Residential Water Conservation During Drought: Experimental Evidence from Three Behavioral Interventions[1]

Katrina Jessoe,[2] Gabriel E. Lade,[3] Frank Loge,[4] and Edward Spang[5]

April 2, 2021

Abstract

This paper deploys a framed field experiment and uses high frequency data to evaluate the short- and long-run effects of three behavioral interventions on residential water use during extreme drought. Our study of the effects of Home Water Reports (HWRs) on hourly water use yields three main results. First, even when layered on top of a 25% drought conservation mandate, HWRs led to conservation effects of 4 to 5%. Second, across all three treatments the profile of water conservation is similar, suggesting that households did not respond to the messaging or recommendations contained in the HWRs. Third, the water conservation effect of all interventions dissipated five months after the intervention ended. In our setting, these behavioral interventions aligned with utility incentives to achieve immediate but temporary water conservation in response to drought.

Keywords: Social Norms; Water Use; Long-Run Effects; Randomized Controlled Trial; High-Frequency Data

JEL:

---

[1]We gratefully acknowledge the staff at WaterSmart and Burbank Water and Power for their important roles in developing and implementing this project. We thank Daniel Brent, Casey Wichman, Kevin Novan, participants at the 2017 AERE conference, Andreas Lange, and two anonymous referees for valuable comments. Natalie Popovich and Jack Gregory provided excellent research assistance. Research support for this project was provided by the Hellman Foundation, Southern California Gas, and the E2e Project. All errors are our own.

[2]Department of Agricultural and Resource Economics, University of California, Davis, One Shields Ave, Davis, CA 95616; Phone: (530) 752-6977; Email: `kkjessoe@ucdavis.edu`

[3]Department of Economics, Macalester College, 1600 Grand Ave, St Paul, MN 55105; Phone: (651) 696-6900; Email: `glade@macalester.edu`

[4]Department of Civil and Environmental Engineering, University of California, Davis, One Shields Ave, Davis, CA 95616; Phone: (530) 754-2297; Email: `fjloge@ucdavis.edu`

[5]Department of Food Science and Technology, University of California, Davis, One Shields Ave, Davis, CA 95616; Phone: (530) 754-5447; Email: `esspang@ucdavis.edu`

1 Introduction

The growth in the application of behavioral interventions has been tremendous. Governments and policymakers across the world use default settings, social norms comparisons, commitment devices and salience signals to encourage retirement savings, college enrollment, vaccinations, and participation in job-training programs (Benartzi et al., 2017). An expansive literature has evaluated the impacts of these nudges on behavior and in particular the use of social norms comparisons as a tool to influence choice (Allcott, 2011; Beshears et al., 2015; Croson and Shang, 2008; Duflo and Saez, 2003; Frey and Meier, 2013; Gerber and Rogers, 2009). While these studies demonstrate that social norms comparisons alter behavior, the mechanisms underlying the response to these nudges remain elusive.

This paper deploys a randomized controlled trial and uses high-frequency data to explore some potential channels through which agents may respond to social norms comparisons. We lean on the experimental design and ask what features of a popular behavioral intervention induce a response? In practice, social norms comparisons typically comprise a bundle of treatments, and experiments are rarely designed to map each treatment to a behavioral effect. The behavioral intervention that we study shares this first commonality - it consists of social norms comparisons, informational messaging and personalized recommendations - but by design our experiment directly tests which of these components induces a response. Second, we combine granular data on water use with detailed information on longer-run decisions to empirically examine the margins by which agents respond to nudges, and the persistence of these effects. Often, data limitations make it challenging to understand how agents respond to treatment. The availability of high-frequency data spanning eighteen months post-treatment and rich supplemental data on participation in rebate programs allow us to overcome some of these hurdles.

We study these questions in the economically interesting and policy-relevant context of residential water use during the longest-recorded drought (2011 to 2016) in California history. The first distinguishing feature of our experiment is its timing. Our study took place during the last full year of the drought when a state-wide policy mandating a 25% reduction in urban water use, as well as many other conservation policies, were already in place (Board, 2015; Browne et al., 2021). The timing of our intervention provides a unique opportunity to test if social norms can induce conservation when layered on top of a suite of existing water conservation policies. A second distinct feature of our setting, and one that generalizes to most urban water utilities, pertains to how urban water is priced. Most urban water districts bundle some fixed costs into volumetric rates (Mitchell et al., 2017). When urban conservation is required, this pricing structure can lead to substantial revenue shortfalls and regulators may respond by raising rates. This was the case in California during this period, when revenue decreased for more than 70% of urban water suppliers (Mitchell et al., 2017). An implication of this pricing structure is that when conservation is mandated, the utility incentive is for temporary rather than permanent reductions in water use. The time span of our data allows us to examine if nudges provide a conservation instrument that aligns with utility incentives for immediate, but temporary, water conservation in response to drought.

This paper uses a large-scale field experiment that provided popular home water reports (HWR) aimed at urban water conservation to a random sample of households. A distinguishing feature of our experimental design is that we randomly varied the content of HWRs to test what features of the report elicit a conservation response. All treatment households received bi-monthly HWRs that compared their own water use to similar and efficient households, offered personalized water conservation recommendations, and shared information about utility-sponsored programs. Some treatment households were randomly assigned to receive personalized recommendations aimed at indoor water conservation. Given that the only difference across the two treatments is the content of the recommendations, a comparison across them allows us to test if households respond to the recommendations contained within these reports. A third treatment arm provided households with an incentive, in the form of a durable good, and informed them of it via the messaging component of the HWR. This treatment offers an opportunity to directly examine if households respond to the messaging portion of HWRs.

We find that the profile and level of water conservation are similar across all three treatments. The provision of HWRs reduced average hourly water use by 4% to 5%. Within- and across-day treatment effects highlight that while there is substantial heterogeneity in when water conservation occurs, conservation patterns are similar across all three treatments. In our setting, neither tailored water recommendations targeting indoor water use nor messaging advertising a durable good incentive had a differential effect on water conservation. This finding adds to recent experimental work that seeks to understand the channels through which home water and energy reports impact customer behavior. To date, the focus has been on the social norms comparison, and how the framing of this comparison in absolute, relative or ordinal metrics or when paired with an injunctive norm affects choices (Allcott, 2011; Byrne et al., 2018; Bhanot, 2017; Brent et al., 2015, 2020). Less is known about the messaging or recommendation components of the report, which may function to reduce cognitive limitations when making choices about water and energy consumption (Wichman, 2017; Brent and Ward, 2019; Attari, 2014; Allcott and Taubinsky, 2015; Jessoe and Rapson, 2014; Carlsson et al., 2019). Similar to Dolan and Metcalfe (2015), we find little evidence that customers respond to these features of the reports. Our finding suggests that households may be responding to the receipt of or the social comparisons component of HWRs.

The magnitude of our estimated treatment effect adds a critical data point to questions about the external validity of behavioral interventions in the context of residential water conservation. Residential households in our setting faced mandatory outdoor watering restrictions, social pressure to conserve water, and a state-wide 25% conservation mandate. These policies translated into large reductions in water use, with control households reducing water use by 26% year-on-year.[1] HWRs were layered on top of these utility-wide conservation policies, and it was unclear ex-ante if any remaining conservation levers existed. Despite the starkly different context, our short-run average treatment effects lie within the range of conservation effects reported from other recent studies of HWRs in California (Brent et al., 2015; Mitchell et al., 2017). For policymakers, this finding suggests that social norms comparisons can be integrated into the drought management toolkit and deployed in conjunction with other conservation instruments (Ferraro and Price, 2013; Brandon et al., 2019).

Another central finding of our study is the absence of a persistent response.The water conservation effect of each intervention is not statistically detectable five months after the end of the experiment. During the treatment period, households likely respond to HWRs through a reduction in indoor water use and outdoor irrigation, and increased compliance with watering restrictions. These conservation behaviors remain for the first four post-treatment months, but decay quickly thereafter. Indoor and outdoor watering behaviors of treatment households mirror those of control households five months after the intervention ended. One reason for this short-lived response may be that, in our setting, households did not respond to treatment through investment in water-efficient capital. Data on the uptake of rebates for water-efficient durables support this hypothesis. We find that HWRs have no impact on participation in water rebate programs. While our results point to potential limitations of HWRs as a long-run water conservation policy, the temporary reduction in water use aligns with the pricing model of many utilities.

The short-lived conservation effect detected in our setting is an exception to the persistent impacts typically documented in the residential water and energy space (Allcott and Rogers,

[1]This water use reduction occurred in many other utilities across the state. Over 43% of utilities met the state-wide conservation mandate. The exact policies leading to this reduction are utility specific (Browne et al., 2021).

2014; Brandon et al., 2017). In the water sector, social norms messaging has led to conservation effects enduring more than two years post-treatment and increased participation in water conservation and rebate programs (Bernedo et al., 2014; Ferraro et al., 2011; Brent et al., 2015). One likely reason for the divergence between our finding and others reported in California is the context in which we study HWRs, specifically, the presence of a historic drought.[2] Households may have responded to the drought or drought-related programs via the uptake of water-efficient durables and rebates, rendering this margin of response unavailable. More broadly, our finding that treatment does not lead to investment in water-efficient durables or lasting habit formation suggests that the geography and time period matter when evaluating the lifespan of treatment effects, and hence the cost-effectiveness of behavioral interventions.

The paper begins by discussing the experimental research design and data. It then compares treatment effects across the three interventions. The paper then evaluates dynamic treatment effects of treatment, leveraging the high-frequency nature of our data to explore the potential margins by which households respond to treatment. Lastly, the paper concludes.

2 Social Norms and Resource Use

Empirical evidence on the effect of social norms comparisons on resource conservation took root in the behavioral psychology literature through the example of towel use among hotel guests. In one hallmark study, Goldstein et al. (2008), randomly placed cards with either a pro-social environmental message or a social norms message in hotel rooms and compared towel reuse in these hotel rooms to reuse in the control group. The authors finding that 44% of guests assigned to the social norms treatment reused towels, perhaps provided inspiration for the widespread dissemination of social norms comparisons throughout water and energy utilities in the U.S.

In the residential water and energy settings, these comparisons primarily occur through the vehicles of Home Energy Reports (HERs) and Home Water Reports, and seek to induce energy and water conservation. From a methodological perspective, what is noteworthy about their deployment is that the companies providing these products typically use randomized controlled trials. This provides a credible research design to evaluate the effect of HWRs and HERs on residential water and energy consumption, respectively.

[2]Bernedo et al. (2014); Ferraro et al. (2011) document long-term conservation effects of a one-time social-norms comparison that was deployed during a drought in the Atlanta, GA area.

The coupling of this research design with utility-specific variation in the content of reports has given rise to a wealth of studies that seek to understand why, for whom, and under what conditions these interventions elicit a response. In the residential energy space, HERs deployed by the company OPower have been shown to induce on average a 1% to 2% conservation effect. However, this average treatment effect masks heterogeneity in the response. Treatment effects vary by baseline electricity use (Allcott, 2011; Byrne et al., 2018), across print and digital reports (Dolan and Metcalfe, 2015), informedness on own electricity use (Byrne et al., 2018), and political ideology (Costa and Kahn, 2013). Dynamics and site selection also influence the magnitude of the response. Backsliding occurs between reports; conservation effects persist once the (e)mailing of reports end, slowly decaying over time; and bias in the sites selected for the deployment of HERs likely overstate treatment effects when HERs are used in other utilities (Allcott and Rogers, 2014; Allcott, 2015). More recent experimental work finds limited evidence of crowd-out effects from the provision of multiple energy nudges, and a differential response when financial incentives are provided in HERs (Brandon et al., 2019; List et al., 2017).

Table 1 shows that researchers have applied the same rigor to the question of social norms messaging and residential water use. The table provides an overview of field experiments since Ferraro et al. (2011) that evaluate the effect of social norms on residential water use. We briefly summarize the empirical setting, sample size, intervention studied, and treatment effects for each study.[3] Collectively, this body of work highlights that social norms messages, on average, induce residential water conservation effects of 3 to 5 percent. These effects have been documented in locations throughout the U.S., during both drought and non-drought conditions, and in response to monthly, bi-monthly, and one-time comparisons (Ferraro et al., 2011; Bernedo et al., 2014; Ferraro and Price, 2013; Goette et al., 2019; Bhanot, 2017, 2018; Brent et al., 2015, 2020).

Experimental studies in the residential water setting have sought to understand the features of the HWRs that induce a conservation effect and the margins by which households respond to these interventions in the short- and long-run. Our paper adds to this literature along three dimensions. A first body of work asks what component - messaging, recommendations, social comparisons - of HERs and HWRs elicit a conservation effect. Most of this literature has zoomed in on the social comparisons piece, and shown that the framing of

[3]The list does not include all studies on non-price conservation instruments in the residential water space. Importantly, it excludes non-price interventions that do not contain a social norms component (e.g. Wichman (2017); Tiefenbeck et al. (2018)), non-experimental studies (e.g. Browne et al. (2021)), and the majority of studies that occurred outside of the U.S. (e.g. Torres and Carlsson (2018)).

the comparison in relative versus absolute metrics, as an ordinal ranking, and as an injunctive norm influences water and energy use (Bhanot, 2017, 2018; Brent et al., 2020; Allcott, 2011). Less well-studied are the two other features of the HWRs - recommendations and messaging - on water use. Our experimental design provides an opportunity to isolate the effect of these treatments on water use. It is most similar to Dolan and Metcalfe (2015), who compare how energy statements with social norms affect energy consumption, relative to energy statements comprised of both social norms and information on energy use. Their study affords the opportunity to isolate the effect of social norms comparisons and the incremental effect of information, while ours provides an opportunity to test the incremental effect of two stalwart features of HWRs - water saving recommendations and messaging.

The conservation effects of these interventions typically persist, with studies citing habit formation and capital investments as the responsible mechanisms. In the residential electricity setting, Allcott and Rogers (2014) show that the energy conservation from HERs decay but remain detectable once treatment ends, and attribute this to habit formation. Brandon et al. (2017) revisit this question using data from dozens of experiments and find that investment in energy-efficient capital partly explains the persistence of the treatment effect. In the residential water setting, (Ferraro et al., 2011; Ferraro and Price, 2013; Bernedo et al., 2014) show that a one-time social norms message leads to conservation effects detectable more than three years post-treatment. Brent et al. (2015) offer up capital investments as one explanation for this persistence, showing that HWRs increased participation in water conservation programs by 8%. High-frequency data on water use offer a new opportunity to understand the margins by which households respond to HWRs in the short- and long-run. We evaluate the effect of HWRs on the daily profile of water consumption and the response to day-of-week outdoor watering restrictions. If HWRs, for example, induce households to reprogram outdoor irrigation, then this will be detectable with high-frequency data, and will suggest that habit formation partly explains the response to treatment. Our study offers a complement to existing work that use more temporally aggregated data to understand the long-run effects of HWRs.

A third contributing feature of our study is the setting in which HWRs were deployed. Previous work has shown that in the context of California, HWRs that compare own water use to that of others lead to conservation effects of 3% to 5% (Brent et al., 2015; Bhanot, 2017, 2018). A shared feature of these studies is that they occur during non-drought periods. In contrast, the timing of our experiment coincides with the most severe drought in California's history. Mandatory outdoor watering restrictions were in place; the state had imposed a 25%

urban conservation mandate; utility-sponsored conservation programs had been deployed; and there was strong social pressure to reduce water use. Perhaps as a result, control households had engaged in year-on-year water reductions of 26%. For these reasons, it was uncertain if the water savings found in earlier studies could be reproduced in our setting.[4] Our study speaks to the external validity of earlier findings on the water savings from HWRs, and to whether this intervention acts as a substitute to coincident conservation programs (Brandon et al., 2019).

Our study shares the same experimental framework as a recent companion paper by Jessoe et al. (2021), but is distinct in the questions asked, outcomes studied, and treatments of interest. In Jessoe et al. (2021), the authors focus exclusively on one treatment arm, "WaterSmart Only", to evaluate the short-run effects of social norms on *electricity* use. The current study focuses on a separate set of questions, and uses different data and treatment arms to answer them. First, by design, our experiment seeks to understand what components on the HWR induce a response. Specifically, we randomly vary the content of the recommendations and the strength of the messages to test if households respond to the recommendations and messaging features of the HWR. Second, this paper evaluates the margins by which households respond to HWRs in the long-run. Third, we ask what is the effect of HWRs on water use during an extreme drought, and when layered on top of mandated watering restrictions.

3 Experimental Design and Data

We deployed a framed field experiment in a service territory with high-frequency water data to evaluate the effect of water conservation instruments on short-run and long-run water use. We implemented the experiment in partnership with WaterSmart and Burbank Water and Power, a municipally-owned utility serving roughly 18,500 single-family homes in the City of Burbank. The experiment spanned March 2015 to May 2016 and included the summer marked by the worst drought in California's history. While treatment ended in May 2016, we continued to collect hourly interval data through December 2017, more than eighteen months post-treatment.

[4]Ferraro et al. (2011) and Brent et al. (2020) show that social-norms comparisons can reduce water use during drought in Georgia and Nevada, respectively. However, the authors used different, utility-based messaging.

3.1 Research Design

Our sample consists of 16,900 single-family homes that had billing records for at least six months preceding treatment. We randomly assigned households to the control group or one of three treatments: 'WaterSmart Only,' 'Hot WaterSmart,' or 'Hot WaterSmart Plus.' We describe each treatment below. The 2,967 households assigned to the control group received no notification that they were in a pilot program.

All treatment households received six bi-monthly HWRs between May 2015 and April 2016.[5] In March 2015, before the arrival of the first report, all treatment households received an introductory letter that explained what HWRs were and when they would be delivered. Both the initial letter and households' first HWR were sent by mail. All subsequent HWRs were sent by mail to households that received their utility bill via mail and by email to those that paid their utility bill online.

In the year preceding the experiment, power calculations were performed using monthly water use data for all eligible households. We assumed that the sample size would consist of 13,200 households, each treatment account would receive five HWRs, and an equal number of households would be assigned to each group. We calculated a minimum detectable effect (MDE) of 1.4%, with 80% power and 5% size for each treatment arm. This implies that our experiment will not be able to statistically distinguish across small differences, less than 1.4%, in the response to the three treatments. It also motivated the provision of a strong durable goods incentive. Power calculations were also performed for a joint treatment indicator in which all three treatment arms were collapsed into a single treatment arm. We calculate a 1.1% MDE, which is smaller than the treatment effects estimated by Ferraro et al. (2011) two years post-treatment (Table 1).

