Are Behavioral Nudges Moral Taxes? Evidence from a Field Experiment on Water Conservation

Working Paper

Daniel A. Brent*[a], Corey Lott[b], Michael Taylor[c], Joseph Cook[d], Kim Rollins[c], and Shawn Stoddard[e]

[a]Department of Economics, Louisiana State University
[b]Department of Economics, University of California Santa Barbara
[c]Department of Economics, University of Nevada Reno
[d]Evans School of Public Policy & Governance, University of Washington
[e]Truckee Meadows Water Authority

December 2016

Abstract

We investigate how the information content in nudges for water conservation drives behavioral change using a large-scale, randomized field experiment in Reno, Nevada. We develop a new social comparison that reduces the correlation between pre-treatment water use and the distance from the peer group, allowing us to plausibly isolate the impact of the normative component of the nudge. The strength of the social sanction, which we define as a household's performance relative to its peers, is a primary driver of social comparisons' efficacy, consistent with social comparisons operating as a moral tax on excess consumption. We introduce an alternative nudge that focuses on financial savings and information about the rate structure that generates a different pattern of behavioral change. The financial nudge is particularly effective with households who did not conserve in response to previous pro-social appeals, generates more persistent water savings, and is less dependent on multiple mailers compared to the social comparison. While both financial and moral motivations are likely driving behavioral changes in response to nudges, there are opportunities to improve welfare by designing nudges that prompt consumers to address internalities due to financial motivations.

Keywords: behavioral interventions, social norms, field experiments, water conservation, water demand
JEL Classification Numbers: D12, C93, H42, L95, Q21, Q25.

*Corresponding Author: Daniel Brent, dbrent@lsu.edu. The authors are grateful to all TMWA staff who contributed to the research through information about water demand issues, data management, and logistical support. We especially appreciate the help of Laine Christman.

1 Introduction

Evidence from several randomized controlled trials has established that normative appeals lead to significant and persistent reductions in residential energy and water consumption.[1] Previous studies have only begun to shed light on the underlying behavioral mechanisms that explain consumer responses to normative appeals (Ferraro and Price, 2013). Moral motivation is a widely accepted behavioral argument: the social norm message raises the psychic cost of consumption, and in response to the self-imposed "moral tax", the consumer reduces consumption. A second behavioral mechanism occurs when the normative message informs the consumer that she may be mis-optimizing as a result of behavioral biases such as, lack of attention, mis-perceptions of operating costs, or confusions about the pricing structure (Allcott and Taubinsky, 2015; Allcott et al., 2014; Ito, 2014; Wichman, 2014). In this case, the message nudges the household to further investigate what may generate the differences between its consumption and that of its peer group. As a result, the household may correct the mis-optimization by reducing consumption. The specific behavioral mechanisms that motivate responses to normative appeals are important from a welfare perspective. Allcott and Kessler (2015) argue that the increased moral cost of consumption reduces the welfare gains from social comparisons, which may be welfare-reducing for a significant portion of households.[2]

If social comparisons operate as a moral tax on consumption, it is reasonable to believe that the moral tax would increase as household consumption increases relative to its comparison group. The literature on social comparisons in charitable giving finds that the level of contribution increases as the reference level moves further above donors' prior contributions (Shang and Croson, 2009). Additionally, Croson and Shang (2008) find that informing a donor that she gave more than her peer group in prior campaigns *decreases* her subsequent giving.[3] These results suggest that the strength of a normative appeal, defined as the distance from the comparison group, is itself a mechanism driving behavior. The mechanism operates either by increasing the salience of the norm or by causing individuals to update their prior belief about the appropriate norm - thereby increasing or decreasing the moral cost of deviating from it.

Social comparisons for water and energy conservation are fundamentally different than in other sectors promoting altruism because consumers receive private benefits for conservation in the form of reduced utility bills.[4] Disentangling the type and magnitudes of

[1] See among others Allcott (2011); Allcott and Rogers (2014); Ayres et al. (2012); Costa and Kahn (2013) for energy and Ferraro et al. (2011); Ferraro and Price (2013); Brent et al. (2015) for water.

[2] Moral costs are only part of the explanation for lower welfare benefits relative to the dollar value of energy savings; the financial and utility costs of achieving those reductions are also important.

[3] Qualitatively similar results in charitable giving are found in Frey and Meier (2004) and Martin and Randal (2008). Schultz et al. (2007) observe a similar pattern of increased energy use when households are informed they are using less energy than their peer group; a problem that is solved by invoking a positive injunctive norm in the form of smiley face emoticon.

[4] An exception is Delmas and Lessem (2014), where the experiment is conducted on students in

these behavioral channels is challenging in studies on water and energy consumption since high users receive both a "stronger" message and are also able to save more money compared to low users (or achieve similar reductions in consumption at a lower utility cost). Unsurprisingly, customers with higher pre-treatment consumption are more responsive to normative appeals in previous water and energy applications (Allcott, 2011; Dolan and Metcalfe, 2015; Ito et al., 2015; Ferraro and Miranda, 2013; Brent et al., 2015). However, heterogeneous treatment effects based on the distance from the peer group have not been estimated due to the strong correlation between pre-treatment consumption and the strength of normative message: high users generally receive messages that they consume more than their peer group while low users are typically informed that they consume less than their peers.[5] Accounting for the high correlation between pre-treatment level of use and the strength of the normative message is important both to explain the previous patterns of heterogeneity, and to address the extent to which social comparisons act as a moral tax on consumption. To our knowledge, our study is the first that is designed to account for this correlation by developing and implementing a new type of normative appeal.

We investigate the content of behavioral nudges in a utility-scale field experiment conducted in summer 2015 in partnership with the Truckee Meadows Water Authority (TMWA) in Reno, Nevada, USA. To induce voluntary conservation, five types of mailers were randomized to single-family home customers in TWMA's service area. Approximately 4,300 households were included in each of the five treatment groups, with 21,552 customers in the control group. Four of the five mailers in our study include a normative message that highlights customers personal contribution to the public good of achieving TWMA's goal of a system-wide 10% reduction in water use to cope with an on-going drought. One of the mailers is a traditional social comparison in gallons. In contrast, a second mailer introduced a new type of social norm as a treatment that compares a household's conservation with its peers' *rate* in percentage terms, thereby decoupling the distance from the peer group (strength of message) from the pre-treatment level of water use. Specifically, we compare a household's progress towards reducing its own consumption by 10% relative to its consumption in 2013, with the percentage reduction achieved by peers. This novel normative appeal allows us to analyze heterogeneity related to the distance from the peer group, while controlling for customers' previous water use. In doing so, we investigate the behavioral channels through which the messages impact customer water use.

In our conservation-rate social comparison, some low users saved less water than their peers and some high users saved more water than their peers. This contrasts to

residence halls who do not pay their bills.

[5] In an regression discontinuity framework Allcott (2011) shows that the particular category (Great, Good, Below Average), and associated emoticon, defined by the household's consumption level relative to the peer group does not impact energy use.

traditional social comparisons in gallons where most high users are informed that they use more water than their neighbors. Additionally, we design a new salient treatment with personalized information that emphasizes financial benefits of saving water by showing where a household lies in the increasing block rate structure.[6] This allows us to compare households that receive treatments emphasizing social versus financial incentives.

We make two primary contributions to the literature on behavioral nudges and social comparisons. First, our new social comparison allows us to plausibly estimate how the distance from the peer group affects the response to treatment. Second, we investigate differences in the behavioral response to nudges focused on social incentives versus financial incentives.

Our results show that the distance from one's peer group is a strong driver of heterogeneity in social comparisons, consistent with the interpretation of social comparisons as a moral tax. There is a monotonic relationship in the conditional average treatment effects based on quartiles of the household's distance from their peer group. This pattern exists in both the traditional social comparison and the new social comparison based on the conservation rate in percentage terms. Importantly, the result holds when we decouple the distance from the peer group from pre-treatment consumption. In fact, the social comparison based on the conservation rate shows the same pattern of response for both high and low users, although the effect is magnified for high users. It is possible that the distance from the peer group for social comparison in gallons may communicate privately beneficial information to help re-optimize as opposed to a moral tax. However, learning about relative conservation rates provides less private information about conservation potential because the initial conditions of the peer group are unknown. This design feature of the conservation rate treatment increases the likelihood that increased moral costs are driving the pattern of heterogeneity based on the distance from the peer group.

Similar to Allcott (2011), we use a regression discontinuity design to show that the results are not due to the discrete effect of consuming above or below one's peer group, but rather due to the *distance* from one's peers. We show a similar, although less consistent, pattern of response based on how far the household was away from their personal 10% goal. The treatment was more effective for households who failed to meet their 10% goal although, similar to the social comparisons, there is no discrete effect of moving beyond the 10% savings threshold.

When comparing the social comparison treatments with our new treatment highlighting financial benefits and the features of the rater structure, which we call the rate treatment, we find no differences in the average treatment effects across treatments. However, we show that social comparisons and the rate treatment differ in three important

[6] Pellerano et al. (Forthcoming) also design an experiment combining treatments about energy rate structures with social comparisons, but their objective is to determine if financial incentives crowds out social incentives.

ways. First, the rate treatment is relatively more effective for households that did not previously respond to pro-social appeals for water conservation. Second, the effect of sending a single rate treatment letter is more persistent than a single social comparison letter. Third, sending an additional letter generates additional savings for the social comparison treatments, but not for the rate treatment. Thus, while the social comparisons largely replicate the pattern of "action and backsliding" in energy (Allcott and Rogers, 2014), the rate treatment appears to generate different types of conservation actions. The different behavioral responses across treatments highlights that policymakers should select the appropriate intervention depending on the policy objectives.[7]

The type of behavioral response is important when evaluating the welfare effect of nudges. There are two economic rationales for how normative messaging campaigns can improve welfare. First, normative messaging campaigns can improve welfare if there are unpriced externalities related to water or energy consumption so that the retail price is below the social marginal cost. When the retail price is below the social marginal cost, there is the potential to improve aggregate welfare if the normative messages are successful at reducing households consumption. Second, normative messaging campaigns can improve welfare if a household's current consumption does not in fact maximize its own welfare. Normative appeals can improve individual welfare if they cause households to address one or more of these biases that lead to mis-optimization. In addition, if addressing mis-optimization leads to an average reduction in consumption and there are unpriced externalities, then there will be both private and external benefits associated with reducing informational and/or attentional biases. Focusing of nudges that address internalities, as opposed to imposing a moral tax, therefore has the potential to increase the welfare benefits of nudges for residential energy and water conservation.

2 Background & Experimental Design

The study was conducted in the metropolitan area of Reno, Nevada, an arid city of approximately one-quarter million people in the western United States. The Truckee Meadows Water Authority (TMWA) is the primary water utility for the Reno metro area. Peak annual water during the summer is typically four times greater than winter water use due to landscape watering. One strategy used by TMWA to manage peak use during the summer is to request voluntary water use restrictions during the months from May through September.

In response to an extended drought in 2015 TMWA launched a major multimedia campaign that included print, radio, TV, social media, and billboard messages requesting TMWA customers to voluntarily use 10% less water during the months from May to

[7] Related research on water conservation by Wichman et al. (2016) shows that some households are more responsive to prices whereas others are more responsive to non-price interventions. Hahn et al. (2016) shows how different behavioral interventions for water conservation map to different policy objectives such as reduced consumption and the adoption of efficient technology.

