareness initiatives (see section 3.2.4)
  5) Installation of water saving devices (see section 3.2.5)

  The evidence gathered looks at **resource conservation in the water and energy sectors**.
- Often, especially for 1 to 3 above, assessments used robust state-of-the-art methods such as randomised controlled trials (RCTs) through collaborations between academics and water suppliers or local authorities. Education initiatives (4 above), instead are often initiated by governments and broadly targeted which makes measurement of effectiveness more difficult.
- A key benefit of alternative interventions compared to TUBs is that they are **not unpopular**. In fact, well designed interventions such as **social norm messages provided on bills can be engaging** and therefore interesting and valued by consumers.
- **Many alternative interventions have been found to be similarly, if not more effective than water restrictions**. This makes them attractive as they can be quick and cost effective to implement while not triggering negative consumer reactions.
- **Messaging interventions**, especially those using strong social norm statements, have been widely shown to significantly reduce water consumption by approximately 1-5% in the short term and by 1.3-2.6% in the long term. This suggests they are **similarly as effective as TUBs.** For energy consumption the effects of this type of intervention were higher, amounting to 3-14% on average.
- **Feedback provision**, for example via (smart) meters or in-home displays, seems to trigger mixed reactions: **sometimes large reductions in consumption are achieved** (larger than the effects of TUBs), but these tend to be short-lived and **fade over time**. On other occasions, feedback provision is not as effective. This is especially the risk for water consumption that is more essential and therefore seems less elastic compared to energy consumption.
- **Changing prices or providing financial incentives** for resource conservation requires feedback provision to be effective. Such measures **could be used to complement TUBs** to increase the effectiveness of water restrictions (see also evidence in section 2.1).
- **Education initiatives alone are unlikely to be effective**. They are, however, found to usefully complement other initiatives as sound consumer understanding is required for

other measures such as messages, feedback or water saving devices to be most effective.

- **Water saving devices** have been found to cause **large and lasting reductions** in water consumption, especially if they are fitted and explained professionally. There is little evidence on how water saving devices are best promoted, distributed and fitted at a large scale. However, it seems that such devices could be most effective at reducing water consumption, although these reductions would not be targeted at periods of drought.

Motivation for considering alternative interventions

While use restrictions and changing prices are the classic tools in resource demand management, there is a growing body of evidence which suggests that other mechanisms, which do not rely on prices or use restrictions, may be similarly effective.

Some such interventions have been informed by behavioural economics and tested in small to large scale academic interventions. **Behavioural economics** systematically studies what drives and influences consumer behaviour. It recognises how **consumers do not always act fully rationally** as assumed in classical economic theory. Systematic **behavioural biases can influence consumers' decision-making** and appealing to these biases may change consumer behaviour without requiring drastic, or costly interventions.

Some of the behavioural biases which have been shown to influence consumers' use of water and other resources include:

- **Social norms**: People have a natural tendency to comply with social norms and take the behaviour of others as examples of appropriate behaviour. As a result, it has been found that providing consumers with messages which compare, or benchmark, their behaviour against that of (similar) peers, such as neighbours, can powerfully influence their choices (Alcott (2011)). For example, Goldstein et al. (2008) had found that hotel guests are more likely to reuse their towels, and therefore reduce hotels' water consumption, when they are given messages such as '*the majority of guests reuse their towels'* rather than appeals to their environmental conscious such as '*Reuse your towel to help conserve environmental resources*'.
  - Social norm messages can be categorised (as defined by Hasting and Rustamov (2015)) into:
    - **Descriptive social norms** which describe the behaviour of others and sub-consciously invite others to imitate this behaviour.
    - **Injunctive social norms** which make suggestions as to whether behaviour will be approved of or not.

  Both types have been shown to influence water consumption.
- **Goal setting and commitment devices**: Encouraging consumers to commit to certain resource saving goals via. E.g. 'pledging to save' can lead to improved engagement with saving efforts.
- **Gamification**: Providing relevant and impactful messages in a vivid and personal way can be powerful (Orland et al., 2014).

- **Loss aversion**: Loss aversion is relevant for how messages are phrased. For example, it can make a difference whether a message is framed as a gain (you could save £10 on your bill) or a loss (you are currently overspending £10 on your bill).[30]

3.1 Overview of alternative interventions

As described in section 1.1, we assessed evidence on interventions from the academic literature and public institutions with the aim of discovering what alternative ways of managing water demand are available in addition to restrictions such as TUBs. What emerged is that **academics have sought to improve the evidence base by conducting randomised controlled trials (RCTs)** testing various types of consumer communication strategies such as **messaging campaigns and feedback provision**.

Governments instead more broadly seem to focus on conducting broad-scale **education and awareness campaigns**. Some evidence was gathered on the impacts of **changing unit prices** or providing **financial incentives** for resource conservation. Finally, some water companies distribute **water savings devices** which aim to reduce domestic water consumption for example as retrofits.

An overview of the various types of interventions we assessed is shown in Table 4. Sections 3.2.1-3.2.5 then explain each type of intervention as well as their effectiveness in turn in more detail.

Table 4 Summary overview of alternative interventions

Type of intervention	Summary	Examples of tools	Range of measured effects
Messaging	Interventions consist of messages inviting/requesting consumers to change their consumption. Tested message content typically contains: ■ Educational content ■ Pro-environmental messages/claims ■ Social norms ■ Goals, targets	Messages are commonly delivered through: ■ Letters / bills ■ Prompts (e.g. cards, displays) at point of consumption ■ Email ■ Websites ■ Social Media	**Short term reductions**: ■ Water: 1 - 4.8% ■ Energy: 3 - 14% **Long term reductions**: ■ Water: 1.3 - 2.6%
Feedback provision on consumption behaviour	Feedback interventions test consumers' responsiveness to the provision of information on current or past consumption. Feedback is typically provided via: ■ In-home displays (IHD) ■ (Smart) meters ■ Letters / bills	■ Smart meter messages on current consumption and feedback on whether it is higher/lower compared to e.g. a reference period, or neighbours. ■ Live feedback on water consumption in shower.	**Average demand reduction in energy meta study was 11.5%.** **Average water demand reduction around 5%.** ■ **IHD**: 22% less energy used in showers; annual water consumption reduced by 7,300L per household. ■ **Meters**: initially high then fading

[30] For example, Brick et al. (2017) use this technique.

			engagement. Short term savings: 2-8% (energy) ■ **Feedback on bills**: 0.3 - 6.3% (energy) ■ **Gamification**: 7-23% (energy)
Changing financial incentives of resource demand	Such interventions change unit prices for example if: ■ There is supply shortage ■ Peak demand ■ A consumption threshold is surpassed	Price changes might form part of a tariff or are announced through e.g. smart meters, IHD or letters.	Difficult to measure in absence of meters. Measured energy demand reductions: 5-22% Most effectively delivered through IHD. Effectiveness for water less clear.
Education, information and awareness initiatives	Such interventions seek to sensitise the public around resource use, in general or particularly to prevent or during droughts.	■ Leaflets with water saving tips ■ Education programmes in schools ■ Public outreach work ■ Public awareness campaigns using pro-environmental claims, reminders to save water etc.	Rarely robustly assessed. Effectiveness of such campaigns in RCTs limited but useful if provided in conjunction with other interventions (see messaging).
Installation of water saving devices	Such interventions comprise the advertisement, distribution and installation of water saving devices as (free or low cost) upgrades for domestic water consumers.	Typical water saving devices are: ■ Dual flush cisterns ■ Tap inserts ■ Economic shower heads ■ Flush reducers	Substantial decreases by between 10-40%, with the exact magnitude depending on a variety of factors, such as location, household size and type of intervention.

Notes: Overview prepared by London Economics based on evidence cited in following sections.

3.2 Description and effectiveness of alternative interventions

The following sub-sections describe in turn each type of intervention presented in Table 4 in detail. For each intervention, we present an **overview**, an **assessment of the effectiveness and robustness** of the resulting impact on resource consumption as well as a **comparison** of this effectiveness to that of TUBs.

3.2.1 Messaging interventions

Overview

The most prominent alternative interventions to use restrictions consist of **companies delivering messages to customers** for example using:

- simple letters;
- messages embedded within bills;
- messages delivered over the internet (e.g. websites, social media); or
- cards/prompts at the point of consumption (e.g. in hotels, public buildings).

Messages do not only vary in the way in which they are delivered but also in the type of content they provide. Messages typically contain information regarding a conscious use of the resources, such as water or energy, appeals to reducing consumption and protecting the environment, technical information on how resource use could be reduced, or using the resource wisely without waste.

In line with classical economic assumptions it was widely perceived that providing individuals with factual information such as '*Letting the tap run while you brush your teeth is harmful to the environment. Only use what you need.*' could induce a change in behaviour. A large body of evidence however suggests that this is not the case[31] and academics have tested various interventions to find what drives conservation especially in the water and energy sectors.

