

What causes heterogeneous responses to social comparison messages for water conservation?

Working Paper

Daniel A. Brent^{*a}, Corey Lott^b, Michael Taylor^c, Joseph Cook^d,
Kimberly Rollins^e, and Shawn Stoddard^f

^aDepartment of Agricultural Economics, Sociology, & Education, Pennsylvania State University: dab320@psu.edu

^bEconomist, Laurits R Christensen Associates, Inc: corejeana@gmail.com

^cDepartment of Economics, University of Nevada Reno: taylor@unr.edu

^dSchool of Economic Sciences, Washington State University: joe.cook@wsu.edu

^eDepartment of Agricultural and Resource Economics, University of Connecticut:
kimberly.rollins@uconn.edu

^fTruckee Meadows Water Authority: sstoddard@tmwa.com

May 2020

Abstract

Social comparisons for water conservation are often implemented in conjunction with a broader set of drought management policies. We investigate the interaction of social comparisons with prior responses to voluntary appeals for water conservation using a large-scale field experiment in Reno, Nevada. We develop a new social comparison framed as performance toward a conservation goal, in contrast to the traditional comparison made in gallons. Our new social comparison decouples the performance relative to the peer group from pre-treatment water use, allowing us to investigate the role of the peer comparison independently from pre-treatment water use. We find that pre-treatment water use drives treatment heterogeneity in the traditional social comparison, while prior conservation effort drives treatment heterogeneity in the new social comparison. The results confirm that the strength of the social norm (i.e. under-performance relative one's peers) is critical in generating water conservation. Simple targeting of both types of social comparisons could have increased aggregate savings by 38% in our study site because the new social comparison generates conservation among a different set of households than the traditional social comparison.

Keywords: behavioral interventions, social norms, field experiments, water conservation, water demand

JEL Classification Numbers: D12, C93, H42, L95, Q21, Q25.

*Corresponding Author: Daniel Brent, dab320@psu.edu. The authors are grateful to the Truckee Meadows Water Authority (TMWA) for financial support and 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. Brent acknowledges funding from the Cooperative Research Centre for Water Sensitive Cities (CRC grant number 20110044).

1 Introduction

Severe droughts facing urban populations are unfortunately becoming an increasingly common event. Many water utilities facing droughts cannot raise prices due to political or regulatory constraints, leading them to embrace a variety of non-price water conservation policies. A popular non-price policy is a social comparison message (SCM), which compares a household’s water use to the water use of a peer group. A number of randomized controlled trials have established that SCMs consistently lead to conservation in the range of 3-5%, and similar approaches have been used to reduce energy consumption as well.¹

A strikingly consistent feature of SCMs is that they are more effective among high-use households in both water and energy (Allcott, 2011; Ayres et al., 2013; Ferraro and Price, 2013; Brent et al., 2015). In fact, many randomized controlled trials find no savings among users in the bottom quartiles of pre-treatment consumption (Brent et al., 2015; Torres and Carlsson, 2018). One characteristic of these SCMs is that the comparison shown to high-use households is, not surprisingly, more likely to show that they use more water than their peer group. This is true even when peer groups are constructed of seemingly homogeneous households based on neighborhoods and household characteristics. We posit the “strength” of the normative message in SCMs depends on the distance between a household’s performance (e.g. water use) and the performance of the relevant comparison group. When SCMs inform these high-use households that they are badly under-performing relative to their peers, they receive “strong” normative messages. The argument that the strength of the normative message is a key component of SCMs is consistent with utility-theoretic models utility of SCMs (Taylor et al., 2018; Allcott and Kessler, 2019), as well as empirical evidence from other settings such as charitable giving (Croson and Shang, 2008; Shang and Croson, 2009).

Combining the empirical pattern that high users are more responsive to SCMs with the design feature that high users are more likely to receive strong normative messages leads

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

to our primary research question. What explains the pattern of treatment heterogeneity in SCMs? Is it the type of household or the content of the normative message? There are logical arguments for whether households or messages matter. The prior literature has not been able to disentangle these factors due to the high degree of correlation between pre-treatment consumption and the strength of the normative message.

The notion that the content of the social comparison a household receives matters is well developed in the literature on SCMs. Schultz et al. (2007) show that adding an injunctive norm to a SCMs for energy eliminates the “boomerang effect” where households consuming below their peer group increase consumption. The injunctive norm, usually in the form of a smiley emoticon for low users, is now standard practice in commercial applications that use SCMs for energy and water conservation.² However, Allcott (2011) found that the injunctive norm combined with SCMs did not affect energy consumption in a regression discontinuity design around the thresholds for the assignment of different injunctive norms. In the literature on charitable giving Croson and Shang (2008) find that SCMs were able to increase donations for donors who previously contributed less than the comparison level but decreased donations for donors who previously contributed above the level. A complication in SCMs for water and energy is that conservation also generates private benefits in reduced bills. Allcott and Kessler (2019) show that households consuming above their peer group are more likely to expect that the SCM saves them money but also feel pressured and experience guilt from the SCMs.

The type of household might matter for several reasons. Some households might have more margins of adjustment in water use, lower opportunity costs of conservation, or more ability to simply scale up similar actions that low-use households perform. Furthermore, SCM campaigns for water conservation are increasingly implemented subsequent to, or concurrently with, other drought management policies. The presence of these additional policies likely changes a household’s opportunity costs of conservation, which is one of the factors that might make a household more responsive to a SCM. Customers who have already reduced their water use in response to prior conservation policies may be

²Opower, which is now part of Oracle, and WaterSmart are two companies that used emoticons in SCMs for energy and water conservation, respectively.

less responsive to SCMs relative to those who have not. This is particularly relevant for locations suffering through multi-year droughts, where utilities persistently ask for additional conservation. In this context, *prior conservation* may be a stronger predictor of customer response to SCMs than pre-treatment water use, which is commonly used in targeted SCM campaigns for cost-effectiveness. As utilities continue to implement SCMs within a suite of conservation policies it is important to understand how these nudges fit within a holistic water conservation policy.³ The effect of prior conservation on future conservation is a question of broader interest if prior conservation efforts diminish the ability of households to conserve in the future.⁴

We address these questions by implementing a randomized field experiment with the Truckee Meadows Water Authority's (TMWA) in Reno, Nevada. Our experiment had two important features. First, the experiment was conducted in the context of a multi-year drought in conjunction with TMWA's broader drought management plan. In the summer of 2014, drought conditions required TMWA to make a voluntary appeal that each customer reduce consumption by 10% relative to the summer of 2013, implemented primarily through a large-scale public outreach campaign.⁵ The drought persisted, and in the summer of 2015 TMWA replicated and extended the public appeal for a 10% reduction across the whole summer season. Second, we developed a new SCM that compares households' percentage change in consumption relative to the same month in 2013 to the corresponding percentage reduction of similar neighbors. This new social comparison leverages the utility-wide public appeal for conservation and decouples the type of message a household receives from their pre-treatment water consumption. We compare this novel treatment to a separate treatment arm that received the traditional SCM in gallons. In this paper we refer to the traditional SCM in gallons as T1 and the new SCM in percentage reductions as T2. The voluntary appeal for conservation in 2014

³For example, WaterSmart software operates or operated multiple SCM campaigns in California utilities during the recent drought ending in 2015. In addition to the SCM campaigns, there was a statewide call for conservation that included voluntary conservation, mandatory outdoor watering restrictions, and water efficiency rebate programs such as converting grass to xeriscape.

⁴In related research Brent (2018) shows how converting to drought-resistant landscapes affects demand elasticity parameters.

⁵During this first year of drought the need for conservation was less urgent; TMWA did not make the request for voluntary conservation until late July.

also allows us to explore treatment heterogeneity based on households' response to the 2014 appeal, which we call "prior conservation". We define prior conservation as the percentage change in consumption from summer 2013 to summer 2014.

Our research design makes two primary contributions to the literature. First, we examine how the content of the normative message impacts the response to SCMs. Prior studies using conventional social comparisons in gallons (or kilowatt hours) are unable to disentangle whether high-use households are more responsive because they have lower opportunity costs of conservation or because they are more likely to receive an SCM with a strong signal. Consistent with previous studies, in our traditional SCM in gallons (T1), pre-treatment consumption is highly correlated with the strength of the normative message. High users are more likely to be informed that they use more water than their peer group. The key feature of our study design is that our new SCM in percentage terms (T2) generates the peer comparison using *percentage changes* in consumption rather than the level of water use. Our new peer comparison based on changes in consumption has a much lower correlation between pre-treatment water use and the strength of the normative message. Consequently, relative to the traditional SCM, low-use households in our new SCM have a higher probability of receiving a "strong" message informing them that they are performing worse than their peers. In T2, prior conservation is in fact more strongly correlated with the strength of the normative message, because households that have undertaken less conservation relative to 2013 are more likely to be informed that they are contributing less toward the 10% reduction goal than their peer group.

Second, we examine how prior conservation is related to the opportunity cost of conservation. Examining the effect of prior conservation requires both the existence of conservation policies prior to the implementation of SCMs and an identification strategy to estimate the effect of prior conservation separately from pre-treatment consumption. Our setting is ideal to address the effect of prior conservation given that the experiment took place in the second year of a multi-year drought where the utility called for a voluntary appeal in the first year. From an econometric perspective, we can identify the effect of both pre-treatment water use and prior conservation on treatment heterogeneity

because they have a relatively low correlation and our treatment assignment is balanced across both variables. The role of prior conservation can best be described through an example. Consider a household that adjusted their irrigation controller in response to the 2014 appeal. They will likely do so again (or the controller only required a one-time adjustment) in response to the 2015 appeal, regardless of whether they receive a SCM. This household will have a higher opportunity cost for *additional* conservation *beyond* what they did in 2014 compared to a household who did not adjust their irrigation controller in 2014. This example shows that when general conservation policies exist, such as TMWA’s utility-wide voluntary reductions, the conservation generated from a SCM may depend on households’ responses to pre-existing conservation policies. Therefore, we argue that prior conservation is related to the opportunity cost of *further* conservation, which may also be an important driver of heterogeneity when analyzing behavioral nudges for water conservation.

Each of the SCM treatments generates statistically significant average treatment effects (ATEs) of roughly 1.5%. We investigate treatment heterogeneity based on both pre-treatment consumption and prior conservation. The strength of the normative message is a major driver of customer response. Prior conservation generates treatment heterogeneity when it is closely correlated with the strength of the normative signal, as is the case in our SCM in percentage terms. By contrast, prior conservation has no significant effect on treatment heterogeneity for the traditional SCM in gallons. Pre-treatment water use increases the effectiveness of both SCMs (T1 and T2), but to a greater extent when it is correlated with strong normative signals (T1).