September, relative to their water use during the same months in 2013. The comparison year was 2013, as TMWA had asked for a 10% reduction the prior summer (July through September of 2014) in response to similar drought conditions. Prior to 2014 and 2015, the last time TMWA had requested their customers to reduce water use to address a drought was in 1992.

To complement the media campaign, we worked with TMWA to design an experiment testing the effect of sending customers letters with different types of information. The letters were aimed at increasing the salience of water use, particularly outdoor use, in household decision-making. The experiment generated five conservation-related mailers sent to residential single family households.

2.1 Description of Treatments

Table 1 summarizes the five experimental treatments (denoted T1 through T5). The appendix contains reproductions of the components of the five mailers. Every letter began: "Because of the extended drought in Northern Nevada, we are asking all of our customers to reduce water use by at least 10% this summer compared to summer 2013 - the last summer before TMWA started asking for summer water use reductions. All letters also included the statement: "Since TMWA customers use on average about four times more water in summer than in the winter, we expect that for most customers the easiest way to achieve this reduction is to adjust outdoor watering." Treatments 2 - 5 included customized information about the customer's water use, with a title introducing the letters that read: "Below is your customized water use report." Treatment 1 provided households with six tips for how to reduce outdoor water use, similar to Ferraro et al. (2011). This letter was not customized for individual households. The six tips were included on the reverse side of the other four mailers (T2-T5).

Treatment 2 (T2) included a figure that displayed the customer's water use in thousands of gallons (kgal) for May through September of 2013 and also their water use in 2015 for each month from May up to the last month billed (Appendix Figure A.1 shows the mailer).[8] Although somewhat similar to the "weak" social norm in the Cobb County experiments (Ferraro et al., 2011), this comparison of household consumption against a utility-wide goal to reduce water differs because it mentions a specific household-level target: to reduce use by 10% compared to the same month in 2013. This figure and accompanying descriptive text was included with T2-T5.

Treatment 3 (T3) contained the same components as T2, with the addition of the message, "Saving water saves you money", and a figure displaying (a) the rate structure with tiers and price for each tier, (b) the customer's water use in kgal within TMWAs increasing-block rate structure for the last month billed in 2015, (c) the upcoming month's target of 10% less water than the same month in 2013 within the rate structure, and the

[8]Thousands of gallons (kgal) is the unit used on TMWA customers' bills.

monetary savings from meeting this goal (see Appendix Figure A.2). T3 thus introduces monetary information that could increase the salience of water use.

Treatment 4 (T4) provided the same information as T2 with the addition of the message, “How does your water use compare?”, and a figure comparing the customer’s total water use in kgal for the last billed month to the median water use of a peer group consisting of single-family residences in their neighborhood with similar yard size and number of bedrooms. This essentially replicates the standard social comparison used in the OPower studies on energy and the Cobb-County (Ferraro et al., 2011) and Watersmart (Brent et al., 2015) experiments in water. See Appendix Figure A.3.

Treatment 5 (T5) included the same information as T2 with the addition of a social norm message that provided a similar comparison between households as in T4, but instead expressed in terms of percent performance towards achieving 10% goals (Appendix Figure A.4). This treatment decouples the strength of the normative message from the level of water use (in kgals), since low water users may not have saved water and high users may have saved a large amount, in percentage terms. This specification enables us to isolate the impact of the strength of the message after conditioning on water use. The accompanying figure displayed the customer’s percent change in water use from 2013 to 2015 with the percent change for the median water user for other single-family residences in their neighborhood with similar yard size and number of bedrooms.

We also include injunctive norms for T4, and T5. The message “Keep up the good work” was included for residences that had met their 10% goal in the last month billed. Households that did not meet their 10% goal in the previous month received the message “As a reminder TMWA is asking all customers to do their best to save at least 10% this summer. Please do your part to help with drought.” We did not include emoticons or “smiley faces” in any of the messages as in Schultz et al. (2007).

Table 1: Information Included in the Five Treatments

Information Components Included in Mailers*	Treatment (Mailer)				
	T1	T2	T3	T4	T5
The message "Helping our region deal with drought: What you can do" with a sheet showing 6 low cost tips to reduce outdoor water use	X	X	X	X	X
The message "What is your 10% Goal?" with a water use bar graph showing in 1000's of gallons the home's: • 2013 water use for May through September, • Target water use for each month in 2015 as 10% less, • 2015 actual monthly water use, up to last month billed.		X	X	X	X
The message "Saving water saves you money" with rate structure ** graphic showing home's water use in 1000's of gallons by tier/price: • for last month billed in 2015, • for same month in 2013, • for target goal of 10% less water used relative to 2013			X		
The message "How does your water use compare?" with a comparison of customer's 2015 last month billed water use in 1000's of gallons with similar neighborhood homes.				X	
The message "Are you doing your part?" with a comparison of customer's 2015 last month billed water use in terms of % change from 2013 with similar neighborhood homes.					X

*See Figures A.1 - A.4 in the Appendix an example of each information component of each treatment. T2 through T5 Mailers included the title: Below is your customized water use report.
**TMWA had recently merged with two other small regional utilities, Washoe County and South Truckee Meadows Groundwater Irrigation District. In 2015 all customers were subject to the rate structures that they had before the merger. Customers receiving this treatment were shown the tariff structure relevant to them.

2.2 Randomization

Our sample frame included single-family homes that (i) had metered water service; (ii) used enough water during at least one month of the 2013 irrigation season to exceed the tier 1 limit (6 kgals), roughly indicating some outdoor water use; (iii) had lived at their current residence since April 2013, and therefore had a 2013 bill for comparison; (iv) had a billing address that corresponded with the residential service address to eliminate rental occupants and other users who may not pay for water or have limited control over water use at the residence; (v) had a 2-inch service main or smaller, excluding large water users; (vi) live within one of the targeted bill cycle regions (some regions were excluded because they had a low number of single-family households, see appendix); and (vii) had nonzero water use during each month of the 2013 irrigation season (May-September) and pre-treatment months during the 2015 irrigation season (May-July), to exclude homes that were receiving bills but were unoccupied for an extended period of time.

These criteria left 42,703 eligible households. We then randomly assigned each of these households to either the control group or one of the five treatment groups. In total, 21,151 households were assigned to receive mailers (Table 1). In addition, we randomized whether households received one or two mailers. A total of 7,086 households were assigned to receive a single mailer in July (using June water use as the last month billed), 7,051 received a single mailer in August, and 7,014 received mailers in both July and August.

The randomization used a procedure of quasi-pairwise matching within blocking groups. This method first defines a set of blocks within which the randomization occurs. The blocking procedure ensures that assignment to treatment is balanced within certain

Table 2: Total Treated Households by Month/Treatment

Treatment Type	July Only	August Only	Both Months	Total
T1: Tips Sheet Only	1,420	1,410	1,402	4,232
T2: Tips + Water Use History	1,414	1,411	1,412	4,237
T3: Tips + History + Rate Information	1,413	1,411	1,410	4,234
T4: Tips + History + Social Norms (Gallons)	1,419	1,416	1,396	4,231
T5: Tips + History + Social Norms (Percent)	1,420	1,403	1,394	4,217
Total	7,086	7,051	7,014	21,151

groups of interest. Our study utilized billing cycles, frequency of meter data (i.e. monthly, daily, or hourly), and rate schedule. Within each block we ordered all observations on average water use in summer 2013 and in sets of five households. We randomly assigned each household to one of five experimental samples that correspond to the five letters. This ensures that the households receiving each of the five letters had similar water consumption in 2013.

Next, within each of the five experimental samples we repeated the procedure to assign households to one of three possible timing treatments (single letter in July, single letter in August, or two letters repeated in July and August), or the control group. The same blocking structure was used within each experimental sample and then households were re-ordered based on summer 2013 water consumption and in sets of 12: two households are randomly assigned to each of the three timing treatments and six households are assigned to the control.[9]

Table 3 shows that in aggregate the randomization achieved very strong balance on observables. Additionally, Tables A.1-A.3 in the Appendix show that the experimental sample is balanced on pre-treatment water use for each treatment, across treatments, and within deciles of baseline water use. Figure A.5 graphically displays the densities of the natural log of pre-treatment water use for the pooled treatment, each of the five individual treatments, and the control group. In addition to achieving balance on average pre-treatment water use, Figure A.5 shows the treatments are balanced across the full distribution of pre-treatment water use. The graphical evidence is formalized by nonparametric Kolmogorov-Smirnov tests (Table A.4 in the Appendix) that fail to reject the null of equality of distributions for pre-treatment water use across the control and the pooled treatment as well as each treatment individually. Our sample is well balanced by design, which allows us to make valid inferences for the conditional average treatment effects within subgroups, particularly subgroups that are functions of the normative message, which depends on pre-treatment water use.

[9] Due to the unequal size of the blocking groups, some timing treatments were oversampled, thereby creating some balance issues. We corrected these by identifying and dropping the oversampled observations (3,677 households: 2,025 control, 1,652 treatment) within the randomization script. All balance tables and regression results reflect the corrected sample.

Table 3: All Treatments Balance on Observables

	Control Mean	Treatment Mean	Difference	(p-value)
2013 Water	23.56	23.55	0.02	0.89
2015 Water	16.90	16.99	-0.09	0.37
Summer Water	21.63	21.70	-0.07	0.52
Winter Water	8.15	8.15	-0.01	0.88
Year Built	1,987.61	1,987.67	-0.05	0.77
Appraised Value	214.85	214.87	-0.02	0.99
Bedrooms	3.37	3.37	0.01	0.43
Lot Acre	0.27	0.27	-0.00	0.91
Yard Acre	0.22	0.22	-0.00	0.95
Built Sq. Ft.	1,985.10	1,991.07	-5.98	0.41
Bathrooms	2.19	2.20	-0.00	0.56

Note: 21,552 Control Observations and 21,151 treated observations, p-value is based on two-sided t-test

3 Methodology

The primary variable of interest is monthly household water use, obtained from TMWA billing records, expressed in average gallons per day (GPD). We calculate GPD by dividing total billing cycle usage by the number of days in that billing period to avoid problems with different lengths of billing periods for different customers. The specification in our regression analysis uses "normalized GPD" as the main dependent variable; every customer's GPD is divided by the average control group consumption during the experimental period (July-September 2015) following Allcott (2011). This allows the regression coefficients to be interpreted as the average percent change in water use, while preserving the treatment effect of very high water users, which the logarithmic transformation of water use would dampen. Our specification is:

$$y_{it} = \alpha + \gamma_l T_{i,l} + \beta \mathbf{x_{it}} + \epsilon_{it} \tag{1}$$

where y_{it} is normalized GPD, $T_{i,l}$ is a dummy variable for the pooled treatment and each of the five treatment letters ($l = Pooled, 1, 2, ..., 5$), and x_{it} is a vector of control variables. We restrict our sample to the post-intervention period, which comprises the billing months of August, September, and October 2015. While treatment is exogenous by virtue of the randomization, including control variables increases the precision of the estimates. All regressions therefore include average water use during irrigation seasons prior to the intervention, billing cycle and month fixed effects, average daily temperature, and days of precipitation in the billing cycle. We matched daily weather data from the NOAA weather station at Reno-Tahoe Airport to the exact dates of each customer's water bill to calculate the weather variables. Robust standard errors are clustered at the household level.