In the academic literature, there is a growing body of messaging interventions which combine factual information with various types of other tools to increase consumer engagement with the messages and their reactions. Typical tools include for example:

- **Social norm messages**[32] which can vary in the **strength** and **degree of personalisation** such as:
  - **Weak social norms**: '*Cobb County residents consume almost one of out every ten gallons of Georgia's public water supply. As a result, our water use has a large impact on the ability of Georgia's waterways to protect wildlife and dilute pollutants that threaten human health. … We need your help. […] We all have to do our part to protect Cobb County's precious water resources. […] Please don't waste water. Remember: every drop counts!*' (Ferraro et al. 2011, p. 319)
  - **Strong social norms with comparison to peers**: '*As we enter the summer months, we thought that you might be interested in the following information about your water consumption last year:* ***Your own total consumption June to October 2006: 52,000 gallons.*** *Your neighbors' average (median) consumption June to October 2006: 35,000 gallons. You consumed more water than 73 percent of your Cobb County neighbors.*' (Ferraro et al. 2011, p. 319)
- **Use of visuals**: These could contain graphs, charts or even images such as smiling/frowning faces to indicate good/bad behaviour. (Brent et al. 2016, Tiefenbeck et al. 2016, Allcott 2011, Brick et al. 2017, websites and sources of the Government of South Africa)
- **Appeals to responsibility**: '*HELP SAVE THE ENVIRONMENT. You can show your respect for nature and help save the environment by reusing your towels during your stay.*' (Goldstein et al. 2008, p. 2)
- **Goal setting and commitment devices:** This would combine any type of other messages with a target of the desired behaviour change e.g. '*Because of the extended drought […], we are asking all of our customers to reduce water use by at least 10% this summer […]*' (Brent et al. 2016, p 6.). In some settings, such goals have been combined with customer self-generated commitments, verbal or written, to complying with the request which can

[31] See for example Ferraro et al. (2011, 2013), Brent et al. (2016). In these studies, the baseline treatment often used purely factual information and the measured effects were mostly insignificant.

[32] Social norm messages were used by the following reviewed sources: Ferraro et al. (2011, 2013), Allcott (2011), Goldstein et al. (2008), Brent et al. (2016), Osbaldiston and Schott (2012), Abrahamse et al. (2005), Asensio and Delmas (2014), Klege et al. (2018), Brick et al. (2017), Visser and Smith for the Water Research Commission (2013), South African Government websites, Brandon et al. (2017), Hastings and Rustamov (2015), Hahn et al. (2016).

increase the effectiveness of the goal setting.[33] In a different context, this has for example led to an astounding 9.1 percentage points increase in voter turnout in the 2008 Obama election campaign (Nickerson and Rogers 2010).

- **Highlighting of private benefits/costs**: Such messages would highlight the advantages, such as saving money on monthly bills, the individual customer could achieve by adjusting their behaviour (Brick et al. 2017). For example: '*Saving water saves you money*' (Brent et al. p. 6)
- **Highlighting social benefits/costs**: Instead of highlighting costs to the individual, messages may make statements regarding the effect of behaviour on society at large such as: '*Last week, you used 66% more/less electricity than your efficient neighbors. You are adding/avoiding 610 pounds of air pollutants which contribute to health impacts such as childhood asthma and cancer*', (Asensio, Delmas (2014, Table S1, appendix)

Effectiveness and robustness of existing messaging interventions

In the short-term[34] the effectiveness of messaging interventions ranged from 1 to 5% reductions in domestic water consumption for the average household. Only few studies assessed whether the interventions had a lasting effect on water consumption. Those that did found that consumption following the messaging interventions remained significantly below the pre-intervention levels, with savings achieved of around 2.6% and 1.3% after respectively 1 and 2 years (Ferraro et al. 2011).

In the literature which looks at similar interventions in the energy sector, the achieved reductions in resource demand were higher and amounted to 3-14% on average.

The vast majority of the cited interventions were tested through randomised controlled trials (RCTs) which produce highly robust results. Often, multiple types of messages were trialled with the same customer base. What emerges is that more **specific** and **personalised messages** as well as messages containing **stronger social norm with peer comparisons** compared to weaker, generic ones were most effective (Ferraro et al. 2011, 2013, Osbaldiston and Schott 2012, Abrahamse et al. 2005). Social norm messages typically seem to cause the strongest reactions, although there is evidence that some households remain unaffected by such messages. Interestingly, these types of households seem to respond better to messages which highlight their **private benefits** such as savings (Brent et al. 2016). Framing such benefits as losses rather than gains sometimes marginally increased responsiveness (Hahn et al. 2016).

Furthermore, it emerges that the effectiveness of messaging interventions can be increased by targeting messages to consumer groups which are likely to be most reactive. These comprise for example **households with above average consumption,** generally **wealthier** and **owner-occupied households** as these seem to best respond to comparative social norm messages which encourage them to lower their consumption (Ferraro et al. 2013, Brent et al. 2016, Brick et al. 2017).

Comparing the effectiveness of messaging to that of TUBs

Overall, the evidence on messaging interventions suggests that they can be **similarly effective as TUBs**, especially if the aim is to achieve short term reductions in consumption. At the same time, messaging campaigns, if combined with pertinent and brief educational content, can trigger a more

[33] See for example: Osbaldiston, Schott (2012), Abrahamse et al. (2005).

[34] The short term here stands for effects which were measured immediately following the receipt of messages. Measurement in most studies lasted for a few weeks or months.

conscious consumption in the long run. Messaging campaigns are therefore **a robust and cost-effective** way of influencing water demand without triggering negative customer reactions, and restricting water use.

3.2.2 Feedback provision on consumption behaviour

Overview

One reason why consumers may not manage their consumption of scarce resources such as energy or water is that it is not obvious to them how much they use or the impacts of their consumption on their private outcomes (such as their bank balances) or social outcomes (like the environment). Hence, providing feedback on consumption behaviour may help consumers to manage their usage by making the consumption decision more salient to them. Such feedback is typically provided via letters, bills, in-home displays (IHD) and (smart) meters.

A number of factors seem to make feedback provision especially effective in encouraging consumers to reduce usage:

- Using **striking visual imagery**: previous research shows that using visual aids to illustrate the impact of consumption decisions on private or social costs encourages consumers to reduce usage, since people tend to respond more to images than to text ('picture superiority effect' (Curran & Doyle (2011)). Consumers cut down their energy consumption by 22% when they were given information on how to manage their energy usage accompanied by a tool in the form of a polar bear animation to visualise the impact of their energy usage and efficiency (Tiefenbeck et al. (2006)).
- **Higher frequency** of feedback, since more frequent communication heightens a sense of urgency. Communicating feedback more often also further increases the prominence of resource consumption/conservation in consumers' memories which in turn encourages behavioural change in favour of conservation (otherwise known as the 'availability heuristic' (Schwarz et al. (1991)). Research in a range of contexts including energy (Jessoe & Rapson (2014); Abrahamse (2005)) and water (Wichamn (2017) conservation has shown that consumers are likelier to reduce consumption with more frequent feedback relating to their usage.
- **Inducing cognitive dissonance**: People often try to behave in a way that is consistent with their stated intentions or beliefs to avoid the feeling of holding mutually contradictory ideas or opinions (known as 'cognitive dissonance'[35]). Feedback can encourage consumers to reduce energy consumption by pointing out that although they state a commitment to saving energy, they are actually high consumers of electricity (Kantola et al. (1984)).
- **Gamification**: which uses people's love of games and competition to encourage them to engage in prosocial behaviours such as voting[36] or natural resource conservation. Previous research shows that participants reduced energy consumption when they played a game whose outcomes (keeping virtual 'chickens' healthy) was linked to their energy consumption choices (Orland et al. (2014)).

Careful bill design may also improve the effectiveness of feedback on encouraging behavioural change. Recent research conducted for the European Commission investigated the impact of bill

[35] See Festinger, L. (1957)

[36] https://www.wired.com/story/samantha-bee-app-voting/

design on consumers' reported intention to manage their energy consumption (European Commission (2018)). The research found that consumers were more likely to report that they planned to manage their energy consumption in future if they received a simple, concise message about the impact of energy costs on their energy bill.

Effectiveness and robustness of existing feedback interventions

Evidence on the **effectiveness of feedback-based interventions was mixed**. Energy consumption reduction ranged from 3 to 23%[37] while the **effectiveness of feedback interventions in the water context was approximately 5%**[38].

It can be difficult to compare the effectiveness of different interventions, since individual studies frequently focus on a single intervention at a time. Comparison across studies is complicated, since studies frequently employ different methodologies (natural experiments, randomised controlled trials, ANOVA etc.) However, a number of themes emerge:

- The following forms of feedback were found to be especially effective:
  - **Increasing billing and feedback frequency** was effective in encouraging consumers to manage consumption of energy (Abrahamse (2005)) or water (Wichman (2017)).
  - **Feedback inducing cognitive dissonance** (Abrahamse (2005)).
  - **Combining feedback with other mechanisms**[39] such as goal setting (e.g. setting consumption reduction targets) or commitment mechanisms (see also section 3.2.1).
  - **Feedback combined with visualisation tools** proved to be especially effective in the energy context, with a large (22%) and apparently persistent reduction in energy use (Tiefenbeck et al. (2016)).
  - Gamification also showed strong initial impacts on energy reduction, but these effects did not persist (Orland et al. (2014)).
- Otherwise, feedback interventions did not often have persistent effects. Consumers' engagement with feedback tools often faded over time and they often had difficulty understanding the information the devices/tools communicated (Buchanan et al. (2015)).
  - However, the finding of small and temporary effects is contested by other meta-studies: for example, Karlin et al. (2015) find that feedback tools reduce energy consumption, on average, by 11.5%. They suggest that previous meta-analyses may not find large effects of feedback tools on energy consumption because they do not take into account that consumers may act on feedback devices/mailouts etc. in a different timeframe than when they are sent out.