The results support two conclusions. First, since the patterns of heterogeneity align closely with the strength of the normative signal, we argue that SCMs likely need to convey that a household is under-performing relative to their peers to generate significant conservation. Since pre-treatment water use appears to increase the magnitude of treatment effects regardless of its correlation with the normative signal, we cannot rule out that opportunity costs also play a role in the response to SCMs. Second, prior conservation is a less important proxy for opportunity costs of conservation than pre-treatment

water use in our setting.

One implication of our results is that the new SCM in percentage terms can be effective among low users who have not responded to earlier non-SCM conservation policies. Achieving conservation from low users contrasts with prior studies finding almost no treatment effects for low-use households, since they would be unlikely to receive a strong normative message in a gallons-based SCM. Sending strong normative messages to low users generates conservation among customer classes that were not previously receptive to SCMs, which may be particularly important when significant water curtailments are necessary during extreme droughts. A simple targeting rule based on prior conservation leads to a 38% increase in the aggregate gallons saved, showing the importance of searching for the mechanisms behind established patterns of heterogeneity.

This research is related to the work on the heterogeneous impacts of SCMs for water and energy conservation. Nemati et al. (2019) finds that low users in the bottom two quintiles of the pre-treatment to not conserve any water when treated with SCM and water data analytics in Fulsom City, California. In fact, the bottom quintile of treated households has a statistically significant increase in water use. Goette et al. (2019) shows that low users have almost no response to a water conservation treatment including an SCM that generated savings of roughly 6 liters per day. Ferraro and Price (2013) and Ferraro and Miranda (2013) show that among three different information treatments for water conservation in Cobb County, Georgia, the SCM generated the largest differential effect by pre-treatment water use. While low-use households still have statistically significant CATEs, high users save 2-3 times more water than low users. Bhanot (2018) finds similar effects for SCM messages in California; most of the conservation is concentrated in the higher pre-treatment water use deciles and most lower deciles have CATEs close to zero. Torres and Carlsson (2018) examine the spillover effects of SCMs in Columbia and find insignificant CATEs for low users and statistically significant CATEs three times larger for high users. Brick et al. (2017) conduct a variety of randomized treatments, including a SCM, for water conservation to cope with severe drought in Cape Town, South Africa. They also find that the lowest quintile of the pre-treatment are not responsive

to the SCM while higher quintiles to conserve in response to treatment. Similar effects have been observed in energy SCMs; Allcott (2011) shows a nearly monotonic relationship between pre-treatment energy consumption on the effectiveness of SCMs across 17 experimental samples.

This research is also related to the work on how behavioral nudges interact with alternative existing policies. Pellerano et al. (2015) interacts social comparisons with features of electricity tariffs and finds evidence for crowding out of intrinsic motivations. West et al. (2019) examines the interaction of social comparisons with water restrictions policies, finding little evidence of crowding out. Brent et al. (2015) find the social comparison increase the probability of signing up for alternative utility water conservation policies. Brent and Wichman (2020) find little interaction between prices and social comparisons in a large southern California utility. We contribute to this research by highlighting how social comparisons can be integrated into utility-wide voluntary appeals for conservation to produce additional conservation, including a new SCM format that draws upon the prosocial contribution to the regional drought.

We also contribute to research on the mechanisms through which behavioral nudges operate. A simple dichotomy is that SCMs can impose a "moral tax" on consumption (Levitt and List, 2007) or provide valuable privately beneficial information to assist in optimizing a household production function (Becker, 1965). The re-optimization mechanism is related to research on "internalities" - the failure of consumers to fully account for all the *private* costs of consumption (Allcott and Kessler, 2019; Allcott et al., 2014; Allcott and Sunstein, 2015; Allcott and Taubinsky, 2015).⁶ Allcott and Kessler (2019) argue that the moral mechanism of response can reduces the welfare gains from SCMs, which may be welfare-reducing for a sizable portion of households.

⁶Prominent examples of internalities in the water and energy sectors include: imperfect information about the costs of water/energy consumption (Allcott and Taubinsky, 2015); dynamic inconsistencies in decision-making (Allcott et al., 2014); cognitive constraints (Brent and Ward, Forthcoming); lack of salience of infrequent or automatic billing (Sexton, 2015; Wichman, 2016); and confusion about nonlinear price structures (Ito, 2014; Wichman, 2014; Lott, 2017).

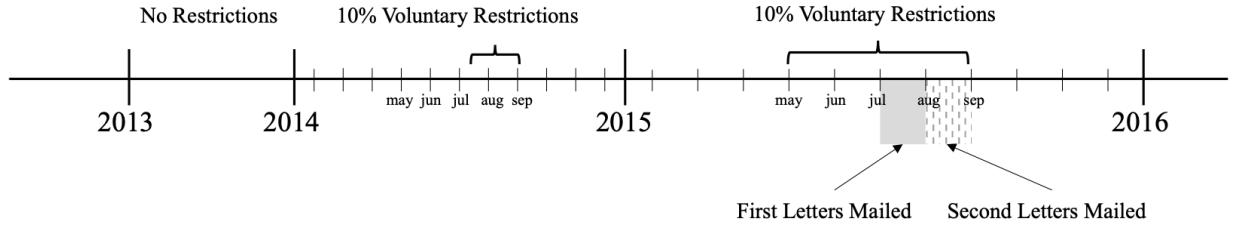
2 Background & Experimental Design

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. Water supply to the Reno metro area is primarily provided by the Truckee Meadows Water Authority (TMWA). The Truckee River is the primary source for TMWA's water supply, which relies on snowmelt from the Sierra Nevada as well as storage provided by Lake Tahoe and Pyramid Lake. Water demand is highly seasonal, with the peak demand period coming in the summer to meet demand for residential irrigation. In 2015, in response to an expected drought, TMWA launched a major media campaign to reduce water use during the summer irrigation season by requesting that each TMWA customer use 10% less water from May through September 2015, relative to their water use during the same months in 2013. TMWA used bill inserts as well as a wide variety of media including print, radio, TV, social media, and billboard messages to publicize the conservation message. TMWA used 2013 as the comparison year for the 10% reduction because they had also asked for conservation during the summer of 2014. Such a conservation request was uncommon in the region prior to 2014; the last time TMWA made a request was during a drought in 1992.

We employ our new SCMs in a utility-scale randomized field experiment conducted in partnership with TMWA. Single-family customers received one of five mailers, with approximately 4,300 households included in each of five treatment groups, and 21,552 in the control group. The control group did not receive any informational mailers, but both treatment and control groups were exposed to the drought messaging asking them to voluntarily reduce consumption by 10% from the media, billboards, and messages printed on monthly bills. Two of the five treatments include a SCM: one is a traditional SCM in terms of total gallons used by the household relative to a peer group; the other is our new SCM using percentage differences as described above. Both SCMs describe TMWA's goal for each household to use 10% less water for each month of the summer of 2015 relative to the summer of 2013 to cope with the drought. In this paper, we focus primarily on the two SCMs; we refer to the traditional SCM in gallons as T1 and the new

SCM in percentage reductions as T2. Figure 1 provides a timeline of water conservation policies TMWA imposed to provide context of where our experiment fits into broader utility water management policy.

Figure 1: Timeline of Water Conservation Policies



Note: The figure presents the timing of voluntary restrictions in the years prior to the drought as well as the timing of our experiment.

2.1 Description of Treatments

This article focuses on the two SCMs (T1 and T2) out of the five employed in the overall field experiment. Since the treatments of interest contain components of the first two treatments we present a brief description of all five treatments. Table 1 summarizes the information in all treatments and the Appendix includes example 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.” We reference the three treatments that are not analyzed in this paper as A1, A2, and A3.

Treatment A1 provided households with six tips that the TMWA media campaign publicized for how to reduce outdoor water consumption, similar to Ferraro et al. (2011). This letter was not customized to report on individual household water use. The six tips were printed on the reverse side of mailers for all other treatments including the two SCMs. The tips were the same for all households an example is provided in the Appendix (Figure A.1),

Treatment A2 augmented the generic tips with personalized information about the customer’s water use, with a title introducing the letters that read: “Below is your customized water use report.” The A2 letter 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 before the letter was sent out (Appendix Figure A.2 shows the mailer). This figure and accompanying descriptive text was also included with A3, T1, and T2. Therefore, the SCM treatments also include water conservation tips and personalized historical water use information.

Treatment A3 contained the same components as A2, with the additional message “Saving water saves you money”, a figure displaying (i) the rate structure with tiers and price for each tier, (ii) the customer’s water use in kgal within TMWA’s increasing-block rate structure for the last month billed in 2015, and (iii) the upcoming month’s target of 10% less water than the same month in 2013 within the rate structure. The letter also provided the monetary savings that the customer could expect from meeting this goal (see Appendix Figure A.3).

In addition to the tips and personalized information, T1 contained a social comparison message under the header “How does your water use compare?” The core of T1 is a graphic 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.⁷ In essence, T1 reproduces the standard SCM 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.4).

T2 was similar to T1 except that the SCM was framed in terms of relative performance toward achieving the 10% goal and the comparison graphic was based on the percentage change in water use in the previous billed month relative the same month in 2013 (Ap-

⁷Households were compared to other households in their same billing cycle, which the utility uses to divide customers into contiguous neighborhood units. Furthermore, customers within the same billing cycle were divided into above/below median yard size (square feet) and above/below median number of bedrooms. In a few cases in which comparison groups were sparse, customers were segmented based on above/below the median yard size only.

pendix Figure A.5). The comparison group was identified in the same way as T1. We also included injunctive norms in the form of a message rather than emoticons or “smiley faces” as in Schultz et al. (2007). Residences that had met their 10% goal in the last billed month received the message “Keep up the good work.” Residences 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.” The injunctive norms for T1 and T2 were both based on whether households had met the 10% goal and were not based on the peer comparison.

Importantly, T1 and T2 contain the same information on an individual household’s performance towards their 10% goal and the same injunctive norm. The only difference in T1 and T2 mailers is that the peer comparison in T1 is based on total gallons and the T2 peer comparison is based on changes in consumption. Figures A.1 - A.5 show what mailers one specific household would have received had they been assigned to each of the five treatments. Importantly, as seen in Figures A.4 and A.5 our example household used less water than their peers but simultaneously had not reduced water use by as large a percentage as their peers. Therefore, this household would have received a stronger normative message being assigned to T2 than T1 without changing any household characteristics. The reverse situation could also be true: a household that used more water than its peers but reduced by a larger percentage would receive a stronger normative message if they were assigned to T1 rather than T2.

Table 1: Information Included in the Five Treatments

Information Components Included in Mailers*	Treatment (Mailer)				
	A1	A2	A3	T1	T2
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: <ul style="list-style-type: none"> • 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: <ul style="list-style-type: none"> • 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

Note: See Figures A.1 - A.5 in the Appendix for examples of each information component of each treatment. A2, A3, T1, and T2 included the title: “Below is your customized water use report”. The treatments in the first three columns shaded grey are not used in this paper, although components of A1 and A2 are present in T1 and T2.