WaterSmart Only: The 4,470 accounts randomly assigned to the WaterSmart treatment received HWRs comprised of personalized conservation recommendations, information on utility-sponsored conservation programs and water use, and social comparisons. Figure A.2 provides an example of a report - the social comparison appears on the top left, utility announcements on the top right, and personalized water-saving actions on the bottom of the report. The former compares own water use in the previous billing cycle to water use of similar and efficient households, and contains an injunctive norm comprised of a smiling,

[5] Appendix A and Figure A.1 provide a detailed timeline on the deployment of HWRs. They also include information on coincident utility conservation programs, data availability, and the timing of the mailer.

indifferent, or frowning water drop. Water recommendations include projected water savings and the value of those water savings calculated using utility water rates. The impact of similar HWRs and Home Energy Reports (HERs) on water and electricity use, respectively, has been the focus of several studies (Brent et al., 2015; Ayres et al., 2013; Allcott, 2011; Schultz et al., 2007).

Hot WaterSmart (HWS): We assigned 4,709 households to a 'Hot WaterSmart' treatment to test if recommendations aimed at hot water use could reduce indoor water use. The treatment is the same as the WaterSmart treatment, with two exceptions. First, half of the personalized water-savings recommendations focus on actions that could reduce indoor water use. Second, in addition to quantifying the expected water savings, all recommendations quantified expected natural gas savings and the cumulative dollar value attributable to these savings. Figure A.3 provides an example of a 'Hot' HWR. In this HWR two of the three water-saving recommendations are tailored towards indoor water use (i.e. reduce water heater temperature and fill bathtub three-quarters of the way).[6] Since the only distinction across this treatment and the 'Water Smart Only' treatment is the recommendations contained in the report, our experiment is designed to examine if households respond to the recommendation component of HWRs.

Hot WaterSmart Plus (HWS+): The 4,701 households randomly assigned to this treatment had a durable goods incentive layered on top of the Hot WaterSmart treatment. This incentive took the form of a water and natural gas conservation contest. A household would win a water- or energy-efficient durable if (i) it enrolled in the contest and (ii) predetermined conservation targets were met.[7] To increase program enrollment, we sought to minimize enrollment costs and offer a prize of meaningful value. Enrollment only required visiting the contest website to enter a name and email address, and to agree to the terms of conditions. Conditional on enrollment, all households that met the targets were guaranteed a prize. The twenty-five households with the greatest water use reductions would win a high-efficiency clothes washer ($850 retail value). The next one hundred enrolled households would win low-flow shower heads ($30 retail value). All remaining households would win an energy efficiency kit ($10 retail value).

[6]Figure A.4 shows all hot water saving recommendations used over the treatment year.

[7]We set two conservation targets: a 27% year-on-year reduction in water use and a 3% year-on-year reduction in natural gas use. The water conservation target corresponded to the conservation mandate imposed by the state of California on BWP. The natural gas conservation target was set based on discussions with our natural gas partner.

Importantly, contest information, the enrollment procedure, and progress towards meeting the conservation targets were conveyed exclusively through messages in the HWRs. Households were informed of this contest in the messages of three HWRs - the program was introduced to households in the third HWR (Figure A.5), and households were provided with individualized progress updates in their fourth and sixth HWRs (Figure A.6). Given that the distinguishing feature across this treatment and Hot WaterSmart is the content of the messages, we can directly test if households respond to the messaging component of HWRs.

3.2 Mapping the Experimental Design to Behavior

A theoretical literature has shown and continues to explore the circumstances under which cognitive limitations such as self-control problems or inattention may introduce errors into the consumer decision-making process (Thaler, 2018; Allcott and Kessler, 2019). Nudges, defined broadly as interventions that seek to alter behavior without changing prices or choice sets, have been proposed as a tool to address classic market failures and behavioral biases (Loewenstein and Chater, 2017). Home Water Reports describe one nudge. They aim to induce water conservation by comparing own water use to that of peers, providing recommendations on ways to reduce water use, and through the actual receipt of the report. Adhering to the taxonomy defined in Carlsson et al. (2019), these reports may function as both a cognitive and moral nudge.

The social norms component of HWRs, in which own water use is compared to water use among peers, comprises a textbook moral nudge. Assume that some consumers derive moral (dis)utility from water consumption. As in Levitt and List (2007), one can conceptualize this as a moral tax on water. The possibility of a moral tax is particularly likely in our setting since our study coincided with the most severe drought in California's history, and strong social pressure from local and state governments to reduce water use. HWRs may increase or decrease this tax depending on how one compares to peers (Allcott, 2011; Byrne et al., 2018; Ferraro and Price, 2013). If HWRs induce conservation by increasing the moral tax on water, then they operate as a moral nudge. Our experiment is not designed to isolate the impact of the moral nudge component of the HWR on behavior.

Cognitive nudges seek to reduce or leverage cognitive limitations to induce a socially desirable response. Water savings recommendations contained in HWRs may reduce the difficulty in converting water-consuming actions into water use and expenditures (Kahneman, 2003; Brent and Ward, 2019; Attari, 2014; Allcott and Taubinsky, 2015). They may also provide new information on the financial savings from water conservation. Our experimental design directly tests if recommendations alter decisions, and in doing so, explicitly examines one cognitive channel through which HWRs may impact consumer behavior. If the margins by which customers respond to HWRs differ based on the recommendations received, then this provides empirical support that HWRs partly operate as a cognitive nudge. However, the failure to detect an effect does not rule out HWRs as a cognitive nudge. Receipt of these reports may draw attention to water conservation and, therefore, reduce households' water use.

The messaging component of HWRs may operate as a cognitive or moral nudge depending on the message a utility chooses to display. However, messages will only influence customer behavior if customers are attentive to them. The contest arm of our experiment seeks to make messages more salient by using them as the exclusive vehicle to display and advertise a durable goods incentive. If households assigned to this treatment disproportionately reduce water use, they may be responding to the durable goods incentive or because messages have become more salient (Myers and Souza, 2020; List et al., 2017). The absence of a differential effect implies that customers may overlook this component of the HWR, disregard it because the incentive is too low or of little value, or perhaps respond to HWRs for intrinsic reasons (Brent and Wichman, 2020).

3.3 Data

Hourly water consumption data serve as our primary outcome of interest. BWP provided water use data for all single-family homes in its service territory between April 1, 2014, a year before the intervention, and December 31, 2017, over a year and a half after the experiment ended. We supplement these data with information on customer participation in BWP's rebate programs, assessor data on housing characteristics, and hourly temperature and precipitation data from a nearby weather monitor.

Table 2 compares baseline characteristics across control and treatment households. The table highlights that, before the intervention, treatment and control households are balanced in seasonal water use, rebate applications and household characteristics.[8] As is typical in California, water use exhibits large seasonal fluctuations, with use peaking in the dry and warm summer months when outdoor irrigation is most prevalent. Hourly water use measures at 11.5 and 19.5 gals/hour in the winter and summer months, respectively. Water use is balanced in all 2015 pre-treatment months. To provide additional evidence on the quality of the randomization, we plot out the distribution of average daily water use across control and treatment households. Figure 1 which illustrates these distributions, makes clear that the distribution of daily water use is balanced across treatment and control.

Detailed data on the uptake of utility-sponsored water-efficient rebates provides an opportunity to test the hypothesis that treatment households respond to WaterSmart by increasing investment in water efficient durables. The utility provided information on the rebate type, rebate amount and rebate date for all utility sponsored water efficiency programs between January 1, 2014 and June 30, 2016. As shown in Table 2, participation in existing water rebate programs is relatively low, with only 3% of households applying for a rebate between January 1, 2014 and June 30, 2016. The value of the rebate is also low, with the mean rebate value amounting to $2.

4 A Comparison Across Interventions

To isolate the impact of each treatment on water use and test for differential effects across the three treatments, we compare average hourly water use across control and treatment households during the twelve-month treatment period. We estimate the following regression:

$$y_{iht} = \sum_j \beta_j \mathrm{T}_{ij} + \mathbf{X}'_{iht}\gamma + \delta_h + \delta_t + \epsilon_{iht}. \tag{1}$$

The dependent variable y_{iht} is household i's water use in hour h of calendar date t. The regressors of interest are T_j, which equal one if a household is assigned to treatment $j \in \{\text{WS, HWS, HWS+}\}$, and zero otherwise. In some specifications, we augment equation (1) to improve the precision of our estimator, conditioning on pre-treatment seasonal water use and weather data ($\mathbf{X}'_{iht}$), as well as hour-of-day (δ_h) and calendar date fixed effects (δ_t).

[8]We use the historical water use data to construct three seasonal average water use statistics that are used as controls in some regression specifications. We define two broad seasonal classifications, summer defined as April to October, and winter defined as November to March. These seasonal classifications are also used to construct other variables. These seasonal designations align with BWP's seasonal outdoor watering restrictions.

Our first set of results, reported in Table 2, makes clear that all three HWRs led to a reduction in water use. As shown in columns (1-2), hourly water use reduced by -0.49 to -0.60 gallons, or 3.8% to 4.8% over the treatment year.[9] The remainder of this table separately highlights treatment effects in the first six months of the treatment period (columns 3 and 4) and the last six months of the treatment period (columns 5 and 6). The level reductions in water use are similar when we break out the results by the first and second half of the intervention, despite baseline use changing substantially across seasons.[10]

While the magnitude of our reported treatment effects mirrors those reported in other studies, the setting is unique. WaterSmart has deployed HWRs throughout utilities in California, and recent work has evaluated the effect of these reports on water use. Under normal precipitation conditions, HWRs induce average water savings of 4% to 5% (Bhanot, 2018; Mitchell et al., 2017; Brent et al., 2015). The unanswered question is how would perform during droughts. We show that HWRs continue to deliver savings of 2.8 to 3.7% during extreme drought, even when layered on top of other conservation instruments. While we cannot weigh in on whether this result would have occurred in the locations that are the focus of Bhanot (2018); Mitchell et al. (2017); Brent et al. (2015), our finding highlights that in our setting other conservation instruments such as outdoor watering restrictions or a conservation mandate do not crowd out the savings from HWRs.[11] More generally, it suggests that behavioral approaches may not operate as a substitute to other conservation instruments during times of water scarcity. The absence of a crowding out effect echos the finding of Ferraro and Price (2013) and Brent et al. (2020), who also find that social norms comparisons deliver water conservation on top of outdoor watering restrictions and during times of drought.

We find no differential impact of the three treatments on water use despite differences in the content of the HWRs. A comparison of the WaterSmart and HotWaterSmart treatment effects highlights that modifying recommendations to include natural gas conservation and the accompanying financial savings led to no changes in average hourly water use. The

---

[9] Comparing results across columns 1 and 2 highlights the stability of our results to the inclusion of controls, suggesting that the addition of these covariates serves to increase the precision of our estimates.

[10] Testing the impact of multiple treatments on multiple outcomes leads to natural concerns that standard errors are incorrect due to a failure to account for multiple hypothesis testing. Appendix C presents adjusted p-values that account for multiple-hypothesis testing across several dimensions. P-values naturally increase with these adjustments, though all primary results remain statistically significant.

[11] We cannot quantify the water conservation that would have occurred if the deployment of HWRs preceded the drought. The ideal experiment would have staggered the introduction of HWRs, randomly assigning HWRs to one treatment group before the drought and to another treatment group during the drought.

addition of recommendations aimed at hot water conservation also did not alter water use patterns across days of the week and hours of the day (Appendix B.2). This finding aligns most closely with Dolan and Metcalfe (2015), who find no differential response to social norms comparisons in the presence and absence of information. One caveat in our comparison of treatment arms is that we cannot rule out the possibility that small (less than 1.4%) differences exist across treatment groups.

Similarly, altering the content of messages in HWRs to showcase and advertise a durable goods incentive for water conservation induced no change in average hourly water use or water usage patterns. One reason why households may not have responded to the incentive program is that it was conveyed exclusively through messages in HWRs, and households may have been inattentive to these messages. Only 68, or 1.4%, of customers enrolled in the "Conserve and Win" program, despite 253 households achieving the water and natural gas targets necessary to receive a durable good prize. Low enrollment rates, including for those who met the contest requirements, suggest that households may simply have overlooked the messaging portion of HWRs.

Collectively, our experimental results provide evidence that households do not respond to the messaging or recommendations components of the HWRs. These results are consistent with a framework in which the receipt of or the social norms component of the report alters behavior. Our findings align with experimental work that demonstrates the framing of social norms messages in absolute, relative, and ordinal metrics, and its pairing with an injunctive norm impact consumption (Brent et al., 2020; Allcott, 2011; Byrne et al., 2018; Bhanot, 2017). Given the similar response across treatment groups, moving forward, we combine all three treatments into a single treatment that we refer to as "WaterSmart," and focus on the short- and long-run impacts of treatment as well as the mechanisms underlying the response to treatment.

5 Dynamic Effects

Hourly water data spanning twenty months post-treatment allow us to identify impacts of HWRs on water use patterns and, indirectly, households' behaviors. We first investigate the short- and long-run average treatment effects of HWRs. This allows us to evaluate persistence in response to treatment. We then take advantage of our hourly water use data to study patterns in treatment effects across hours of the day and days of the week.

We leverage utility-designated watering restrictions and seasonal patterns in water use to understand the margins along which households respond to treatment. In particular, we explore treatment effect heterogeneity along these two dimensions to indirectly test whether treatment affected both indoor and outdoor water use behaviors. Last, we explore the extent to which households respond to HWRs by investing in water-efficient capital.

5.1 Short and Long-Run Effects

We first compare hourly water use across control and treatment households in calendar-months before, during, and after treatment to identify the duration of treatment effects. We estimate,

$$y_{iht} = \sum_{\tau} \beta_{\tau} \left(\mathbf{1}[t_{\tau} = \tau] \times \mathrm{T}_i\right) + \mathbf{X}'_{iht}\gamma + \delta_h + \delta_t + \epsilon_{iht}. \tag{2}$$

As before, the dependent variable y_{iht} is household i's hourly water use (gals/hour) in hour h of calendar date t. T_i is an indicator for our joint treatment effect, and the indicators $\mathbf{1}[t_{\tau} = \tau]$ equal one if date t is in month τ. We also condition on the covariates previously described. We include data from March 2015, two months before the intervention, through December 2017, twenty months after the intervention ended.

The coefficients of interest, β_{τ}, measure the effect of assignment to treatment on hourly water use in calendar month τ, relative to control households. Figure 3 plots the coefficient estimates and the corresponding 95% confidence intervals. The shaded area A corresponds to the treatment year. We classify the four months post-treatment as the 'backslide' period, and label this as shaded area B. The third area, denoted by C and which we refer to as the 'convergence' period, corresponds to post-treatment months five to twenty.

The figure illustrates that, while HWRs led to a near-uniform reduction in water use levels during the treatment months, the conservation effect quickly decays. In the twelve months during which households received HWRs, water use reduced by 0.5 to 0.6 gals/hour.[12] However, the treatment effect is short-lived. We continue to observe conservation effects of 0.5 to 0.6 gals/hour in the first two post-treatment months, but these effects decay by roughly a third by the third post-treatment month. Five months after treatment ends, there is no

[12] Figure B.5 replicates Figure 3 using the log of water use as the dependent variable. Percentage reductions are largest in winter months when average water use is lower.

statistically discernible difference in water use across control and treatment households.[13] The convergence in water use between treatment and control households remains through the end of our sample.

Our finding that HWRs have short-lived impacts contrasts some previous experimental work. In the residential water setting, the conservation effects of a similar one-time intervention remained detectable six years after treatment (Bernedo et al., 2014). The impacts of HERs on residential energy use decayed but persisted five to ten years post-treatment (Allcott and Rogers, 2014; Brandon et al., 2017). One explanation for the lack of persistence in our setting is differences in the research design. In our experiment, treatment stopped at roughly the same time for all households. This set-up allows us to detect treatment effects in each post-treatment month relative to control households. In contrast, Allcott and Rogers (2014) randomize when one block of households in each treatment site stopped receiving HERs. This allows them to also quantify treatment effects in each post-treatment month relative to households that continue to receive reports.

Differences in the empirical setting, as opposed to differences in the research design, may also explain the absence of long-run effects. In the residential water setting, experimental studies have detected persistent effects by comparing treatment to control households in the post-treatment period (Bernedo et al., 2014; Ferraro et al., 2011). While we replicate this approach, we find that treatment effects do not endure beyond five months. A standout feature of our experiment is that it coincided with extreme drought, and it ended just as many of the drought restrictions were removed. During our treatment period, a statewide 25% mandatory conservation mandate was in place, outdoor watering restrictions increased in stringency, and water-efficient rebate programs had been deployed. These policies all relaxed following the winter of 2016. By June 2016, the regulator removed the statewide conservation mandate and BWP increased the number of allowable outdoor watering days. Looking at Figure 2 we see that water use is higher beginning in June 2016 relative to the corresponding month in the previous year, when stringent drought measures were in place. Given that our treatment period aligned with a period of extreme conservation measures, and our post-treatment period coincided with the removal of many of these measures, the absence of a persistent effect may be partly attributable to the end of a historic drought.

[13] Our experiment is powered to detect treatment effects of up to 1.1%, so we cannot definitively rule out the possibility of smaller long-run treatment effects. Our minimum detectable effects are smaller than the effects estimated in Ferraro et al. (2011), which evaluate treatment effects in the two years following treatment.

5.2 Hour-of-Day Treatment Effects

We take advantage of the temporal granularity in our data to better understand the absence of a persistent effect in our setting. We begin by evaluating the impact of assignment to treatment on the within-day water use profile. We also take advantage of seasonal differences in water use to indirectly examine how indoor- versus outdoor-water use behaviors respond to treatment. We do this for each of the three periods in Figure 3 – treatment, backslide, and re-convergence – to examine how these patterns of behavior change over time.

We estimate a fixed effects model where we interact assignment-to-treatment with indicator variables for each hour-of-day:

$$y_{iht} = \sum_{\eta} \beta_{\eta} \left(\mathbf{1}[h = \eta] \times \mathrm{T}_i \right) + \mathbf{X}'_{iht}\gamma + \delta_t + \epsilon_{iht}. \tag{3}$$

The regression mirrors equation (1) except that assignment-to-treatment is interacted with a vector of indicators $\mathbf{1}[h = \eta]$ that equal one when hour-of day h equals η. We estimate equation (3) for five distinct sub-samples: (i) summer treatment months, (ii) winter treatment months, (iii) the 'backslide' months; (iv) summer 'convergence' months; and (v) winter 'convergence' months. Recall that the treatment months refer to when households received bi-monthly HWRs; the 'blackslide' period defines the first four months post-treatment spanning May 2016 to August 2016; and the 'convergence' period characterizes post-treatment months 5 to 20. We estimate separate specifications for summer and winter months to understand household behavior during months when demand for outdoor watering is high.

