E2e Working Paper 033

Spillovers from Behavioral Interventions: Experimental Evidence from Water and Energy Use

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

December 2017

This paper is part of the E2e Project Working Paper Series.

E2e is a joint initiative of the Energy Institute at Haas at the University of California, Berkeley, the Center for Energy and Environmental Policy Research (CEEPR) at the Massachusetts Institute of Technology, and the Energy Policy Institute at Chicago, University of Chicago. E2e is supported by a generous grant from The Alfred P. Sloan Foundation.

Spillovers from Behavioral Interventions: Experimental Evidence from Water and Energy Use[1]

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

December 6, 2017

Abstract

This paper provides experimental evidence that behavioral interventions spill over to untreated sectors by altering consumer choice. We use a randomized controlled trial and high-frequency data to test the effect of social norms messaging about residential water use on electricity consumption. Messaging induces a 1.3 to 2.2% reduction in summertime electricity use. Empirical tests and household survey data support the hypothesis that this nudge alters electricity choices. An engineering simulation suggests that complementarities between appliances that use water and electricity can explain only 26% of the electricity reduction. Incorporating the cross-sectoral spillover increases the cost-effectiveness of the intervention by 62%.

Keywords: Social Norms; Spillovers; Randomized Controlled Trial; Energy Use; Water Use
JEL: C93, D91, L95, Q40

---

[1]We gratefully acknowledge the UC Davis Center for Water–Energy Efficiency (CWEE), the staff at WaterSmart and Burbank Water and Power for their important roles in developing and implementing this project. We thank Lucas Davis, Paul Ferraro, Meredith Fowlie, Garth Heutel, Kendra Olmos, Kevin Novan, Michael Price, Ivan Rudik, Aaron Smith, Casey Wichman, and seminar participants at Resources for the Future, UCI and the University of Illinois, and the AEA, Advances with Field Experiments, AERE and AAEA conferences for valuable comments. Natalie Popovich provided excellent research assistance. Research support for this project was provided by the Southern California Gas Company, Burbank Water and Power, Metropolitan Water Agency, Hellman Foundation, 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 and CARD, Iowa State University, Ames IA; 50011; Phone: (515) 294-5619; Email: gelade@iastate.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

Firms and governments are increasingly relying on behavioral instruments that seek to alter consumer behavior without changing prices or choice sets.[1] The U.S. and U.K. governments have established divisions tasked with incorporating insights from behavioral sciences into government programs. Private firms also deploy nudges – Bank of America's 'Keep the Change' program rounds debit card purchases to the nearest dollar and deposits the difference into customer savings accounts. The widespread adoption of behavioral policies is informed by an economics and social psychology literature documenting their ability to alter behavior. Studies demonstrate the importance of default settings as a lever to increase participation in retirement savings, organ donations, and dynamic electricity pricing programs (Fowlie et al., 2017; Johnson and Goldstein, 2003; Madrian and Shea, 2001). Altering price salience changes behavior - individuals are less responsive to taxes excluded from posted prices, shipping fees on eBay, earning statements appearing earlier in the week, and automatic bill payments (Chetty et al., 2009; Brown et al., 2010; DellaVigna and Pollet, 2009; Sexton, 2015). And in applications ranging from personal savings decisions to energy conservation, commitment devices and social comparisons have been deployed to alter behavior (Ashraf et al., 2006; Allcott, 2011). Common to all these studies is a shared focus on the evaluation of an intervention on the *targeted* outcome or goal.

Less well understood is whether, and to what extent, behavioral interventions extend beyond the targeted outcome, affecting unexposed margins of behavior for treated individuals. The possibility that interventions spill over is not novel. Studies have investigated if exposure to conditional cash transfers, deworming, energy efficiency programs, and therapy affect outcomes of untreated agents (Angelucci and Giorgi, 2009; Boomhower, 2016; Miguel and Kremer, 2003; Fletcher and Marksteiner, 2017). However, the literature has largely overlooked whether exposure to treatment affects untreated margins of behavior for treated individuals, a spillover we define as a "cross-sectoral" spillover. For example, do default programs for retirement savings alter other forms of savings? Do nudges aimed at encouraging healthy eating impact rates of exercise? Do encouragements to enroll in automatic bill payment for the cable bill increase enrollment in other automated payment programs? If behavioral policies extend beyond their intended objective or sector, this meaningfully affects their scope, cost-effectiveness, welfare impacts, and potentially the sign of a simple

[1] "Policymakers around the world are embracing behavioural science. *The Economist*, May 18, 2017, https://www.economist.com/news/international/21722163-experimental-iterative-data-driven-approach-gaining-ground-policymakers-around.

cost-benefit analysis. This paper uses a randomized controlled trial to examine whether exposure to a nudge influences behavior outside of the stated outcome, and explores potential mechanisms underpinning the presence of cross-sectoral spillovers.

Our empirical context – the effect of social norms messaging about residential water use on household electricity consumption – presents an ideal and policy relevant setting to probe for the existence of cross-sectoral spillovers. Social norms are perhaps the most frequently deployed and studied nudge and have been shown to alter consumer behavior in charitable giving, retirement savings, and voting behavior (Beshears et al., 2015; Croson and Shang, 2008; Duflo and Saez, 2003; Frey and Meier, 2013; Gerber and Rogers, 2009).[2] Nowhere has the study of social norms been more widespread than in the context of energy and water use (Allcott, 2011; Ayres et al., 2013; Costa and Kahn, 2013; Brent et al., 2015; Dolan and Metcalfe, 2015; Ferraro and Price, 2013; Mitchell and Chesnutt, 2013). Public and private utilities have responded by incorporating social pressure as a key component of their water and energy saving efforts. One notably absent feature in the deployment and evaluation of these programs is the possibility that social norms may spill over to affect choices beyond the targeted outcome.