2.2 Randomization

Our sample frame included 42,703 eligible⁸ single family homes, which we randomly assigned to either the control group or one of the five treatment groups. Randomization blocks were defined by billing cycles, rate schedule, and frequency of recorded meter data (i.e. monthly, daily, or hourly, though all customers only receive monthly usage totals). The Appendix provides more details about our randomization procedure and the process of generating the mailers. In total, 21,151 treatment households were assigned to receive mailers of which 4,231 were selected for T1 and 4,217 were selected for T2. The control groups consists of 21,552 households. We also randomized whether households received one or two mailers. A total of 2,839 households were assigned to receive a single mailer in July (using June consumption as the last month billed), 2,819 received a single mailer in August (using July consumption as the last month billed), and 2,790 received mailers

⁸Specifically, we included 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 (6,000 gals), indicating some outdoor water use; (iii) had lived at their current residence since April 2013, and therefore had summer 2013 bills 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 unusually 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 unoccupied for an extended period of time.

Table 2: Total Treated Households by Month/Treatment

Treatment Type	July Only	August Only	Both Months	Total
T1: Tips + History + Social Norms (Gallons)	1,419	1,416	1,396	4,231
T2: Tips + History + Social Norms (Percent)	1,420	1,403	1,394	4,217
Total	2,839	2,819	2,790	8,448

in both July and August. Table 2 shows the number of observations by treatment and timing of mailers.

Table 3 shows that each treatment is balanced relative to the control and that treatments are balanced relative to each other. Additionally, Tables A.1-A.5 in the Appendix show that the whole experimental sample is well balanced and that each treatment is balanced relative to each the control and each other within quartiles of pre-treatment consumption. Figure A.6 graphically displays the densities of pre-treatment consumption for the pooled treatment, each individual treatment, and the control group. In addition to achieving balance on average pre-treatment consumption, Figure A.6 in the Appendix shows the treatments are balanced across the full distribution of pre-treatment consumption. The graphical evidence is formalized by nonparametric Kolmogorov-Smirnov tests (Table A.6 in the Appendix) that fail to reject the null of equality of distributions for pre-treatment consumption 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 pre-treatment consumption and prior conservation.

Attrition in our setting is primarily due to utility accounts closing for some reason, such as a household moving (we had only five households out of the full treated sample of 21,151 call in to the utility to opt out of the experiment, which we consider inconsequential). Eighty-three of 4231 (2.0%) households assigned to T1 dropped out of the sample for administrative reasons, while 87 of 4217 (2.1%) households assigned to T2 dropped out. Tests for equalities of proportions fail to reject differences in attrition across treatment groups. Attrition in the control group was slightly higher at 2.8% and we do reject equality of proportions between both treatments and the control. We are not sure why the attrition rate was slightly higher in the control, but we are comforted by the fact that attrition is roughly the same even if there are detectable statistical differences. And in

any case, we do not believe attrition is a validity threat since a household's decision to move homes is very unlikely to be related to the information treatment they received in the experiment.

Table 3: Balance on Observables**(a) Treatment 1 vs. Control**

	Control Mean	Treatment Mean	Difference	(p-value)
2013 Water	23.56	23.49	0.08	0.73
2015 Water	16.90	16.95	-0.05	0.76
Summer Water	21.63	21.66	-0.03	0.87
Winter Water	8.15	8.27	-0.13	0.32
Year Built	1,987.61	1,987.96	-0.35	0.26
Appraised Value	214.85	216.22	-1.37	0.64
Bedrooms	3.37	3.37	0.00	0.76
Lot Acre	0.27	0.27	0.00	0.60
Yard Acre	0.22	0.22	0.00	0.51
Built Sq. Ft.	1,985.10	2,003.09	-17.99	0.16
Bathrooms	2.19	2.21	-0.02	0.12

(b) Treatment 2 vs. Control

	Control Mean	Treatment Mean	Difference	(p-value)
2013 Water	23.56	23.53	0.03	0.88
2015 Water	16.90	16.98	-0.08	0.63
Summer Water	21.63	21.70	-0.07	0.73
Winter Water	8.15	8.15	-0.00	0.98
Year Built	1,987.61	1,987.58	0.04	0.91
Appraised Value	214.85	218.05	-3.19	0.28
Bedrooms	3.37	3.37	0.00	0.76
Lot Acre	0.27	0.27	-0.00	0.47
Yard Acre	0.22	0.23	-0.00	0.53
Built Sq. Ft.	1,985.10	2,002.28	-17.18	0.18
Bathrooms	2.19	2.19	0.00	0.92

(c) Treatment 1 vs. Treatment 2

	Control Mean	Treatment Mean	Difference	(p-value)
2013 Water	23.53	23.49	0.04	0.88
2015 Water	16.98	16.95	0.03	0.89
Summer Water	21.70	21.66	0.04	0.88
Winter Water	8.15	8.27	-0.12	0.34
Year Built	1,987.58	1,987.96	-0.38	0.34
Appraised Value	218.05	216.22	1.82	0.63
Bedrooms	3.37	3.37	0.00	1.00
Lot Acre	0.27	0.27	0.00	0.33
Yard Acre	0.23	0.22	0.00	0.31
Built Sq. Ft.	2,002.28	2,003.09	-0.82	0.96
Bathrooms	2.19	2.21	-0.02	0.20

Note: There are 21,552 Control observations and 8,448 treated observations. Water is measured in thousands of gallons (kgal), Year Built indicates the year of construction, the Appraised Value is in thousands of dollars, Lot Acre and Yard Acre are measures in acres, Built Sq. Ft. is the square footage of the home, and Bedrooms and Bathrooms are the counts. The p-value is based on two-sided t-test for equality of means.

3 Methodology

The primary variable of interest is monthly household water consumption, 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 billing periods of different lengths. Our regression analysis uses “normalized GPD” as the main dependent variable, in which each customer’s GPD is divided by the average control group consumption across the experimental period (July-September 2015) following Allcott (2011). This allows the regression coefficients to be interpreted as the average percent change in consumption, while preserving the treatment effect of very high water users, which the logarithmic transformation of consumption would dampen. Our specification is:

$$y_{it} = \alpha + \gamma_l T_{i,l} + \beta \mathbf{x}_{it} + \epsilon_{it} \quad (1)$$

where y_{it} is normalized GPD, $T_{i,l}$ is an indicator variable for the pooled treatment and each of the two treatment letters ($l = Pooled, 1, 2$), and x_{it} is a vector of control variables. We restrict our sample to the post-intervention period (consumption in July-September), 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 consumption during irrigation seasons prior to the intervention (pre-treatment water)⁹ billing cycle and month fixed effects, and average daily temperature and days of precipitation during the billing cycle. Average pre-treatment consumption is used interchangeably with “baseline consumption” or “baseline water” throughout the paper. 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.

⁹Average pre-treatment consumption is equal to average water use during May, June, August, and September in 2013 and 2014, and May and June for 2014.

3.1 Identifying the Effect of Difference from the Peer Group and Prior Conservation

Figure 2 illustrates the existing correlation between baseline water use and the distance from the peer group for the traditional SCM in gallons (T1) and the percentage change SCM (T2).¹⁰ For each treatment, we partition households into quartiles of baseline water use (displayed on the x-axis), and within each of those quartiles we further partition households into quartiles of the difference between a household's level of consumption (or conservation rate) and that of its peer group. Since comparisons are based on median consumption (or percent conservation) within the peer group, the first two quartiles (Q1:Much better and Q2:Better) are households who are doing better than their peers, and the upper two quartiles (Q3:Worse and Q4:Much Worse) represent households who are doing worse.

Panel (a) of Figure 2 shows that for the traditional SCM in gallons, most low water users (Q1 on the x-axis) unsurprisingly consume less water than their peer group: roughly 90% of the consumers with the lowest baseline water consumption were informed that they used less than their neighbors (Q1 + Q2 of Difference from Peer group (kgal)).¹¹ Likewise, most high users (Q4 on the x-axis) received a message telling them they used more water than their peers.

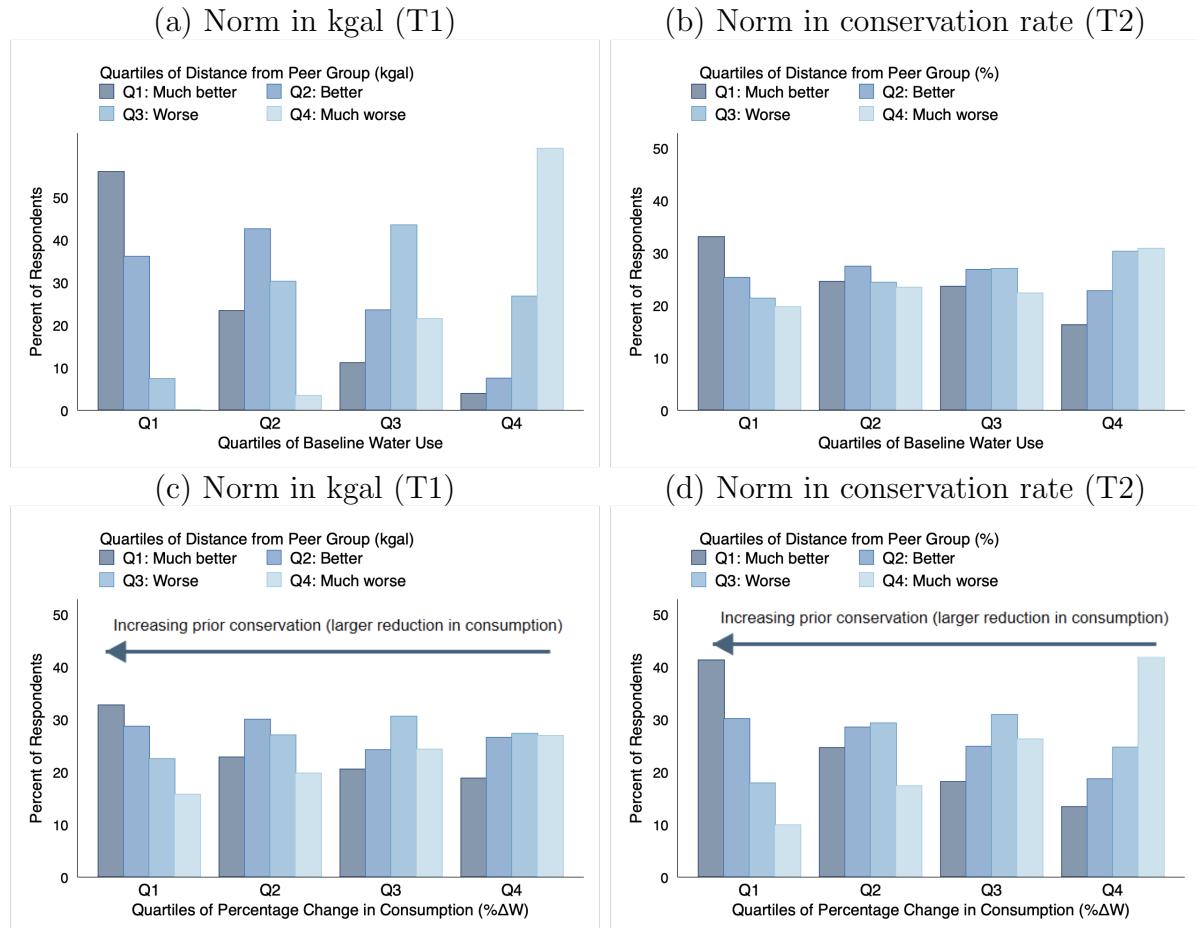
This is not the case for our SCM in percentage terms (T2): a substantial fraction of low users conserved less than their peers and many high users conserved more than their peers (Figure 2, Panel (b)). A substantial proportion of households in the bottom quartile of baseline water consumption received a strong normative appeal. Likewise,