We first focus on the effects of HWRs on water use during summer and winter treatment months (Panels A1 and A2 of Figure 4).[14] The figure suggests that treatment impacted both outdoor and indoor water use. During most summer treatment months, households were only allowed to water outdoors on Tuesdays and Saturdays after 6 PM and before 9 AM (Figure A.1). During winters months, households were only allowed to water outdoors on Saturdays during the same hours. In the summer months, households exhibit the largest water use reductions during the early morning and evening hours when outdoor watering was allowed. The largest reduction in winter water use also occurs during these hours, though the early morning and evening treatment effects are much smaller relative to the summer response. This may occur because demand for outdoor watering is lower in the winter, or

[14] Tables with hour-of-day and day-of-week regression results are available in Appendix B. Appendix C presents p-values that (partially) correct for multiple hypothesis testing.

because outdoor watering was limited to only one day a week. The timing of the largest summer and winter reductions in water use coincides with the hours when outdoor irrigation is permitted, and suggests one margin of response is outdoor irrigation

The within day response to treatment suggests that households are also responding to treatment by reducing indoor water use. We estimate treatment effects, around -0.4 gals/hour, from 10 AM to 6 PM when outdoor watering is prohibited. These treatment effects are present in both summer and winter months. Conservation effects during these hours suggest that households are also responding to treatment through indoor water use.

The remaining panels in Figure 4 plot treatment effects for the backslide and convergence periods.[15] Consider the backslide period (Panel B). Relative to the previous summer, the treatment effects diminish, most notably in the early morning and evening hours. Standard errors increase appreciably. However, the treatment effect profile remains similar to the one observed in Panel A1, suggesting that household water conservation habits persisted, but to a lesser extent, in the four months following treatment. Little evidence of a treatment impact remains in the convergence period (Panels C1 and C2). Both the winter and summer daytime treatment effects are precisely estimated and indistinguishable from zero, suggesting that indoor water use is nearly identical across control and treatment households. During the summer, some households still appear to conserve water in the early morning and evening hours. However, the standard errors are large, and the point estimates are smaller than those estimated during the treatment period, suggesting that the outdoor watering impacts of HWRs do not persist. Collectively, these results suggest that while HWRs induced meaningful changes in outdoor and indoor watering behaviors during the treatment period, these effects are relatively short-lived.

5.3 Day of Week Treatment Effects

We estimate day of week treatment effects to indirectly test if households respond to HWRs through the margins of outdoor water use and habit formation. In response to the historic drought, BWP imposed utility-wide outdoor watering restrictions that designated the days of the week and the hours on these days when outdoor irrigation was permitted. The deployment of HWRs may have induced systematically different responses during time periods with and without watering restrictions in place, both during the treatment period and in

[15]Note that all backslide months occurred over the summer of 2016.

the post-treatment months. For example, if treatment households responded to HWRs by re-programming their irrigation systems to run for fewer minutes on the hours and days when irrigation was permitted (e.g., 6-7 AM on a Tuesday), we would observe a large treatment effect on watering days during hours where outdoor watering was permitted. This would provide indirect evidence that households respond to treatment via outdoor irrigation. Further, if these settings remained in place after the last HWR was sent, we would continue to observe the same treatment effect pattern in both the backslide and re-convergence periods. This would provide indirect evidence in support of habit formation. To test for these margins of response, we estimate day-of-week effects both in the treatment and post-treatment periods, and do so separately for hours when outdoor watering was permitted and prohibited using the following regression:

$$y_{iht} = \sum_{\delta} \beta_\delta \left(\mathbf{1}[t_\delta = \delta] \times \mathrm{T}_i \right) + \mathbf{X}'_{iht}\gamma + \delta_h + \delta_t + \epsilon_{iht}. \tag{4}$$

We estimate equation (4) for four distinct periods. As with our hour-of-day results, we estimate the equation for winter and summer treatment months because different watering restrictions were in place in the summer (6 PM to 9 AM on Tuesday and Saturday) and winter (6 PM to 9 AM on Saturday) treatment months. We then estimate the equation for the backslide period, when the summer treatment irrigation rules were reinstated and outdoor watering was allowed on Tuesday and Saturday between 6 PM and 9 AM. Last, we estimate the equation for the entire post-convergence period, when outdoor watering restrictions were relaxed and irrigation was permitted from 6 PM to 9 AM on Tuesday, Thursday and Saturday.

Figure 5 presents our results, and plots the day of week treatment effect as well as the corresponding 95% confidence interval. The shaded gray area designates days when outdoor watering is permitted.[16] Red diamonds display treatment effects during the hours of the day when outdoor watering is permitted, and blue circles represent treatment effects during hours when outdoor watering is banned.[17] Panels A1 and A2 present the summer

[16] A comparison across plots highlights that watering restrictions changed throughout our study period. BWP (mostly) allowed outdoor irrigation on Tuesdays and Thursdays in the summer of 2015 and on Saturdays during winter 2015-2016. The utility returned to a Tuesday/Saturday schedule after the experiment ended, and instituted a permanent Tuesday/Thursday/Saturday schedule in August 2016 that remained in effect through the end of our sample. For brevity, we include results for the most prominent watering regimes. We show results for all watering regimes in Appendix B.3.

[17] Even on days when outdoor watering is permitted, it is only allowed during certain hours of the day. To account for this estimate we separate treatment effects for the hours spanning 6 PM to 9 AM, and the hours spanning 9AM to 6PM.

and winter treatment effects, respectively, panel B presents treatment effects during the 'backslide' period, and panel C depicts treatment effects in the re-convergence period.

Panel A1 illustrates that during the summer treatment period, the largest conservation effects occur on irrigation days during hours when households are allowed to water their lawns. Summertime treatment effects during these hours are about three times as large as the average treatment effects reported in Table 3. Treatment effects during non-watering hours on watering days (i.e. blue circle on gray shaded days) are similar to those estimated during all hours on non-irrigation days. We see similar results during winter months (Panel A2). Interestingly in the winter months, we still estimate a large treatment effect on Tuesdays during early morning and evening hours, though outdoor watering is prohibited. The results suggest that treatment caused households to reduce lawn irrigation relative to control households and that these habits persisted through the treatment year. We also find a stable reduction in water use on days and hours when households are not allowed to irrigate, suggesting that outdoor water use behavior is not the only margin driving our ATE.

The remaining panels illustrate that the conservation patterns developed during treatment persist into the first four post-treatment months, but erode in the convergence period. As shown in Panel B, we continue to observe large treatment effects in the backslide period during hours and days when outdoor watering is permitted. The response mirrors the results in Panel A1, though the estimates are slightly noisier. The results provide evidence of short-lived habituation to treatment, both in the daily patterns of water conservation and the magnitude of this conservation effect. These treatment effects are no longer present in the convergence period. We observe minimal differences across control and treatment households for all days and all hours. These results demonstrate that in the short- to medium-run, treatment households form indoor and outdoor water conservation habits, but that this suite of conservation habits decays entirely five months after treatment ends.

5.4 Capital Investments

One reason for the temporary response may be that, in our setting, households did not respond to treatment by investing in water-efficient capital. To examine this hypothesis, we use rebate application data to test for the effect of assignment to treatment on households' durable investments. BWP provided household-level information for all residential rebate applications between July 2014 and June 2016. We focus on rebate applications for water-efficient durables, including clothes washers, dishwashers, outdoor rain-water barrels, and turf replacement, and consider two outcomes. First, we create indicators for whether a household applied for any rebate in a pre-treatment month (July 2014 to February 2015), treatment month (March 2015 to May 2016), or post-treatment month (June 2016). Our second measure aggregates the dollar value of all rebates received by households in each of these three periods.

We estimate a simple difference-in-differences model, comparing rebate outcomes across control and treatment households before and after treatment,

$$r_{ip} = \gamma T_i + \sum_{\rho} \delta_\rho(\mathbf{1}[p = \rho]) + \sum_{\rho} \beta_\rho \left(\mathbf{1}[p = \rho] \times \mathrm{T}_i\right) + \mathbf{X}_i'\gamma + \epsilon_{ip}. \tag{5}$$

The dependent variable, r_{ip}, is either the indicator variable for whether household i applied for a water efficiency rebate in period p or the rebate amount ($) household i received in period p. We include an indicator variable, T_i, that denotes if a household was assigned to the WaterSmart treatment, and indicator variables for the treatment and post-treatment periods, ρ. Our regressors of interest are the treatment assignment by period interactions, and our coefficients of interest β_ρ, measure the effect of WaterSmart in the treatment and post-treatment periods on rebate participation relative to control households in the respective periods.

Table 4 reports our results along with rebate participation summary statistics for control households. Baseline program participation is low; less than 3% of control households applied for any rebates in the treatment year. The dollars received are also low, with households receiving just $5 on average. Most rebates are for appliances (clothes washers or dishwashers).[18] We find that HWRs had no detectable impact on rebate applications or the dollar amount in rebates a household receives. The absence of an effect is notable given that HWRs advertise rebate programs. The results provide another piece of supporting evidence that: (i) households did not respond to the messaging or recommendation components of HWRs; and (ii) assignment to treatment did not increase investment in hard capital.

6 Conclusion

This paper uses high-frequency data to evaluate the short and long-term effects of three behavioral interventions on household water use during a historic drought. We focus on the

[18] BWP offers generous rebates; the median rebate pays $70. The largest rebates over this period were for outdoor lawn replacement. Appliance rebates range from $25 to $170, while turf rebates range from $440 to $2,500.

impacts of Home Water Reports, an oft-deployed behavioral nudge comprised of social norms comparisons, messaging, and individualized conservation recommendations. We first use our experimental design to isolate the extent to which customers respond to each component of this bundled treatment. Neither the content of water recommendations or personalized messaging advertising a durable goods incentive alters the response to HWRs, suggesting that households respond either to the social comparisons portion of the treatment or the receipt of the HWR.

With climate change, droughts are expected to become more frequent and more severe. Given that our study coincides with the most extreme drought in California's history, it offers a preview into the potential of social norms comparisons to induce conservation under future weather conditions. We deployed our study during a period when households had already engaged in substantial conservation efforts. Year-on-year water reductions for control households amounted to 26%, and a statewide conservation mandate of 25% was in place. During peak drought conditions HWRs delivered average water savings of 2.8 to 3.5% on top of existing conservation efforts, with households reducing both indoor and outdoor water use. The magnitude of the estimated effects are similar to those reported under non-drought conditions, and provide a new data point on the suitability of behavioral interventions as a short-run drought conservation instrument.

These conservation effects of HWRs are short-lived and fully dissipate five months post-treatment. We find that, during the treatment period, households change indoor and outdoor water use patterns. These habits persist in the first four-months post treatment, but consumption patterns across treatment and control households are identical five months post-intervention. One reason for the temporary effect in our setting may be that shortly after our treatment ended, the drought was declared over and conservation policies were removed or relaxed.

The ability of HWRs to quickly elicit a 3% to 4% reduction in water use during times of drought may align well with the pricing model used by many urban water utilities. The rules governing the allocation of water in California at times require immediate and large water conservation among urban users. This was the case in the summer of 2015. One unintended consequence of the urban water conservation experienced during this drought is that it led to reduced revenues for more than 70% of urban utilities. This is because of how residential water is priced in California: most water utilities bundle some fixed costs into volumetric rates. The degree to which costs exceeded revenues depended in part on the magnitude and duration of water conservation. Many utilities responded to the end of the drought by raising volumetric rates to cope with resulting revenue shortfalls. In our setting, HWRs provide an immediate, but short-lived, reduction in water use that is compatible with the incentive structure of utilities during droughts.

References

Allcott, H. (2011). Social Norms and Energy Conservation. *Journal of Public Economics 95*(9), 1082–1095.

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics 130*(3), 1117–1165.

Allcott, H. and J. B. Kessler (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics 11*(1), 236–76.

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

Allcott, H. and D. Taubinsky (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review 105*(8), 2501–38.

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

Ayres, I., S. Raseman, and A. Shih (2013). Evidence from Two Large Field Experiments that Peer Comparison Feedback can Reduce Residential Energy Usage. *Journal of Law, Economics, and Organization 29*(5), 992–1022.

Benartzi, S., J. Beshears, K. Milkman, C. Sunstein, R. Thaler, M. Shankar, W. Tucker-Ray, W. Congdon, and S. Galing (2017). Should Governments Invest More in Nudging? *Psychological Science 28*(8), 1041–1055.

Bernedo, M., P. Ferraro, and M. Price (2014). The Persistent Impacts of Norm-Based Messaging and Their Implications for Water Conservation. *Journal of Consumer Policy 37*, 437–452.

Beshears, J., J. Choi, D. Laibson, B. Madrian, and K. Milkman (2015). The Effects of Providing Peer Information on Retirement Savings Decisions. *The Journal of Finance 70*(3), 1161–1201.

Bhanot, S. P. (2017). Rank and response: A field experiment on peer information and water use behavior. *Journal of Economic Psychology 62*, 155–172.

Bhanot, S. P. (2018). Isolating the effect of injunctive norms on conservation behavior: New evidence from a field experiment in california. *Organizational Behavior and Human Decision Processes*.

Board, S. W. R. C. (2015). Resolution no. 2015-0032 to adopt an emergency regulation for statewide urban water conservation.

Bonferroni, C. (1935). Il calcolo delle assicurazioni su gruppi di teste. *Tipografia del Senato*.

Brandon, A., P. J. Ferraro, J. A. List, R. D. Metcalfe, M. K. Price, and F. Rundhammer (2017). Do The Effects of Social Nudges Persist? Theory and Evidence from 38 Natural Field Experiments. NBER Working Paper No. 23277.

Brandon, A., J. A. List, R. D. Metcalfe, M. K. Price, and F. Rundhammer (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences 116*(12), 5293–5298.

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

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

Brent, D. A. and M. B. Ward (2019). Price perceptions in water demand. *Journal of Environmental Economics and Management 98*, 102266.

Brent, D. A. and C. J. Wichman (2020). Do behavioral nudges interact with prevailing economic incentives? pairing experimental and quasi-experimental evidence from water consumption. *Find this resource*.

Browne, O. R., L. Gazze, and M. Greenstone (2021). Do conservation policies work? evidence from residential water use. *Environmental and Energy Policy and the Economy 2*(1), 190–225.

Byrne, D. P., A. L. Nauze, and L. A. Martin (2018). Tell me something i don't already know: Informedness and the impact of information programs. *Review of Economics and Statistics 100*(3), 510–527.

Carlsson, F., C. Gravert, O. Johansson-Stenman, V. Kurz, et al. (2019). *Nudging as an environmental policy instrument.* Department of Economics, Göteborg University.

Clarke, D., J. P. Romano, and M. Wolf (2020). The Romano–Wolf multiple-hypothesis correction in Stata. *The Stata Journal 20*(4), 812–843.

Costa, D. L. and M. E. Kahn (2013). Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association 11*(3), 680–702.

Croson, R. and J. Shang (2008). The Impact of Downward Social Information on Contribution Decisions. *Experimental Economics 11*(3), 221–233.

Dolan, P. and R. Metcalfe (2015). Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation. Becker Friedman Institute for Research in Economics Working Paper No. 2589269.

Duflo, E. and E. Saez (2003). The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. *Quarterly Journal of Economics 118*(3), 815–842.

Ferraro, P., J. J. Miranda, and M. Price (2011). The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment. *American Economic Review: Papers and Proceedings 101*(3), 318–322.

Ferraro, P. and M. Price (2013). Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-scale Field Experiment. *Review of Economics and Statistics 95*(1), 64–73.

Frey, B. and S. Meier (2013). Social Comparisons and Pro-Social Behavior: Testing 'Conditional Cooperation' in a Field Experiment. *American Economic Review 94*(5), 1717–1722.

Gerber, A. and T. Rogers (2009). Descriptive Social Norms and Motivation to Vote: Everybody's Voting and So Should You. *Journal of Politics 71*(1), 1–14.

Goette, L., C. Leong, and N. Qian (2019). Motivating household water conservation: A field experiment in singapore. *PloS one 14*(3), e0211891.

Goldstein, N. J., R. B. Cialdini, and V. Griskevicius (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of consumer Research 35*(3), 472–482.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics 6*(2), 65–70.

Jessoe, K., G. E. Lade, F. Loge, and E. Spang (2021). Spillovers from behavioral interventions: Experimental evidence from water and energy use. *Journal of the Association of Environmental and Resource Economists 8*(2), 315–346.

Jessoe, K. and D. Rapson (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review 104*(4), 1417–38.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review 93*(5), 1449–1475.

Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives 21*(2), 153–174.

List, J. A., R. D. Metcalfe, M. K. Price, and F. Rundhammer (2017). Harnessing policy complementarities to conserve energy: Evidence from a natural field experiment.

Loewenstein, G. and N. Chater (2017). Putting nudges in perspective. *Behavioural Public Policy 1*(1), 26.

Mitchell, D., E. Hanak, K. Baerenklau, A. Escriva-Bou, H. McCann, M. Perez-Urdiales, and K. Schwabe (2017). Building drought resilience in california's cities and suburbs. Public Policy Institute of California.

Myers, E. and M. Souza (2020). Social comparison nudges without monetary incentives: Evidence from home energy reports. *Journal of Environmental Economics and Management 101*, 102315.

Romano, J. P. and M. Wolf (2005a). Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing. *Journal of the American Statistical Association 100*(469), 94–108.

Romano, J. P. and M. Wolf (2005b). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica 73*(4), 1237–1282.

Romano, J. P. and M. Wolf (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics and Probability Letters 113*, 38–40.

Schultz, P. W., J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science 18*(5), 429–434.

Thaler, R. H. (2018, June). From cashews to nudges: The evolution of behavioral economics. *American Economic Review 108*(6), 1265–87.

Tiefenbeck, V., L. Goette, K. Degen, V. Tasic, E. Fleisch, R. Lalive, and T. Staake (2018). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management science 64*(3), 1458–1476.

Torres, M. M. J. and F. Carlsson (2018). Direct and spillover effects of a social information campaign on residential water-savings. *Journal of Environmental Economics and Management 92*, 222–243.

Wichman, C. J. (2017). Information provision and consumer behavior: A natural experiment in billing frequency. *Journal of Public Economics 152*, 13–33.