3.1 Identifying the Effect of Distance from the Peer Group

To test the role of normative information in social comparisons we assume that the magnitude of the moral tax depends on the distance between a household's consumption and the relevant comparison group. This is one of the cases nested within the consumer utility model of social comparisons presented in Allcott and Kessler (2015). In their model consumers face a constant marginal moral tax for "inappropriate" consumption. One specification of "inappropriate" consumption is that all water use exceeding the level of one's peer group is inappropriate, and these excess units are subject to a moral tax. Households consuming below the peer group receive a moral subsidy while households consuming above the peer group face a moral tax. We use the terms "magnitude of the moral tax" interchangeably with the "distance from the norm", "performance relative to the norm", or "strength of normative message". Households consuming further above the norm are performing worse than their peers, are subject to a higher moral tax, and receive a stronger normative message.

Traditional social comparisons force the majority of low users to receive messages indicating that they are performing better than their peers, while high users typically receive messages indicating that they are performing worse than their peers. This correlation between pre-treatment consumption and the strength of normative message creates a challenge for disentangling the behavioral mechanisms behind the robust results reported in previous studies using traditional social comparisons, that higher users save more energy (Allcott, 2011) and water (Ferraro and Price, 2013; Brent et al., 2015). These users typically receive stronger normative messages from learning that they are using more water than their peers, but they are also likely to have more low-cost options to reduce consumption. Our social comparison in percentage terms is designed to reduce the correlation between pre-treatment water use and the strength of the normative message, allowing us to isolate the impact of the strength of the normative message they receive, as indicated by the distance from the peer group.

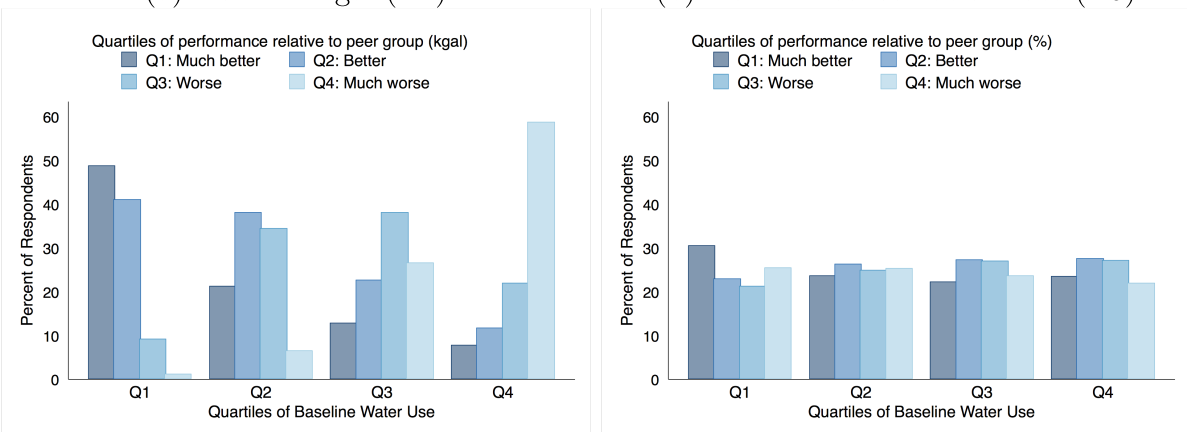
To demonstrate the variation our design introduces between water use and the distance from of the peer group we plot the percentage of households receiving a letter who are performing better or worse than their peers by quartiles of baseline water use (Figure 1). We define the performance of a household based on quartiles of the distance between a household's water use, or conservation rate, and that of its peer group. Since the social comparisons are based on median consumption or conservation rate within the peer group, the first two quartiles (Q1:Much better and Q2:Better) are households who consumed less than their peer group (or reduced consumption by a higher percentage than peers) and thus received moral subsidies. The upper two quartiles (Q3:Worse and Q4:Much Worse) represent households who consumed more than their peer group (or reduced consumption by a smaller percentage), and received moral taxes. Panel (a) of Figure 1 shows that most low users (Q1 on the x-axis) in the traditional social comparison in gallons (Treatment

4) used less water than their peer group: roughly 90% of the consumers with the lowest water consumption were informed that they used less than their neighbors (Q1 + Q2 of Distance from Peer group (kgal)).[10] Likewise, most high users (Q4 on the x-axis) received a message telling them they used more water than their peers.

This is not the case for our conservation rate social comparison (Treatment 5): a substantial fraction of low users conserved less than their peers and many high users conserved more than their peers (Figure 1, Panel (b)). Even among households in the bottom quartile of baseline water consumption, some reduced their water consumption by less than the median conservation rate in their peer group and thus received a strong normative message. Likewise, some households with high pre-treatment water use reduced their consumption by a larger percentage than their peer group and received a moral subsidy. The distribution of norms within each quartile of baseline water use is remarkably balanced for the conservation rate comparison treatment. This is an interesting result itself: the utility-wide voluntary conservation messaging program in Reno achieved conservation across a range of low- and high-water users.[11]

Figure 1: Strength of Normative Message by Quartiles of Baseline Water Use

(a) Norm in kgal (T4) (b) Norm in conservation rate (T5)



Note: The graph displays the percentage of respondents receiving messages divided up by quartiles of the performance relative to the peer group within each quartile of baseline water use. The x-axis displays the quartiles of baseline water use and the y-axis displays the percentage of respondents receiving a given message. The performance relative to the norm are designated by the different colored bars. The performance relative to the peer group is defined based on quartiles of the difference between a household's water use (panel (a) - T4) or conservation rate (panel (b) - T5) and it's peer group's water use.

[10]The only reason why there are some low users who are above their peer group in the traditional social comparison (T4) is that the norm is based on a peer group - defined by households in the same meter route who have similar number of bedrooms and yard size (above/below the median). (By comparison, Ferraro and Price (2013) compare household consumption to the sample median, producing a treatment where the strength of the descriptive norm is perfectly correlated with pre-treatment water use.) Therefore a household with a high-water-use peer group can be above the median or 75th percentile of the sample-wide distribution of baseline water use, but still consume less than the peer group.

[11]From a water supply planning perspective, of course, an 8% reduction in water use from a high baseline water user (e.g. with lots of outdoor water use) yields more water conservation than an 8% reduction from a household with only small, indoor water use.

We incorporate the information content of the mailers by estimating conditional average treatment effects (CATEs), where we condition on the distance between a household's consumption and its peer group. While the content of the mailers depends on water use, the first mailer only depends on water use prior to the intervention. We drop observations for households treated with a second letter since the content of the mailer becomes endogenous to water use. Since treatment is randomized across the distribution of baseline water use, CATEs provide valid inference. In this sense the results can be interpreted as causal treatment effects in the same style as studies that condition on pre-intervention consumption (Allcott, 2011; Ferraro and Miranda, 2013; Brent et al., 2015).

The CATE model is defined as

$$y_{it} = \alpha + \sum_{c=1}^{k} \gamma_{l,c} T_{i,l} \times C_{i,c} + \sum_{c=1}^{k} \theta_c C_{i,c} + \beta \mathbf{x_{it}} + \epsilon_{it} \quad (2)$$

In this model we are concerned with $\gamma_{l,c}$, which is the CATE for letter l in subgroup c. $T_{i,l}$ is an indicator for whether a household was treated with letter l and $C_{i,c}$ is an indicator for whether a household falls into subgroup c of the conditioning variable $C_{i,c}$. The presence of $C_{i,c}$ accounts for the sample-wide differences in water use for subgroup c. When we condition based on the performance relative to the peer group $C_{i,c}$ equals a set of dummies indicating distance between a household's water use (or conservation rate) and that of its peers in the previous month. Since control households did not actually receive any treatment, we calculate the distance from their peer group normative message that they would have received had they been treated. This accounts for unobserved factors within each subgroup that are common to the treatment and control groups. In the case of the distance from the comparison group in kgal (T4), $C_{i,c}$ controls for the fact that high users generally receive comparisons above their peer group in the gallons comparison (T4) and issues such as mean reversion in the conservation rate comparison (T5). Since treated households only receive one realization of the distance from their peer group, based either on their July or August water consumption, we randomly assign each control household the distance from their peer group based on either July or August in the same proportion as the treated households.[12] We also run specifications where the distance variables for the control households change over time based on the realization in the previous month that produce very similar results.

[12] One third of the treatment group received a single mailer in July, one third received a single mailer in August, and one third received mailers in July and August. Since we drop all observations after the second mailer two thirds of the treatment received the first mailer in July and one third in August.

4 Results

4.1 Base Results

We begin by reporting the average treatment effects pooling all mailers, and then briefly discuss each treatment individually. Column 1 of Table 4 shows that the average treatment effect (ATE) pooling all treatments is slightly greater than a 1% reduction in water use. This is a smaller effect than commonly reported for social comparisons: Opower's interventions typically reduced energy consumption by about 2%, and both Ferraro and Miranda (2013) and Brent et al. (2015) find average reductions in water use of approximately 5%. This should be considered in the context of an extensive utility-wide water conservation campaign in a second year of drought conditions. Additionally, given that the aforementioned studies on water examine some of the first interventions using social comparisons for water conservation, the lower treatment effects are consistent with finding of Allcott (2015) that initial sites often have higher average treatment effects than subsequent sites.

Column (2) breaks out each of the treatments individually. Sending conservation tips alone (T1) was not effective, similar to results found in the Cobb County experiments (Ferraro et al., 2011). The other four treatments all generated statistically significant reductions in water use. Though the point estimate of the "Tips+History treatment (T2) is roughly 33% smaller in magnitude than T3-T5, the differences between any combination of T2-T5 are not statistically significant. Columns (3)-(5) reproduce the ATE for each letter in separate regressions using the individual treatment group and the control. This is simply to display that both the point estimate and the precision are almost identical whether we use the entire sample with five treatment dummies or restrict the sample to one treatment and the control. Restricting the sample simplifies the presentation of the results. All subsequent regressions also include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and baseline water use.