Furthermore, the interventions frequently had **methodological concerns** which make it difficult to interpret their results. For example, low sample sizes (Kerracher et Torriti (2013)) or sample selection bias if feedback devices were not randomly allocated across the sample. In one intervention the feedback devices were offered as a gift (Tiefenbeck et al. (2016)). Therefore, there may be concerns that the results of the study are confounded by characteristics of participants. The individuals choosing to ask for a feedback device may be systematically different from individuals

[37] See Orland, B. et al., (2014), Tiefenbeck et al., (2016) Buchanan et al., (2015), Kerracher, C., Torriti,J. (2013), .

[38] See Wichman, C. (2017), Hastings and Rustamov (2015)

[39] See Osbaldiston, R., Schott, P. (2012), Abrahamse, W., et al., (2005) Environmental sustainability and behavioural science: meta-analysis of proenvironmental behaviour experiments, Environment and Behaviour 2012 44: 247

who did not ask for the device in some dimension that is correlated with pro-environmental behaviour.

Reviews of the literature on feedback devices (Buchanan et al. (2015)) have also pointed out that studies often did not disentangle the effects of the feedback devices on behaviour from the 'Hawthorne effect'[40], where participants in an experiment may modify their behaviour in line with how they think the experimenter would like them to behave. Therefore, impacts on energy consumption may have been incorrectly attributed to the feedback devices rather than to the participants knowing they were being observed.

Comparing the effectiveness of feedback provision to that of TUBs

Overall, it appears that well-designed feedback provision can have a **substantial and possibly larger impact compared to TUBs** on resource consumption. At the same time, feedback provision requires that water consumption is well measured, for example through meters, or ideally, smart meters with IHD. This likely makes these types of interventions **more expensive compared to messaging campaigns** which may require less investments in the infrastructure.

Furthermore, it appears that the effectiveness of feedback information can wear out over time. Water companies would therefore likely need to be innovative in delivering new types of feedback messages if a change in behavioural patterns is required, for example prior to or during periods of drought.

3.2.3 Interventions which change financial incentives of resource demand

Overview

Instead of using interventions which are grounded in insights from behavioural economics, water companies also have classical economic tools at their disposal to manage customer demand. The most obvious mechanism to manage demand involves changing unit prices, or more generally the financial incentives at stake. A few sources have adjusted financial incentives in order to manage water consumption during droughts. Such interventions for example comprised:

- Increasing prices for consumption above a certain threshold;
- Increasing prices during peak demand hours; and
- Provision of (monetary) rewards for resource saving efforts.

As before, the evidence in this section draws from evidence in the water and energy sectors.

Effectiveness and robustness of existing interventions which change unit prices

Across all studies, **reductions in energy consumption ranged from 5 to 22%**. For water consumption there was less evidence and the results of the evidence assessed was mixed.

In a meta-study of existing experimental evidence, Abrahamse et al. (2005) conclude that providing monetary incentives for reducing domestic energy consumption was effective in reducing demand by approx. 5-12%. These effects were mostly targeted to affect short-term demand over periods of

[40] The Hawthorne effect is defined in Monahan, Torin; Fisher, Jill A. (2010). Benefits of 'Observer Effects': Lessons from the Field. U.S. National Library of Medicine, National Institutes of Health.

a few weeks to few months and most interventions were either not measured for longer, or it was found that the initial effects were not sustained unless additional incentives were provided.

In the energy sector, for example, Ofgem (2017) had commissioned a study to assess the distributional effects of introducing time of use tariffs with half-hourly settlements. The study estimated that most consumer groups would benefit from switching to half-hourly settlements as they would be expected to adjust their demand by shifting consumption to cheaper periods in the day. However, the study also concluded that adjustments were lowest during domestic peak demand periods (e.g. evenings) and that within each consumer group there would likely also be losers due to low responsiveness.

In an RCT with 437 participating households, Jessoe and Rapson (2014) provided residential electricity customers with real-time information about their energy usage which significantly increased the price elasticity of demand. They observed responses to short term price increases ranging from 200 to 600% of base prices.[41] Only changing prices, without informing households timely of the surges through in-home displays (IHD), led to reductions of 0-7%. When IHDs provided feedback, reactions increased to 8-22% which seems to be driven by learning through information provided through the IHD about how adjustments are most effective in response to price changes.

Lu et al. 2017 instead concluded that the international evidence on **the evidence on the effects of Increasing Block Tariffs (IBTs) on water demand was mixed** in the sense that **such pricing practices sometimes lead to reductions in water demand but not always**. Moreover, these types of prices are fairly complex to administer and difficult to understand for consumers which may further reduce their appeal for water companies.

Comparing the effectiveness of changing the financial incentives of resource demand to that of TUBs

It is well-documented and grounded in the law of supply and demand that consumers respond to changes in prices. Therefore, it is not surprising that in this section, we find that consumers mostly adjust their behaviour in line with the provided financial incentives by shifting demand away from periods with peak prices and reducing demand when they may gain a reward.

In energy consumption, the effect sizes were sometimes sizeable reaching a maximum of 22% reductions in demand. For water however, the evidence was less clear. This seems to point to the fact that **water is an essential service**, more so compared to energy. Therefore, it is not surprising that the **price elasticity of demand for water is lower compared to that for energy**. This makes financial incentives possibly less suitable for managing water demand, though financial incentives could possibly be put in place to accompany TUBs in order to increase compliance with reducing non-essential water use.[42]

[41] The authors observed 2 and 4-hour long surges.

[42] See also evidence from Mini et al. 2014 summarised in section 2.1 on combining water restrictions with changes in pricing structure.

3.2.4 Education, information and awareness initiatives

Overview

Many water companies in the UK and abroad are seeking to **increase a conscious and wise use of water** among their customers. One method that is regularly adopted is seeking to provide educational content to customers via messages, public outreach campaigns or alike.

Effectiveness and robustness of existing education initiatives

The main challenge in assessing the effectiveness of such education initiatives is that **programmes which broadly target the public are rarely assessed using robust methodologies** which measure follow-up reactions in household consumption.

As mentioned in section 3.2.1, **messaging campaigns which uniquely provided informational, or educational content mostly remained without an effect**, according to robust RCT studies.

Nevertheless, there is some evidence which suggests that educational content is worthwhile providing. As discussed in section 2.1, the Environment Agency (2013) for example assessed water demand following the issuing of generic water saving messages prior to the 2012 UK drought and found reactions of +0.5 to -1.7 % on average, though these estimates were not statistically significant. Ferraro et al. (2011) and (2013) find that educational content provided via a 2-page 'tip sheet' was insignificant. The information was only effective at influencing consumer demand when it was combined with social norm messages, in particular when these contain a strong and personalised social comparison to a peer group (neighbours).

Details of education initiatives run on a large scale in schools, businesses and households in Zaragoza (Spain) were detailed by Ecopeace Middle East (2009). These initiatives were combined with the creation of goals around adopting good practices and municipal orders to save water resulted in reductions in domestic water consumption from 113L per person per day in 1996 to 96L in 2007 (the lowest in Spain). Yet, it is impossible to assess which share of this effect can be attributed to the educational initiatives.

Last but not least, Israel has conducted broad education campaigns to sensitise the public around water scarcity to purse an '*immediate potential saving of approximately 15% of the domestic water consumption through a national explanatory campaign*' (Lev 2012, p. 23). However, this figure appears to be an estimate and, in the meantime, public awareness campaigns and appeals to restrict domestic water use have been suspended following significant increases in water supply due to technological innovations including improved irrigation systems and desalination efforts.

Comparing the effectiveness of education, information and awareness interventions to that of TUBs

Due to the lack of reliable measurements of the effectiveness of general education, information and awareness campaigns, it is difficult to compare them to TUBs.

There appears to be consensus that information and education alone are unlikely to be effective. Yet, many other types of interventions seem to benefit if they are accompanied with information materials.

3.2.5 Installation of water saving devices

Overview

A number of studies have been carried out, examining water saving reductions stemming from the installation of **water saving devices** (WSD). Such devices include for example:

- shower aerators;
- tap inserts;
- shower timers;
- dual flush converters;
- cistern displacement devices;
- trigger hose guns; and
- efficient shower heads.

These studies tend to fall into one of three categories,

1) retrofitting by a trained plumber with water saving advice;
2) retrofitting by a trained plumber without advice; and
3) self-installation of WSD.

There is a concern in the literature that the installation of WSD is insufficient to cause reductions in water consumption. This is because it is viewed that **WSD need to be combined with educational programmes to inform users on effective use of WSD**. WSD alone may even lead to undesired outcomes, such as customers taking longer showers or leaving (efficient) taps running when not being used (e.g. Omambala, 2010).

To test this, Ross (2015) looked at the impact of messaging provided as part of a UK water company's WSD retrofit programme. Of the sample of customers receiving the retrofit, the treatment group was given standard behaviour change information whilst the control group was not given such advice. The treatment group saw a water saving of 38% or 18L per property per day more than the control group, thereby highlighting the benefit of combining educational information with water saving devices. Omambala et al. (2011) observed that consumption falls immediately following the installation of WSD and rebounds by a small amount thereafter before stabilising at a higher but still significantly lower level compared to pre-WSD installation in the long run.