Table 1: Experimental Evidence of Social Norms Comparisons in Residential Water Use

Paper	Intervention	Relevant Estimate	Setting
Bernedo et al. (2014)	One-Time, Utility I+ SC + PS	-1.7% (Yr 3), -1.3% (Yr 4)	100,000 HHs, Atlanta, GA 2007-2013; Drought
Bhanot (2017)	Bi-monthly, WS HWR + Rank	+1% to +3.6%[a]	4,000 HHs, Castro Valley, CA 2012-2013
Bhanot (Forthcoming)	Bi-monthly, WS HWR + IN	-2.2% to -3.1%[b]	45,000 HHs, Bay Area, CA 2014-2015
Brent et al. (2015)	Bi-monthly, WS HWR	-4.9% to -7.7%; no effect	7,400 HHs; 3 CA utilities
Brent et al. (2020)	Two-Time, Utility SC	-1.5%[c]	4,300 HHs, Truckee, NV 2015, Drought
Brent and Wichman (2020)	Bi-monthly, WS HWR + Price	-3.8%[d]	46,000 HHs, Southern CA 2014-2017
Ferraro et al. (2011)	One-Time, Utility I+ SC + PS	-2.5% (Yr 1), -1.25% (Yr 2)[e]	100,000 HHs, Atlanta, GA 2007-2009; Drought
Ferarro and Price (2013)	One-Time; Utility I+ SC + PS	-2.7% to -4.8%[f]	100,000 HHs, Atlanta, GA 2007; Drought
Goette et al. (2019)	Bi-monthly, Researcher PS + SC + F + $$	-2.2% to -2.6%[g]	1,000 HHs; Singapore 2016-2017
Kazukauskas et al. (2021)	Continuous; Private SC + RTF	0%	525 HHs; Umea, Sweden; 2016

Notes: The table summarizes relevant comparison studies to our own. The second column describes the frequency and mode of the intervention in the top row, and the type of the intervention in the second row. For the mode of intervention, we indicate the source of messaging: 'Private' is a municipality-owned rental company, 'Researcher' indicates researcher-run experiment, 'Utility' indicates a utility-designed, and 'WS' indicates WaterSmart designed intervention. For the type of intervention, 'F' indicates feedback, 'HWR' indicates a Home Water Report (often with an injunctive norm and information), 'IN' indicates an injunctive norm, 'I' indicates information, 'PS' indicates pro-social comparison, 'Price' indicates different pricing structures, 'Rank' indicates ranked-peer comparisons, 'RTF' indicates real-time-feedback, 'SC' denotes the study used a social comparison, '$$' indicates monetary prize. The third column presents the primary and most relevant findings from the studies. Where possible, we convert estimated treatment effects to percentage changes. The last column describes the study setting.

[a]Treatment effects are impacts of HWRs with two ranked comparison treatments relative to conventional HWRs.

[b]Injunctive norms included variations of HWRs with and without versions of the water droplet. HWRs with no drop induced the smallest percentage decrease while HWRs with the injunctive drop caused the largest decrease.

[c]Treatments included two variants of social comparisons, which had very similar average impacts. The authors find strong normative messaging likely leads to larger reductions.

[d]The authors find HWRs are not more effective for customers with higher marginal prices and do not make customers more price sensitive.

[e]The authors further explore the same experiment in previous studies, showing most impacts are driven by high water users and the impacts wane over time.

[f]The authors find a weak social norm had a smaller impact than a strong social norm that included a social comparison. Technical information alone had no statistically significant impact on water use.

[g]Impacts consisted of leaflets left at households with different messaging on the back of the leaflet. The largest treatment effects are for the post-like campaign, though there is little difference between each treatment.

Table 2: Balance Tests and Summary Statistics

	Control	WS	HWS	HWS+
Baseline Water Use April '14 (gals/hour)	17.00	16.77	17.06	16.92
Difference		-0.22	0.06	-0.08
p-value		(0.38)	(0.82)	(0.75)
Baseline Water Use Summer '14 (gals/hour)	19.60	19.39	19.56	19.47
Difference		-0.21	-0.04	-0.13
p-value		(0.44)	(0.89)	(0.62)
Baseline Water Use Winter '15 (gals/hour)	11.67	11.76	11.67	11.74
Difference		0.09	-0.00	0.07
p-value		(0.64)	(0.99)	(0.72)
2015 Pre-treatment Water Use (gals/hr)	12.34	12.34	12.35	12.38
Difference		-0.00	0.01	0.04
p-value		(0.99)	(0.97)	(0.84)
Rebate (Indicator)	0.03	0.03	0.03	0.03
Difference		0.00	0.01	0.00
p-value		(0.94)	(0.10)	(0.58)
Rebate ($)	1.99	1.96	2.64	2.28
Difference		-0.03	0.65	0.29
p-value		(0.93)	(0.13)	(0.43)
Year Built	1945	1945	1944	1944
Difference		-0.06	-1.05	-0.61
p-value		(0.86)	(0.14)	(0.40)
Square Feet	1630.42	1,634.24	1,634.09	1,643.19
Difference		3.82	3.67	12.77
p-value		(0.82)	(0.82)	(0.44)
Bedrooms	2.91	2.90	2.92	2.92
Difference		-0.01	0.01	0.01
p-value		(0.64)	(0.66)	(0.53)
Bathrooms	1.93	1.93	1.95	1.94
Difference		-0.00	0.01	0.01
p-value		(0.98)	(0.55)	(0.57)
Pool(Indicator)	0.23	0.23	0.23	0.22
Difference		-0.00	0.01	-0.01
p-value		(0.78)	(0.62)	(0.36)

Notes: The table presents average characteristics for control versus treatment households for our three treatments. WS, HWS, and HWS+ indicate the WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus treatment groups, respectively. 'Difference' is the difference in means relative to the control group and 'p-value' is the p-value from the relevant regression coefficient from running an OLS regression of the outcome on the three treatment indicator. For water use, standard errors are clustered at the household. For all other outcomes, standard errors are robust to residual heteroskedasticity.

Table 3: Water Intent-to-Treat Effects (Dependent Variable: Water Use (gals/hour))

	Full Sample (5/15-4/16)		First Six Months (5/15-10/15)		Last Six Months (11/15-4/16)	
	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.539***	-0.508***	-0.507***	-0.446***	-0.582***	-0.569***
	(0.171)	(0.101)	(0.196)	(0.116)	(0.163)	(0.112)
Hot WaterSmart	-0.603***	-0.591***	-0.524***	-0.497***	-0.701***	-0.684***
	(0.167)	(0.097)	(0.192)	(0.113)	(0.158)	(0.108)
Hot WaterSmart Plus	-0.508***	-0.485***	-0.461**	-0.401***	-0.571***	-0.567***
	(0.168)	(0.097)	(0.194)	(0.114)	(0.159)	(0.106)
H_0: WS=HWS	0.65	0.34	0.92	0.59	0.39	0.24
H_0: WS=HWS+	0.83	0.78	0.78	0.64	0.93	0.99
Mean Control Use	12.6	12.6	14.3	14.3	10.9	10.8
Observations	139,470,003	136,932,213	70,545,877	68,894,068	68,924,126	68,038,145
Weather Controls	No	Yes	No	Yes	No	Yes
Date, Month FEs	No	Yes	No	Yes	No	Yes
Pre-Treatment Use Controls	No	Yes	No	Yes	No	Yes

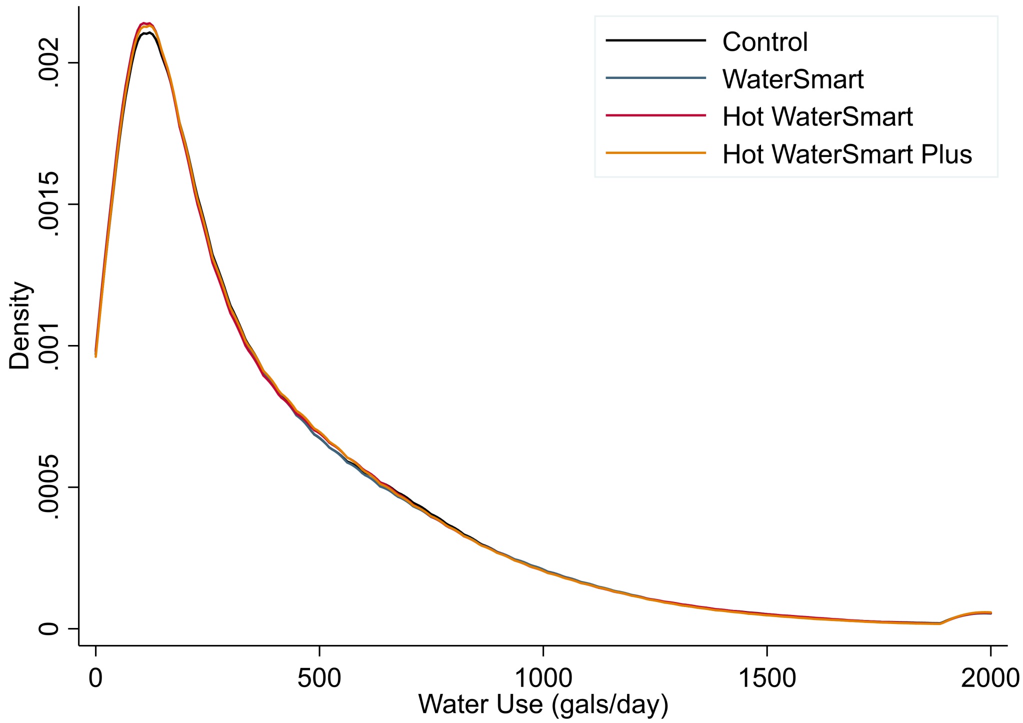
Notes: The table reports intent-to-treat estimates from an OLS regression of hourly water use on assignment to WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus treatments. Columns 1 and 2 include all observations from May 1, 2015 to April 30, 2016. Columns 3 and 4 restrict the sample to the first half of the intervention, and columns 5 and 6 restrict the sample to the second half of the treatment period. H_0: WS=HWS presents the p-value from a two-sided equivalency test between the WaterSmart and Hot WaterSmart treatments, and similarly for H_0: WS=HWS+. Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

Table 4: Water-Related Rebate Applications

	Any Rebate		Appliance Rebate		Turf Rebate	
Panel A: Rebate Application Indicator Outcome						
	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.002	-0.002	-0.002	-0.002	0.000	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
WaterSmart X Post-Treatment	0.001	0.001	0.001	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Mean Control Participation (%)	0.028	0.029	0.025	0.026	0.0030	0.0031
Panel B: Rebate Amount Outcome						
	Any Rebate		**Appliance Rebate**		**Turf Rebate**	
	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	1.608	1.230	-0.008	0.024	1.616	1.206
	(1.406)	(1.477)	(0.243)	(0.250)	(1.386)	(1.458)
WaterSmart X Post-Treatment	-0.480	-0.491	0.067	0.082	-0.547	-0.572
	(0.554)	(0.585)	(0.092)	(0.098)	(0.547)	(0.577)
Mean Control Amount ($)	4.88	5.07	1.80	1.81	3.08	3.26
Observations	54,441	48,324	54,441	48,324	54,441	48,324
Pre-Treatment Controls	No	Yes	No	Yes	No	Yes

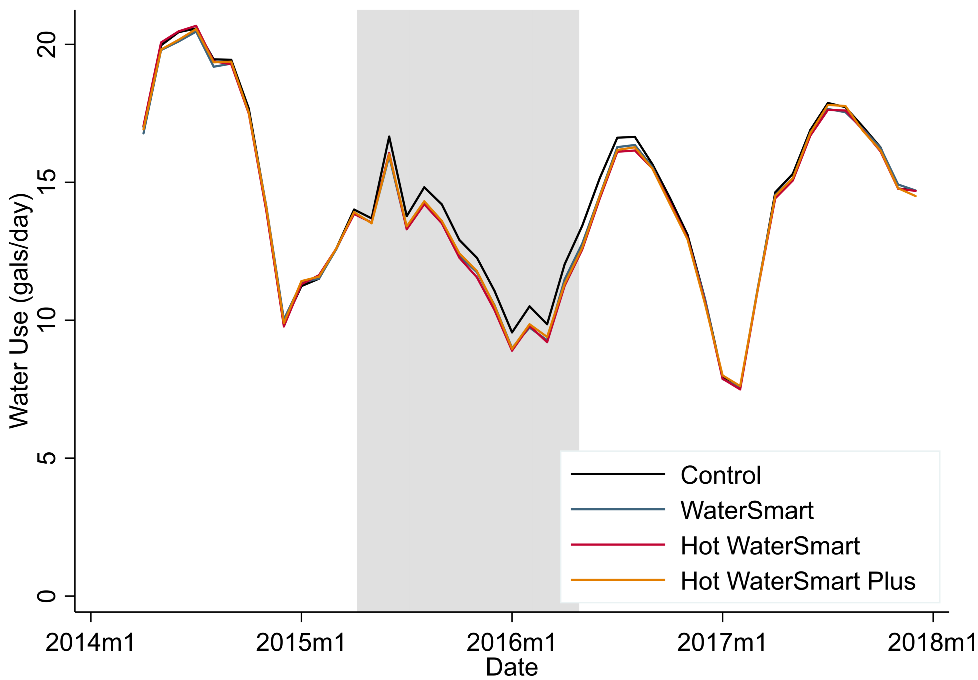
Notes: The table reports intent-to-treat results from an OLS regression of rebate application indicators (Panel A) and rebate payment amounts (Panel B) on assignment to WaterSmart, Hot WaterSmart, or Hot WaterSmart Plus over the treatment and post-treatment periods. The period spans May 2015 through Jun 2016. Pre-Treatment Controls include baseline household average water use for Winter 2014, Summer 2015, and Winter 2015. Standard errors are robust to arbitrary heteroskedasticity. *, **, *** denote significance at the 10%, 5%, and 1%level.

Figure 1: Pre-Treatment Water Use Balance



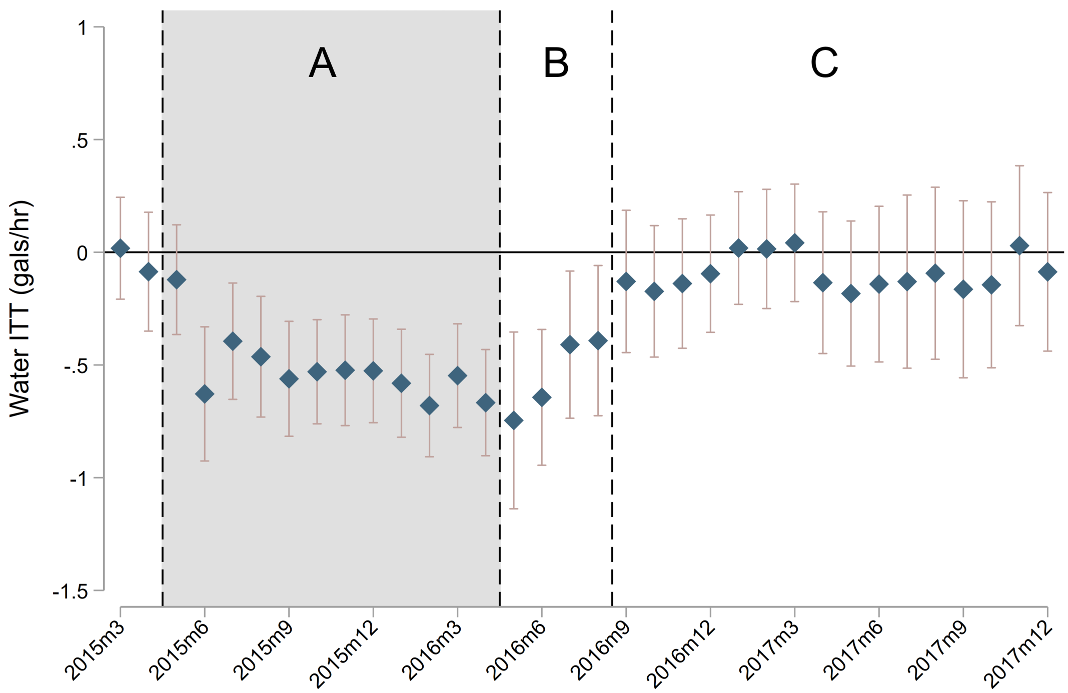
Notes: Figure 1 kernel density functions of daily average, pre-treatment water use for the control and treatment groups. The pre-treatment period includes April 2014 through February 2015. The distribution is truncated at 2,000 gals/day.

Figure 2: Average Water Use By Treatment



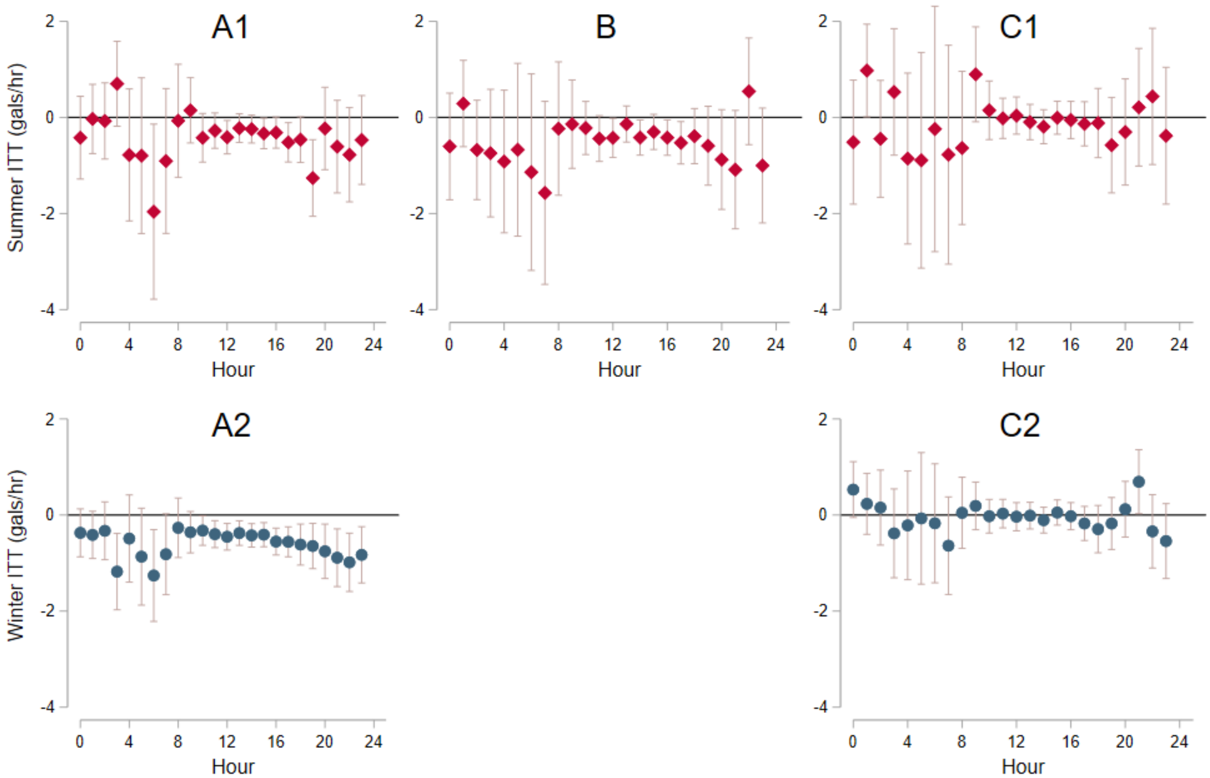
Notes: Figure 2 graphs average hourly water use by month, broken up by control and treatment households. The shaded area is the treatment period.