Next we largely replicate the well-known finding that social comparisons generate highly heterogeneous treatment effects, with most of the conservation coming from the largest users. The results are based on the same model in equation 2 but the conditioning variables are quartiles of baseline water use. Figure 2 estimates CATEs by quartiles of baseline water use for the pooled treatment and each individual treatment. Each panel reflects the results of one regression of the CATEs; the shaded bars are the point estimates and the error bands are the 95% confidence intervals. The pooled treatment (the leftmost panel in the figure) produces CATEs that conform with the previous literature: higher users save more water in response to treatment, and most of the effects are concentrated in households with baseline consumption above the median. Treatments 2 through 4 have similar patterns with the largest savings coming from the highest quartile consumers. Next we analyze whether this pattern of heterogeneity for the social comparison

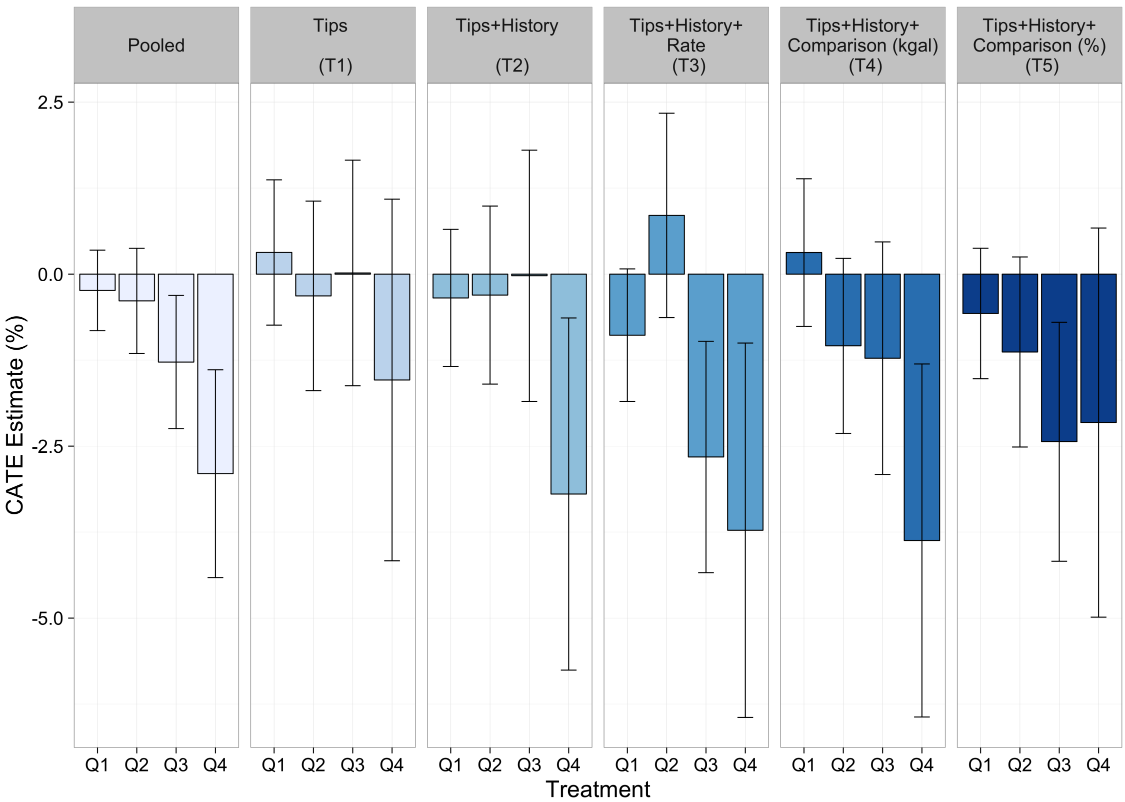
Table 4: Base Regression

	(1) Pooled	(2) All Letters	(3) Letter 1	(4) Letter 2	(5) Letter 3	(6) Letter 4	(7) Letter 5
All Treatments	-1.200***						
	(0.260)						
Tips		-0.381	-0.382				
		(0.455)	(0.455)				
Tips+History		-0.977**		-0.973**			
		(0.453)		(0.453)			
Tips+History+Rate		-1.601***			-1.596***		
		(0.466)			(0.466)		
Tips+History+Comparison (kgal)		-1.452***				-1.451***	
		(0.446)				(0.446)	
Tips+History+Comparison (%)		-1.587***					-1.591***
		(0.473)					(0.473)
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bill Cycle FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Water	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Households	42,357	42,357	25,520	25,534	25,532	25,529	25,510
Observations	119,025	119,025	74,458	74,522	74,530	74,494	74,449

Note: The dependent variable is normalized average daily water use; the coefficients can be interpreted as a percentage change in water use. Robust standard errors clustered at the household level are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

treatments (T4 and T5) is due to baseline water use or the distance from the peer group.

Figure 2: Conditional Average Treatment Effects by Quartiles of Baseline Consumption



Note: Each bar graph represents the output of one regression where the dependent variable is normalized average daily water use. The bars are the point estimates of the CATEs for each quartile of baseline consumption, and the error bars are 95% confidence intervals constructed from robust standard errors clustered at the household level. All regressions include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and baseline water use.

4.2 Distance from the Peer-Group Descriptive Norm

In order to separate the effects of the strength of the normative message and baseline water use we run regressions for the norm-based CATEs. Our first specification defines strong normative messages as mailers that informed households that they were *above* their peer group: either using more water than their peer group (T4) or a lower year-on-year percentage reduction than their peer group (T5). Our second specification estimates CATEs based on quartiles of the distance from the peer group, providing more variation in the performance relative to the peer group.

Column (1) of Table 5 shows the CATEs based on being above or below the peer group for the social comparison in gallons. All of the savings essentially come from households who are informed that they were consuming more than their peers. The same pattern holds when we estimate CATEs within quartiles of distance from comparison group (Column (2) of Table 5). There is a monotonic relationship between the consumption relative to the peer group and the CATEs. As described above, it is difficult to interpret the results for the normative message expressed in gallons as due to the normative message or due to the fact that households consuming above their peers are high users with more available margins to conserve water.

Columns (3)-(4) of Table 5 estimate the CATEs based on the distance from the comparison group framed as a conservation rate, which separates the performance relative to the peer group from baseline water use. The pattern of results is similar: households treated with a message telling them they conserved less than their peers save water, whereas those who conserved more than their neighbors do not. In fact, the results in column (4) using quartiles of the distance from the peer group are even starker. Among households treated with the conservation rate social comparison, those who were reducing their consumption by a much smaller percentage than their peers before receiving the letter (Treat*Q4 Norm) saved over 5.7%. This is compared to a 1.4% increase (though not statistically significant) among households who reduced consumption by a much larger percentage than their peers (Treat*Q1 Norm). The results are consistent with social comparisons working as a moral tax, with higher moral taxes generating larger reductions in consumption. The results are also consistent with the second behavioral channel: stronger messages indicate more room for re-optimization, with possible financial gain. While we cannot eliminate either explanation we provide two reasons why higher moral costs is a more plausible interpretation of the results for the conservation rate treatment. First, the conservation rate treatment provides less privately beneficial information on re-optimization than the traditional social comparison because the reference level (peer group's use 2013 consumption) is unknown. Second, we find it unlikely that re-optimization would cause a household to increase consumption after learning that it saved more water than its peers. While the increase in consumption for the lowest quartile is not consistently statistically significant, it provides suggestive evidence for the moral cost explanation.

We also analyze heterogeneity in the effect of the content of the normative message by dividing the sample based on baseline water use. Columns (5) and (6) restrict the sample to households below and above the median baseline water use respectively. We do not perform the analysis for the social comparison in gallons due to the strong correlation between the distance from the peer group and baseline consumption. The pattern of results is the same for both low and high pre-treatment users. Baseline water use acts to magnify the results; the savings for households reducing by a smaller percentage than their peers (Treat*Q3 Norm and Treat*Q4 Norm) are roughly twice as large for high users compared to low users. The increase in consumption among those who drastically outperform their peers is similar in magnitude for both low and high users, and is statistically significant at the 10% level for low users. From a utility's perspective, this last result is counter-productive as high-performing households slid backwards when told they were doing more than their peers. (Schultz et al., 2007) found a similar result in energy conservation that was ameliorated by providing emoticons as a positive normative injunction. We did not include emoticons in our mailers, only text-based injunctions to "keep up the good work". Overall, these data suggest that social comparisons do appear to function as a moral tax for all users, but high users have more margins of adjustment.

Table 5: Distance from Peer Group

	Norm in Gallons (T4)		Norm in % (T5)			
	(1) All	(2) All	(3) All	(4) All	(5) Low Users	(6) High Users
Treat*Below Peer	0.339		0.226			
	(0.487)		(0.594)			
Treat*Above Peer	-2.255***		-3.613***			
	(0.701)		(0.653)			
Treat*Q1 Norm		1.078		1.409	1.509*	1.417
		(0.665)		(0.912)	(0.777)	(1.651)
Treat*Q2 Norm		-0.352		-0.134	0.552	-0.721
		(0.698)		(0.707)	(0.633)	(1.186)
Treat*Q3 Norm		-1.769**		-2.463***	-1.534**	-3.239***
		(0.743)		(0.693)	(0.631)	(1.119)
Treat*Q4 Norm		-2.555**		-5.778***	-4.408***	-7.113***
		(1.216)		(1.005)	(1.031)	(1.690)
Below Peer	-15.32***		-19.58***			
	(0.307)		(0.237)			
Q1 Norm		-30.36***		-32.14***	-23.41***	-41.69***
		(0.540)		(0.366)	(0.336)	(0.651)
Q2 Norm		-24.26***		-20.10***	-14.64***	-25.57***
		(0.489)		(0.345)	(0.320)	(0.590)
Q3 Norm		-17.03***		-12.52***	-9.704***	-15.78***
		(0.436)		(0.350)	(0.326)	(0.594)
Observations	68,933	68,870	68,929	68,822	33,126	35,468

Note: The dependent variable is normalized average daily water use, and the sample is restricted to the month after a household's first mailer. CATEs are estimated based on the distance (above/below) of household consumption relative to its peer group. Column headers All represents the whole sample, and Low/High Users restrict the sample to households above or below median pre-treatment water use. All regressions include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and baseline water use. Robust standard errors clustered at the household level are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

4.3 Effects of meeting the community-wide 10% Goal

Next we analyze heterogeneity in the treatment effect based on meeting the personal 10% savings goal. All the treatments with personalized information (T2-T4) notify the household whether they met their personal 10% savings goal for the previous month. We run two estimate two specifications of CATEs to learn about the impact of how far away the household was from their 10% savings goal. The first specification is based on equation 2 where the conditioning variable is an indicator for whether the household met the 10% goal in the previous month. In the second specification the conditioning variables are indicators for the quartiles of distance between a household and its 10% goal in the previous month. One key distinction between quartiles based on the distance from the peer group and the distance from the personal goal is the interpretation within each quartile. The peer group is based on the median consumption or conservation rate so the quartiles are centered at zero. However, the median households conserved by *more* than 10% so the third quartile in the difference from the goal contains some households than saved more than 10% and some households that saved less than 10%. We examine the discrete effect of moving just below the 10% goal with a regression discontinuity below.

For this analysis the conservation rate for households within: *Q1 Goal* $\gg$ 10%, *Q2 Goal* $>$ 10%, *Q3 Goal* $\lesssim$ 10%, *Q4 Goal* $<$ 10%. Similar to the analysis on the distance from the peer group, we assign control households values for the conditioning variables even though they did not receive a letter, and we drop all observations following a household's second mailer. Whether a household met the goal is balanced across treatment and control since it is based on pre-treatment water use.[13]

The first specification in panel (a) of Table 6 the 10% goal is an important driver of water savings from the treatment, as shown by the results in panel (a) Table 6. Across all treatments with the historical information (T2-T4), households that met their personal 10% goal saved roughly 1%, and households who failed to meet the goal saved over 3% compared to control households. Not surprisingly, households who met the 10% goal (*Met Goal*) used less water regardless of whether they were treated; however, meeting the goal is not simply a proxy for baseline water use. There are no significant differences in baseline water use based on whether a household met the 10% goal. The same pattern of results holds for all of the treatments, where the treatment is larger for households who did not meet the goal. The differences are statistically significant for all treatments that include the historical information except Treatment 3 (Tips+History+Rate). The large differences in Treatment 5 is likely due to the fact that those who were further above their own personal 10% goal were also likely above their peer group's conservation rate.