Methods for encouraging uptake of water saving devices include letters, telephone calls, door-to-door, community events and the establishment of water saving community groups. The literature reveals that the door-to-door approach is met by customers with hostility, perhaps due to the association with door-to-door sales. Letter drops need to have a clear message and be accompanied with easy methods of responding. Telephone calls should be used as a follow-up after a letter is sent and it is vital that the calling staff are well trained (Bremner et al. 2012).

Effectiveness and robustness of existing initiatives around water saving devices

The findings from these studies confirm that installation of WSD results in lower water consumption, however the exact drop in consumption varies across studies. Overall, **indoor water use decreased substantially by between 10-40%** where the exact magnitude of effects depends on a variety of characteristics such as location, household size and type of intervention.[43] The most robust methods were used in the UK and found effect sizes of **reduced water consumption per property per day of 18-26L, or 20.6% over a period of 2.5 years** (Ross 2015, Omambala 2010, 2011).

According to the evidence, **the most effective intervention is retrofitting by a trained plumber supplemented by advice on how to utilise the water saving devices**.

The measured effects for groups of households, or even individual households are highly reliable as the measurements are usually taken from meters and caused by mechanical changes in the water supply of an individual tap, shower or toilet through the WSD. However, only a few studies looked and tested how WSD are best advertised and distributed (Ross (2015)). It is therefore difficult to conclude what the overall effect size would be if WSD were offered to all water customers in a given area, or how they might be offered most effectively.

Comparing the effectiveness of water saving devices to that of TUBs

Comparing the effectiveness of WSD to TUBs, it seems that **WSD can be a highly effective way of reducing water consumption in a lasting way**. The individual reductions in consumption seem to promise that WSD could cause substantial water savings and that these savings are fairly permanent and once stabilised remain constant over time. A remaining issue however regards the fact that WSD quality and therefore effectiveness degrades over time and **many devices require replacement over a number of years (8-10 years) and savings are only attained while devices are properly fitted and not removed**.

Therefore, if WSD were distributed to and properly installed in large numbers of households, then it is likely that the **effects could be much larger compared to TUBs**.

4 How to pilot interventions for managing water demand

Following the assessment of various types of interventions to manage water demand ranging from TUBs to alternative interventions, it is useful to assess how water companies or other authorities could find out what would work best in their specific circumstances. New interventions could then be piloted such that the interventions that are finally implemented are known to work and lead to the desired effects. Hence, this section provides:

- An overview in section 4.1 of various piloting methods, such as:
  - Consumer surveys and stated behaviour research;
  - Laboratory-based behavioural experiments;
  - Online behavioural experiments; and
  - Field behavioural experiments (also known as randomised controlled trials (RCTs))
- A brief outline of how online experiments can be used to pilot water demand interventions in section 4.2; and
- Finally, a detailed methodology for using randomised controlled trials in section 4.3.

[43] See for example: Mayer et al. (2013), Lev (2012), Roccaro et al. (2011), Tsai et al. (2011), Hastings and Rustamov (2015), Manouseli et al. (2016), Omambala et al. (2011). Despite the sizeable effects of the literature, there are several methodological shortcomings of some individual studies such as: small samples (Mayer et al. (2003), Roccaro et al. (2011)); and lack of control groups. Therefore, it cannot be excluded that the measured effects, of experiments without a control group, are due to confounding factors such as the weather.

4.1 Overview of methods

In the table below, we describe a non-exhaustive list of methods which could be used, and which would be useful and cost effective for assessing the effectiveness of TUBs, or alternative interventions. We explain survey-based methods followed by various types of behavioural experiments.

Behavioural economic experiments are a collection of inductive scientific methods, designed to **evaluate how consumers behave in different (sometimes simulated) situations** (OECD, 2012). Experiments study the effect of a specific variable or feature on consumer decision-making, in a controlled, stylised environment, which nonetheless mirrors the principal characteristics of the choice and market environment being investigated. The results of experiments provide fundamental and often causal insights on consumer behaviour, which inform policy or business decisions.

Behavioural experiments can be used by water companies to test foreseen changes to water management interventions (pre-testing in an online or laboratory experiment), or to **evaluate the effectiveness of interventions** directly (in-field testing in an RCT). The main advantage compared to consumer surveys is that experiments can be incentivised and therefore provide more robust and realistic results.

Method	Description	Strengths	Weaknesses	Use for managing water demand
Consumer surveys - stated behaviour research	• A range of survey questions are asked to a broad, ideally representative group of consumers. • Questions are hypothetical in nature and ask individuals about their hypothetical reactions, or understanding of materials such as information messages, TUB warnings provided.	• Quick and cost-effective, especially if conducted online. • Can use representative consumer group with broad demographic backgrounds.	• Low external validity (Reactions to interventions are hypothetical, may not translate into real-life behaviour. • No counterfactual comparison. Reach limited to consumers who have internet access • Telephone / face-to-face interviews are lengthy and costly	• Useful as pre-research to inform development of interventions or to complement experimental research
Laboratory-based behavioural experiments	• Behavioural experiment conducted in laboratories - typically at universities • Often used to test specific design features of a policy or intervention.	• Quick and cost-effective to implement. • A counterfactual is assessed by allocating participants to control and treatment groups. • Maximum control over what actions and choices are available to participant. • Easy recruitment via laboratory subject pools. • Strongest internal validity. • Monetary incentives possible.	• Limited external validity because: ○ Highly stylized environments. ○ Participants are typically university students or limited to geographic reach of laboratory. ○ Monetary incentives can be low.[44]	• Could provide insights into how consumers might respond to various pieces of communication regarding water management interventions.
Online behavioural experiments	• Online experiments are conducted on a virtual interface, generally across a wide geographical area with a participant sample that matches/mirrors the specific population of interest. • Market research firms typically recruit representative consumer samples from dedicated panels	• Quick and cost-effective • A counterfactual is assessed by allocating participants to control and treatment groups. • High-level of control and internal validity • Easy recruitment of large, representative samples from curated panels. • Monetary incentives possible	• Need to be complemented with other methods in areas/regions where internet penetration is low. Limited external validity like laboratory experiments. • Tested scenarios can be incentive compatible but usually remain of hypothetical nature. Lower level of control compared to laboratory experiments	• Could provide insights into how consumers might respond to various pieces of communication regarding water management interventions.
Field behavioural experiments (aka	• Field experiments are distinct in that they do not simulate an environment but are set in reality.	• Considered the gold standard in experimental methods.	• Limited internal validity - given the potential low level of control over the decision environment, it	• Could trial information campaigns about TUBs provided in certain

randomised controlled trials (RCTs))	• There are two types of field experiments: randomised controlled trials & natural experiments. • The former refers to a behavioural experiment conducted in an everyday/real-life environment, wherein the experimenter manipulates the success measures by introducing a targeted intervention but has no control over other environmental factors. • In a natural experiment, the experimenter has no control over any measures, as they occur and vary naturally across the sample for example due to a change of law externally implemented.	• Counterfactual is established by randomising how/where the intervention is trialled. A control group remains unchanged. • Maximum external validity as behaviour is observed in naturally occurring environment. Appropriate for testing various sorts of interventions to which other experiments (online/laboratory) may only provide limited evidence. • Large samples. • Participants are unaware that they are involved in an experiment, which controls for systemic biases.	can be difficult to conclude on causality. • Confounding factors might make especially null results difficult to interpret. • Can be complex and time-consuming to implement Sometimes long measurement periods are required.	community centres, or on water bills for just a sub-group of all of customers and then measure how behaviour changes following the introduction of changes.

[44] Levitt & List (2005) discuss how low stakes, a small/non-representative sample pool and the artificial laboratory environment, mean that the results of such experiments do not extrapolate to the real world. However, there have been many case studies wherein laboratory experiments do carry over accurately into more realistic settings. Evidence from highly stylised laboratory experiments has previously been used to inform regulation and public policy.

4.2 Outlining online experiments

In this section, we provide more detail on how online experiments are conducted. We chose to focus on online experiment because, as visible in section 4.1, they have many advantages and relatively few disadvantages. They can therefore be **a cost-effective way of measuring the likely effectiveness of various types of interventions** or serve as pre-testing evidence prior to conducting an RCT (explained in more detail in section 4.3).

Online experiments are especially **useful to test customer communication strategies** to understand how different types of messages are understood and perceived.

An online behavioural experiment is typically embedded in a consumer survey. Instead of asking standard survey questions, a behavioural experiment asks respondents to take decisions or complete tasks which are shown on screen. Such tasks could mimic the steps individuals take to sign-up for an online account with their water company or show a water bill and consequently ask for opinions, understanding of content, or actions to take following the exposure to information. Next and unknown to participants, the experimenter systematically varies how information is presented to different groups of respondents, or which options are available for respondents to take. These variations are called 'treatments' and allow the experimenter to test the effects of specific interventions compared to a baseline condition.