¹⁰Figure 2 uses data for households assigned to the relevant treatment. Panels (a) and (c) use households assigned to T1 and panels (b) and (d) use households assigned to T2. We have also generated the figures with all households in the sample (treatment and control) and the figures look almost identical. Using the full sample shows that patterns in Figure 2 are not sensitive to the particular randomization assignment; the patterns hold more generally.

¹¹The only reason why there are some low users who are above their peer group in the traditional SCM (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 full sample median, producing a treatment where the strength of the descriptive norm is perfectly correlated with baseline consumption. Therefore, a household with a high-water-use peer group can be above the 75th percentile of the sample-wide distribution of baseline consumption, but still consume less than the peer group.

some households with high baseline water use reduced consumption by a larger percentage than their peer group and received a weak normative appeal. The distribution of norms within each quartile of baseline consumption is well balanced for the conservation rate comparison treatment.

Figure 2: Strength of Normative Message by Quartiles of Baseline Consumption & Prior Conservation



Note: The graph displays the percentage of households receiving messages divided up by quartiles of the performance relative to the peer group within each quartile of baseline (pre-treatment) consumption and prior conservation ($\% \Delta W$). The x-axis displays the quartiles of baseline consumption or prior conservation and the y-axis displays the percentage of households receiving a given message. The performances 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 consumption (panel (a) – T1) or conservation rate (panel (b) – T2) and the peer group's consumption.

Panels (c) and (d) in Figure 2 repeat this exercise based on prior conservation instead of baseline consumption. Our primary measure of prior conservation is $\% \Delta W$, which is the percentage change in consumption during the 2014 irrigation season relative to the 2013 irrigation season. Low values of $\% \Delta W$ (negative, and high in absolute value) repre-

sent households who significantly reduced consumption in 2014, and therefore have high prior conservation. Households with higher values of $\% \Delta W$ either reduced consumption by smaller amounts, or increased consumption, in 2014. The cutoffs for the quartiles of $\% \Delta W$ are as follows: Q1 is below -16%, Q2 is between -15% and -4%; Q3 is between -3% and 8%, and Q4 is above 8%. Therefore the change in consumption in 2014 relative to 2013 was negative for all households in Q1 and Q2, positive for some Q3 households and negative for others, and positive for all Q4 households. We will use $\% \Delta W$ and prior conservation interchangeably. It is important to reiterate that lower values of $\% \Delta W$ corresponds to *higher* levels of prior conservation.

The bottom two panels of Figure 2 show that prior conservation is highly correlated with the normative message for the SCM based on percentage reductions (T2), and less correlated with the normative message for the traditional SCM in gallons (T1). In fact, T1 and T2 show opposite relationships when comparing the strength of the normative message to baseline water use and prior conservation. This allows us to evaluate the impact of baseline water use and prior conservation for different SCMs, where the strength of the message is targeted at different types of households.

Importantly, prior conservation and baseline water use are not highly correlated: the correlation coefficient is 0.04, and the treatment groups are well balanced across prior conservation.¹² These two features of our data allow us to examine treatment heterogeneity across both baseline water use and prior conservation.

We analyze heterogeneity based on pre-treatment water use and prior conservation by estimating conditional average treatment effects (CATEs). We focus on CATEs based on subgroups above and below the median baseline water use and prior conservation as well as quartiles of each variable. Since treatment is randomized across the distribution of pre-treatment water use and prior conservation, CATEs provide valid inference - 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;

¹²The p-values of a t-tests for of prior conservation across the pooled treatment, T1, T2, and T1 vs. T2, are 0.39, 0.85, 0.26, and 0.46 respectively. A visual depiction of the correlation between baseline consumption and prior conservation is shown in Figure A.7.

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 treatment 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 consumption for subgroup c . The regressions used to generate Figure 3 define $C_{i,c}$ as the four quartiles of pre-treatment consumption for panel (a) and prior conservation for panel (b).

4 Results

4.1 Base Results

We begin by reporting the average treatment effects pooling the two treatments of interest, and then briefly discuss each treatment individually. Column 1 of Table 4 shows that the average treatment effect (ATE) pooling all three treatments is slightly greater than a 1.5% reduction in consumption.¹³ Overall, our pooled ATE is smaller than commonly reported for SCM: 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 consumption of approximately 5%. However, these results should be considered in the context of an extensive utility-wide water conservation campaign during the second year of a severe drought. Additionally, given that the aforementioned studies on water examine some of the first interventions using SCM for water conservation, the lower treatment effects are consistent with the findings of Allcott (2015) that initial sites often have higher average treatment effects than subsequent sites.

Column (2) breaks down the treatments individually. Each treatment generated statistically significant reductions in consumption and the point estimates are all very close

¹³For reference, in the full field experiment the generic tips treatment had no statistically-significant impact on conservation, and the ATE of the tips plus historical information treatment was slightly less than 1% and statistically significant.

to each other. Columns (3)-(4) reproduce the ATE for each letter in separate regressions using the individual treatment group and the control. This simply demonstrates that both the point estimates and the standard errors are almost identical whether we use the entire sample with three treatment indicator variables or restrict the sample to one treatment and the control. Restricting the sample simplifies the presentation of the results. Column (5) presents an interaction of the pooled treatment effect with the T2 indicator. The interaction term tests for a differential effect between T1 and T2; this term is economically and statistically insignificant, indicating no differences in the ATEs for T1 and T2. All subsequent regressions also include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and average pre-treatment consumption.

Table 4: Base Regression

	(1) Pooled	(2) Both	(3) kgal (T1)	(4) % (T2)	(5) Interaction
SCM pooled	-1.502*** (0.349)				-1.433*** (0.447)
SCM kgal (T1)		-1.433*** (0.447)	-1.428*** (0.447)		
SCM % (T2)		-1.570*** (0.473)		-1.569*** (0.473)	
SCM pooled*T2					-0.137 (0.600)
Weather Controls	Yes	Yes	Yes	Yes	Yes
Bill Cycle FEs	Yes	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes	Yes
Baseline Water	Yes	Yes	Yes	Yes	Yes
Households	29,722	29,722	25,529	25,510	29,722
Observations	85,586	85,586	74,494	74,449	85,586

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

4.2 Treatment Heterogeneity

Next, we show that results for all of our treatments are consistent with prior research which found that treatment effects are concentrated among households with high pre-treatment water use. Panel (a) of Figure 3 displays CATE results by quartiles of baseline consumption for T1 and T2 based on the same model in equation 2, where the conditioning variables are quartiles of baseline consumption. Each graph reflects the results of one regression of the CATEs; the shaded bars are the point estimates and the error bands are

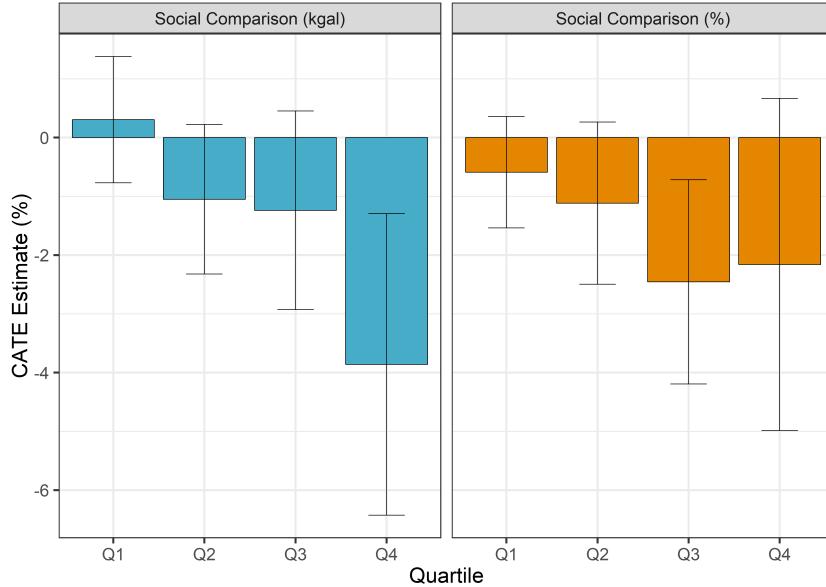
the 95% confidence intervals. There is a positive relationship between pre-treatment consumption and the estimated CATE for both T1 and T2. Households with higher baseline consumption responded more strongly to the traditional SCM treatment in gallons (T1) as well as the treatment in percentage reduction (T2), although the estimated effects are only significantly different from zero for the fourth consumption quartile for T1 and the third consumption quartile for T2. The CATE for the highest quartile of baseline consumption (Q4) is almost twice as large (3.9% vs. 2.2%) for the SCM in levels (T1) as the treatment in percentages (T2). Again, one explanation for this is that Q4 households were much more likely to receive a stronger normative message under T1 than T2 because of the correlation patterns shown in Figure 2. Over 65% of Q4 households received a T1 message stating that they were performing much worse than their peer group, compared to only 30% of the Q4 customers who received T2 (see Figure 2).

Panel (b) of Figure 3 displays CATE results by quartiles of prior conservation ($\% \Delta W$) for each treatment based on the equation 2. Recall that Q1 represents the highest level of prior conservation and Q4 the lowest. Prior conservation does not appear to drive any treatment heterogeneity in the SCM in gallons; all CATEs are similar in magnitude and not significantly different from zero. Conversely, the SCM in percentage terms appears to have CATEs that are monotonically increasing with lower levels of prior conservation - similar to the pattern of CATEs for the SCM in gallons based on pre-treatment consumption. Since prior conservation is highly correlated with type of messages that households receive in T2, the pattern of heterogeneity appears linked to the normative message in the SCM.