The second specification, presented in column (b), estimates CATEs based on the difference from the personal 10% goal. Although the same general pattern holds, there are differences across treatments. There is not a consistent monotonic between the distance from the goal and the CATEs. This likely reflects that in T3-T5 there are other factors affecting treatment effect heterogeneity. Since the historical information contains information for several months we also generate a variable for whether the mailer indicated that the household always met the goal, sometimes met the goal, or never met the goal. The results are shown in Appendix Table A.5 and are largely consistent with the results presented in Table 6: most of the saving coming from households who never met the 10% goal during summer 2015. One interpretation is that meeting the 10% goal provides a social sanction that the household is doing their part for the community-wide water conservation.

[13]The p-values based on the two-sided t-tests for differences in the July goal and August goals across treatment status are 0.54 and 0.35 respectively.

Table 6: Conditional Average Treatment Effects: 10% Goal

(a) : Met or Failed the Goal

	(1) Pooled	(2) Treatment 2	(3) Treatment 3	(4) Treatment 4	(5) Treatment 5
Met Goal*Treat	-1.083***	-0.543	-1.650***	-1.219**	-0.967*
	(0.305)	(0.546)	(0.551)	(0.548)	(0.575)
Failed Goal*Treat	-3.779***	-3.124***	-3.447***	-3.548***	-5.139***
	(0.389)	(0.694)	(0.734)	(0.624)	(0.686)
Met Goal	-19.61***	-19.61***	-19.61***	-19.57***	-19.60***
	(0.252)	(0.253)	(0.252)	(0.253)	(0.252)
Observations	85,693	68,948	68,954	68,933	68,929

(b) Conditional Average Treatment Effects: Quartiles of Distance from the Goal

	(1) Pooled	(2) Treatment 2	(3) Treatment 3	(4) Treatment 4	(5) Treatment 5
Treat*Q1 Goal	-1.000**	-1.293	-0.867	-1.871**	0.0210
	(0.495)	(0.902)	(0.873)	(0.869)	(0.997)
Treat*Q2 Goal	-1.670***	0.190	-1.616**	-2.386***	-2.947***
	(0.431)	(0.774)	(0.790)	(0.771)	(0.805)
Treat*Q3 Goal	-3.192***	-3.436***	-4.530***	-1.836**	-3.152***
	(0.397)	(0.693)	(0.683)	(0.752)	(0.642)
Treat*Q4 Goal	-4.138***	-3.569***	-2.499**	-4.493***	-6.244***
	(0.525)	(0.930)	(1.039)	(0.820)	(0.906)
Q1 Goal	-30.96***	-30.93***	-30.95***	-30.89***	-30.94***
	(0.355)	(0.356)	(0.356)	(0.356)	(0.356)
Q2 Goal	-18.60***	-18.59***	-18.61***	-18.57***	-18.61***
	(0.347)	(0.348)	(0.348)	(0.348)	(0.348)
Q3 Goal	-10.63***	-10.64***	-10.65***	-10.63***	-10.66***
	(0.354)	(0.355)	(0.355)	(0.355)	(0.355)
Observations	85,595	68,856	68,860	68,840	68,834

Note: The dependent variable is normalized average daily water use, and the sample is restricted post treatment data dropping all observations after the second mailer. In panel (a) the CATEs are based on whether the household met its personal 10% goal in the previous month. In panel (b) The CATEs are based on quartiles of the distance from the 10% goal. All regressions include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and baseline water use. Robust standard errors clustered at the household level are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

4.4 Prior Conservation

The CATEs in the previous sections suggest that the social comparisons and historical information relative to the 10% goal operate through a moral mechanism, such as a moral tax on excess consumption. The results are also consistent, though perhaps less likely, with consumers re-optimizing if households consider the excess consumption relative to the peer group is due to inefficient water use. In order to determine whether a moral behavioral mechanism is likely present, we analyze heterogeneity by households who responded to the utility's appeal to save water in 2014. As mentioned earlier, the utility first called for all households to reduce their use by 10% relative to 2013 midway through the summer of 2014.[14] In this specification, we interact the treatment dummies with an

[14] The appeal in 2014 is the reason why the utility, and our treatments, reference the 10% goal to 2013 consumption since this was the most recent year without voluntary restrictions in place.

indicator variable for whether the household reduced water use in summer 2014 (defined as the average of May-October water use) relative to 2013 by at least 10%. Since we are not incorporating any of the content of the mailers (i.e. strength of normative message), we include all post-treatment data, including observations for households that received two mailers.

There are two primary hypotheses for why households who responded to the utility's appeal in 2014 would react differently to the treatment. First, prior conservation may serve as a signal that a household is sensitive to responding to pro-social appeals for water conservation, and thus be more responsive to treatment. This hypothesis is consistent with studies that find environmentalists and political liberals (Costa and Kahn, 2013) and households that vote (Bolsen et al., 2014) are more responsive to social comparisons for energy and water conservation. If prior conservation magnifies the treatment effect (Treat*Goal 2014 ¡ 0) then it is likely that households are responding to treatment due to moral motivations. An alternative hypothesis is that households who conserved in 2014 will be less responsive to the treatment since they likely exhausted low cost conservation activities in 2014, raising the marginal cost of conservation in 2015. If prior conservation dampens the treatment effect (Treat*Goal 2014 ¿ 0) it is likely that treatment induces households to search for low-cost conservation options. It is possible that both of these mechanisms are operating simultaneously.

The results are presented in Table 7 along with the base specification in columns (1) and (3) for reference. Column (2) interacts the pooled treatment dummy with a dummy for whether the household met the 2014 10% goal. In aggregate the interaction term is positive and significant: households who conserved in 2014 are *less* responsive to treatment compared to households that did not reduce consumption in 2014. The treatment coefficient in column (2), corresponding to households who did not meet the goal in 2014, is roughly 50% larger than the ATE in column (1). Columns (3)-(4) of Table 7 show the results based on interacting the 2014 goal indicator with each of the five treatment dummies. Four out of the five treatments have a positive interaction term, but only Treatment 3, which emphasizes financial savings, is statistically significant and again implies that water conservation in 2014 weakened the treatment effect. This is consistent with different mechanisms operating across treatments. The rate treatment is *more* effective with households who had *not* met the 2014 goal, which could indicate that they were either not engaged with previous pro-social appeals or that they had ample low-cost conservation options still available to them. It should be noted that the interactions of prior conservation with the social comparison treatments (T4 and T5) are also positive, but not statistically significant. Additionally, the interaction term for the rate treatment (T3*Goal 2014) is statistically different from the interaction term for the history only treatment (T2*Goal 2014), but not from either of the two social comparison interaction terms (T4*Goal 2014 & (T5*Goal 2014). Although there is not one definitive interpretation of the prior conservation results, it is clear that the rate treatment is more effective among households that did not meet the 2014 goal.

Table 7: Prior Conservation (2014)

	Pooled		Letters	
	(1)	(2)	(3)	(4)
	Base	Goal 2014	Base	Goal 2014
All Treatments	-1.178***	-1.770***		
	(0.260)	(0.403)		
Treat (All)*Goal 2014		1.047**		
		(0.523)		
Treatment (T1)			-0.360	-0.975
			(0.455)	(0.709)
Treatment (T2)			-0.956**	-0.773
			(0.453)	(0.708)
Treatment (T3)			-1.579***	-2.833***
			(0.466)	(0.732)
Treatment (T4)			-1.430***	-1.931***
			(0.446)	(0.684)
Treatment (T5)			-1.567***	-2.370***
			(0.473)	(0.756)
T1*Goal 2014				1.016
				(0.916)
T2*Goal 2014				-0.387
				(0.912)
T3*Goal 2014				2.320**
				(0.944)
T4*Goal 2014				0.954
				(0.898)
T5*Goal 2014				1.370
				(0.958)
Goal 2014		-5.926***		-5.926***
		(0.358)		(0.358)
Households	42,357	42,357	42,357	42,357
Observations	119,025	119,025	119,025	119,025

Note: The dependent variable is normalized average daily water use. Columns (1)-(2) show regressions pooling all treatments and columns (3)-(4) estimate effects for each of the five individual treatments. Goal 2014 is a dummy if the household reduced water use in 2014 relative to 2013 by at least 10%. All regressions include controls for baseline water use, weather, month fixed effects, and billing cycle fixed effects. Robust standard errors clustered at the household level are reported in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

4.5 Persistence & Additionality

The results analyzing heterogeneity by prior conservation suggest that emphasizing financial information may generate a different behavioral mechanism than the traditional social comparisons. We build on this result by analyzing variation across treatments in the persistence of the savings over the summer demand period and the effect of sending an additional letter. In order to understand the implications of the timing and number of treatments we run the following regression,

$$
\begin{aligned}
y_{it} = \alpha + & \gamma_{1,l} 1\{First\ Mailer, First\ Month_{it,l}\} + \\
& \gamma_{2,l} 1\{First\ Mailer, Second\ Month_{it,l}\} + \\
& \gamma_{3,l} 1\{First\ Mailer, Third\ Month_{it,l}\} + \\
& \gamma_{4,l} 1\{Second\ Mailer_{it,l}\} + \beta \mathbf{x_{it}} + \epsilon_{it}
\end{aligned} \tag{3}
$$

where the treatments are now broken down by the months since the first letter was received in addition to a dummy for whether there was a second mailer ($\gamma_{4,l}$). We only analyze three months from the first letter since this corresponds to the end of the summer demand season (October). Therefore, γ_2 and γ_3 test whether the treatment was persistent and γ_4 represents the additionality of the second mailer.[15] Table 8 presents the results for the treatments pooled together as well as each individual treatment. Examining the pooled treatment shows that the treatment is reasonably persistent but does decrease in magnitude as time from the first mailer increases. In aggregate the second mailer generates roughly 0.6% in additional conservation, though this is only significant at the 10% level.

The most intriguing insights into persistence and additionality are how they vary across treatments. T2 and T3 are the most persistent, and for these treatments the second mailer generates no additional savings. In fact the second mailer seems to increase consumption in T2, though the estimate is not statistically significant. Both of the social comparison treatments (T4 and T5) are not persistent and achieve significant additionality from the second mailer. The savings from the first treatment decreases by roughly 50% or more per month. Adding another mailer increases conservation by almost 2%. The results for T4 and T5 are consistent with the Allcott and Rogers (2014) results where the pattern of action and backsliding stems from a model of consumption cues.

The differences in persistence and additionality across treatments can be interpreted in the framework of intent-oriented actions and impact-oriented actions developed by Attari (2014). Households perform intent-oriented actions, generally behavioral changes such as turning off the water when brushing teeth, with the intention to help the environment while impact-oriented actions, investments such as replacing a toilet, are focused on making a large difference in water use. The results suggest a divergence of impact-oriented actions in T2 and T3 compared with intent-oriented actions in T4 and T5. This could be due to the social comparisons operating through moral motivations leading to intent-oriented actions. Conversely, the persistent results from the treatment emphasizing the rate structure and financial savings could operate through financial motivations and lead to impact-oriented actions.