The following are key design components which need to be considered for any online behavioural experiment:

- **Defining the experimental environment:** The experiment environment is the setting in which consumers are observed as they take decisions/actions. For example, the experimental environment could mimic the way in which consumers access and view their water bills.
- **Devising the intervention as experimental treatments:** The treatments are a key part of any experiment design. These introduce controlled variation in the testing to examine the impact of, for example, different information frames (i.e. information content and presentation) and stimuli on individuals' behaviour, targeting the research questions and issues at stake. Respondents are randomly (and equally) allocated across the different treatments in any experiment. The effectiveness of a specific intervention – i.e. treatment - would be assessed by comparing behaviour and success criteria (see below) of respondents assigned to different treatment groups, or between a control and a treatment group. Multiple treatments can be tested within the same online experiment. How many treatments can be tested exactly depends on the types of treatments, their expected effectiveness and the overall sample size of the experimental test.
- **Defining success criteria:** Success criteria are the performance measures that are used to measure the effectiveness of an intervention. The success criteria will be designed to examine the specific behaviours the intervention is trying to target. For example, if the objective is to test messages which are most effective at communicating the introduction and functioning of TUBs, possible success criteria could be measuring:
  - The **level of understanding** of the messages. This is to assess whether consumers understand what is communicated to them, or asked of them.
  - **Engagement with the message**. This could be assessed using follow-up questions prompting individuals on elements they did (or did not) see on the water bill, ranking elements they remembered or appraising the interest/relevance of shown content.
  - **Actions in response to messages**. This could be done in different ways. For example, either by observing actions and decisions in a follow-up experimental task, or by inviting

participant to state which actions they would likely take following reading the shown message on their water bill. For example, one could ask participants to indicate how likely they would respect and follow the rules indicated by the TUB. The realism of such questions can be enhanced by using realistic scenarios in which actions are based.

- **Considering monetary incentives**: Monetary incentives can increase the realism and saliency of experimental tasks as they simulate the stakes involved in similar real-life decisions. Whether and how monetary incentives should be included in any specific online experiment needs to be decided during the design phase – there are valid arguments for and against using incentives. When incentives are included, these are typically paid to respondents in 'survey points' which can be converted into cash, or cash vouchers for popular shops in the UK, or online shops.
- **Analysing the data**: The obtained data from the online experiment (and survey) can be analysed in a multitude of ways. Firstly, descriptive statistics give an overview of how respondents behaved in the experiment in general. Next, inductive statistical methods such as inference testing (e.g. t-tests, z-tests) and/or regression analyses are performed. These tests seek to assess the effectiveness of the tested interventions (i.e. treatments). For example, to assess whether a specific information message was effective, it would be tested whether the level of understanding, or level of engagement with the message in the experimental group who saw the message was significantly higher compared to the control group who did not see the message. If a statistical test shows significance, the effectiveness of the intervention would be concluded. (see further explanations of how to analyse experimental data in section 4.3.4).
- **Sampling and recruiting:** While online experiments can be designed by behavioural specialists, they are typically hosted by market research firms on proprietary platforms. Such firms curate panels which consist of consumers with various socio-demographic characteristics. Depending on the objectives of the interventions which should be tested, the client may be able to define the sample of consumers which should take part in the experiment. For example, consumers in a certain geographic region or age group could be targeted. The closer the target sample matches characteristics of the actual population who may later be exposed to the intervention, the higher will be the external validity of the results. This means that it is highly likely that the effects observed in the online experiment will closely correspond to effects expected in reality.

4.3 Implementing a pilot intervention via a randomised controlled trial – a detailed methodology

This section contains a detailed methodology for piloting interventions for managing water demand using randomised controlled trials (RCTs) to robustly assess the interventions' effectiveness.[45] RCTs could test various types of interventions[46] which aim at influencing water demand, during droughts or in general, such as:

- Temporary use bans or other types of water restrictions;
- Messaging campaigns containing for example information on water saving, calls to save water or other types of messages (e.g. on customer bills, on the internet, social media, in local newspapers, radio etc.);

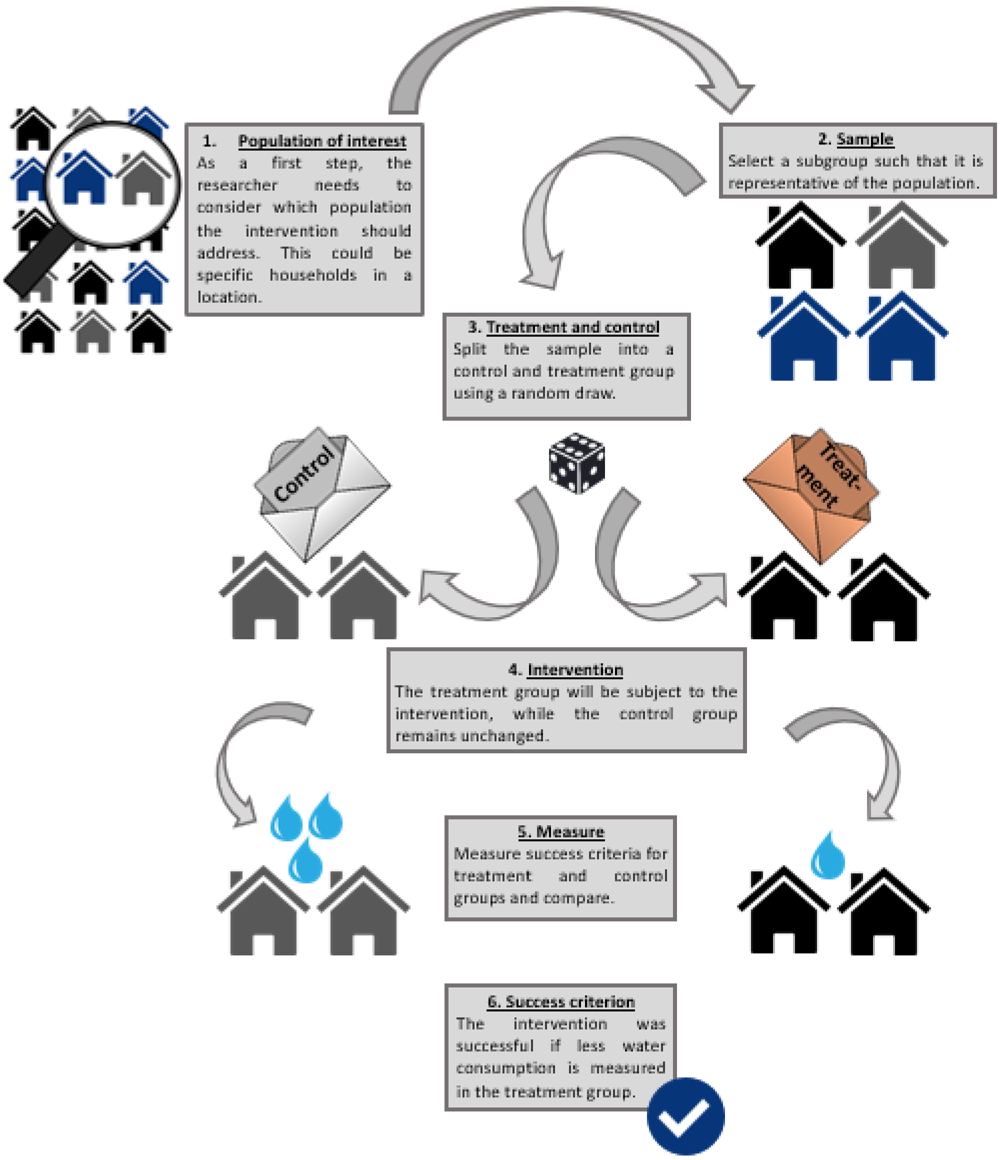
[45] London Economics have written a similar methodology for the UK Department for Transportation to tackle roadside littering. See Department for Transport (2018) 'Reducing roadside litter using randomised controlled trials'.

[46] For more information on various types of interventions see section 3.

- Feedback provision for example through meters, smart meters or other types of in-home displays;
- Interventions in the water infrastructure (e.g. installation of water saving devices); and
- Interventions which change the financial incentives of water demand.

As seen before, **RCTs are not the only methodology available for piloting interventions, but they are one of the most robust approaches**. An overview of the methodology is provided in the infographic here below:

Figure 3 The key steps in RCTs



Source: London Economics

One of the key challenges in assessing the effectiveness of water demand interventions lies in establishing a control group, or counterfactual scenario. This means establishing *what would have happened if the intervention had not been introduced*. As seen in previous sections, often, the control group, or counterfactual is of questionable quality. For example, the water demand of the same months in a previous year are used to measure how water demand in the current year has changed. However, such comparisons can have weaknesses due to so-called 'confounding factors'. The difference between last year's water consumption and consumption this year might not only be different due to the introduction of a new water demand intervention. Instead, it could be different because of different meteorological circumstances, other things happening in the region (e.g. sports events, holiday periods) or other shocks to water demand (e.g. due to households purchasing new appliances which consume water).

The key advantage of RCTs therefore is that they establish a counterfactual by splitting the overall population into treatment and control groups. Only the treatment group would be subjected to the intervention while the control group remains completely untouched and so provides the counterfactual. By measuring simultaneously water consumption in the control and treatment groups, we can therefore **establish the pure and causal effect of the intervention**.

RCTs can be relatively simple, quick, and thus cost-effective to set up. Nonetheless, an experienced researcher who is familiar with the limitations and risks of the methodology should accompany the design and implementation.