Figure 3: Intent-to-Treat Effects over Time



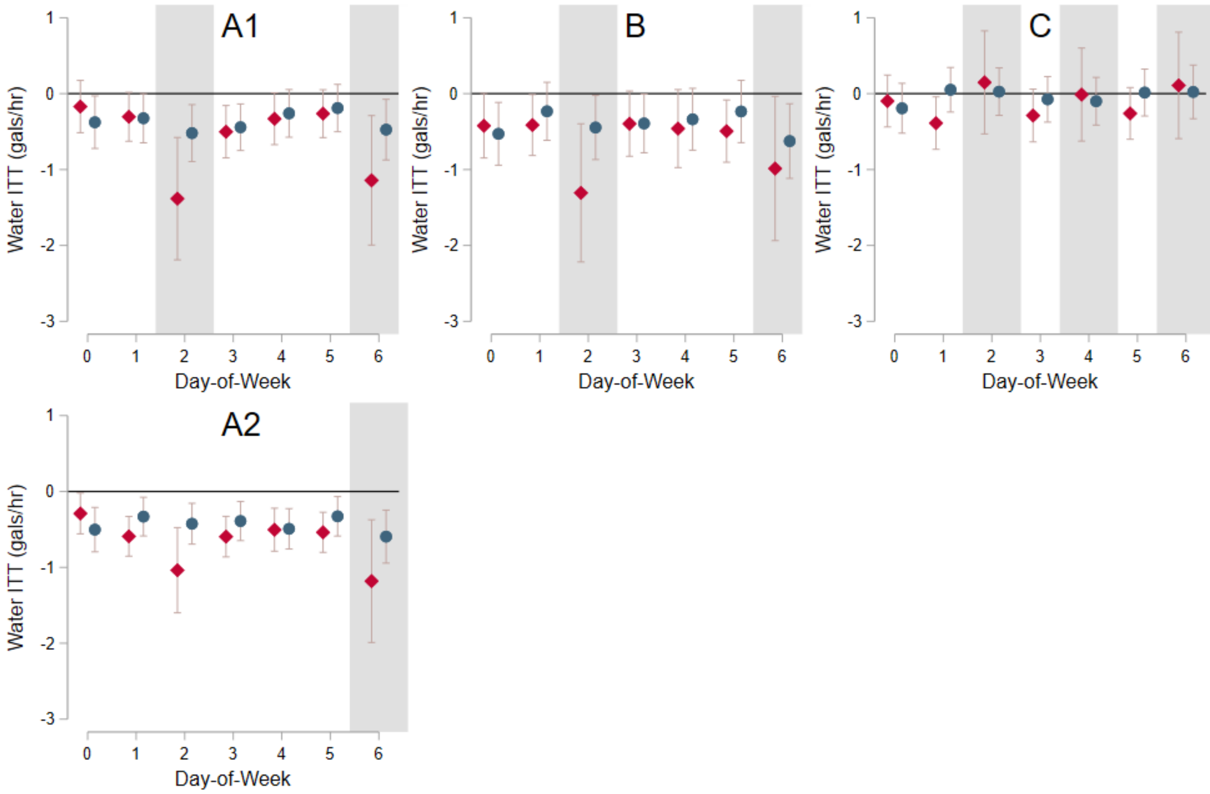
Notes: Figure 3 graphs monthly intent-to-treat effects and 95% confidence intervals over time for the joint treatment indicator (WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus). The shaded area A corresponds to the treatment period, area B corresponds to the 'backslide' period, and area C corresponds to the 're-convergence' period. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure 4: Intent-to-Treat Effects by Hour-of-Day



Notes: Figure 4 graphs intent-to-treat effects and 95% confidence intervals over each hour-of-day for the joint treatment indicator (WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus). The top row presents results during summer months (May to October) in red diamonds and the bottom row presents results during winter months (November to April) in blue circles. A is the treatment period, B is the 'backslide' period, and C is the 're-convergence' period. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure 5: Intent-to-Treat Effects by Day-of-Week



Notes: Figure 5 graphs intent-to-treat effects and 95% confidence intervals over each day-of-week for the joint treatment indicator (WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus). A1 is the first half of the treatment period (June 2015 to October 2015), A2 is the second half of the treatment period (November 2015 to March 2016), B is the 'backslide' period (May 2016 to August 2016), and C is the 're-convergence' period (September 2016 to December 2017). Shaded bars indicate days when BWP permitted outdoor watering. Blue circles and red diamonds denote treatment effects for non-watering and watering hours, respectively. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Online Appendix for
Residential Water Conservation During Drought: Experimental Evidence from Three Behavioral Interventions

Katrina Jessoe, Gabriel E. Lade, Frank Loge, and Edward Spang

April 2, 2021

A Experiment Details

Timeline of the Intervention. Figure A.1 provides a timeline of our data coverage, treatment, and outdoor watering restriction regimes. Our hourly water use coverage span April 1, 2014, to December 31, 2017. We define the baseline water use period as April 1, 2014 to February 28, 2015. We use these data to construct three baseline measures of water use: April 2014 (April 2014), Summer 2014 (May 2014 to October 2014), and Winter 2015 (November 2014 to February 2015). Each variable is household specific, and measured as average hourly water use. Welcome letters arrived in households' mailboxes the last week of March 2015. Between early May and early June, Water Smart sent the first HWR. Reports were sent on a staggered schedule, and based upon an account's billing cycle. Figure A.1 shows the start date tor the first HWR and each subsequent mailer.

Outdoor Watering Restrictions. The bottom of Figure A.1 describes the timing and days of BWP's various outdoor watering restrictions over this period. BWP had no restrictions before July 2014. The outdoor watering schedule changed with seasons and the severity of the drought. In the earliest and latest dates, households were allowed to water Tuesdays, Thursdays, and Saturdays. The stringency of the watering restrictions and the enforcement of these restrictions intensified during our treatment period. During the treatment year, the primary restrictions only allowed outdoor watering from 6 PM to 9 AM on Tuesdays and Thursdays in the summer. Winter restrictions limited outdoor watering to Saturdays only. Figure A.6 provides an example of the messaging on HWRs that households received about these restrictions. It highlights that households could only water on specific days, for 15 minutes per station, after 6 PM and before 9 AM, and could not water for two days after it rained.

Home Water Reports. Figures A.2 to A.6 provide examples of the HWRs households received through the treatment year. Figure A.2 shows an example of a conventional HWR, our first treatment arm. The reports consist of a household comparison and injunctive norm, water saving action recommendations, and utility and program messaging. Figures A.3 and A.4 show the Hot WaterSmart HWRs. The only difference were the water saving messages, which were tailored towards indoor water use activities and included energy savings for each recommended action. Figures A.5 and A.6 show Hot WaterSmart Plus HWRs. Figure A.5 shows the initial notification for the 'Conserve and Win' program and Figure A.6 shows an example of a progress update we provided to all households in the treatment.

Figure A.1: Experiment Timeline - Key Dates and Watering Restrictions

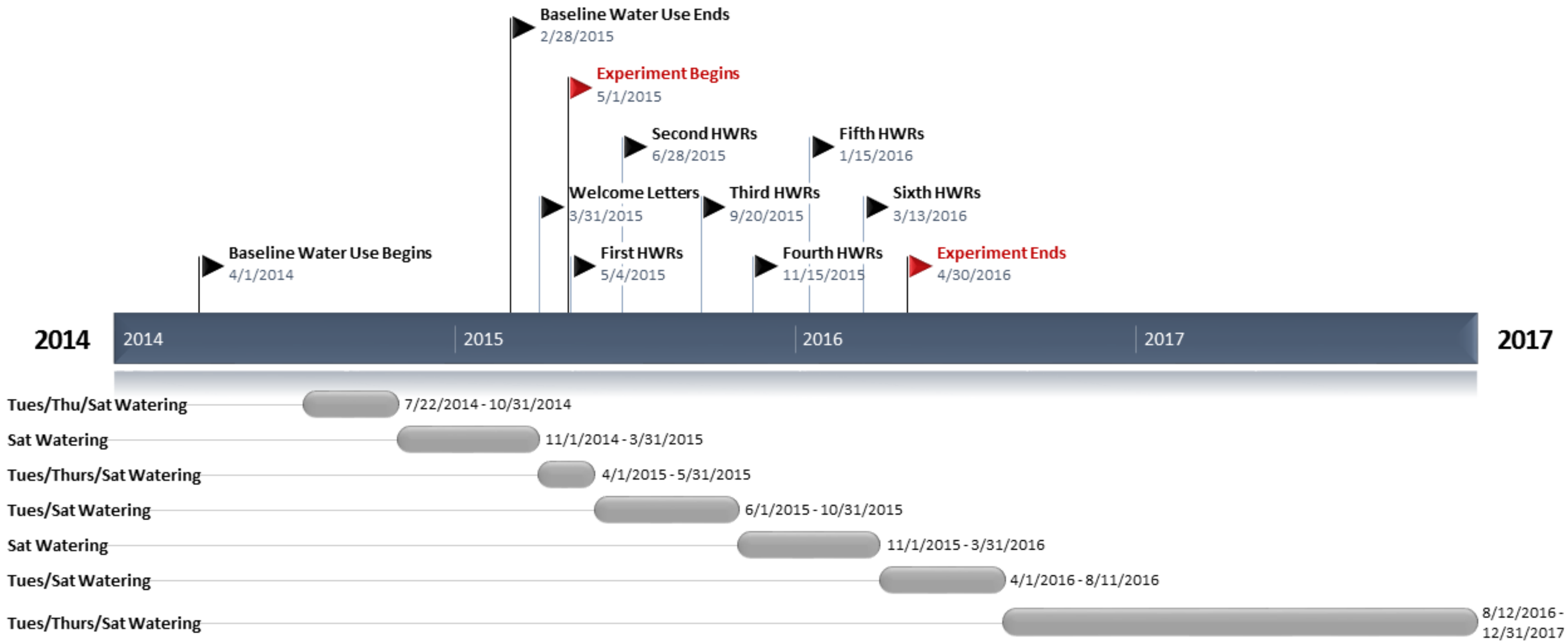


Figure A.2: WaterSmart Home Water Report

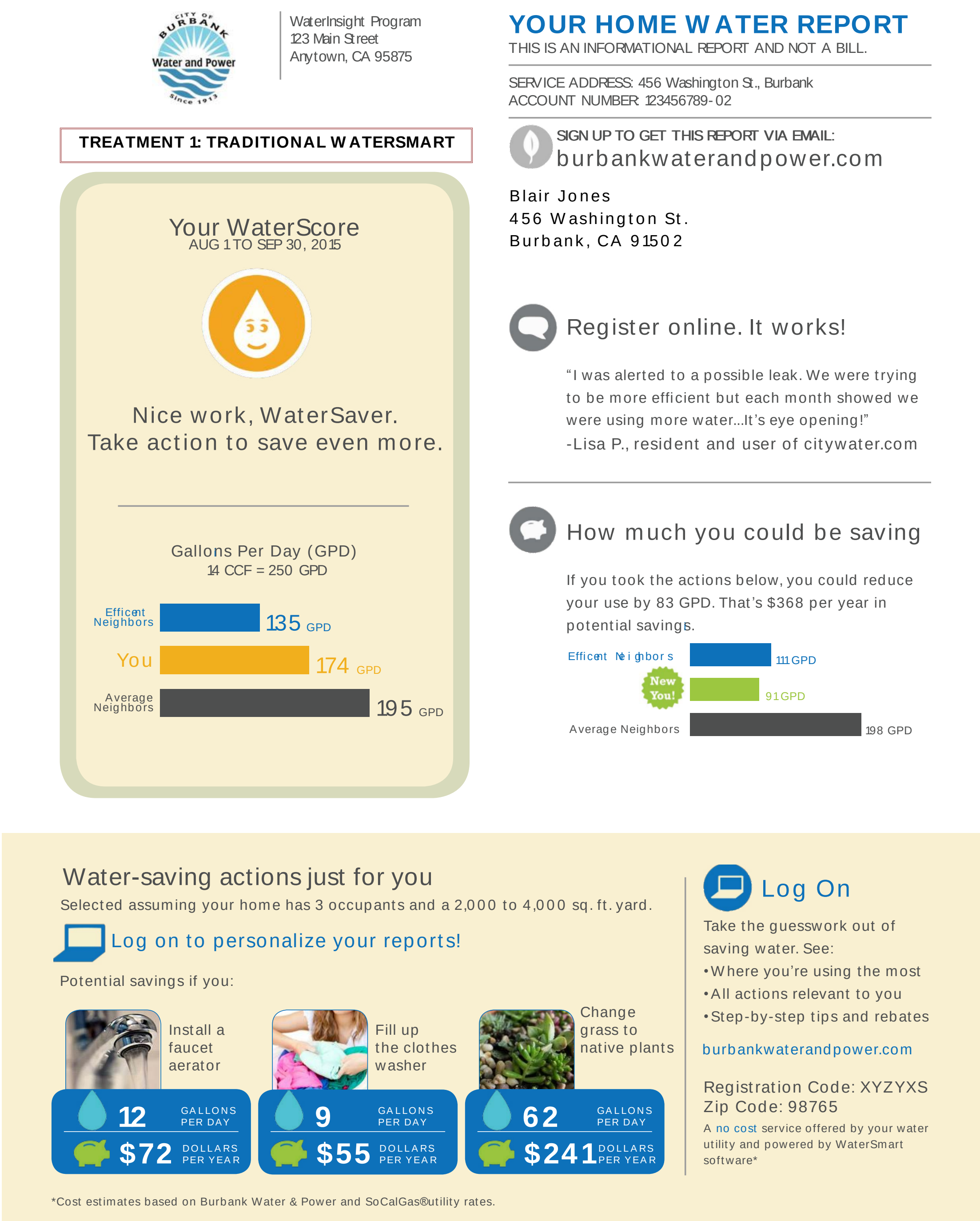
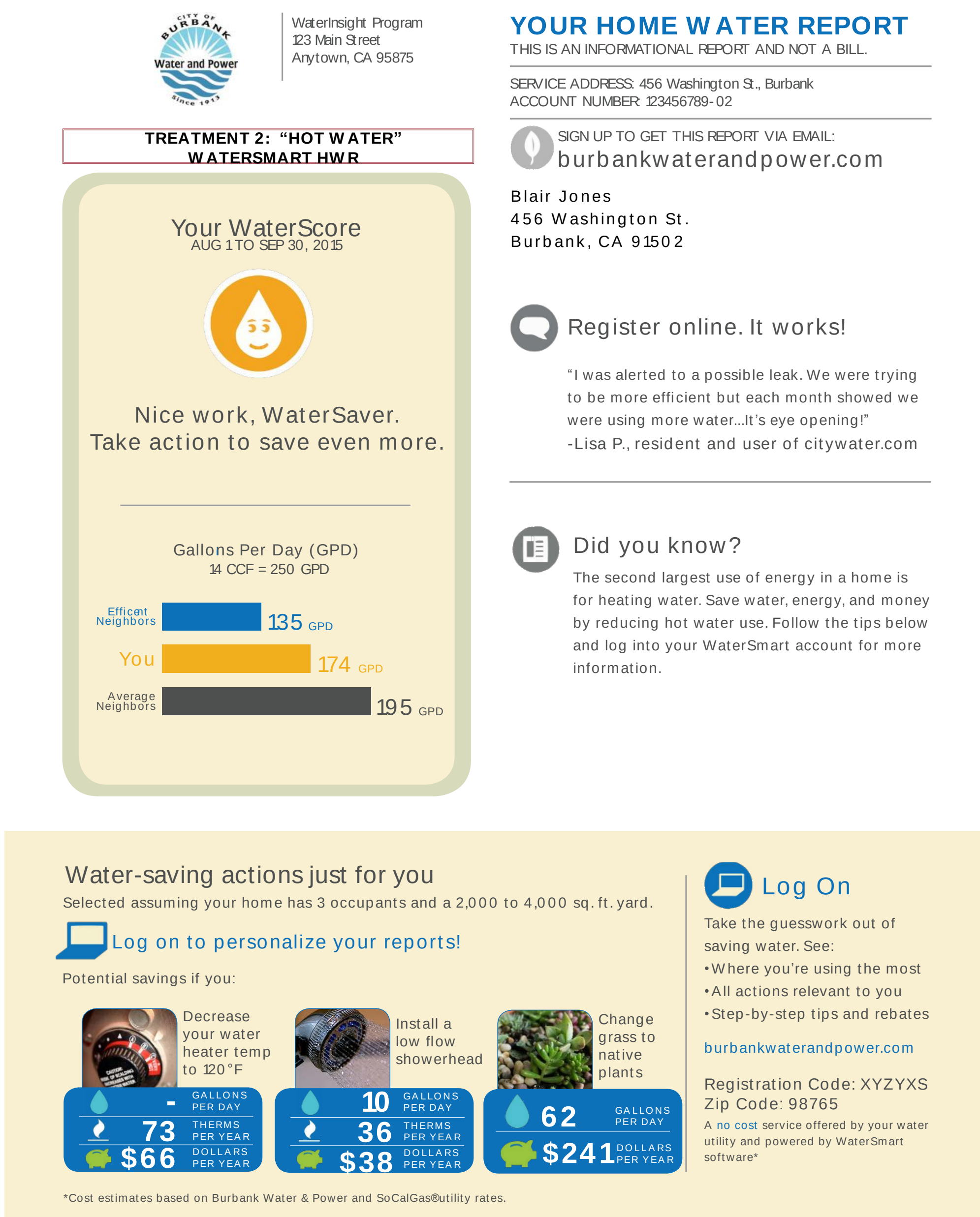


Figure A.3: Hot WaterSmart Home Water Report



WaterInsight Program
123 Main Street
Anytown, CA 95875

TREATMENT 2: "HOT WATER" WATERSMART HWR

Your WaterScore
AUG 1 TO SEP 30, 2015

Nice work, WaterSaver.
Take action to save even more.