[15] The second mailer affects consumption in two months: September and October.

Table 8: Persistence & Additionality of Second Mailers

	(1) Pool	(2) T1	(3) T2	(4) T3	(5) T4	(6) T5
1st Letter, 1st Month	-1.136***	-0.416	-0.991*	-1.395**	-1.434***	-1.528***
	(0.295)	(0.531)	(0.517)	(0.542)	(0.517)	(0.552)
1st Letter, 2nd Month	-1.067***	-0.137	-1.030*	-1.899***	-1.093*	-1.098*
	(0.314)	(0.586)	(0.594)	(0.592)	(0.591)	(0.623)
1st Letter, 3rd Month	-0.806**	-0.293	-1.568**	-1.499**	-0.168	-0.479
	(0.391)	(0.687)	(0.729)	(0.739)	(0.717)	(0.765)
2nd Letter	-0.653*	-0.356	0.810	0.158	-1.817**	-1.963**
	(0.356)	(0.754)	(0.806)	(0.759)	(0.741)	(0.825)
Households	42,357	25,520	25,534	25,532	25,529	25,510
Observations	119,025	74,458	74,522	74,530	74,494	74,449

4.6 Robustness

We run several robustness tests to assure that our results are not actually being driven by spurious correlation or the discrete effects of being above or below the norm. The treatment period started in August 2015 utilizing July 2015 data on the mailers. Our first robustness check runs falsification tests by using the content generated by June 2015 data. We generate the distance from the peer group using June 2015 for both treated and control households to see if the types of households that were above or below the peer group were actually driving the results as opposed to the content of the mailers. Table 9 shows that results using the simulated mailer content. The results are substantially different than Table 5. Households above and below the false peer group have statistically significant CATEs in the conservation rate treatment (column (3)), and neither of the CATEs are statistically significantly different from each other. The CATE for households below the peer group in the traditional social comparison is not statistically significant but this is likely due to the fact that they are low users and their real letter also indicated that they were below their peer group. The specification with quartiles of the norm in percentage terms (column (4)) shows that three out of the four CATEs generate savings over 2% and are significant at the 5% level. In general households above the false peer group do save more water, but this is likely because there is still some serial correlation in water use; households above the peer group in June are more likely to be above the peer group in July and August.

We also test for the discrete effects for moving above a peer group or failing to meet the 10% goal in a regression discontinuity design. We find no effect of moving above the peer group in either gallons or percentage terms. This is consistent with the findings of Allcott (2011) for social comparisons in energy and with unpublished results from the project reported in Brent et al. (2015). We also find no effect of moving slightly above the 10% goal. The Appendix describes the regression discontinuity design in more detail and presents both graphical evidence and the regression discontinuity estimates based on Calonico et al. (2015).

Table 9: Falsification Test for Distance from Peer Group

	Norm in Gallons (T4)		Norm in % (T5)	
	(1)	(2)	(3)	(4)
Treat*Below Peer	-0.660		-1.184**	
	(0.491)		(0.589)	
Treat*Above Peer	-1.794**		-2.181***	
	(0.707)		(0.684)	
Treat*Q1 Norm		-0.148		-2.274***
		(0.719)		(0.848)
Treat*Q2 Norm		-1.166*		-0.238
		(0.667)		(0.800)
Treat*Q3 Norm		-2.197***		-1.629**
		(0.803)		(0.808)
Treat*Q4 Norm		-1.290		-2.708**
		(1.167)		(1.118)
Below Peer	-7.384***		-11.20***	
	(0.290)		(0.245)	
Q1 Norm		-13.62***		-18.32***
		(0.490)		(0.365)
Q2 Norm		-10.34***		-10.93***
		(0.445)		(0.357)
Q3 Norm		-6.816***		-6.444***
		(0.432)		(0.360)
Observations	68,933	68,930	68,929	68,896

Note: The dependent variable is normalized average daily water use, and the sample is restricted to the month after a household's first mailer. CATEs are estimated based on the quartiles of the distance of household consumption relative to its peer group. Robust standard errors clustered at the household level are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

5 Conclusions

In aggregate we find that four out of the five nudges caused households to reduce water consumption by 1-1.5% in an environment where the utility was publicly asking for a 10% voluntary reduction relative to 2013 among all customers. While there may be some concern that the 10% goal reduces the external validity of our experimental results, calls for uniform reductions are a common utility practice as evidenced by Governor Brown's 2015 call for reducing water consumption in California by 25%. We also note that there was a utility-wide appeal for conservation in the well-known water conservation experiments in Cobb County examined by Ferraro and Price (2013). None of the treatments that generated statistically significant ATEs are statistically different from each other.

Our design decouples pre-treatment water use from the distance from the peer group by presenting a social comparison framed as savings over time in percentage terms. This allows us to explore whether the finding that higher users are more responsive to social comparisons is due to the distance from the peer group or the lower opportunity costs of water savings for high users. For both social comparisons, including the treatment that decouples the distance from the peer group from water consumption, the distance from the peer group is a strong driver of treatment effect heterogeneity. In fact, the distance from the peer group in our conservation rate social comparison is prevalent for both low and high users, with the level of baseline water consumption appearing to magnify the effect of being further away from the comparison group. The results are consistent with social comparisons operating as a moral tax on consumption, where the size of the moral tax depends on how far a household is from its peer group. The increased moral costs of consumption decrease the welfare benefits compared to simply analyzing the dollar value of water savings. It is possible that the distance from the peer group also reveals conservation potential in a re-optimization framework. However, this is less likely in our conservation rate social comparison because a savings rate communicates less privately beneficial information since the reference level of the peer group is unknown. While it is likely that both moral and financial incentives are operating in social comparisons, we provide evidence that the moral mechanism is a significant driver of treatment effect heterogeneity.

We find a different pattern of heterogeneity for a new rate treatment that highlights financial savings and provides information on the household's place within the increasing block rate structure. The rate treatment is more effective among households who did not respond to a pro-social appeal from the utility in the previous year. Additionally, the rate treatment does not follow the pattern of "action and backsliding" shown by Allcott and Rogers (2014) to be typical of social comparisons and that we replicate in our study. Rather, a single mailer for the rate treatment is persistent over the study period, and an additional mailer does not generate more water savings. This is consistent with households responding to the social comparisons with intent-oriented actions that fade over time, whereas the financial treatment may lead to impact-oriented actions that are more persistent (Attari, 2014). Although we do not currently have access to daily meter data, such data might allow one to plausibly estimate daily outdoor water use and thus examine whether the rate treatment triggers changes in how households program irrigation systems or whether they fallow lawns for the remainder of the season, two of the most impactful and persistent municipal water conservation actions available in the arid Western U.S.

The behavioral mechanisms underlying the response to nudges are important when evaluating nudges in terms of economic welfare as opposed to water savings. Generating nudges that prompt consumers to re-optimize consumption to address internalities (Allcott et al., 2014) will generally produce higher welfare gains than nudges that impose a moral tax on consumption.

References

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

_ , "Site Selection Bias in Program Evaluation," *Quarterly Journal of Economics*, 2015, *130* (3), 1117–1165.

_ **and Dmitry Taubinsky**, "Evaluating Behaviorally-Motivated Policy: Experimental Evidence from the Lightbulb Market," *American Economic Review*, 2015, pp. 1–80.

_ **and Judd B. Kessler**, "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons," *NBER Working Paper Series*, 2015, *21671.*

_ **and Todd Rogers**, "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *American Economic Review*, 2014, *104* (10), 3003–3037.

_ **, Sendhil Mullainathan, and Dmitry Taubinsky**, "Energy policy with externalities and internalities," *Journal of Public Economics*, apr 2014, *112*, 72–88.

Attari, Shahzeen Z, "Perceptions of water use," *Proceedings of the National Academy of Sciences*, 2014, *111* (14), 5129–5134.

Ayres, Ian, Sophie Raseman, and Alice Shih, "Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage," *Journal of Law, Economics, and Organization*, August 2012, *29* (5), 992–1022.

Bolsen, Toby, Paul J. Ferraro, and Juan Jose Miranda, "Are Voters More Likely to Contribute to Other Public Goods? Evidence from a Large-Scale Randomized Policy Experiment," *American Journal of Political Science*, jan 2014, *58* (1), 17–30.

Brent, Daniel A., Joseph H. Cook, and Skylar Olsen, "Social Comparisons, Household Water Use, and Participation in Utility Conservation Programs: Evidence from Three Randomized Trials," *Journal of the Association of Environmental and Resource Economists*, 2015, *2* (4), 597–627.

Calonico, Sebastian, Matias D. Catteneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

_ **,** _ **, and** _ , "Optimal Data-Driven Regression Discontinuity Plots," *Journal of the American Statistical Association*, 2015, *110* (512), 1753–1769.

Costa, Dora L. and Matthew E. Kahn, "Energy Conservation Nudges and Environmentalist Ideology: Evidence From a Randomized Residential Electricity Field Experiment," *Journal of the European Economic Association*, jun 2013, *11* (3), 680–702.

Croson, Rachel and Jen Yue Shang, "The impact of downward social information on contribution decisions," *Experimental Economics*, 2008, *11* (3), 221–233.

Delmas, Magali A. and Neil Lessem, "Saving power to conserve your reputation? The effectiveness of private versus public information," *Journal of Environmental Economics and Management*, 2014, pp. 1–18.

Dolan, Paul and Robert Metcalfe, "Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation," 2015.

Ferraro, Paul J. and Juan José Miranda, "Heterogeneous Treatment Effects and Mechanisms in Information-Based Environmental Policies: Evidence from a Large-Scale Field Experiment," *Resource and Energy Economics*, 2013, *35* (3), 356–379.

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

_ **, Juan Jose Miranda, and Michael K. Price**, "The Persistence of Treatment Effects with Norm-Based Policy Instruments : Evidence from a Randomized Environmental Policy Experiment," *American Economic Review*, 2011, *101* (3), 318–322.

Frey, Bruno S. and Stephan Meier, "Pro-social behavior in a natural setting," *Journal of Economic Behavior and Organization*, 2004, *54* (1), 65–88.

Hahn, Robert, Robert D Metcalfe, David Novgorodsky, and Michael K Price, "The Behavioralist as Policy Designer: The Need to Test Multiple Treatments to Meet Multiple Targets," *National Bureau of Economic Research Working Paper Series*, 2016, *22886.*

Imbens, Guido W. and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, feb 2008, *142* (2), 615–635.

Ito, Koichiro, "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing," *American Economic Review*, February 2014, *104* (2), 537–563.

_ **, Takanori Ida, and Makoto Tanaka**, "The Persistence of Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand," *E2e Project Working Paper Series*, 2015, *017.*

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (June), 281–355.

Martin, Richard and John Randal, "How is donation behaviour affected by the donations of others?," *Journal of Economic Behavior & Organization*, 2008, *67* (1), 228–238.