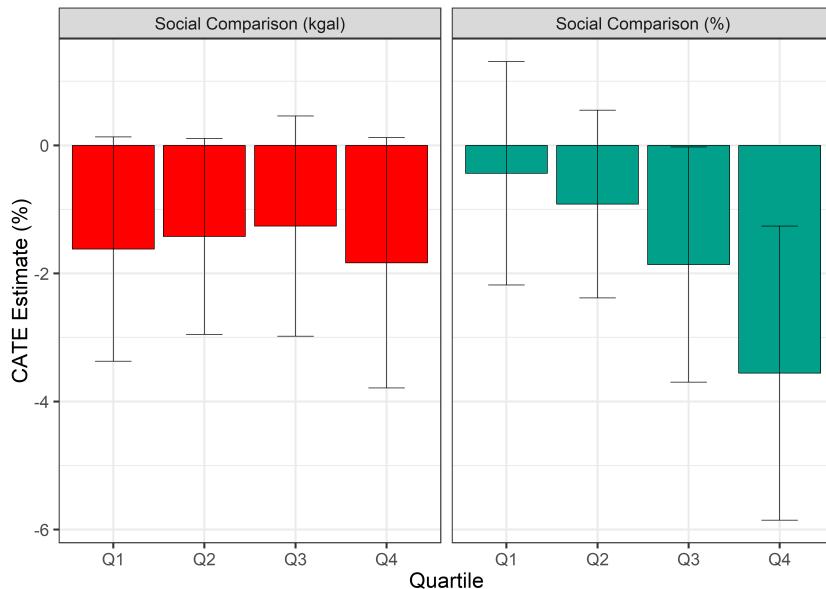
We investigate the impact of both pre-treatment consumption and prior conservation as drivers of heterogeneity in two ways. First, we separately estimate treatment effects for four sub-samples defined according to whether they were above or below the median of pre-treatment consumption and the percentage change in consumption ($\% \Delta W$). Recall that households who substantially decreased consumption during the 2014 drought have low $\% \Delta W$, which is synonymous with high prior conservation. The households in the four sub-samples are specified as follows:

Figure 3: Conditional Average Treatment Effects by Quartiles of Pre-treatment Consumption & Prior Conservation

(a) Pre-treatment consumption



(b) Prior Conservation



Note: Each of the four graphs represents the output of one regression where the dependent variable is normalized average daily water consumption. The bars are the point estimates of the CATEs for each quartile of pre-treatment consumption in panel (a) and prior conservation in panel (b). 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 pre-treatment consumption.

1. **Low-Low:** below median pre-treatment water use and below median $\% \Delta W$ (low water use & high prior conservation),
2. **Low-High:** below median pre-treatment water use and above median $\% \Delta W$ (low water use & low prior conservation),
3. **High-Low:** above median pre-treatment water use and below median $\% \Delta W$ (high water use & high prior conservation),
4. **High-High:** above median pre-treatment water use and above median $\% \Delta W$ (high water use & low prior conservation).

These categories help delineate households' opportunity costs of conservation, as well as which treatment is likely to send a strong normative message. We expect households with low baseline water use and high prior conservation (Low-Low) to have higher opportunity costs of further conservation than households with high baseline water use and low prior conservation (High-High). Therefore, we expect this latter group to have higher estimated CATEs. For the message households with high water use will likely receive a strong T1 message and households with high $\% \Delta W$ will receive a strong T2 message.

Second, we estimate a model that interacts each treatment with indicators for above-median pre-treatment water use and above-median $\% \Delta W$. This formalizes whether the differences observed in the sub-samples are statistically different.

To assist in the interpretation of Table 5, the coefficients for subgroups that are expected to receive a strong normative message are in bold. Column (2) of Table 5 shows that neither treatment is effective among the subgroup of households that have low use and low $\% \Delta W$ (high prior conservation); both estimated coefficients are small and not significantly different from zero. (Column (1) shows the base ATEs presented in Column (2) of Table 4 for reference.) Among the subgroup of households with low baseline consumption and low prior conservation (high $\% \Delta W$), the norm in percentage terms (T2) is statistically significant and three times as effective as the traditional SCM, which is small and not significantly different from zero (column 3). Among high baseline water users with high prior conservation (low $\% \Delta W$) the results are reversed (column 4). The traditional SCM (T1) is statistically significant and roughly four times as effective on these households as the SCM in percentage terms, which is not significantly different from zero. Column 5 shows that both treatments are statistically significant and highly

effective among the subgroup of households that are both large baseline waters users and who had low prior conservation (high $\% \Delta W$). Finally, the interaction model in column (6) formalizes these patterns of heterogeneity. For the SCM in gallons, only high pre-treatment water use is a significant driver of treatment heterogeneity. Conversely, the effectiveness of the SCM in percentage terms is driven primarily by high $\% \Delta W$ (low prior conservation). For the SCM in percentage terms high pre-treatment water use has an estimated coefficient that is similar magnitude, although not statistically significant.

Table 5: Heterogeneity by Baseline Water Use and Prior Conservation

	(1) Base	(2) Low-Low	(3) Low-High	(4) High-Low	(5) High-High	(6) Int
SCM kgal (T1)	-1.433*** (0.447)	-0.288 (0.558)	-0.534 (0.664)	-2.919*** (1.077)	-2.290** (1.081)	-0.182 (0.576)
SCM % (T2)	-1.570*** (0.473)	-0.734 (0.534)	-1.674** (0.684)	-0.690 (1.034)	-3.453*** (1.269)	0.0610 (0.590)
T1*High Water						-2.687*** (0.878)
T1*High $\% \Delta W$						0.145 (0.892)
T2*High Water						-1.401 (0.919)
T2*High $\% \Delta W$						-1.994** (0.945)
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bill Cycle FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Water	Yes	Yes	Yes	Yes	Yes	Yes
Households	29,722	7,830	6,584	7,686	7,622	29,722
Observations	85,586	22,589	18,915	22,152	21,930	85,586

Note: The dependent variable is normalized average daily water consumption. Column (1) replicates the base specification presented in column (2) of Table 4. Columns (2) - (5) repeat the base specification on subsets of the sample partitioned by both above/below median pre-treatment consumption and above/below median $\% \Delta W$ and contain the following subgroups: Column (2) Low-Low: households with low water use and low $\% \Delta W$ (high prior conservation), Column (3) Low-High: households with low water use and $\% \Delta W$ (low prior conservation), Column (4) High-Low: households with high water use and low $\% \Delta W$ (high prior conservation), Column (5) High-High: households with high water use and high $\% \Delta W$ (low prior conservation). $\% \Delta W$ is defined as the percentage change in consumption from 2013 to 2014. Column (6) presents an interaction model on the full sample where High Water is above the median of pre-treatment consumption and High $\% \Delta W$ is above the median of $\% \Delta W$. All regressions include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and pre-treatment consumption. Robust standard errors clustered at the household level are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

One explanation for these results is that the primary source of treatment heterogeneity is the strength of the normative message. In T1, households with high pre-treatment consumption were more likely to receive messages that they were consuming more than their peers (see Figure 2 Panel (a)) , and these households responded with larger decreases in

consumption. Conversely, for T2, households with high $\%ΔW$ (low conservation reduction in the prior drought) were more likely to receive messages that they were conserving less than their peers in the current drought (see Figure 2 Panel (d)). These households responded to the stronger normative message with larger decreases in consumption.

Collectively this evidence confirms that the content of the normative message in SCMs is an important source of heterogeneity in consumer response. One important policy implication is that water service providers could target households with customized SCMs based on both baseline consumption and prior conservation, both of which are readily observable in billing data. While targeting has been brought up in the literature previously (Ferraro and Miranda, 2013), this has primarily been viewed as a way to make such programs more cost-effective by only sending SCMs to high-use households. We investigate targeting in more detail in Section 4.4.

4.3 Robustness

We present several robustness tests to assess the sensitivity of our results to our definition of prior conservation in Table 6. All regressions take the form of the interaction model presented in column (6) of Table 5. Recall that our primary measure of prior conservation is based on the percentage change in consumption from 2013 to 2014, which we replicate in column (1) of Table 6 for reference. In column (2) we define $\%ΔW$ as the change in gallons from 2013 to 2014. This is more appropriate if the absolute change in a household's consumption is more relevant than the percentage change. Next, in column (3) we define high $\%ΔW$ as an indicator equal to one if the household did not conserve at all or increased their consumption in 2014 . For reference, the mean percentage change in the 2014 irrigation season was -2.2% and 70% of households reduced their consumption. Therefore, 30% of the sample have high $\%ΔW$ in this definition. Lastly, in column (4) we define high $\%ΔW$ as a dummy that equals one for the 45% of households who did not meet the 10% goal in 2014. The results in Table 6 are quantitatively and qualitatively similar. The interaction terms are stable for both treatments and the interaction of T2 with high $\%ΔW$ is similar across specifications. The 10% goal metric presented in column (4) is not statistically significant, but it is of similar magnitude. Overall the results are

robust to various definitions of $\% \Delta W$.

Table 6: Robustness to Prior Conservation Definition

	(1)	(2)	(3)	(4)
	Base (%)	Gallons	Conserve 2014	10% 2014
SCM kgal (T1)	-0.182 (0.576)	-0.281 (0.609)	-0.134 (0.546)	-0.145 (0.675)
SCM % (T2)	0.0610 (0.590)	0.128 (0.627)	0.0295 (0.559)	0.0223 (0.682)
T1*High Water	-2.687*** (0.878)	-2.637*** (0.876)	-2.568*** (0.878)	-2.577*** (0.881)
T1*High $\% \Delta W$	0.145 (0.892)	0.281 (0.888)	-0.0580 (0.917)	0.0118 (0.911)
T2*High Water	-1.401 (0.919)	-1.536* (0.927)	-1.402 (0.920)	-1.368 (0.921)
T2*High $\% \Delta W$	-1.994** (0.945)	-1.936** (0.946)	-2.214** (0.993)	-1.478 (0.935)
Weather Controls	Yes	Yes	Yes	Yes
Bill Cycle FEs	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes
Baseline Water	Yes	Yes	Yes	Yes
Households	29,722	29,722	29,722	29,722
Observations	85,586	85,586	85,586	85,586

Note: The dependent variable is normalized average daily water consumption. The regressions take the same form as column (6) of Table 5, which is replicated in column (1). High $\% \Delta W$ is defined as above the sample median of the change in gallons from 2013 to 2014 in column (2), an indicator for households who increased consumption in 2014 relative to 2013 in column (3), and an indicator for households who conserved less than 10% in 2014. All regressions include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and pre-treatment consumption. Robust standard errors clustered at the household level are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

In addition to the normative message of performance relative to a peer group both treatments also include an injunctive norm based on whether the household's previous was 10% lower to the same month in 2013. This discrete norm may also have an impact on the treatment effects and could be correlated with pre-treatment consumption and/or prior conservation. We 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, nor do we find any effect of moving slightly above the 10% goal. This is consistent with the findings of Allcott (2011) for SCMs in energy. The Appendix describes the regression discontinuity design in more detail and presents both graphical evidence and regression discontinuity estimates based on Calonico et al. (2015).