Gallons Per Day (GPD)
14 CCF = 250 GPD

Efficient Neighbors	135 GPD
You	174 GPD
Average Neighbors	195 GPD

YOUR HOME WATER REPORT

THIS IS AN INFORMATIONAL REPORT AND NOT A BILL.

SERVICE ADDRESS: 456 Washington St., Burbank
ACCOUNT NUMBER: 123456789-02

SIGN UP TO GET THIS REPORT VIA EMAIL:
burbankwaterandpower.com

Blair Jones
456 Washington St.
Burbank, CA 91502

Register online. It works!

"I was alerted to a possible leak. We were trying to be more efficient but each month showed we were using more water...It's eye opening!"
-Lisa P., resident and user of citywater.com

Did you know?

The second largest use of energy in a home is for heating water. Save water, energy, and money by reducing hot water use. Follow the tips below and log into your WaterSmart account for more information.

Water-saving actions just for you

Selected assuming your home has 3 occupants and a 2,000 to 4,000 sq. ft. yard.

Log on to personalize your reports!

Potential savings if you:

	Gallons per day	Therms per year	Dollars per year
Decrease your water heater temp to 120°F	-	73	$66
Install a low flow showerhead	10	36	$38
Change grass to native plants	62		$241

*Cost estimates based on Burbank Water & Power and SoCalGas®utility rates.

Log On

Take the guesswork out of saving water. See:
- Where you're using the most
- All actions relevant to you
- Step-by-step tips and rebates

burbankwaterandpower.com

Registration Code: XYZYXS
Zip Code: 98765

A no cost service offered by your water utility and powered by WaterSmart software*

Figure A.4: Hot WaterSmart Recommendations

Burbank Water and Power Since 1913

WaterInsight Program
123 Main Street
Anytown, CA 95875

YOUR HOME WATER REPORT
THIS IS AN INFORMATIONAL REPORT AND NOT A BILL.

SERVICE ADDRESS: 456 Washington St., Burbank
ACCOUNT NUMBER: 123456789-02

SIGN UP TO GET THIS REPORT VIA EMAIL:
burbankwaterandpower.com

Blair Jones
456 Washington St.
Burbank, CA 91502

RECOMMENDATIONS

Water-saving actions just for you
Selected assuming your home has 4 occupants.

Log on to correct us!

Potential savings if you:

	Recommendation	Gallons per day	Therms per year	Dollars per year
R1	Decrease your water heater temp to 120°F or lower.	-	73	$66
R2	Install a low flow showerhead	10	36	$38
R3	Install low-flow faucet aerators	24	59	$66
R4	Take a 5 minute shower	19	68	$61
R5	Replace dishwasher with an ENERGY STAR certified model	9	46	$41
R6	Decrease pool temp by 2 degrees	-	675	$607
R7	Replace gas tank water heater with tankless gas water heater	-	115	$103
R8	Replace clothes washer with an ENERGY STAR certified model	11	15	$14
R9	Fill up clothes washer	6	7	$10

Figure A.5: Hot WaterSmart Plus Home Water Report (Report 1)

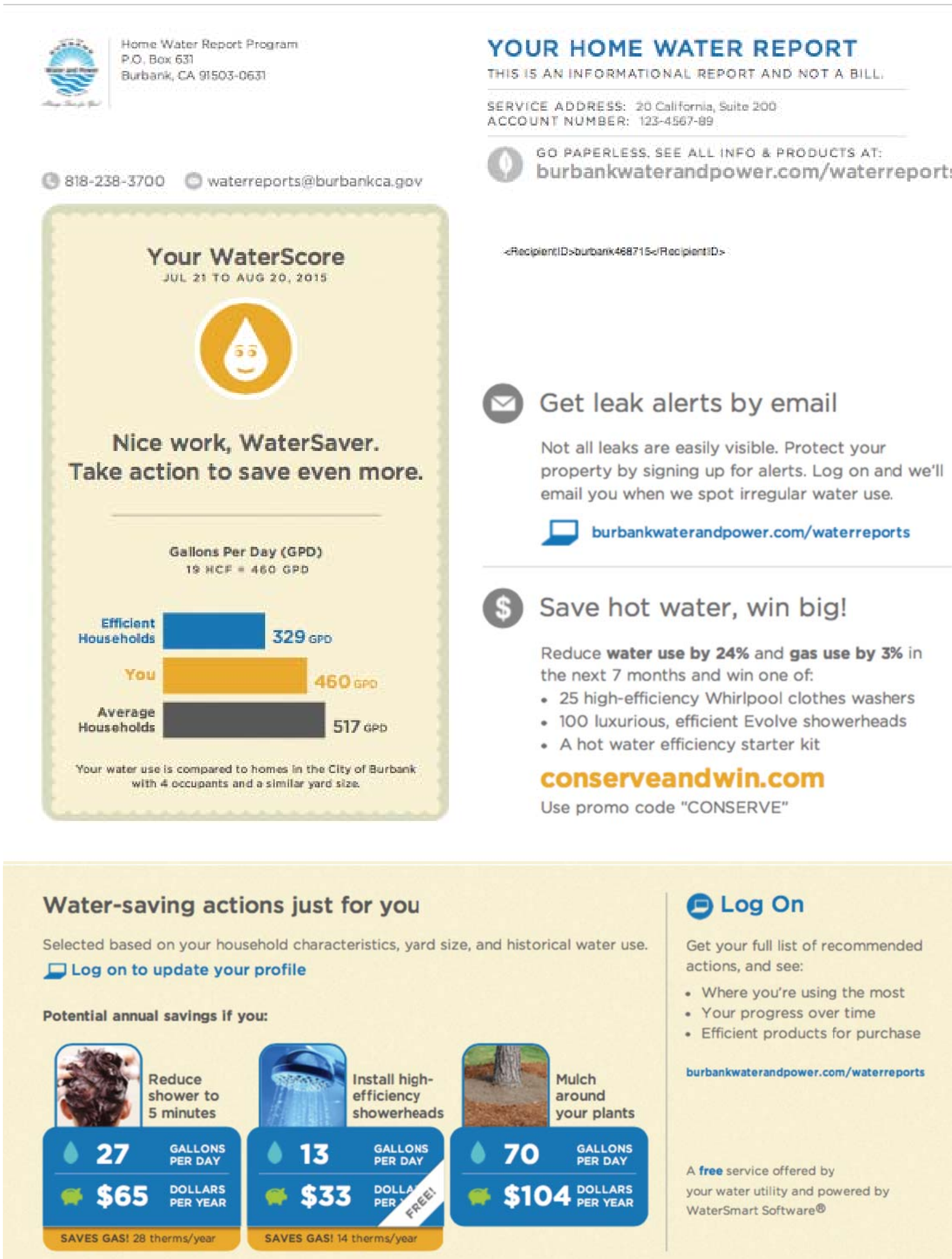


Figure A.6: Hot WaterSmart Plus Home Water Report (Report 5)

B Additional Results

B.1 Hour-of-Day and Day-of-Week Tables

Tables B.1 and B.2 present the estimated treatment effect coefficients graphed in Figures 4 and 5, respectively. Column headers correspond to each panel in the respective figures.

B.2 Treatment Comparisons

We now compare treatment effects for our hour-of-day and day-of-week regressions across the three treatment arms. The results here further justify our focus on a joint treatment indicator that we use in Section 5 of the paper.

Hour-of-Day Treatment Effects. Figure B.1 graphs point estimates for hour-of-day treatment effects broken out by treatment group. We present results for summer (top panel) and winter (bottom panel) months over the treatment year. The corresponding joint treatment effects in the main text are Panels A1 and A2 in Figure 4. The treatment effects profiles are nearly identical to our main result. The effects are not statistically different across the three groups. The one exception may be the impact of Hot WaterSmart, where we estimate slightly smaller treatment effects in early morning hours (4 AM to 8 AM) and larger treatment effects in evening hours (8 PM to 11 PM) relative to other treatment groups, particularly in the summertime.

Day-of-Week Treatment Effects. Figure B.2 graphs point estimates for day-of-week treatment effects by treatment group for the treatment year.[19] We present results for summer (A1) and winter (A2) months, and further break out treatment effects by hours of the day where watering is allowed (left panel) and prohibited (right panel). The corresponding figures in the main text are Panels A1 and A2 in Figure 5. As in the main text, we see that the largest treatment effects occur on days of the week and hours of the day when outdoor watering is allowed. Treatment effects are nearly identical, and not statistically significant, across any specification.

[19] We do not present corresponding results for the backslide and re-convergence period to save space. Results are very similar, showing that treatment effects converge back to zero across all treatment groups.

B.3 All Watering Regimes

BWP had four different outdoor watering regimes in place during the treatment year and two during the backslide period (Figure A.1). BWP allowed outdoor watering on Tuesday, Thursdays, and Saturdays from May 1, 2015 to May 31, 2015. From June 1, 2015 to October 31, 2015, households could water their lawns on Tuesdays and Saturdays. From November 1, 2015 to March 31, 2016, households could water only on Saturdays. From April 1, 2016 to August 11, 2016, Tuesday and Saturday watering resumed. BWP adopted a permanent Tuesday, Thursday, Saturday watering schedule on August 12, 2016. In Section 5.3, we restrict our attention to the Tuesday/Saturday and Saturday regimes during the treatment period, and the Tuesdays/Saturday regime during the backslide period since they covered most of the sample.

Figures B.3 and B.4 compare day-of-week treatment effects across all watering regimes. As in Figure 5, we show treatment effects for hours of the day when outdoor watering was permitted and those when it was not. We see no discernible difference in households' water use during the first treatment regime (5/1/15 to 5/31/15, Panel A1). The result is unsurprising since we find no treatment effect in May 2015 since households were beginning to receive their first HWRs (Figure 3). The largest treatment effects are on Tuesdays and Thursdays during the remaining treatment regimes (Panels A2 to A4). Even when wintertime restrictions allowed watering on only Saturday (11/1/2015 to 3/31/2016, Panel A3), we continue to find a large treatment effect (≈ 1.5 gals/hour) during morning and evening hours on Tuesdays. The results suggest that habits developed in the first half of the treatment period persisted through the winter months. The treatment effect increased on Tuesday watering hours when BWP returned to a Tuesday/Saturday schedule (April 1, 2016 to April 30, 2016, Panel A4), suggesting treatment households watered their lawns less than control households when they were permitted to water again.

As in Figure 5, we continue to see the largest treatment effects during watering hours on Tuesdays and Thursdays in the backslide period (5/1/2016 to 8/10/2016, Panel B1). The treatment effects are no longer detectable by the time that BWP instituted its Tuesday, Thursday, Saturday schedule mid-August 2016 (Panel B2).

B.4 Log Results

Short- and Long-Run Impacts Figure B.5 reproduces Figure 3 using the log of household water use as our dependent variable.[20] Two differences are apparent. First, the ITT impacts increase over the treatment year, while they are relatively constant after the second treatment month in Figure 3. Second, treatment effects persist 10 months (as opposed to four months) post-treatment. These differences highlight a key distinction between the specifications. Log impacts depend on households' response to treatment and the level of control households' water use. A 0.5 gal/hr treatment effect is larger in percentage terms in the wintertime because water use is lower in these months.

Patterns of Water Use: Hour-of-Day Figure B.6 presents hour-of-day results using log water use as the dependent variable. Focusing on the treatment period, Column A, we find a flatter response to treatment reflecting changing patterns of water use across hours of the day. The largest treatment effect, around a 4% reduction, is still observed at 7 AM; however, middle of the day impacts are larger, around 3%, due to the low baseline water use in those hours. As before, we see that the overall patterns of the treatment effect remain in the 'backslide' period (column B), and dissipate in the 're-convergence' period (column C).

Patterns of Water Use: Day-of-Week Figure B.7 presents our day-of-week results using log water use as our dependent variable. Again, we see a flatter treatment response in watering and non-watering hours of the day, reflecting changing baseline water use. We also see a flatter response across days of the week, hiding heterogeneity since average water use increases on Tuesday and Thursdays. We see similar features in panel A2. The treatment effect remains significant and nearly identical in the summer after treatment, panel B. The treatment effect is detectable during watering hours (evening to early morning) for some days of the week even later after treatment ended, panel C, consistent with our findings in Figure B.5.

[20] We transform households hourly water use y as $\log(y+1)$ because there are many observations when households water consumption is zero. Results are similar if we use other transformations like the inverse hyperbolic sine.

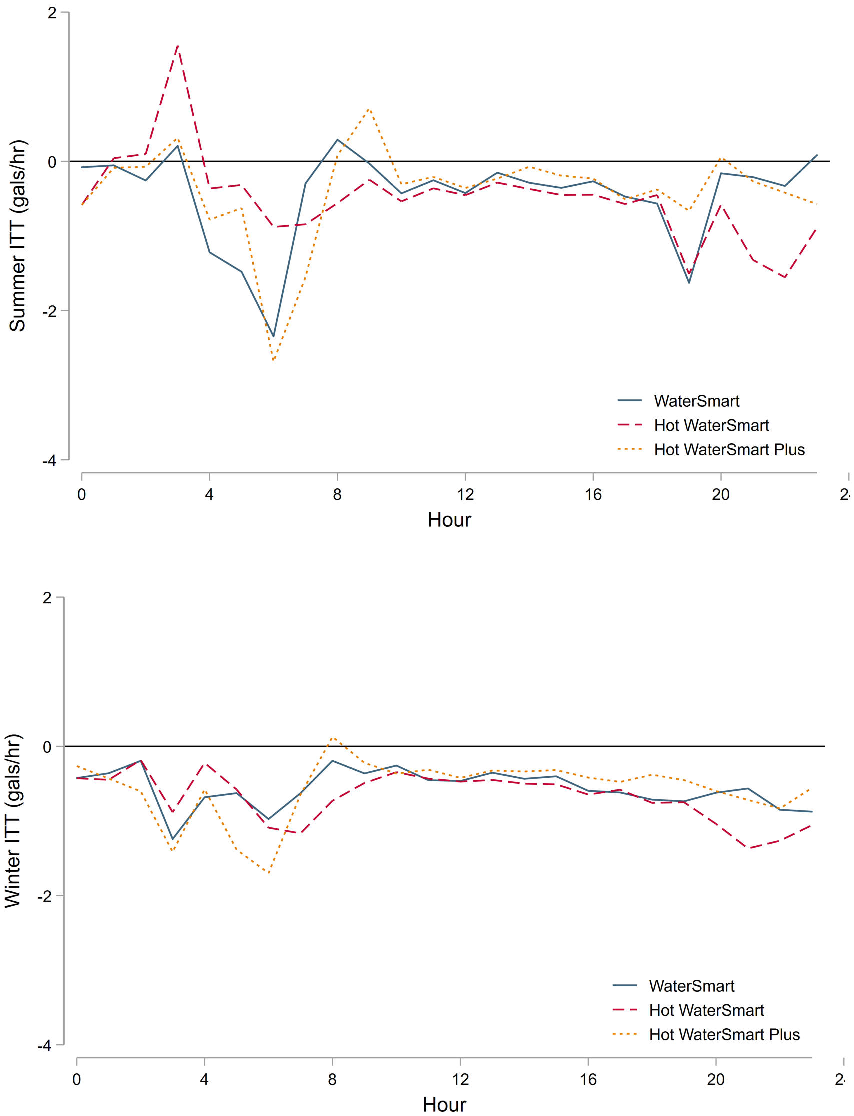
Table B.1: Intent-to-Treat Effects by Hour-of-Day

	A1	A2	B	C1	C2
	(1)	(2)	(3)	(4)	(5)
Hour 1	-0.422	-0.371	-0.604	-0.515	0.529*
	(0.439)	(0.254)	(0.566)	(0.659)	(0.297)
Hour 2	-0.032	-0.414*	0.289	0.975**	0.232
	(0.367)	(0.251)	(0.458)	(0.492)	(0.324)
Hour 3	-0.074	-0.329	-0.676	-0.446	0.154
	(0.405)	(0.307)	(0.528)	(0.621)	(0.398)
Hour 4	0.701	-1.178***	-0.743	0.530	-0.382
	(0.451)	(0.405)	(0.677)	(0.672)	(0.472)
Hour 5	-0.780	-0.490	-0.917	-0.856	-0.216
	(0.700)	(0.463)	(0.757)	(0.908)	(0.577)
Hour 6	-0.796	-0.867*	-0.673	-0.892	-0.071
	(0.827)	(0.515)	(0.917)	(1.145)	(0.698)
Hour 7	-1.962**	-1.259***	-1.139	-0.240	-0.172
	(0.929)	(0.488)	(1.043)	(1.303)	(0.632)
Hour 8	-0.907	-0.817*	-1.570	-0.776	-0.639
	(0.769)	(0.429)	(0.971)	(1.162)	(0.518)
Hour 9	-0.070	-0.267	-0.233	-0.635	0.045
	(0.600)	(0.316)	(0.708)	(0.813)	(0.377)
Hour 10	0.148	-0.358	-0.140	0.897*	0.189
	(0.347)	(0.219)	(0.469)	(0.504)	(0.253)
Hour 11	-0.424	-0.324**	-0.220	0.149	-0.026
	(0.258)	(0.157)	(0.282)	(0.311)	(0.179)
Hour 12	-0.275	-0.399***	-0.440*	-0.019	0.029
	(0.188)	(0.143)	(0.243)	(0.214)	(0.152)
Hour 13	-0.412**	-0.453***	-0.424**	0.038	-0.036
	(0.177)	(0.141)	(0.208)	(0.198)	(0.149)
Hour 14	-0.224	-0.376***	-0.137	-0.096	-0.013
	(0.152)	(0.131)	(0.192)	(0.188)	(0.142)
Hour 15	-0.243	-0.424***	-0.419**	-0.193	-0.107
	(0.148)	(0.125)	(0.187)	(0.181)	(0.137)
Hour 16	-0.332**	-0.410***	-0.302	-0.006	0.051
	(0.163)	(0.128)	(0.187)	(0.176)	(0.134)
Hour 17	-0.316*	-0.553***	-0.419**	-0.053	-0.025
	(0.165)	(0.143)	(0.188)	(0.197)	(0.144)
Hour 18	-0.519**	-0.558***	-0.528**	-0.133	-0.176
	(0.209)	(0.159)	(0.225)	(0.235)	(0.181)
Hour 19	-0.463*	-0.616***	-0.389	-0.117	-0.294
	(0.243)	(0.218)	(0.293)	(0.368)	(0.251)
Hour 20	-1.261***	-0.645***	-0.590	-0.577	-0.177
	(0.407)	(0.241)	(0.419)	(0.507)	(0.276)
Hour 21	-0.230	-0.755***	-0.876*	-0.303	0.121
	(0.437)	(0.289)	(0.530)	(0.564)	(0.296)
Hour 22	-0.608	-0.889***	-1.086*	0.212	0.692**
	(0.492)	(0.306)	(0.629)	(0.624)	(0.340)
Hour 23	-0.776	-0.985***	0.543	0.437	-0.341
	(0.500)	(0.310)	(0.566)	(0.723)	(0.389)
Hour 24	-0.470	-0.829***	-1.000	-0.381	-0.542
	(0.472)	(0.299)	(0.610)	(0.727)	(0.398)
Observations	68,894,068	68,038,145	46,071,940	85,357,023	84,264,247