All of the following steps assume that an intervention is ready for testing (see section 3 for examples of interventions which have previously been assessed).

4.3.1 Defining a sample for the testing

An early step in the setup of an RCT is the definition of the test sample. The sample is drawn from the overall '*population'* for which the pilot is trying to assess the effectiveness of a water management intervention. The population could, for example, consist of all **water consumers in a certain region** (e.g. in all of Yorkshire). The test sample should be a representative subgroup of this overall population. For example, this could mean that the sample would consist of water consumers from only one or a couple of distribution areas only.

For being representative, **the test sample should share key characteristics with the full population**. As illustrated in the Figure below. For example, water demand and the demographic profile of customers in the test region should be similar to that of the overall population. For example, if most of Yorkshire's customers live in rural areas with large garden surfaces, a pilot test should also be conducted in a rural area and not solely focus on interventions in an urban setting. If a test were conducted only in urban areas (e.g. in the city of Leeds), the results may be misleading and show limited (or more pronounced) effectiveness of the intervention compared to when it is rolled out across the whole of Yorkshire in rural and urban areas alike.

Figure 4 Selecting a representative sample for testing

The sample should be representative of the underlying population. This means that it should have similar characteristics to the rest of the population.

Source: London Economics. All icons taken from Pixabay.com licensed CC0.

Any sampling characteristics should be observable (i.e. measurable) and ideally be linked to a) the level of water demand (e.g. size of household, garden) and b) the expected effectiveness of the intervention (e.g. demographic of the household).

A number of potential sampling characteristics are described here below:

Box 3 Possible sampling characteristics

The following provides guidance on the types of characteristics that could be used to select a representative sample for an RCT to test a water demand intervention. If, for example, only households with large gardens were targeted by the intervention, it would be beneficial to characterise also the geographic location of these households and meteorological conditions in which such households are located.

The actual selection of sampling criteria will depend on the availability of data on those criteria for potential members of the sample. Nevertheless, it is useful to initially assess as many characteristics as possible before narrowing the criteria based on their feasibility.

- **Region**:
  - Rural vs. urban
  - Nearby presence of rivers, and reservoirs
- **Household composition**:
  - Number of people living in the household
  - Size of the home
  - Size of the garden

- Socio-demographics of the household (e.g. presence of children, age profiles, income, indicators for potential vulnerability)
- How long customer has been paying bills
- Customer owns or rents the home

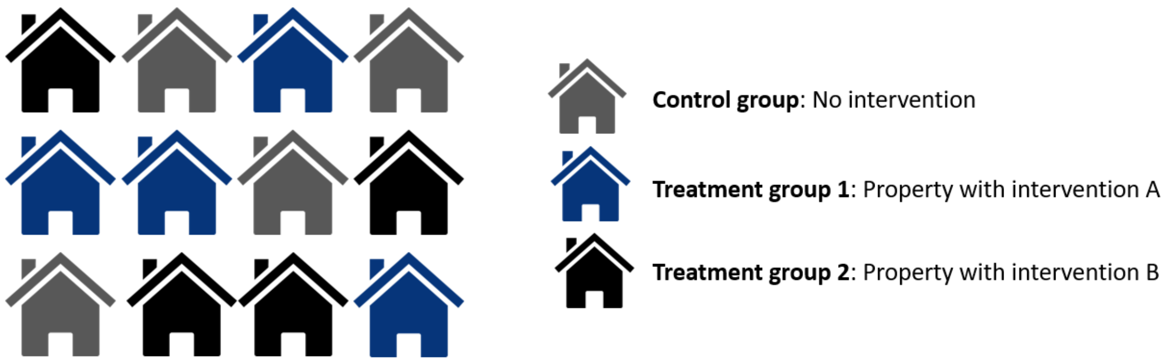
- **Water demand characteristics**:
  - Water consumption is measured (i.e. (smart) metered, measurable via other devices) at household level
  - Water consumption is measured at street, regional or other level
  - Tenure of customers
- **Other customer characteristics**:
  - Billing frequency
  - Payment methods
  - Access to information via internet, social media, apps

4.3.2 Defining Treatment and Control Groups

Once the test sample is defined, it needs to be split into treatment and control groups. **Only the treatment group(s) will receive the intervention(s)**. All circumstances must remain unchanged for the control group because it uniquely serves for measurement purposes. This also means that the control group should remain as isolated as possible from the treatment group(s) to avoid contamination effects from the intervention.

There will usually be only one control group while there could be multiple treatment groups to test different interventions, or different variations of the same intervention (see Figure 5). To keep the later analysis of the RCT (see section 4.3.4) as simple as possible, the control and treatment group(s) should be of similar size.

Figure 5 Assigning treatment and control status to groups



The selected sample of households will be split into groups: some groups will receive the intervention(s) – the treatment group(s) – at least one group will not receive the intervention – the control group (or baseline).

Source: London Economics

Key to splitting the sample into control and treatment groups is that the split should be done using **a random draw** (e.g. a coin flip). However, there are alternative ways to make this randomisation, which mainly depend on the size of the sample and are described in turn in the following subsections.

Large sample RCTs

Most commonly, as seen also in many of the presented interventions in section 3, RCTs have large samples sizes. This means, they typically comprise observations in the hundreds, or even thousands in each test group. This would be the case, for example, when an intervention is tested at the household/customer level.

To setup the RCT, the sample would be split into two or more groups, and each group would be assigned control or treatment status. This could, for example, be done using a random number generator, tossing a coin, or rolling a dice.

Such **randomisation of a large sample will ensure that the different groups resemble each other on observable characteristics** (see Box 3 for examples). This is true because of the law of large numbers, which says that the outcome of a random draw will resemble the population average if repeated many times. As a result, we would expect the different test groups to be very similar, as long as they are sufficiently large. The researcher may do so-called randomisation checks to verify whether the samples are indeed similar. This could, for example, imply comparing the average household size across the different groups.

Small sample RCTs using matched samples

In case a large sample RCT is not feasible, fewer observations (i.e. a small sample) can be arranged to mimic the large sample setup. This could be the case when measurement information is not available at the household level but is instead available on a more aggregate level (e.g. at street, or wider regional level), for example, when water demand is not individually metered for every customer.

With large sample RCTs, the law of large numbers guarantees that the different test groups are similar. For small sample RCTs, this is not the case. It is instead **necessary to construct the control and treatment groups such that they are as similar as possible on specified characteristics** (see Box 3 for possible characteristics) through careful selection and assembly of the groups.

For instance, if a test should consist of two different regions/districts, they should resemble each other in key characteristics such as:

- Household density;
- Meteorological conditions;
- Socio-demographic composition of the population, etc.

Ideally, we would observe that in **historic data**, the regions track each other in terms of movements in consumer water demand. This means, that it is not a problem if one region has a generally higher consumption compared to another, **what is important is that any fluctuations**, for example due to weather events, **appear similarly in the data**.

This together with general similarity between the regions is key in order to conclude later in the analysis that the only differences which occur between groups are causal to the tested intervention.

In a final step, the assignment of which group will receive the treatment, and which will be the control group should be done at random. Assignment should **not** be chosen deliberately based on any characteristics as this could affect the results and interpretation of the RCT. For instance, the treatment should not be assigned to the region which is expected to respond more strongly to the intervention as this would produce a biased result of the overall effectiveness of the pilot compared to a full roll-out.

4.3.3 Success Criteria

Alongside designing the RCT and the interventions, the researchers need to decide how the success of the intervention will be measured. To this end, success criteria, or performance measures need to be put in place.

In water demand interventions, the **outcome measures should ideally involve water consumption** of the affected control and treatment groups during the intervention.

Other auxiliary measures may be collected such as:

- Consumer awareness of details of the intervention (e.g. are consumers aware that TUBs have been issued, that smart meters are being rolled out, water saving devices offered for installation);
- Consumer understanding of the interventions (e.g. do consumers understand the details of TUBs, do consumers understand the messages they received as part of the intervention);
- Consumer reactions (e.g. on social media, in complaints registers etc.);
- Consumers' perceived changes in demand (e.g. whether they think they changed their consumption in response to an intervention); and
- Consumers' intention to change (e.g. whether consumers indicate that they would likely change their water consumption in response to an intervention).

Most importantly, the same measures or methods of measurement should be used in the control and treatment group to ensure unbiased final results.

4.3.4 Analysis

Once the RCT has been implemented and data is collected, the researchers need to analyse the data to draw conclusions on the effectiveness of the tested interventions.

The most important piece of analysis will be to **measure the differences between the control and treatment group(s)** – i.e. the treatment effect(s). This can be done by measuring the success measures (see section 4.3.3) separately for the control and treatment group(s).

The researchers will be able to conclude that **the intervention was successful if the success measures have improved in the treatment group compared to the control group**.[47] For example, a messaging campaign was successful if following the distribution of the messages water consumption in the treated sample households/regions was lower compared to areas without the intervention.

[47] More detail on statistical significance can be found in Annex 1.