Since we also randomized the timing and number of letters that a household receives, we also analyze differences between T1 and T2 in these experimental design features (Ta-

ble 2). The results are shown in Table 7. The treatment effects from social comparisons wane significantly over time with a roughly 50% decline in conservation each month from when the first letter was received. Sending a second mailer has a strong and significant effect on conservation. Column (4) tests for any differences in treatment effects between the two SCMs by interacting the pooled variables with a T2 indicator. All the interaction terms are close to zero and not statistically significant. While timing and the number of mailers matter, they appear to affect conservation similarly for both social comparisons.

Table 7: Number of Mailers and Treatment timing

	(1) Both	(2) T1	(3) T2	(4) Interaction
1st Letter, 1st Month	-1.474*** (0.403)	-1.434*** (0.517)	-1.528*** (0.552)	-1.417*** (0.517)
1st Letter, 2nd Month	-1.108** (0.449)	-1.093* (0.591)	-1.098* (0.623)	-1.111* (0.591)
1st Letter, 3rd Month	-0.304 (0.554)	-0.168 (0.717)	-0.479 (0.765)	-0.171 (0.716)
2nd Letter	-1.922*** (0.558)	-1.817** (0.741)	-1.963** (0.825)	-1.848** (0.740)
1st Letter, 1st Month*T2				-0.113 (0.704)
1st Letter, 2nd Month*T2				0.00619 (0.817)
1st Letter, 3rd Month*T2				-0.268 (0.981)
2nd Letter*T2				-0.148 (1.099)
Households	29,722	25,529	25,510	29,722
Observations	85,586	74,494	74,449	85,586

Note: The dependent variable is normalized average daily water consumption. All regressions include controls for temperature, precipitation, bill cycle fixed effects, month fixed effects, and pre-treatment consumption. Robust standard errors clustered at the household level are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

4.4 Targeting

In this section we explore how our results could allow utilities to better target social comparisons to specific households. Our exercise differs from others (Ferraro and Miranda, 2013; Alcott and Kessler, 2019) since we consider targeting of different treatments to different households as opposed to the optimal households to receive a single treatment. We run the interaction regression reported in column (6) Table 5 on total consumption rather than normalized gallons per day. This better places our results in the context of the TMWA’s aggregate drought policies, where the total gallons saved

is critical. The treatment as implemented saved roughly 0.31 kgal per person, and in aggregate the nudge saved over 7145 kgals. If TMWA would have optimally targeted the same sample by sending the traditional social comparison (T1) to households who had high prior conservation and sending the new social comparison (T2) to households with low prior conservation, the aggregate savings would have been 9865 kgals, 38% higher.¹⁴ Statistical tests show that the differences in the ATEs between the experiment as implemented and a targeted version are statistically significant at the 10% level.¹⁵ A key cause for the increase in effectiveness is the ability for the social comparison in percentage terms to generate savings even among households with below-median baseline water use but who had achieved relatively little conservation, a relatively large group.

An attractive feature of our targeting approach is that it is relatively simple to implement with only water billing records. We present these steps to inform water managers or researchers who want to implement a targeted SCM campaign.

1. Calculate prior conservation for all households. This is simply the change in consumption relative to a base period.
2. Calculate the median of the prior conservation.
3. Divide the sample into two groups:
 - Group 1: Above median of prior conservation.
 - Group 2: Below median of prior conservation.
4. Assign the SCM in following way:
 - Group 1 receives the SCM comparing gallons to a peer group (our T1).
 - Group 2 receives the SCM comparing the change in consumption (our T2).

When considering the external validity of the targeting results, we should acknowledge that the targeting gains depend on the correlation between average pre-treatment consumption and prior conservation. The benefits of targeting that we find are likely due to having more households receive strong normative messages. Our targeting approach

¹⁴The hypothetical targeted ATE is calculated by taking the weighted average of linear combinations of the treatment effect and relevant interaction for T1 and T2 presented in column (6) of Table 5 weighted by the proportion of the population in each subgroup.

¹⁵The formal tests rely on a Wald test for equality of the weighted average of the coefficients in column (6) of Table 5 using the weights as implemented versus the targeted weights. The difference in targeted vs. actual ATEs is 120 gallons and the p-value for the Wald test for the difference being zero is 0.059.

will increase aggregate savings when there are two distinct groups: (1) households with high baseline consumption and low prior conservation (who receive the SCM in gallons), and (2) households with low baseline consumption but high prior conservation (who receive the SCM in percentage terms). While these two distinct groups had a significant number of households in our sample, this may not be true in other settings.¹⁶

Another consideration for the external validity of targeting is how the heterogeneous effects of SCMs interact with alternative water conservation policies. The recent literature finds that high-use households are less responsive to prices (Baerenklau et al., 2014; Klaiber et al., 2014). So, while price increases may cause a significant decrease in consumption for low-use households, it may not be desirable due to concerns over affordability. There is less heterogeneity in responses to voluntary and mandatory watering restrictions (Wichman et al., 2016).¹⁷ While there may be some interactions between SCM treatments and alternative policies, Brent and Wichman (2020) show little interaction between prices and SCMs. Therefore, sending a variety of messages targeted to specific households presents an attractive option for increasing the efficiency of SCMs.

5 Discussion

The use of nudges in public policy has exploded in recent years in a variety of sectors ranging from healthy eating to paying taxes on time. Nudges, and particularly social comparisons, are now heavily utilized to improve environmental outcomes including energy and water conservation. Previous research has established some consistent features of SCMs in water and energy; they typically generate small aggregate treatment effects that are concentrated among high-use households. These findings have led to policy recommendations to improve the cost-effectiveness of nudges by targeting high-use households (Ferraro and Miranda, 2013). More sophisticated targeting rules also attempt to optimize welfare improvements as opposed to conservation per dollar (Allcott and Kessler, 2019). However, there is still uncertainty about *why* high-use households are more responsive.

¹⁶We are also considering the benefit of targeting an experiment with two SCM treatments, as opposed to the benefit of adding and targeting T2 to a an existing campaign that only uses the traditional SCM in gallons.

¹⁷One exception is that households with irrigation systems – who are more often high users – have a greater response to mandatory restrictions.

With traditional SCMs, high-use households may have lower opportunity costs of conservation and they are also more likely to receive strong normative messages conveying that inappropriate use relative to their peers. Additionally, most of the discussions of targeting have focused on the traditional social comparison as opposed to potentially new information treatments. Without disentangling the mechanisms it is difficult to consider more sophisticated targeting rules that allow nudges to be effective among lower-use households.

We develop a new SCM where the comparison is in terms of percentage reduction in water use as opposed to absolute water use. This SCM allows more low-user households to receive messages that they are performing worse than their peers. This SCM is also related to the opportunity cost of conservation, whereby a household's capacity to respond to new conservation policies depends on their past water conservation actions. Similar to other studies, we find high pre-treatment water use is the primary driver of treatment heterogeneity for the standard SCM in gallons, however prior conservation is the dominant form of heterogeneity for the new SCM in percentage terms. One explanation is that where high pre-treatment water use best explains who receive strong normative message (doing worse than one's peers) for traditional SCMs, prior conservation is a better predictor of the strength of the normative message in the new SCM in percentage terms.

The findings have important policy implications for utilities using nudges to help manage drought. By targeting our two nudges to households based on pre-treatment water use and prior conservation - both of which are easily calculated with billing data - the aggregate treatment effect increases by 38%. This increase is possible because while the traditional SCMs are typically not effective among below-median water users, our new SCM is effective among the subset of this group who have not responded to prior conservation efforts.

One caveat of our research is that we are unable to fully disentangle the mechanisms that explain our results. We believe the most compelling explanation is that consumers are responding to the strength of the normative message. This explains why prior conservation affects treatment heterogeneity in the SCM in percentage terms but not the

traditional SCM. However, the possibility still exists that prior conservation itself dictates the effectiveness of conservation policies. Future research should help disentangle these results by developing better metrics for the opportunity cost of conservation using high frequency metering data (Mayer, 2016).

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 **Cass R Sunstein**, “Regulating internalities,” *Journal of Policy Analysis and Management*, 2015, 34 (3), 698–705.
- 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,” *American Economic Journal: Applied Economics*, 2019, 11 (1), 236–76.
- 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.
- 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 2013, 29 (5), 992–1022.
- Baerenklau, Kenneth A, Kurt A Schwabe, and Ariel Dinar**, “The residential water demand effect of increasing block rate water budgets,” *Land Economics*, 2014, 90 (4), 683–699.
- Becker, Gary S.**, “A Theory of the Allocation of Time,” *Economic Journal*, 1965, pp. 493–517.
- Bhanot, Syon P**, “Isolating the effect of injunctive norms on conservation behavior: New evidence from a field experiment in California,” *Organizational Behavior and Human Decision Processes*, 2018.
- Brent, Daniel A.**, “Estimating Water Demand Elasticity at the Intensive and Extensive Margin,” Technical Report, Department of Economics, Louisiana State University 2018.
- and **Casey J. Wichman**, “Do behavioral nudges interact with prevailing economic incentives? Pairing experimental and quasi-experimental evidence from water consumption,” 2020.
- and **Michael B. Ward**, “Energy Efficiency and Financial Literacy,” *Journal of Environmental Economics and Management*, Forthcoming.

- , **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.
- Brick, Kerri, Samantha De Martino, and Martine Visser**, “Behavioural nudges for water conservation: Experimental evidence from Cape Town,” 2017.
- Calonico, Sebastian, Matias D. Cattaneo, 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.
- 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.
- Goette, Lorenz, Ching Leong, and Neng Qian**, “Motivating household water conservation: A field experiment in Singapore,” *PloS one*, 2019, 14 (3).
- 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.
- Klaiber, H Allen, V Kerry Smith, Michael Kaminsky, and Aaron Strong**, “Measuring price elasticities for residential water demand with limited information,” *Land Economics*, 2014, 90 (1), 100–113.
- Lee, David S. and Thomas Lemieux**, “Regression Discontinuity Designs in Economics,” *Journal of Economic Literature*, 2010, 48 (June), 281–355.

Levitt, Steven D. and John A. List, “What do laboratory experiments measuring social preferences reveal about the real world?,” *The Journal of Economic Perspectives*, 2007, 21 (2), 153–174.