Table B.2: Intent-to-Treat Effects by Day-of-Week

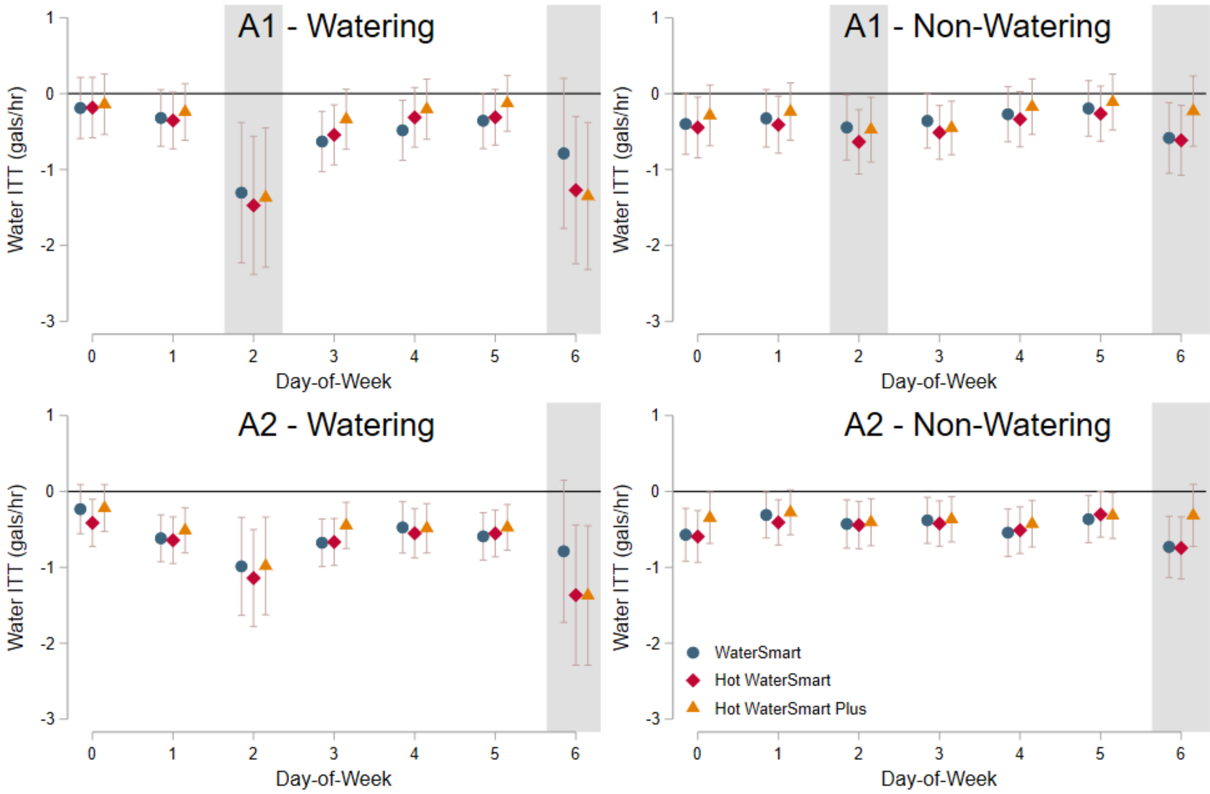
	A1	A2	B	C
	(1)	(2)	(3)	(4)
Day 1 (Water Hours)	-0.170	-0.290**	-0.422*	-0.097
	(0.176)	(0.137)	(0.217)	(0.174)
Day 2 (Water Hours)	-0.305*	-0.591***	-0.413**	-0.388**
	(0.165)	(0.134)	(0.205)	(0.177)
Day 3 (Water Hours)	-1.383***	-1.037***	-1.307***	0.149
	(0.411)	(0.287)	(0.465)	(0.346)
Day 4 (Water Hours)	-0.502***	-0.595***	-0.396*	-0.288
	(0.176)	(0.136)	(0.219)	(0.177)
Day 5 (Water Hours)	-0.331*	-0.504***	-0.460*	-0.012
	(0.174)	(0.145)	(0.262)	(0.313)
Day 6 (Water Hours)	-0.264	-0.539***	-0.493**	-0.261
	(0.161)	(0.135)	(0.209)	(0.174)
Day 7 (Water Hours)	-1.142***	-1.181***	-0.986**	0.109
	(0.435)	(0.412)	(0.485)	(0.358)
Day 1 (Non-Water Hours)	-0.377**	-0.503***	-0.529**	-0.193
	(0.177)	(0.149)	(0.211)	(0.168)
Day 2 (Non-Water Hours)	-0.324*	-0.331**	-0.232	0.051
	(0.165)	(0.130)	(0.195)	(0.150)
Day 3 (Non-Water Hours)	-0.520***	-0.425***	-0.446**	0.026
	(0.191)	(0.137)	(0.215)	(0.160)
Day 4 (Non-Water Hours)	-0.442***	-0.390***	-0.393**	-0.075
	(0.157)	(0.131)	(0.197)	(0.153)
Day 5 (Non-Water Hours)	-0.260	-0.492***	-0.338	-0.100
	(0.160)	(0.136)	(0.208)	(0.160)
Day 6 (Non-Water Hours)	-0.190	-0.327**	-0.235	0.014
	(0.159)	(0.133)	(0.209)	(0.158)
Day 7 (Non-Water Hours)	-0.474**	-0.595***	-0.625**	0.023
	(0.204)	(0.178)	(0.251)	(0.180)
Observations	56,936,571	56,852,066	38,612,709	169,621,270

Figure B.1: Hour-of-Day Results by Treatment Group



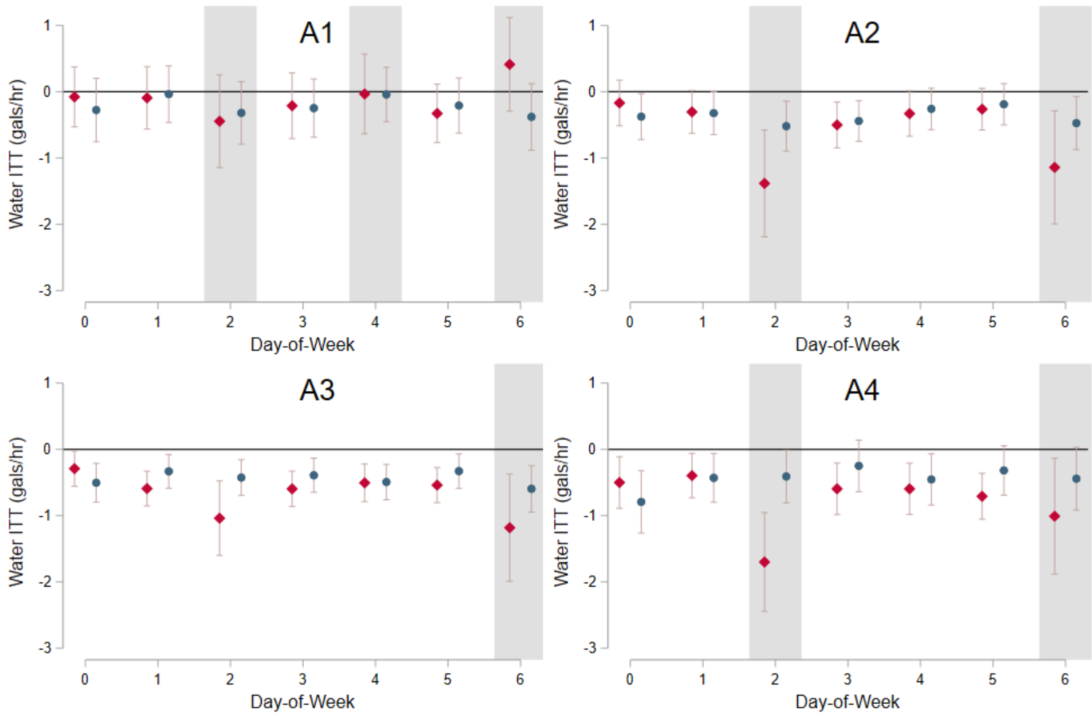
Notes: The figure graphs point estimates for hour-of-day treatment effects by treatment group for the summertime (the top figure) and the wintertime (the bottom figure). The corresponding figures in the main text are Panels A1 and A2 in Figure 4. For simplicity, no 95% confidence intervals are presented. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure B.2: Day-of-Week Results by Treatment Group



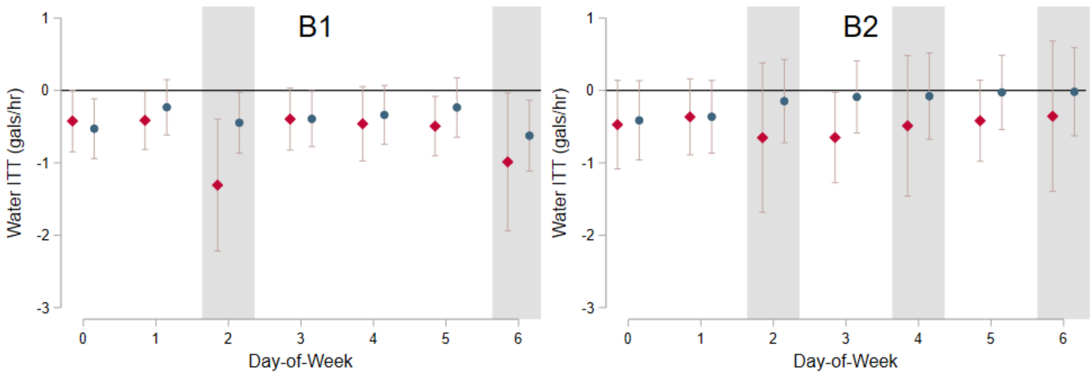
Notes: The figure graphs point estimates for day-of-week treatment effects by treatment group for the summertime (A1) and the wintertime (A2). We further break the treatment effects out by hours of the day where watering is allowed (left panel) and prohibited (right panel). The corresponding figures in the main text are Panels A1 and A2 in Figure 5. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure B.3: Day-of-Week Results: All Watering Regimes (Treatment)



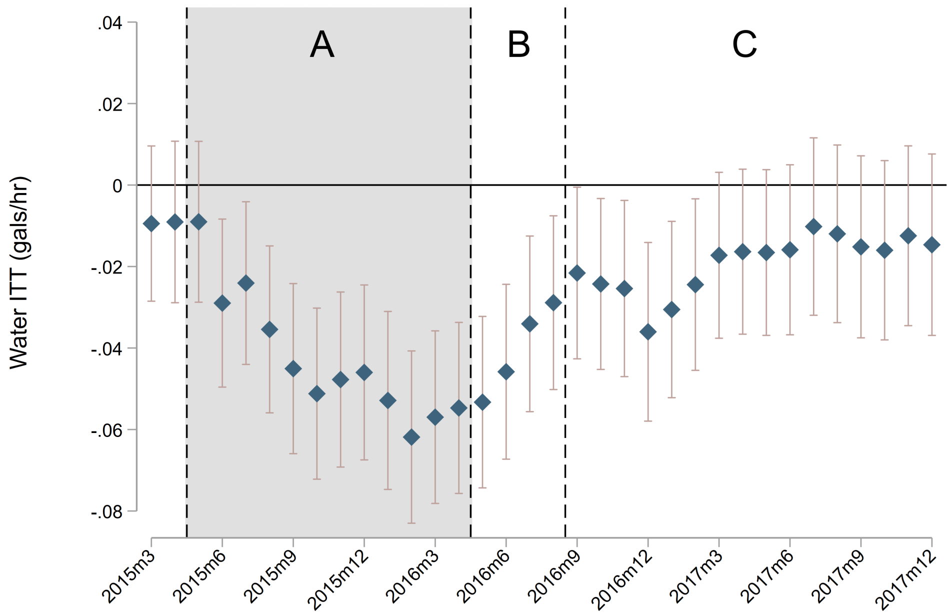
Notes: The figure graphs intent-to-treat effects and 95% confidence intervals for each day-of-week for the joint treatment indicator during the treatment year. A1 is the first outdoor watering regime (May 2015), A2 is the second watering regime (June 2015 to October 2015), A3 is the third watering regime (November 2015 to March 2016), and A4 is the fourth watering regime (April 2016). Shaded bars indicate days when BWP permitted outdoor watering. Blue circles and red diamonds denote treatment effects for non-watering and watering hours, respectively. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure B.4: Day-of-Week Results: All Watering Regimes (Backslide)



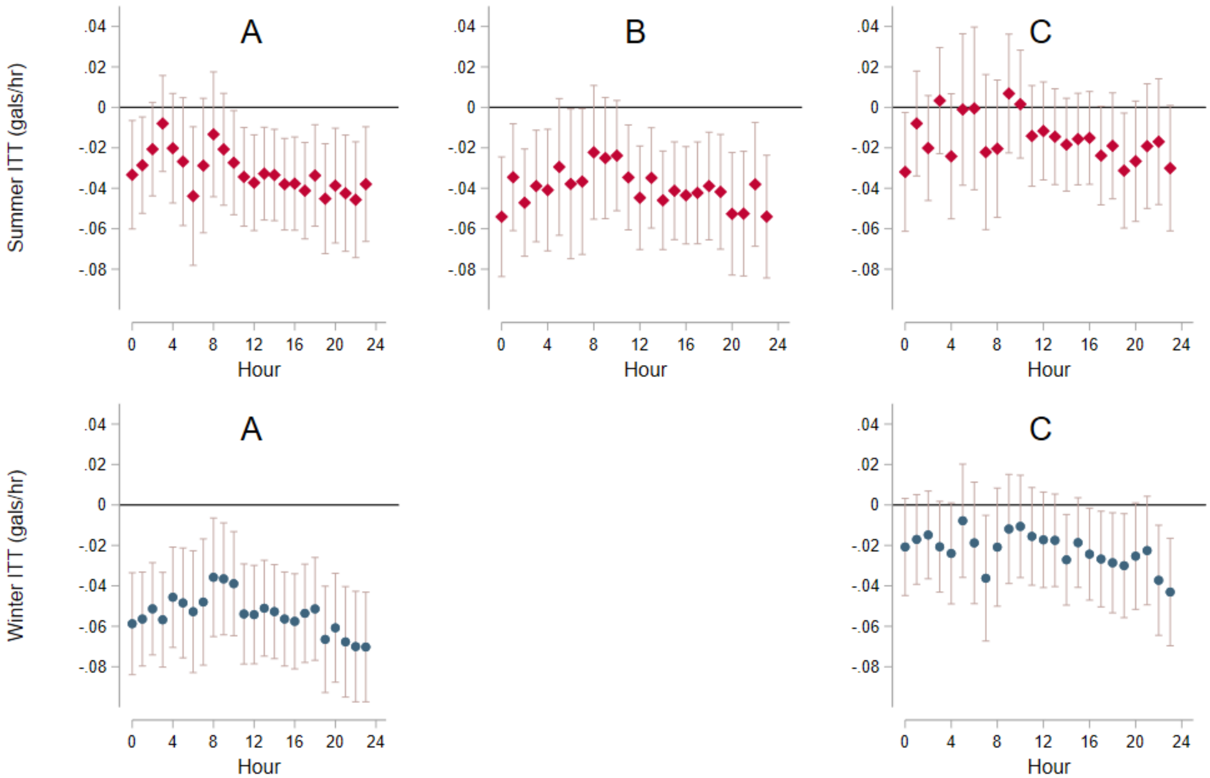
Notes: The figure graphs intent-to-treat effects and 95% confidence intervals for each day-of-week for the joint treatment indicator during the backslide period. B1 is the first outdoor watering regime (May 2016 to August 10, 2016), and B2 is the second watering regime (August 11 2016 to August 31, 2016). Shaded bars indicate days when BWP permitted outdoor watering. Blue circles and red diamonds denote treatment effects for non-watering and watering hours, respectively. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure B.5: Intent-to-Treat Effects over Time
(Log Water Use)



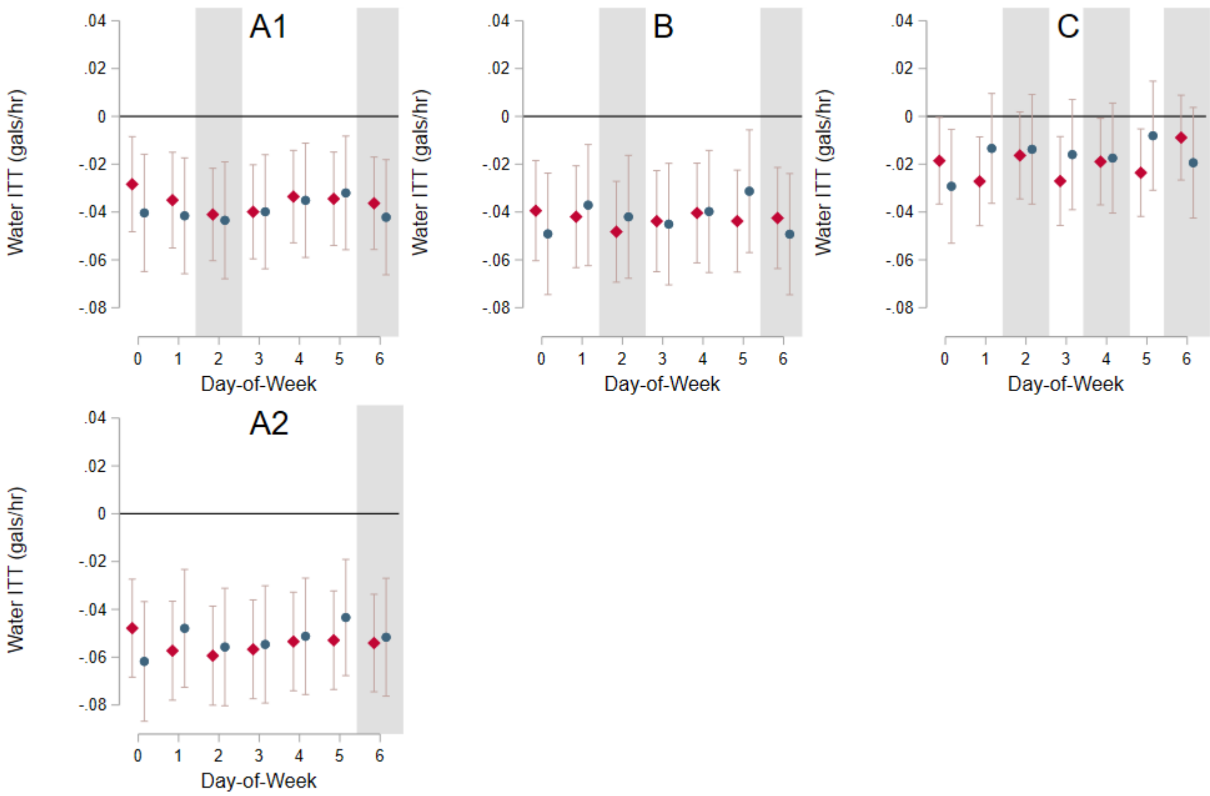
Notes: The figure graphs monthly intent-to-treat effects and 95% confidence intervals over time for the joint treatment indicator (WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus). The shaded area A corresponds to the treatment period, area B corresponds to the 'backslide' period, and area C corresponds to the 're-convergence' period. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure B.6: Intent-to-Treat Effects by Hour-of-Day
(Log Water Use)



Notes: The figure graphs intent-to-treat effects and 95% confidence intervals over each hour-of-day for the joint treatment indicator (WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus). The shaded area A corresponds to the treatment period, area B corresponds to the 'backslide' period, and area C corresponds to the 're-convergence' period. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

Figure B.7: Intent-to-Treat Effects by Day-of-Week
(Log Water Use)



Notes: The figure graphs intent-to-treat effects and 95% confidence intervals over each day-of-week for the joint treatment indicator (WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus). The shaded area A corresponds to the treatment period, area B corresponds to the 'backslide' period, and area C corresponds to the 're-convergence' period. Treatment effects are presented separately for watering hours (red diamonds) and non-watering hours (blue circles), where the shaded days are days where outdoor watering restrictions bind. All regressions include baseline household water use, local weather controls, as well as date and hour-of-day fixed effects.