Box 4 Case study: Results of a hypothetical messaging RCT

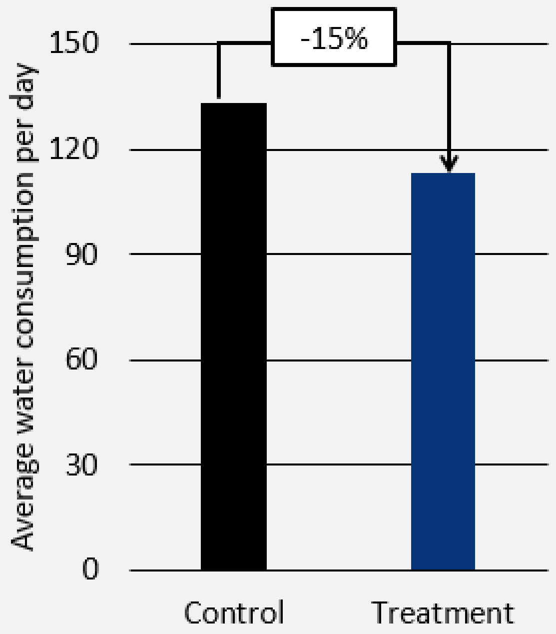
A new messaging campaign has been tested in two different distribution areas (e.g. towns). There were roughly 2,000 households in each of the two distributions areas. Prior to the test, the households in the two distribution areas resembled each other on numerous observable characteristics, such as socio-demographics, and geography, and followed similar seasonal water demand patterns in historic measurements.

A coin flip has assigned one area as the control, and the other as the treatment group which was later exposed to the new messaging campaign.

Water demand was measured in both areas for 2 months following the distribution of messages. During this period, water consumption in the control area corresponded closely to trends in the previous year and was around 133L per household per day. In the treatment area instead, water consumption was only around 113L per household per day. This represents an 15% reduction in water consumption in the treated area compared to the control area which is attributable to the tested new messaging intervention.

The tested intervention has therefore significantly reduced water consumption as desired and it can be concluded as successful.

Figure 6 Illustration of the measured treatment effect of reduced water consumption in the measurement period



5 Conclusions

This white paper has assessed existing evidence on the effectiveness of water restrictions such as TUBs. This was done using a rapid evidence assessment of a variety of sources such as academic papers, government agency reports and research conducted by utility providers.

Following the assessment of water restrictions, evidence on alternative interventions, especially those motivated by behavioural economics have been evaluated. This was done in order to uncover how water demand may be best managed given the pros and cons of various methods.

Finally, the white paper reviewed and summarised piloting methodologies for testing the effectiveness of interventions in order to contribute to the evidence base on how water demand could be managed most effectively.

5.1 Main findings

Evidence on the effectiveness of water restrictions

The white paper found that the **evidence** on the effectiveness of temporary use bans (TUBs) **is relatively scarce**. At the same time, the **quality and robustness** of some available assessments are **questionable**, or at least highly circumstantial.

The most reliable sources come from the Environment Agency (2013) and UKWIR (2007). The former found that **the 2012 TUBs had no measurable impact on water consumption**. This was likely because the weather during the TUB was very wet and therefore required low compliance.

The **2006 TUBs significantly impacted consumption levels** according to both sources. According to UKWIR (2007) in the absence of wet weather, as seen in 2012, TUBs would be expected to reduce water use at peak times 5-10%. Yet the authors caveat that these effects were likely circumstantial and might not translate to other TUBs.

Evidence from the USA and Australia showed that **voluntary restrictions are unlikely to be effective**. Mandatory water restrictions reduced consumption by 4-23% and were most effective when combined with financial incentives. The largest water savings were achieved by large and wealthier households.

Evidence on how consumers react to water restrictions

In terms of how consumers react to water restrictions it can be concluded that **consumers widely seem to misunderstand TUBs**. Most consumers are **unable to pinpoint which types of water uses are restricted** exactly.

Qualitative evidence suggests that **consumers would broadly accept TUBs especially in severe weather conditions**. However, this **acceptance falls drastically** when consumers feel that water levels have fallen to critical levels because of **mismanagement**, or other possibly avoidable causes such as leakages.

Multiple sources, from the UK and internationally, suggest that **consumers prefer TUBs to seeing their bills increase**. This is because they feel that restrictions imposed by TUBs are often not unsurmountable, or regard non-necessary types of usage only.

Academic research has shown that **the strongest driver of compliance with water restrictions is an individual's perception of their capacity and ability to comply**.

Evidence on the effectiveness of alternative interventions

Given the uncertainty about the effectiveness of water restrictions and the fact that they are widely unpopular among consumers, this paper looked at what alternative types of interventions are available to water companies and authorities to manage water demand.

Various types of alternative interventions to TUBs were assessed such as:

1) Messaging interventions (see section 3.2.1)
2) Feedback provision on consumption behaviour (see section 3.2.2)
3) Interventions which change the financial incentives of resource demand (see section 3.2.3)
4) Education, information and awareness initiatives (see section 3.2.4)
5) Installation of water saving devices (see section 3.2.5)

The evidence gathered looked at resource conservation in the water and energy sectors.

Often, especially for 1)-3), assessments used robust state-of-the-art methods such as randomised controlled trials (RCTs) through collaborations between academics and water suppliers or local authorities. Education initiatives (4), instead were often initiated by governments and broadly targeted which made measurement of effectiveness more difficult.

The main benefit of alternative interventions compared to TUBs is that they are **not unpopular**. In fact, well designed interventions such as **social norm messages provided on bills can be engaging and therefore of value to consumers**.

Many alternative interventions have been found to be similarly, if not more effective than water restrictions. This makes them attractive as they can be quick and cost effective to implement while not triggering negative consumer reactions.

Messaging interventions, especially those using strong social norm statements, have been widely shown to significantly reduce water consumption by approximately 1-5% in the short term and by 1.3-2.6% in the long term. This suggests they are **similarly as effective as TUBs**.

Feedback provision, for example via (smart) meters or in-home displays, seemed to trigger mixed reactions: **sometimes large reductions in consumption are achieved** (larger than the effects of TUBs), but these tend to be short-lived and **fade over time**. On other occasions, feedback provision is not as effective. This is especially the risk for water consumption which is more essential and therefore seems less elastic compared to energy consumption.

Changing prices or providing financial incentives for resource conservation requires feedback provision to be effective. Such measures **could be used to complement TUBs** to increase the effectiveness of water restrictions.

Education initiatives alone were found to be unlikely to be effective. They could however usefully complement other initiatives as sound consumer understanding is required for other measures such as messages, feedback or water saving devices to be most effective.

Water saving devices have been found to cause **large and lasting reductions** in water consumption, especially if they are fitted and explained professionally. There is little evidence on how water saving devices are best promoted, distributed and fitted at a large scale. But it seems that they could be most effective at impacting water consumption – though reduction would not be targeted and limited to manage short-lived periods of drought.

How to pilot interventions for managing water demand

Various types of piloting methodologies are available to water companies and other authorities. Each method has its strengths and weaknesses. These relate especially to:

- how quick and effective the method is to implement;
- how realistic the environment is in which the pilot is implemented, which affects how 'externally valid' the resulting findings are (i.e. how close the results are to what we would expect to see in a full roll out of the intervention); and
- how robustly it can be concluded that any measured effects in the pilot were effectively caused by the piloted interventions (this is also related to the 'internal validity' of the methodology).

Based on these criteria it can be concluded that experimental methods – i.e. simulated online experiments and, in particular, so called 'field experiments' (RCTs) with actual consumers – provide the best evidence.

Online experiments have the advantage of being conducted in a fully controlled environments and can be used to (pre-)test various types of interventions. However, **field experiments, or randomised controlled trials, provide the most robust and valid evidence** as they are implemented with real consumers, and are broadly considered as the gold standard for assessing interventions in resource conservation.

Bibliography

Abrahamse et al. (2005), A review of intervention studies aimed at household energy conservation, *Journal of Environmental Psychology*, 25 (3).

Alcott (2011), Social norms and energy conservation, *Journal of Public Economics*, 95 (9-10).

Asensio, Delmas (2015), Nonprice incentives and energy conservation, *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 112(6).

Brennan et al. (2007), The welfare costs of urban outdoor water restrictions., *The Australian Journal of Agricultural and Resource Economics*, 51.

Brent et al. (2017), Are Normative Appeals Moral Taxes? Evidence from a Field Experiment on Water Conservation, *Departmental Working Papers LSU*.

Brick et al. (2017), Behavioural Nudges for Water Conservation: Experimental Evidence from Cape Town, *WRC Working Paper*.

Buchanan et al. (2015), The question of energy reduction: The problem(s) with feedback

Castledine et al. (2014), Free to choose: Promoting Conservation by Relaxing Outdoor Water Restrictions, *Journal of Economic Behavior & Organization,* 107 (A).

Cooper (2016), What drive compliance? An application of the theory of planned behaviour to urban water restrictions using structural equation modelling, *Applied Economics*, 49 (14).

Curran, Doyle (2011), Picture superiority doubly dissociates the ERP correlates of recollection and familiarity, *Journal of Cognitive Neuroscience*, 23 (5).

Department for Transport (2018) Reducing roadside litter using randomised controlled trials.

Dessai, Sims (2011), Public perception of drought and climate change in southeast England, *Environmental Hazards*, 9 (4).

Ecopeace Middle East (formerly FOEME) (2009), Best Practices in Domestic Water Demand Management.

Environment Agency (2017), Drought response: our framework for England.