Pellerano, José A, Michael K Price, Steven L Puller, and Gonzalo E Sánchez, "Do Extrinsic Incentives Undermine Social Norms? Evidence from a Field Experiment in Energy Conservation," *Environmental and Resource Economics*, Forthcoming.

Schultz, P Wesley, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vladas Griskevicius, "The constructive, destructive, and reconstructive power of social norms," *Psychological Science*, 2007, *18* (5), 429.

Shang, Jen and Rachel T.A. Croson, "A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods," *Economic Journal*, 2009, *119* (540), 1422–1439.

Wichman, Casey J., "Perceived price in residential water demand: Evidence from a natural experiment," *Journal of Economic Behavior & Organization*, 2014, pp. 1–16.

_ **, Laura O. Taylor, and Roger H. von Haefen**, "Conservation policies: Who responds to price and who responds to prescription?," *Journal of Environmental Economics and Management*, 2016, *79*, 114 – 134.

Appendix

Implementation

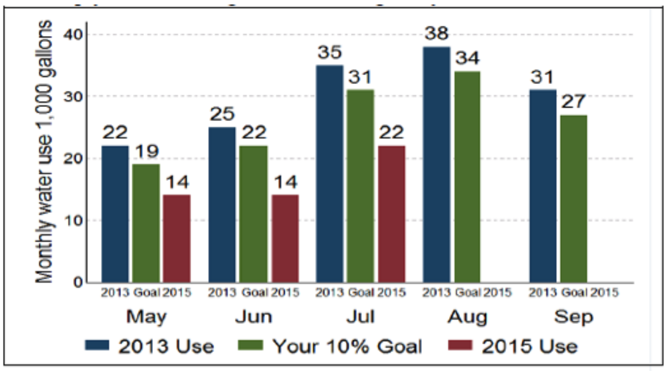
The process for generating and mailing these letters is as follows:

- 1-2 days after the most recent month's water use data is loaded into the billing system we pull this information into Stata and using a set of pre-programmed routines use it to generate the graphics and data for the mail merge.
- A mail merge is performed in Microsoft Word using the generated data and graphics.
- PDFs of the letters are emailed to Digiprint
- Digiprint prints and ships the the letters within 1-3 days of receiving the electronic files.

The average time from the data upload to letters shipment was 2 days with a maximum of 8 days during this study. We had attrition during the study of about 1.5 percent of the treatment customers; 142 customers dropped out the study in July and 211 customers dropped out the study in August. This attrition was likely due to customers closing accounts or billing data (meter reading) errors. Furthermore, the mailers did not generate a very large increase in call center volume; out of the 23,213 customers we attempted to reach with this pilot we estimate that only 43 contacted the call center. Most of the customers who called the call center just wanted to ask clarification questions about the information in their letter; only 26 wanted further assistance beyond what the call center representatives could provide; and only five customers ended being truly upset by the pilot program.

Treatment Figures

Figure A.1: Treatment 2 - Historical water use information



<u>What is *your* 10% goal?</u> The graph to the left shows your household's total water use in the past few months (red) compared to your water use in summer 2013 (blue). Water use on your July 2015 bill was **37% lower** than on your July 2013 bill. To do your part to help the community conserve water this summer, your total water use for August and September should be **at least 10%** lower than in 2013 (green). Your August goal amounts to saving **133 gallons per day** compared to how you used water in August 2013.

Figure A.2: Treatment 3 - Rate structure information

<u>Saving water saves you money</u>. The graph on the right shows how you **could save $13.00** if you reduce your water use on your August bill by 10% compared to 2013. The price you pay for each gallon of water increases as total water use increases. For most homes, indoor use falls within the first tier at the lowest price. Summer outdoor water use brings homes into the higher prices associated with tiers 2 and higher. The graph shows where your water use from your July 2015 bill (red) falls in the price schedule. It also shows your August 2013 water use (blue) and your 10% reduction goal for August 2015 (green).

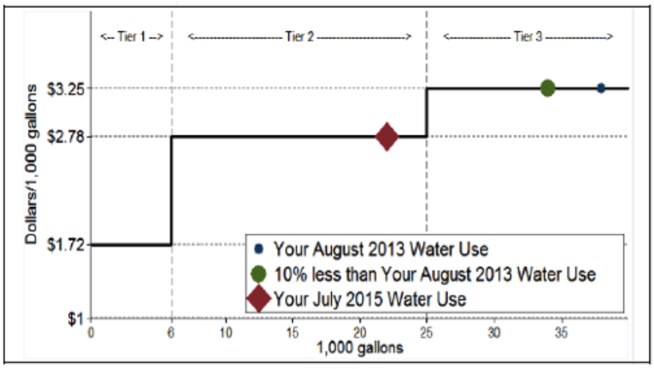


Figure A.3: Treatment 4 - Social comparison, reported in thousands of gallons

<u>How does your water use compare?</u>
The graph on the right shows your water use from your July bill compared to similar properties in your area. You used **1,000 gallons less** than your neighbors with similar properties.

You saved 35% on your July bill compared to 2013.
Keep up the good work!

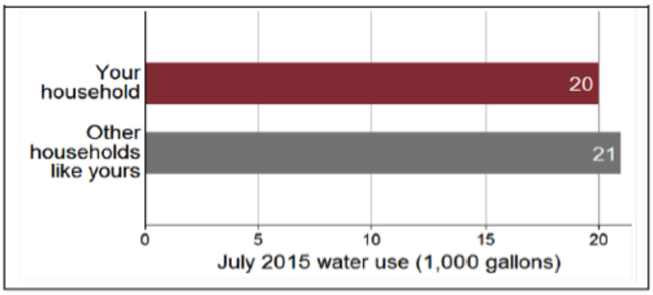
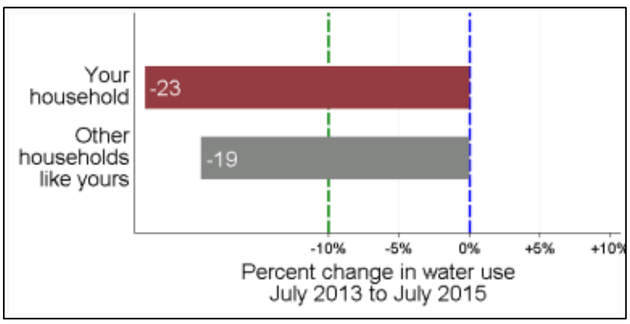


Figure A.4: Treatment 5 - Social comparison, reported as progress towards TMWA's 10% conservation goal

<u>Are you doing your part?</u> The graph on the right shows your change in water use form July 2013 to July 2015 compared to similar properties in your neighborhood. Your neighbors used 19% less water last month compared to 2013.

You saved 23% on your July water bill compared to 2013.
Keep up the good work!



Additional Balance Tests

Table A.1: Balance on Pre-treatment Water Use for each Letter

	Control Mean	Treatment Mean	Difference	(p-value)
Letter 1	21.69	21.81	-0.11	0.66
Letter 2	21.56	21.69	-0.12	0.63
Letter 3	21.63	21.66	-0.03	0.90
Letter 4	21.62	21.66	-0.04	0.86
Letter 5	21.65	21.70	-0.05	0.84

Note: p-values are based on two-sided t-tests

Table A.2: Balance on Pre-treatment Water Use across Letters

	Mean 1	Mean 2	Difference	(p-value)
Letter 1 vs Letter 2	21.75	21.62	0.13	0.49
Letter 1 vs Letter 3	21.75	21.65	0.10	0.58
Letter 1 vs Letter 4	21.75	21.64	0.11	0.55
Letter 1 vs Letter 5	21.75	21.67	0.08	0.68
Letter 2 vs Letter 3	21.62	21.65	-0.02	0.89
Letter 2 vs Letter 4	21.62	21.64	-0.02	0.93
Letter 2 vs Letter 5	21.62	21.67	-0.05	0.78
Letter 3 vs Letter 4	21.65	21.64	0.01	0.96
Letter 3 vs Letter 5	21.65	21.67	-0.03	0.89
Letter 1 vs Letter 5	21.64	21.67	-0.03	0.85

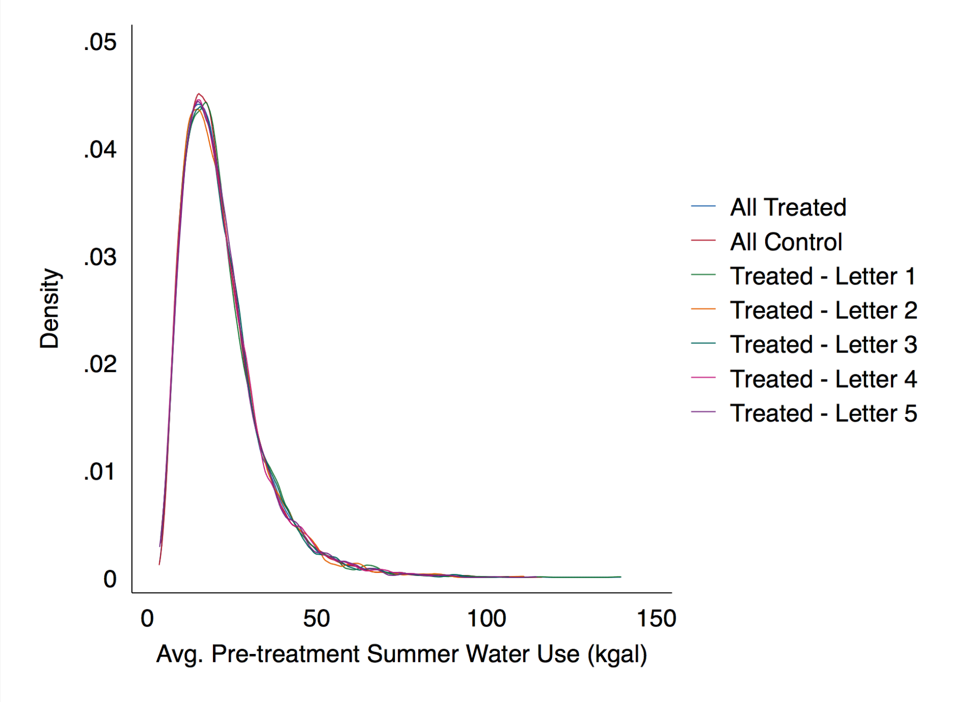
Note: p-values are based on two-sided t-tests

Table A.3: Balance within Deciles of Pre-treatment Water Use

	Control Mean	Treatment Mean	Difference	(p-value)
Decile 1	7.96	7.91	0.05	0.20
Decile 2	11.06	11.06	-0.00	0.97
Decile 3	13.50	13.45	0.04	0.02
Decile 4	15.70	15.68	0.02	0.32
Decile 5	17.95	17.94	0.01	0.59
Decile 6	20.36	20.38	-0.02	0.41
Decile 7	23.25	23.23	0.02	0.38
Decile 8	26.81	26.83	-0.02	0.62
Decile 9	32.39	32.32	0.07	0.27
Decile 10	48.24	48.13	0.11	0.77

Note: p-values are based on two-sided t-tests

Figure A.5: Distributions of Pre-treatment Water Use Across Treatment Status



Note: The lines are kernel density estimates of pre-treatment water use for all treated households, all control households, and households within each of the five treatment groups.