C Multiple Hypothesis Testing

Our study considers the impact of home water reports on many outcomes of interest and for multiple treatments. This leads to natural concerns related to over-rejecting null hypotheses using standard inference procedures. Here, we account for multiple hypothesis testing in a few key ways to explore the sensitivity of our conclusions regarding the impact of the three treatments as well as the impacts of HWRs on water usage patterns and rebate uptake. A key challenge in implementing multiple hypothesis testing (MHT) corrections in our setting are the number of observations. Many procedures that adjust p-values to account for MHT involve bootstrapping the data. This is computationally challenging since our regressions involve estimating models with up to hundreds of millions of observations. Further, some empirical models estimate the impacts of three treatments, while others estimate patterns of behavior using a joint treatment indicator.

We explore the importance of adjusting our inference to account for MHT using the procedure developed in Romano and Wolf (2005a,b, 2016) with code developed by Clarke et al. (2020). The procedure uses re-sampling, bootstrap-based methods that produce adjusted p-values that control for the family-wise error rate (FWER) in a more efficient manner than classical corrections proposed by Bonferroni (1935) and Holm (1979).

We start by considering our main hypotheses. In Section 4, we test whether each of the three interventions affected overall water use and water usage in the first versus last six months of the intervention. Further, in Section 5, we test whether treatment affected participation in utility rebate programs. Given the different time frequencies of the two data sets, we take a conservative approach. We calculate the average water use for each household over the entire treatment year, as well as the average use in the first and last six months of the treatment period. We then apply the Romano-Wolf procedure to these three average, household-level water usage outcomes and households' participation in the three rebate programs considered in Table 4.

Table C.1 presents the results. Even after eliminating all the high-frequency nature of our data, we continue to find similar, and statistically significant, impacts of each treatment on each water usage outcome. We also continue to find no impact of treatment on any rebate usage. Conventional p-values, shown in parentheses, highlight the strong, statistically significant impacts of each treatment on all three water use outcomes. Romano-Wolf adjusted p-values, shown in brackets, adjust for MHT across the six outcomes and three treatments.

As expected, adjusted p-values are larger, but we continue to find statistically significant impacts of each treatment on every water usage outcome.

Tables C.2 and C.3 present similarly adjusted p-values that account for multiple hypothesis testing across the different samples of our data and for different hours of the day and days of the week, respectively. Again, given the sample size, we collapse our data to perform each exercise. Further, given that we observe multiple observations for each household, we use a cluster bootstrap procedure to account for within-household correlations in the data.

Table C.2 presents hour-of-day results. To estimate the Clarke et al. (2020) p-values, we first collapse household water use to the average monthly water use for each hour. For example, we calculate every households' average water use in July 2015 at 4 PM, as well as for every other hour-of-day and month-of-year. We then estimate the impact of the joint treatment indicator interacted with each hour-of-day for each of the samples from Panels A1, A2, B, C1, and C2 in Figure 4. Coefficient estimates are very similar to those in Table B.1, highlighting that we recover similar estimates even with the lower-frequency data. We continue to find the largest and most statistically significant treatment effects in early morning and evening hours during the treatment period, with more modest reductions in the middle of the day. The effects wane in the backslide and re-convergence periods (columns 3 through 5). Adjusted p-values are, on average, two to three times larger than conventional p-values. Nonetheless, the primary outcomes we highlight in our paper remain statistically significant during the treatment year, as seen in columns (1) and (2).

Table C.3 presents similar results for our day-of-week regressions. Each water-usage outcome, in this case, is collapsed to day-of-week by hour-of-day by month. For example, we calculate the average monthly water usage on Tuesdays at 4 PM in July 2015 for each household. Again, coefficients are very similar to our main regression outcomes in Table B.2. The largest treatment effects occur on days that households were allowed to water outdoors during the hours when outdoor irrigation was permitted. Romano-Wolf adjusted p-values are, again, typically two to three times larger than conventional p-values. Our main conclusions remain, and we continue to find treatment had statistically significant impacts during the days and hours that we highlight in our paper.

The above analysis has natural limitations. While we were conservative in not conducting MHT adjustments using our hourly treatment data, none of the results adjust for every outcome we consider in this paper. Such an adjustment would further increase our adjusted p-values.

Table C.1: Multiple Hypothesis Testing Corrections - Main Outcomes

	Water Use	Sum Water Use	Win Water Use	Any Rebate	Appliance Rebate	Turf Rebate
	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.492***	-0.485**	-0.583***	-0.003	-0.003	-0.000
	(0.0029)	(0.0105)	(0.0002)	(0.3962)	(0.3888)	(0.9168)
	[0.0220]	[0.0459]	[0.0040]	[0.6287]	[0.6287]	[0.9002]
Hot WaterSmart	-0.540***	-0.481**	-0.676***	-0.001	-0.000	-0.001
	(0.0010)	(0.0103)	(0.0000)	(0.8787)	(0.9831)	(0.6966)
	[0.0040]	[0.0339]	[0.0020]	[0.9561]	[0.9860]	[0.8962]
Hot WaterSmart Plus	-0.488***	-0.454**	-0.578***	-0.003	-0.004	0.001
	(0.0029)	(0.0154)	(0.0002)	(0.3855)	(0.2813)	(0.6622)
	[0.0120]	[0.0579]	[0.0060]	[0.6188]	[0.5170]	[0.6986]
Observations	16,818	16,766	16,330	16,818	16,818	16,818

Notes: The table presents results from an OLS regression of each respective outcome on indicators for assignment to WaterSmart, Hot WaterSmart, and Hot WaterSmart Plus. Conventional p-values are reported in parentheses, and Romano-Wolf adjusted p-values, based on 500 bootstrap samples, are reported in brackets. *, **, *** denote significance at the 10%, 5%, and 1%level based on conventional p-values.

Table C.2: Hour-of-Day Multiple Hypothesis Testing Corrections

	A1 (1)	A2 (2)	B (3)	C1 (4)	C2 (5)
Hour 1	-0.420	-0.368	-0.661	-0.491	0.522*
	(0.3396)	(0.1482)	(0.2510)	(0.4534)	(0.0784)
	[0.5469]	[0.4212]	[0.5349]	[0.5469]	[0.2894]
Hour 2	-0.060	-0.397	0.325	1.000**	0.229
	(0.8701)	(0.1143)	(0.4770)	(0.0426)	(0.4787)
	[0.8962]	[0.3892]	[0.8363]	[0.1876]	[0.8363]
Hour 3	-0.104	-0.333	-0.680	-0.436	0.158
	(0.7988)	(0.2790)	(0.1965)	(0.4824)	(0.6904)
	[0.9002]	[0.7026]	[0.6068]	[0.8184]	[0.9002]
Hour 4	0.720	-1.211***	-0.726	0.504	-0.389
	(0.1115)	(0.0031)	(0.2829)	(0.4571)	(0.4089.)
	[0.3134]	[0.0160]	[0.5509]	[0.6108]	[0.6108]
Hour 5	-0.727	-0.516	-0.925	-0.898	-0.193
	(0.2979)	(0.2658)	(0.2223)	(0.3214)	(0.7358)
	[0.6547]	[0.6547]	[0.6467]	[0.6547]	[0.7066]
Hour 6	-0.750	-0.850*	-0.649	-0.940	-0.085
	(0.3618)	(0.0977)	(0.4790)	(0.4112)	(0.9025)
	[0.7665]	[0.3313]	[0.7665]	[0.7665]	[0.9022]
Hour 7	-1.964**	-1.274***	-1.167	-0.277	-0.068
	(0.0351)	(0.0097)	(0.2631)	(0.8311)	(0.9138)
	[0.1218]	[0.0499]	[0.4890]	[0.9681]	[0.9681]
Hour 8	-0.961	-0.819*	-1.481	-0.704	-0.563
	(0.2165)	(0.0557)	(0.1277)	(0.5420)	(0.2760)
	[0.4431]	[0.2056]	[0.3333]	[0.5469]	[0.4431]
Hour 9	-0.067	-0.251	-0.200	-0.632	0.072
	(0.9118)	(0.4308)	(0.7770)	(0.4355)	(0.8492)
	[0.9840]	[0.8603]	[0.9800]	[0.8603]	[0.9840]
Hour 10	0.163	-0.378*	-0.109	0.896*	0.185
	(0.6386)	(0.0893)	(0.8156)	(0.0758)	(0.4671)
	[0.8343]	[0.2435]	[0.8343]	[0.2375]	[0.8144]
Hour 11	-0.414	-0.336**	-0.254	0.165	-0.011
	(0.1058)	(0.0331)	(0.3725)	(0.5958)	(0.9487)
	[0.3273]	[0.1297]	[0.6547]	[0.7725]	[0.9421]
Hour 12	-0.266	-0.407***	-0.442*	-0.018	0.029
	(0.1567)	(0.0047)	(0.0698)	(0.9337)	(0.8508)
	[0.3433]	[0.0120]	[0.1956]	[0.9601]	[0.9601]
Hour 13	-0.425**	-0.474***	-0.398*	0.036	-0.046
	(0.0185)	(0.0008)	(0.0552)	(0.8572)	(0.7591)
	[0.0519]	[0.0060]	[0.1178]	[0.9321]	[0.9321]
Hour 14	-0.236	-0.381***	-0.119	-0.108	-0.023
	(0.1214)	(0.0038)	(0.5361)	(0.5664)	(0.8720)
	[0.3253]	[0.0120]	[0.8563]	[0.8563]	[0.8902]
Hour 15	-0.251*	-0.458***	-0.411**	-0.249	-0.120
	(0.0893)	(0.0004)	(0.0276)	(0.1730)	(0.3828)
	[0.2056]	[0.0020]	[0.0679]	[0.2695]	[0.3912]
Hour 16	-0.355**	-0.426***	-0.296	-0.049	0.042
	(0.0310)	(0.0010)	(0.1134)	(0.7815)	(0.7555)
	[0.0878]	[0.0020]	[0.2136]	[0.9102]	[0.9102]
Hour 17	-0.331**	-0.545***	-0.417**	-0.062	-0.014
	(0.0455)	(0.0001)	(0.0272)	(0.7530)	(0.9233)
	[0.1297]	[0.0020]	[0.0798]	[0.9361]	[0.9361]
Hour 18	-0.520**	-0.551***	-0.530**	-0.146	-0.175
	(0.0128)	(0.0005)	(0.0188)	(0.5311)	(0.3311)
	[0.0439]	[0.0040]	[0.0479]	[0.5529]	[0.5409]
Hour 19	-0.487**	-0.614***	-0.388	-0.173	-0.303
	(0.0471)	(0.0047)	(0.1862)	(0.6383)	(0.2237)
	[0.1577]	[0.0220]	[0.4271]	[0.6168]	[0.4271]
Hour 20	-1.273***	-0.657***	-0.584	-0.562	-0.179
	(0.0019)	(0.0063)	(0.1640)	(0.2622)	(0.5140)
	[0.0100]	[0.0140]	[0.3453]	[0.4052]	[0.4790]
Hour 21	-0.243	-0.769***	-0.860	-0.299	0.109
	(0.5806)	(0.0081)	(0.1040)	(0.5931)	(0.7129)
	[0.9062]	[0.0559]	[0.3134]	[0.9062]	[0.9062]
Hour 22	-0.667	-0.899***	-1.012	0.116	0.688**
	(0.1801)	(0.0035)	(0.1059)	(0.8533)	(0.0433)
	[0.3054]	[0.0180]	[0.2774]	[0.8343]	[0.1497]
Hour 23	-0.764	-0.994***	0.557	0.397	-0.337
	(0.1274)	(0.0014)	(0.3239)	(0.5850)	(0.3882)
	[0.3792]	[0.0120]	[0.6088]	[0.6088]	[0.6088]
Hour 24	-0.464	-0.848***	-0.986	-0.416	-0.499
	(0.3257)	(0.0047)	(0.1050)	(0.5657)	(0.2056)
	[0.5190]	[0.0359]	[0.2914]	[0.5868]	[0.4331]
Observations	2,312,554	2,289,118	1,517,933	2,861,685	2,849,460

Notes: Conventional p-values are reported in parentheses. Romano-Wolf adjusted p-values, based on 500 clustered bootstrap samples at the household level, are reported in brackets. *, **, *** denote significance at the 10%, 5%, and 1%level based on conventional p-values.

Table C.3: Day-of-Week Multiple Hypothesis Testing Corrections

	A1 (1)	A2 (2)	B (3)	C (4)
Day 1 (Water Hours)	-0.175	-0.297**	-0.423*	-0.102
	(0.3244)	(0.0316)	(0.0574)	(0.5605)
	[0.5549]	[0.1038]	[0.1617]	[0.5988]
Day 2 (Water Hours)	-0.303*	-0.594***	-0.414**	-0.371**
	(0.0646)	(0.0000)	(0.0484)	(0.0357)
	[0.1218]	[0.0020]	[0.1218]	[0.1218]
Day 3 (Water Hours)	-1.390***	-1.063***	-1.149**	0.149
	(0.0012)	(0.0003)	(0.0145)	(0.6712)
	[0.0040]	[0.0020]	[0.0180]	[0.6966]
Day 4 (Water Hours)	-0.487***	-0.608***	-0.431*	-0.285
	(0.0054)	(0.0000)	(0.0519)	(0.1091)
	[0.0120]	[0.0020]	[0.0978]	[0.1198]
Day 5 (Water Hours)	-0.348**	-0.523***	-0.406	-0.022
	(0.0486)	(0.0003)	(0.1486)	(0.9439)
	[0.1238]	[0.0040]	[0.2395]	[0.9481]
Day 6 (Water Hours)	-0.272*	-0.546***	-0.512**	-0.240
	(0.0920)	(0.0001)	(0.0185)	(0.1674)
	[0.1856]	[0.0020]	[0.0599]	[0.1856]
Day 7 (Water Hours)	-1.125**	-1.205***	-0.863*	0.122
	(0.0114)	(0.0043)	(0.0783)	(0.7325)
	[0.0299]	[0.0160]	[0.1098]	[0.7405]
Day 1 (Non-Water Hours)	-0.396**	-0.521***	-0.509**	-0.199
	(0.0251)	(0.0005)	(0.0208)	(0.2372)
	[0.0599]	[0.0040]	[0.0599]	[0.2455]
Day 2 (Non-Water Hours)	-0.337**	-0.342***	-0.202	0.039
	(0.0409)	(0.0088)	(0.3019)	(0.7978)
	[0.0739]	[0.0160]	[0.4750]	[0.7944]
Day 3 (Non-Water Hours)	-0.488**	-0.418***	-0.462**	0.033
	(0.0111)	(0.0025)	(0.0354)	(0.8383)
	[0.0120]	[0.0060]	[0.0639]	[0.8343]
Day 4 (Non-Water Hours)	-0.453***	-0.392***	-0.362*	-0.081
	(0.0039)	(0.0030)	(0.0667)	(0.5958)
	[0.0080]	[0.0080]	[0.0878]	[0.6188]
Day 5 (Non-Water Hours)	-0.268*	-0.504***	-0.319	-0.104
	(0.0977)	(0.0002)	(0.1283)	(0.5173)
	[0.2355]	[0.0020]	[0.2355]	[0.5130]
Day 6 (Non-Water Hours)	-0.214	-0.328**	-0.246	0.006
	(0.1855)	(0.0150)	(0.2563)	(0.9715)
	[0.3932]	[0.0379]	[0.3932]	[0.9701]
Day 7 (Non-Water Hours)	-0.471**	-0.591***	-0.526**	0.026
	(0.0214)	(0.0010)	(0.0413)	(0.8869)
	[0.0579]	[0.0100]	[0.0599]	[0.8842]
Observations	13,463,651	13,349,540	10,581,076	39,848,080

Notes: Conventional p-values are reported in parentheses. Romano-Wolf adjusted p-values, based on 500 clustered bootstrap samples at the household level, are reported in brackets. *, **, *** denote significance at the 10%, 5%, and 1%level based on conventional p-values.