Lott, Corey, “The Effect of Financial and Social Incentives on Water Conservation.” PhD dissertation, University of California, Santa Barbara 2017.

Mayer, Peter W, “Water Research Foundation Study Documents Water Conservation Potential and More Efficiency in Households,” *Journal-American Water Works Association*, 2016, 108 (10), 31–40.

Nemati, Mehdi, Steven Buck, and Hilary Soldati, “Heterogeneous Effects of Real-time Consumption Analytics on Residential Water Consumption,” Available at SSRN 2877233, 2019.

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*, 2015, pp. 1–16.

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.

Sexton, Steven, “Automatic bill payment and salience effects: Evidence from electricity consumption,” *Review of Economics and Statistics*, 2015, 97 (2), 229–241.

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.

Taylor, Michael H., Kimberly Rollins, and Corey Lott, “Exploring the behavioral and welfare implications of social-comparison messages in residential water and electricity,” *Economics Letters*, 2018, 168, 65–69.

Torres, Mónica M Jaime and Fredrik Carlsson, “Direct and spillover effects of a social information campaign on residential water-savings,” *Journal of Environmental Economics and Management*, 2018, 92, 222–243.

West, Jeremy, Robert W Fairlie, Bryan Pratt, and Liam Rose, “Do Regulatory Policies Crowd Out Social Pressure for Resource Conservation,” Technical Report 2019.

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

— , “Information Provision and Consumer Behavior: A Natural Experiment in Billing Frequency,” *Resources for the Future Discussion Paper*, 2016, DP 15-35-REV.

— , **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

Randomization and Implementation

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 a treatment is balanced within certain groups of interest. Blocks were defined by billing cycles, rate schedule and frequency of recorded meter data (i.e. monthly, daily, or hourly, though all customers only receive monthly usage totals). Within each block we ordered all observations on average water consumption in summer 2013 in sets of five households. We randomly assigned each household to one of five experimental samples that correspond to the five treatments (regardless of the ultimate assignment to treatment group vs. control group). This ensures a similar distribution of 2013 water consumption within each of the five experimental samples.

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.¹⁸

The process for generating and mailing letters was as follows:

- 1-2 days after the most recent month's consumption 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.
- PDF's of the letters are emailed to Digiprint
- Digiprint prints and ships 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 of the study in July and 211 customers dropped 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.

¹⁸Due 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 after the conclusion of the field experiment (3,677 households: 2,025 control, 1,652 treatment). All balance tables and regression results reflect the corrected sample.

Treatment Figures

Figure A.1: Treatment A1 - Tips



Quality. Delivered.

July 1, 2015

1355 Capital Blvd. • P.O. Box 30013 • Reno, NV 89520-3013
P 775.834.8080 • F 775.834.8003



Dear [REDACTED],

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. 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 their outdoor watering. Below are six tips that we hope will help you to meet or exceed your 10% conservation goal.

Six Low-Cost Tips to Reduce the Amount of Water You Use Outdoors

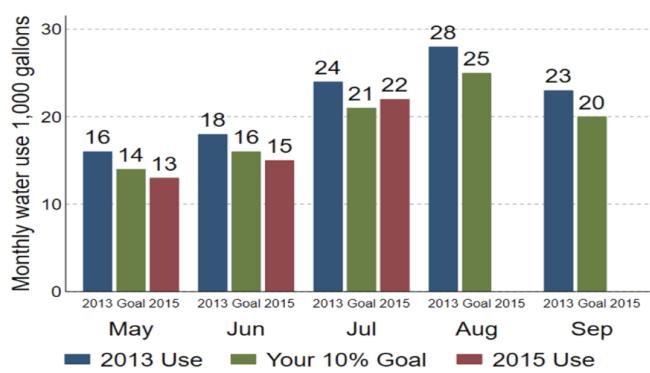
Most households in the region could reduce outdoor watering by using these six low-cost tips. These are ordered with from highest to lowest potential water savings.

- 1. Check your sprinklers, spigots, and hoses regularly for leaks.** Finding and fixing broken sprinkler heads is an easy way to save water and maintain the health of your landscape. Install **low-flow hose nozzles** to save even more water.
- 2. Water when it's cool, between 3 am and 8 am.** Avoid watering between 11am and 7pm, when much water evaporates before your plants can use it. Check your irrigation timers – you can significantly reduce your outdoor water use by watering between 3 am and 8 am. Turn sprinklers off when it's windy or raining.
- 3. Set your lawn mower blades to 3 inches or higher.** Mowing high encourages grass roots to grow deeper and provides shade for the soil, helping your lawn hold moisture better than a closely mown lawn.
- 4. Don't apply fertilizer in the heat of the summer.** Grass naturally slows its growth in the heat of the summer to conserve water and nutrients. Fertilizing in the summer encourages grass to grow, which increases your lawn's water needs. Fertilize your lawn in the spring before it gets hot and in the fall after temperatures cool.
- 5. Mulch helps soil retain water.** Applying shredded wood chips, compost, leaves, and other organic mulch to the base of your plants and trees protects the root zone from water loss. Applying mulch is an easy way to reduce the amount of water you use on your non-turf landscaping.
- 6. Water and wait to avoid overwatering.** When setting up your lawn sprinklers, adjust them so they shut off before runoff occurs or puddles form. Then, wait an hour for the water to soak down into the roots. Repeat this process until the soil is moistened to a depth of 6-8 inches. Set your irrigation timers for all zones accordingly.

For more information on how to reduce your outdoor water use, go to www.tmwa.com/save.

*Truckee Meadows Water Authority is a not-for-profit, community-owned water utility,
overseen by elected officials and citizen appointees from Reno, Sparks and Washoe County.*

Figure A.2: Treatment A2 - Historical water use information



What is your 10% goal? 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 **8% 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 **100 gallons per day** compared to how you used water in August 2013.

Figure A.3: Treatment A3 - Rate structure information

Saving water saves you money. The graph on the right shows how you **could save \$9.75** 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).

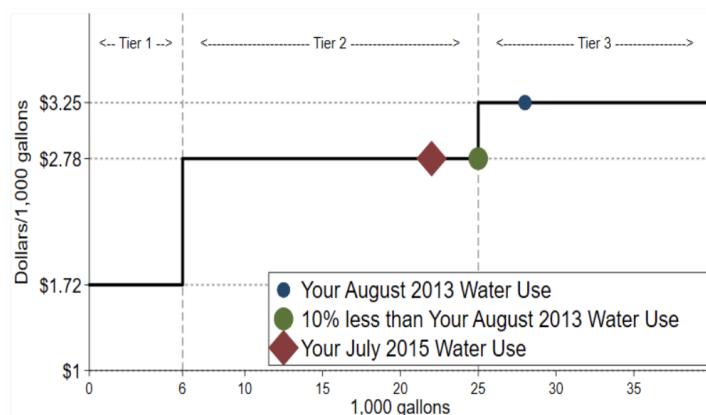


Figure A.4: Treatment 1 - Social comparison, reported in thousands of gallons

How does your water use compare?

The graph on the right shows your water use from your July bill compared to similar properties in your area. You used **2,000 gallons less** than your neighbors with similar properties.

You saved 8% on your July bill compared to 2013.

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 the drought.

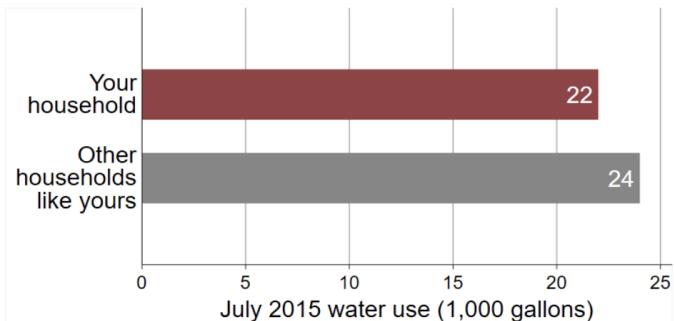
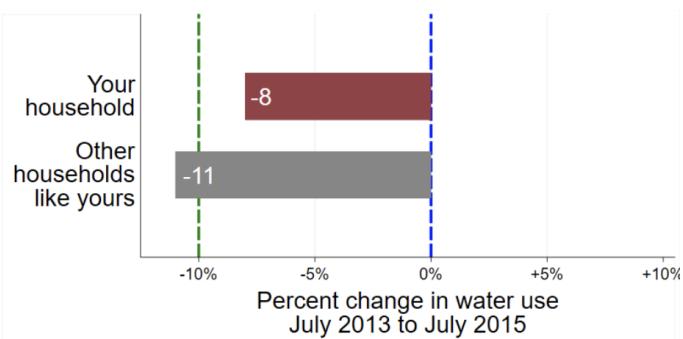


Figure A.5: Treatment 2 - Social comparison, reported as progress towards TMWA's 10% conservation goal

Are you doing your part? The graph on the right shows your change in water use from July 2013 to July 2015 compared to similar properties in your neighborhood. Your neighbors used 11% less water last month compared to 2013.

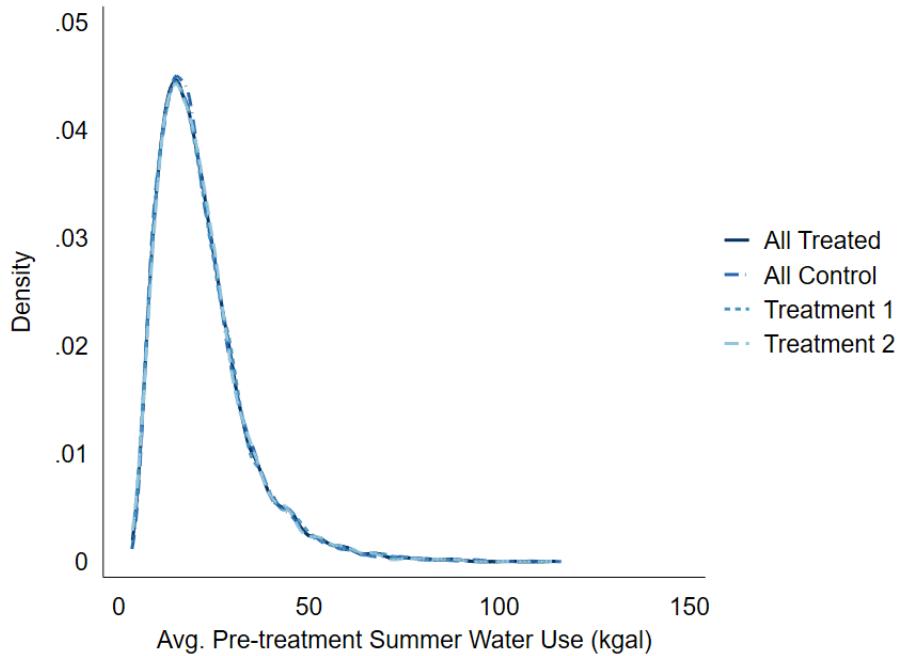
You saved 8% on your July bill compared to 2013.