Table A.4: Kolmogorov-Smirnov Tests

	D-statistic	(p-value)
All Treatment vs Control	0.01	0.88
Letter 1 vs Control	0.01	0.89
Letter 2 vs Control	0.01	0.87
Letter 3 vs Control	0.01	0.86
Letter 4 vs Control	0.01	0.99
Letter 5 vs Control	0.01	0.76

Note: p-values are based on the combined D-statistic

Additional Specifications

Table A.5: Heterogeneity by whether household met 10% goal in all periods

	(1)	(2)	(3)	(4)	(5)
	Pooled	Letter 1	Letter 2	Letter 3	Letter 4
Treat*Always	-0.313	0.293	-1.193	-0.158	-0.879
	(0.443)	(0.837)	(0.911)	(0.810)	(0.865)
Treat*Mixed	-0.994**	-1.197	-1.231	-0.821	-1.557**
	(0.389)	(0.734)	(0.752)	(0.722)	(0.779)
Treat*Never	-3.338***	-0.936	-2.104	-4.197**	-4.703**
	(1.147)	(2.205)	(2.519)	(1.710)	(2.356)
Always Met Goal	-30.79***	-30.80***	-30.81***	-30.80***	-30.80***
	(0.723)	(0.724)	(0.723)	(0.723)	(0.723)
Mixed Met Goal	-15.17***	-15.15***	-15.17***	-15.20***	-15.17***
	(0.707)	(0.709)	(0.708)	(0.708)	(0.708)
Observations	91,280	68,948	68,954	68,933	68,929

Note: The dependent variable is normalized average daily water use, and the sample is restricted post treatment data dropping all observations after the second mailer. Always, Mixed and, Never are indicator variables for how often the household met the goal. All regressions include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and baseline water use. Robust standard errors clustered at the household level are reported in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Discrete Effects Moving Above the Peer Group and 10% Goal

It is important to distinguish the continuous distance from the peer group consumption from the discrete injunctive norm defining appropriate behavior (Schultz et al., 2007). We consider two separate effects. First, we test whether performing slightly worse than one's peers has an effect on consumption, and second we test for the discrete effect of just barely missing the 10% conservation goal. In our setting the descriptive injunctive norm is based on whether a household met the 10% goal. Therefore if a household less than 10% it received the message, "Please do your part to help with the drought.", while a household that saved more than 10% was told, "Keep up the good work!". It is also possible that the household considered their performance relative to their peer group as an additional categorical norm. The results in Table 5 contain *both* the effect of the discrete injunctive norm and the continuous descriptive norm.

To isolate the effect of the injunctive norm, we employ a regression discontinuity (RD) design (Imbens and Lemieux, 2008; Lee and Lemieux, 2010), analyzing behavior on either side of the injunctive category similar to Allcott (2011).[16] In the RD analysis we restrict the sample to households treated with a social comparison (T4 and T5) and examine the effect of being just above the peer group. The running variable is a household's distance from its peer group in gallons for Treatment 4 and in percentage reduction for Treatment 5. The dependent variable in both is residual normalized consumption based on a regression of normalized water use on weather, month fixed effects, and household fixed effects, following the approach of (Allcott, 2011). The RD estimation assumes that factors varying with the distance from the peer group, such as baseline water and the strength of the descriptive norm, are the same for households just above and below their peer group. Since some households were above their peer group but saved more than 10%. Similarly some households were below their peer and saved less than 10%. We repeat the analysis dropping these households and the results are very similar.

We begin with graphical evidence of differences in water use near the peer group (Figure A.6), as is standard in RD approaches.[17] For both social comparisons the graphical evidence in Figure A.6 suggests that moving above the peer group does not affect consumption.

The graphical evidence is corroborated in the RD estimates for moving above the peer group. We use three different RD estimators developed by Calonico et al. (2014): the conventional, bias-corrected, and bias corrected with robust standard errors. In all specifications the impact for moving above the peer group is small and not statistically significant (Table A.6). The RD estimates show that the effects in Table **??** are driven by the *distance* from the peer group as opposed to simply being above or below the peer group. There is either no effect associated with

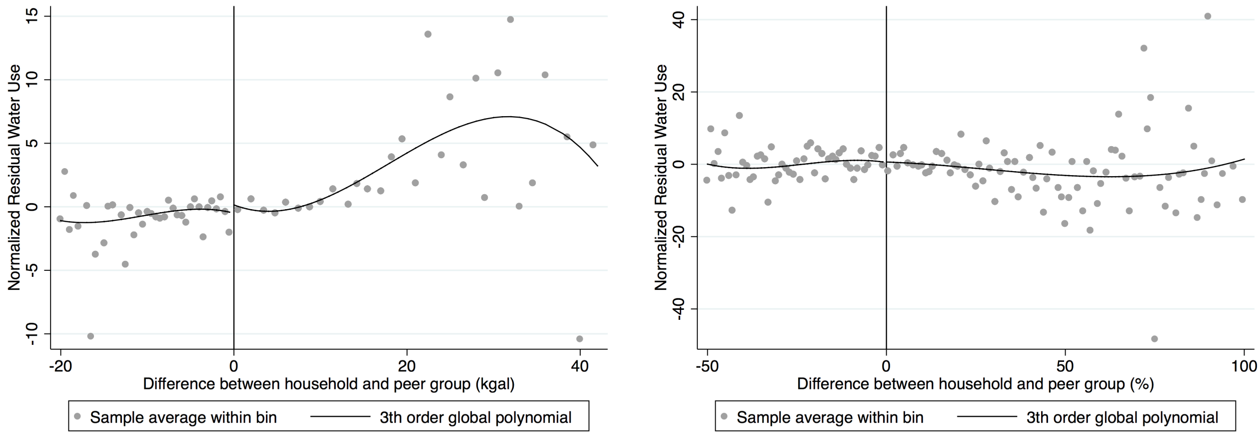
[16] Allcott (2011) showed little impact of moving into one of the three distinct categories in the Home Energy Report ("Great", "Good", or "Below Average") in a regression discontinuity design. In that study a household is assigned the category "Great" if they consume below the 20th percentile of its peers, "Good" if they consume below the average of its peers, or "Below Average" if they consume above the average of its peers.

[17] The graphs are generated with a data-driven approach using spacing estimators to generate the bin sizes in the plots (Calonico et al., 2015). The points on the graph are the average normalized residual consumption within each bin, and the lines are the fitted values of separate third-order polynomial regressions on either side of the distance threshold (zero).

adding a negative injunctive norm.

Figure A.6: Effect of Moving Above the Peer Group

(a) Norm in kgal (T4) (b) Norm in percentage reduction (T5)



Note: The dependent variable is residual normalized consumption and the units are percentage terms. The discontinuity is based on the moving above the peer group's conservation rate.

Table A.6: Regression Discontinuity Estimates of the Moving Above the Peer Group

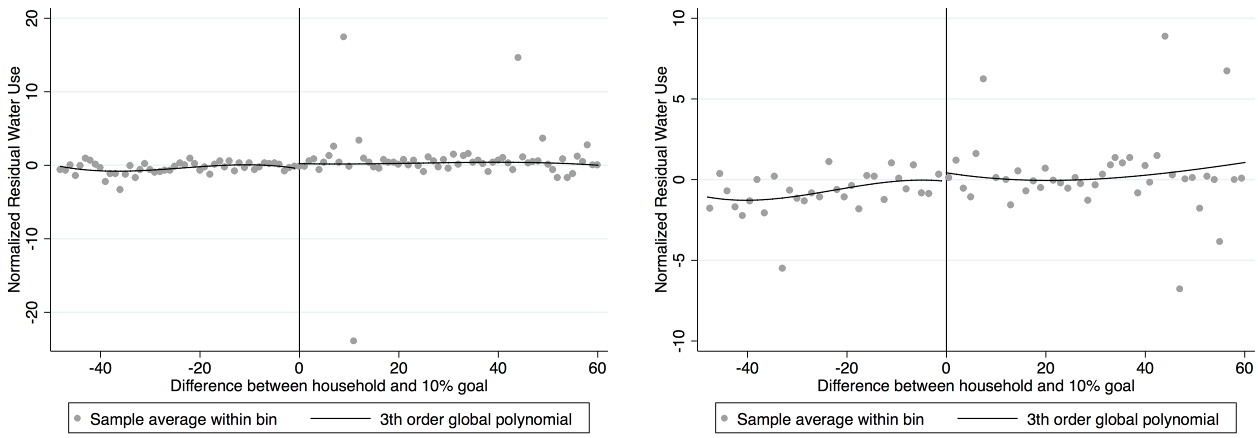
	(1) Kgal (T4)	(2) % (T5)
Conventional	1.326	-0.159
	(0.897)	(1.037)
Bias-corrected	1.329	-0.212
	(0.897)	(1.037)
Robust	1.329	-0.212
	(1.063)	(1.234)
Observations	3,842	2,869

Note: The rows are three separate RD estimators and the appropriate standard errors according to Calonico et al. (2014). The dependent variable is residual normalized consumption. The discontinuity in column (1) is based on the moving above the peer group's consumption (T4), and in column (2) is based on the moving above the peer group's year-on-year change in consumption (4).

We repeat this exercise to see if moving slightly below above the 10% goal influences consumption. The analysis is the same as reported above except the running variable is the year-on-year percentage change in water consumption and the threshold is the -10%. To be consistent with the analysis in the main text we subtract 10% off the running variable and such that the threshold is at zero and year-on-year changes of less than 10% are positive and more than 10% are negative. Figure A.7 graphs residual consumption on the y-axis with the year-on-year percentage change in water consumption (minus 10%) on the x-axis. We show all treatments pooled and Treatment 2, where the historical information relative to the 10% goal is the only personalized information. There is no visual evidence of a change in consumption right at the threshold. This is corroborated with the RD estimates for the pooled treatment group and each of the treatments individually.

Figure A.7: Effect of the Failing to Meet the 10% Goal

(a) All Treatments (b) Tips + History (T2)



The dependent variable is residual normalized consumption and the units are percentage terms. The discontinuity is based on the moving above the consumption threshold that constitutes the household's 10% goal. Data from all treatments are included.

Table A.7: Regression Discontinuity Estimates of the Failing to Meet the 10% Goal

	(1) Pooled	(2) Letter 1	(3) Letter 2	(4) Letter 3	(5) Letter 4	(6) Letter 5
Conventional	0.377	1.029	0.238	-0.308	1.022	-0.047
	(0.543)	(1.147)	(1.104)	(1.193)	(1.253)	(1.146)
Bias-corrected	0.508	1.290	0.338	-0.387	1.262	0.058
	(0.543)	(1.147)	(1.104)	(1.193)	(1.253)	(1.146)
Robust	0.508	1.290	0.338	-0.387	1.262	0.058
	(0.630)	(1.299)	(1.312)	(1.415)	(1.443)	(1.338)
Observations	15,432	3,100	3,472	3,568	3,409	3,269

Note: The rows are three separate RD estimators and the appropriate standard errors according to Calonico et al. (2014). The dependent variable is residual normalized consumption. The discontinuity in column (1) is based on moving above the consumption threshold that constitutes the household's 10% goal. The columns show the pooled treatment and each of the individual treatments.