Environment Agency (2013), Quantifying the Impact of Water Company Drought Measures on Water Demand.

European Commission (2018) prepared by Ipsos-LE-Deloitte consortium, Consumer study on precontractual information and billing in the energy market improved clarity and comparability.

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

Ferraro, Miranda, Price (2011), The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment, *American Economic Review: Papers & Proceedings 2011,* 101 (3).

Festinger (1957), A Theory of Cognitive Dissonance, California: Stanford University Press.

Goldstein, Cialdini, Griskevicius (2008), A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels, *Journal of Consumer Research*, 35, (3).

Hahn et al. (2016), The Behavioralist as Policy Designer: The Need to Test Multiple Treatments to Meet Multiple Targets, *NBER Working Paper No. 22886.*

Hastings, Rustamov (2015), Potential Best Management Practice: Customer Water Use Messaging, Prepared for California Urban Water Conservation Council.

Hensher et al. (2006), Water supply security and willingness to pay to avoid drought restrictions, *The Economic Record,* 82 (256).

Jessoe, Rapson (2014), Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use, *American Economic Review*, 104(4).

Kantola, Syme, Campbell (1984), Cognitive dissonance and energy conservation. *Journal of Applied Psychology*, 69(3).

Klege et al. (2018), The Effectiveness of Competition and Responsibility Assignment in Saving Energy: A Non-Residential example of the Power of the “Nudge”, ERSA working paper 763.

Lev (2012) for the State of Israel Water Authority, Israeli Experience in Water Saving in The Municipal Sector, available at: http://www.water.gov.il/Hebrew/ProfessionalInfoAndData/2012/17-Israel-Water-Sector-Water-Saving.pdf.

Manouseli, Kayaga, Kalawsky (2017), Evaluation of water efficiency programs in single-family households in the UK: a case study. Water Science and Technology: Water Supply, ws2017071.

Mini et al. (2014), The effectiveness of water conservation measures on summer residential water use in Los Angeles, California, *Resources, Conservation and Recycling*, 94.

Monahan et al. (2010), Benefits of 'Observer Effects': Lessons from the Field. U.S. National Library of Medicine, National Institutes of Health.

Mayer et al. (2003), Residential Indoor Water Conservation Study: Evaluation of High Efficiency Indoor Plumbing Fixture Retrofits in Single-Family Homes in the East Bay Municipal Utility District Service Area.

Nicherson, Rogers (2010), Do You Have a Voting Plan? Implementation Intentions, Voter Turnout, and Organic Plan Making, *Psychological Science*, Vol 21, No. 2.

OECD (2012), Enhancing Consumer Policy Making: The Role of Consumer Surveys.

Ofgem (2017) report prepared by Cambridge Economic Policy Associates Ltd ‘Distributional Impact of Time of Use Tariffs’, available at:

https://www.ofgem.gov.uk/system/files/docs/2017/07/distributional_impact_of_time_of_use_tariffs_1.pdf.

Omambala, (2010), Evidence Base for Large-scale Water Efficiency: Water Efficiency Retrofitting in Schools. Phase II Second Report, available at: https://www.waterwise.org.uk/wp-content/uploads/2018/02/Evidence-base-for-large-scale-water-efficiency-in-homes.phase-ii-interim-report.pdf.

Omambala, et al. (2011), Evidence Base for Large-scale Water Efficiency Phase II Final Report. Waterwise, available at: https://www.waterwise.org.uk/wp-content/uploads/2018/02/2011_Waterwise_evidence_base_phase_two_final_report.pdf.

Orland, et al. (2014), Saving energy in an office environment: A serious game intervention, *Energy and Buildings*, Vol 74.

Osbaldiston, Schott (2012), Environmental Sustainability and Behavioral Science: Meta-Analysis of Proenvironmental Behavior Experiments, *Environment and Behavior,* 44 (2).

Roccaro, Falciglia, Vagliasindi (2011), Effectiveness of water saving devices and educational programs in urban buildings. *Water science and technology: a journal of the International Association on Water Pollution Research*, 63(7).

Schwarz et al. (1991), Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*. 61 (2).

Tiefenbeck et al., (2016), Better Get Focused: How Feedback on a Specific Behavior Can Reduce Energy Consumption, *International Energy Policy & Programme Evaluation Conference, Amsterdam, Netherlands, June 7-9, 2016*.

Tsai, Cohen, Vogel (2011), The Impacts of Water Conservation Strategies on Water Use: Four Case Studies. *Journal of the American Water Resources Association*, 47(4), 687-701.

UKWIR (2007), Modelling the impact of restrictions on demand during drought.

UKWIR (2013), Understanding the Impacts of Drought Restrictions.

UKWIR (2014), Managing Through Drought: Code of Practice and Guidance for Water Companies on Water Use Restrictions.

Wichman, (2017), Enduring effects of changes in billing frequency: Evidence from urban water use, unpublished working paper.

Yorkshire Water (2018a), Water Restrictions Research.

Yorkshire Water (2018b), Reducing Demand for Water Research.

Index of Tables, Figures and Boxes

Abbreviations

EA	Environment Agency
IBT	Increasing Block Tariffs
IHD	In-home displays
RBC	Perceived Behavioural Control
RCT	Randomised Controlled Trial
REA	Rapid Evidence Assessment
TUB	Temporary Use Ban
UKWIR	UK Water Industry Research
WSD	Water saving devices
WRZ	Water Resource Zone
WTP	Willingness-to-Pay

Tables

Figures

Boxes

ANNEXES

Annex 1 Statistical analysis of treatment effects

In addition to analysing the effectiveness of the tested intervention using descriptive methods (e.g. measuring water consumption in litres, or comparing the level of awareness of a TUB in the treatment and control groups), statistical methods can be applied. These methods would assess the robustness and credibility of the measured effects of the RCT.

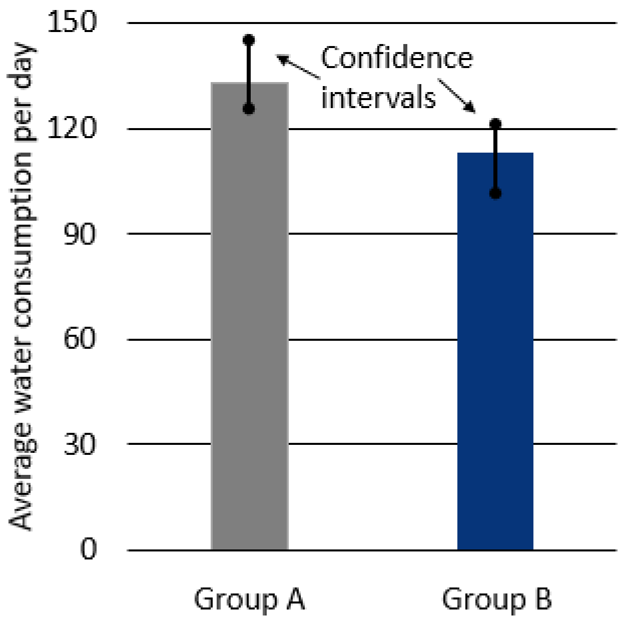
In a first step, this can be done by measuring the mean of the success measure, for example, the average water consumption per household during a specified period, in each of the test groups using descriptive statistics.

A further step then establishes whether any measured treatment effect is also statistically significant. This can be done using confidence intervals and t-tests.

- **Confidence intervals**: Confidence intervals illustrate the margin of error around reported results and are calculated for a chosen confidence level (typically 95%). The confidence level represents how 'sure' we can be that the true result lies within the confidence interval. Generally, the 95% confidence interval is calculated using the formula: Lower bound = $M – Z_{95}*\sigma_M$; Upper bound = $M + Z_{95}*\sigma_M$. Where M is the sample mean, Z_{95} is the number of standard deviations extending from the mean of a normal distribution needed to include 95% of the area, and σ_M is the standard error of the mean. If the sample means of two groups do not have overlapping confidence intervals, it can be concluded that the difference in means is statistically significant.
- **T-tests**: This test will establish whether a resulting treatment effect is statistically significant. The null hypothesis would be $L_1 = L_2$, i.e. that water consumption is the same in both groups, and the test would establish whether the null hypothesis can be rejected in favour of the alternative hypothesis $L_1 \neq L_2$, i.e. the average water consumption is not the same. The result of the test would depend on whether the t-statistic calculated from the data exceeds the relevant critical value.

In RCTs with sufficiently large samples, the analysis can be augmented by assessing treatment effects for different sub-groups within the test samples. For example, it could be assessed whether the effectiveness of the treatment was larger in magnitude in high consumption households, or in owner occupied households, compared to households with generally lower consumption, or in households who rent their home. Such analysis could comprise statistical tools such as linear regression analyses.

Figure 7 Assessing statistical significance of treatment effects



Notes: If the confidence intervals on the averages of the outcome measure in each of the test groups, such as the average water consumption in Groups A and B, are not overlapping, it can be concluded that the measured treatment effect is statistically significant. This means that the difference in measured water consumption is too large to be measured by pure coincidence. Instead it was effectively caused by the tested intervention.

Source: London Economics

LE London Economics

Somerset House, New Wing, Strand,
London, WC2R 1LA, United Kingdom
info@londoneconomics.co.uk
londoneconomics.co.uk
@LondonEconomics
+44 (0)20 3701 7700