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 the drought.



Additional Balance Tests

Figure A.6: Distributions of Pre-treatment Consumption Across Treatment Status



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

Table A.1: 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
Build 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

Table A.2: Balance on Observeables for Pooled Treatments (T1 and T2) by Quartiles

	Control Mean	Treatment Mean	Difference	(p-value)
Quartile 1	10.17	10.20	-0.03	0.61
Quartile 2	16.28	16.25	0.03	0.49
Quartile 3	22.56	22.57	-0.02	0.74
Quartile 4	37.70	37.72	-0.01	0.96

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

Table A.3: Balance on Observeables for T1 by Quartiles

	Control Mean	Treatment Mean	Difference	(p-value)
Quartile 1	10.17	10.19	-0.02	0.74
Quartile 2	16.28	16.27	0.01	0.86
Quartile 3	22.56	22.57	-0.01	0.84
Quartile 4	37.70	37.64	0.06	0.87

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

Table A.4: Balance on Observeables for T2 by Quartiles

	Control Mean	Treatment Mean	Difference	(p-value)
Quartile 1	10.17	10.20	-0.03	0.65
Quartile 2	16.28	16.23	0.05	0.38
Quartile 3	22.56	22.58	-0.02	0.76
Quartile 4	37.70	37.80	-0.09	0.81

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

Table A.5: Balance on Observeables for T1 vs. T2 by Quartiles

	Control Mean	Treatment Mean	Difference	(p-value)
Quartile 1	10.20	10.19	0.01	0.92
Quartile 2	16.23	16.27	-0.04	0.59
Quartile 3	22.58	22.57	0.01	0.94
Quartile 4	37.80	37.64	0.16	0.75

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

Table A.6: Kolmogorov-Smirnov Tests

	D-statistic	(p-value)
All Treatment vs Control	0.01	0.88
Treatment 1 vs Control	0.01	0.99
Treatment 2 vs Control	0.01	0.76
Treatment 1 vs Treatment 2	0.01	0.99

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

Discrete Effects Moving Above the Peer Group and 10% Goal

It is important to distinguish the continuous difference from 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 ?? 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).¹⁹ The running variable is the difference in performance between a household and the peer group in gallons for Treatment 1 and in percentage reduction for Treatment 2. The dependent variable in both is residual normalized consumption based on a regression of normalized consumption on weather, month fixed effects, and household fixed effects, following the approach of (Allcott, 2011). The RD estimation assumes that factors varying with the difference from the peer group, such as pre-treatment 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 consumption near the peer group (Figure A.8), as is standard in RD approaches.²⁰ For both SCM the graphical evidence in Figure A.8 suggests that moving above the peer group does not affect consumption.

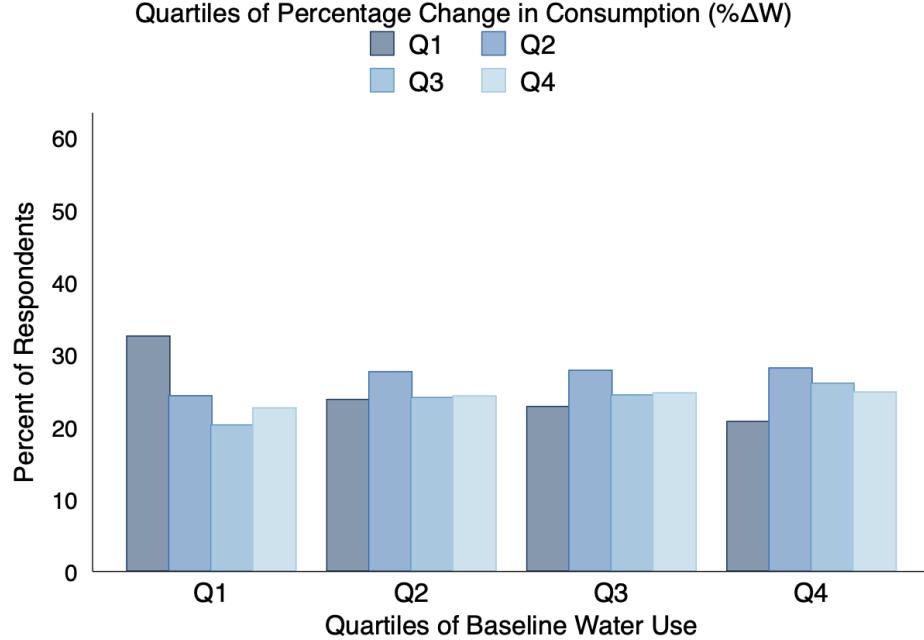
The graphical evidence is corroborated in the RD estimates for moving above the peer group. We use three different RD estimators developed by Calonico et al. (2014): the conventional, bias-corrected, and bias corrected with robust standard errors. In all specifications the impact for moving above the peer group is small and not statistically significant (Table A.7). The RD estimates show that the effects in Table 5 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

¹⁹Allcott (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 peer consumption, "Good" if they consume below the average of peer consumption, or "Below Average" if they consume above the average of peer consumption.

²⁰The 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).

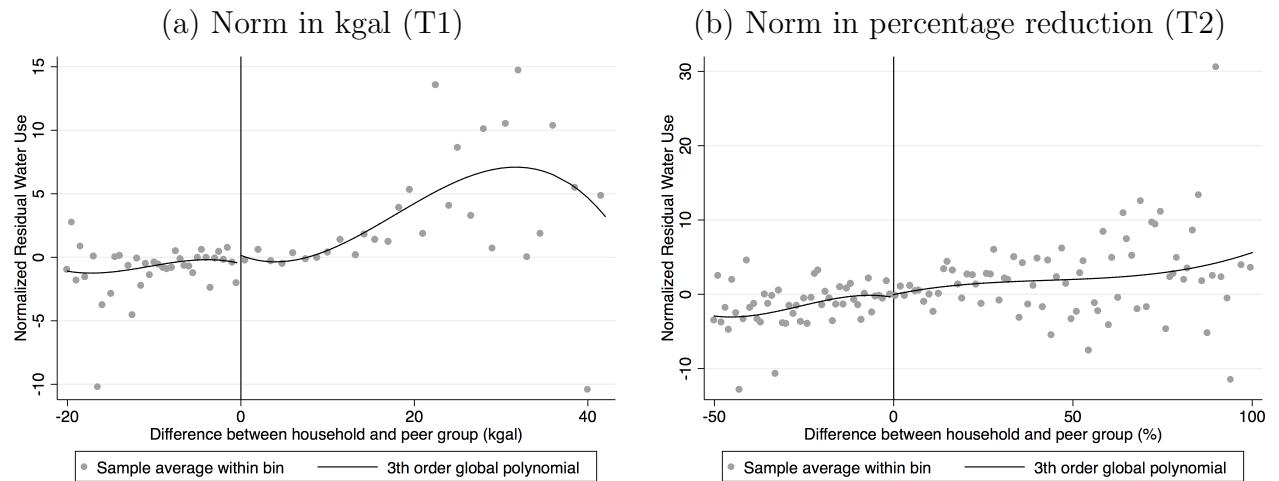
adding a negative injunctive norm.

Figure A.7: Correlation of Baseline Water Use and Prior Conservation



Note: The graph displays the percentage of households in each quartile of prior conservation within quartiles of baseline water use. Prior conservation is defined as the percentage change in water from 2013 to 2014.

Figure A.8: Effect of Moving Above the Peer Group



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.

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.9 graphs residual consumption on the y-axis with the year-on-year percentage change in water consumption (minus 10%) on the x-axis. There is no visual evidence of a change in

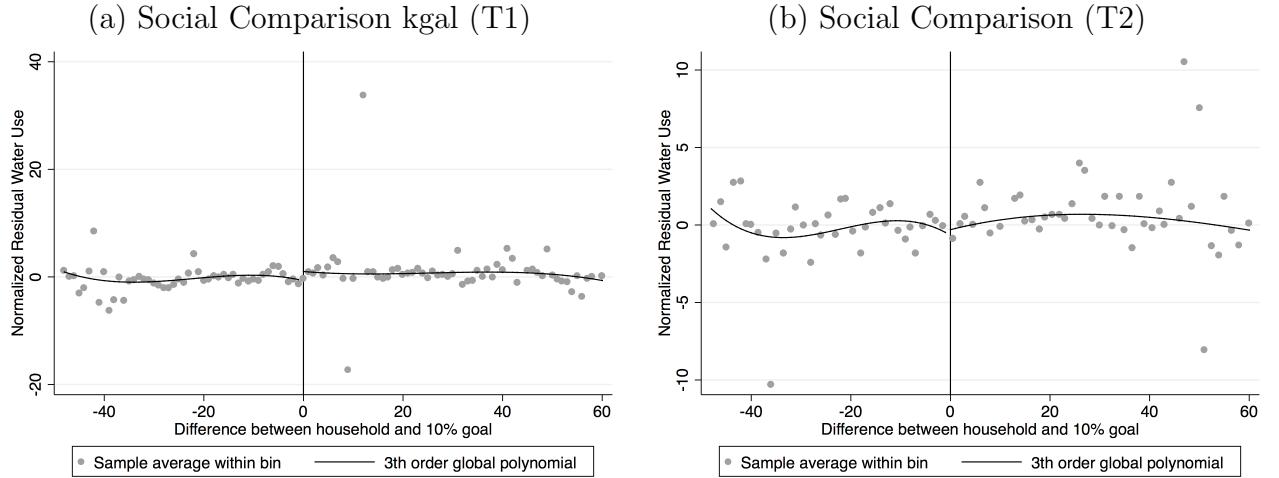
Table A.7: Regression Discontinuity Estimates of the Moving Above the Peer Group

	(1) Kgal (T1)	(2) % (T2)
Conventional	1.331 (0.907)	-0.102 (1.057)
Bias-corrected	1.380 (0.907)	-0.162 (1.057)
Robust	1.380 (1.042)	-0.162 (1.248)
Observations	5,604	5,576

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 (T1), and in column (2) is based on the moving above the peer group's year-on-year change in consumption (T2).

consumption right at the threshold. This is corroborated with the RD estimates (Table A.8) for each of the treatments.

Figure A.9: Effect of the Failing to Meet the 10% Goal



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.8: Regression Discontinuity Estimates of the Failing to Meet the 10% Goal

	(1) kgal (T1)	(2) % (T2)
Conventional	1.009 (1.320)	-0.099 (1.218)
Bias-corrected	1.216 (1.320)	-0.093 (1.218)
Robust	1.216 (1.530)	-0.093 (1.433)
Observations	5,575	5,569

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.