This paper deploys a randomized controlled trial in a service territory with high-frequency water and electricity use data to investigate the effects of home water reports (HWR) on patterns of electricity use. Home water reports compare a household's water use to that of similar neighbors and provide conservation tips and information about water use. Importantly, this messaging does not target, nor mention, electricity use or conservation. In our experimental design, approximately 4,500 households were randomly assigned to receive a HWR, a treatment we refer to as "WaterSmart". The roll out of the experiment in an area with high-frequency water and electricity use data provides unique opportunities to study hourly patterns in response to treatment, and test several hypotheses underpinning the presence of cross-sectoral spillovers.

A central result is that this water conservation instrument leads to a reduction in electricity use. We find an electricity conservation effect of 1.3 to 2.2% in the summer months. The finding is novel not only because a water conservation instrument affects energy use, but because the magnitude of the electricity response rivals that from the deployment of home energy reports (HERs) focused exclusively on electricity conservation (Allcott, 2011;

[2]The use of home reports originated from findings in the social psychology literature showing that simple behavioral interventions that use social norms, frequent feedback, and customized information promote conservation (Hutton and McNeill, 1981; Schultz et al., 2007; Nolan et al., 2008).

Ayres et al., 2013). A second noteworthy element is when the reductions occur. We find that the largest treatment effects occur in the summer between the hours of 3 PM and 7 PM, a period that includes the hours of peak electricity demand, both for the households in our study and the electric grid as a whole. This increases the value of the spillovers since they occur during the hours when wholesale electricity prices are highest.

To understand the mechanisms underpinning this cross-sectoral spillover, we first outline a model that illustrates two channels through which HWRs may impact household energy use, and then empirically test for the presence of these channels. One candidate mechanism is that the cross-sectoral spillovers arises from a mechanical relationship (i.e., complementarity) between appliances that use both energy and water. For example, if households respond to treatment by washing one less load of laundry, this leads to a reduction in electricity use along two margins – less energy is used running the washing machine and the dryer. An alternative or coincident explanation is that HWRs affect electricity use choices more generally. For example, HWRs may increase household attention to utility bills or affect the 'moral utility' associated with water and energy use. This channel may increase electricity conservation beyond what would be expected from mechanical complementarities alone.

Results from three empirical tests and an engineering simulation refute the hypothesis that mechanical complementarities are the sole driver of electricity savings, and are consistent with a framework in which HWRs alter customer choice about water and electricity consumption. A first test posits that if a mechanical relationship explains the cross-sectoral spillover, then water conservation should occur when electricity conservation occurs. Taking advantage of hourly interval data on water and electricity use, we find that the time profile of electricity and water savings are mismatched, with reductions in energy use but no significant reductions in water use between the hours of 3 PM to 7 PM and 8 PM to 10 PM. A second test examines whether households exposed to HWRs alter their consumption of air conditioning, an action that requires electricity but not water as an input. Our finding that electricity conservation increases in temperature but reductions in water use exhibit no temperature response gradient points to the possibility that households may respond to treatment by adjusting their thermostats. Third, results from a post-treatment household survey highlight that exposure to HWRs is positively and significantly correlated with a decrease in some actions that require electricity but not water as an input. Finally, results from a detailed, appliance-level engineering simulation calibrated to Southern California households imply that mechanical complementarities can explain only twenty-five percent of the estimated electricity savings.

Our work highlights the importance of moving beyond a partial equilibrium framework in program evaluation and contributes to a growing literature examining the cost-effectiveness of nudges and their welfare impacts (Allcott and Kessler, 2015). Economists have shown that social norms messaging is a cost-effective water and energy conservation policy (Allcott and Mullainathan, 2010; Brent et al., 2015; Ferraro and Price, 2013). Our results suggest that these cost-effectiveness estimates are understated for two reasons. First, electricity conservation yields private and social savings. Incorporating the private savings and social benefits from the electricity conservation spillover increases the net benefits of HWRs by 39% from net benefits of 2.91 to net benefits of 4.04. Second, electricity savings occur during peak hours when wholesale costs are highest, and the marginal emissions from electricity generation are larger. Accounting for when reductions in electricity use occur further increases the benefit to cost ratio to 4.70. While these cost-effectiveness estimates demonstrate the importance of accounting for spillovers in policy analysis, this valuation metric should not be interpreted as a welfare analysis. Notably, our analysis remains silent on possible disutility or costs households experience from the receipt of HWRs, and the reduction in surplus attributable to inefficient pricing of electricity and water (Allcott and Kessler, 2015). The former may occur due to effort costs involved with conservation or moral costs from consumption. The latter reduction in surplus will arise because in most hours marginal prices for electricity and possibly water exceed the marginal social costs of production. For these reasons, the welfare gains from the deployment of HWRs will likely be less than the savings reported using our cost-benefit framework.

As behavioral instruments grow in popularity and breadth of application, our results point to the possibility for cost-sharing and collaboration in their deployment. This is particularly relevant for the context of water and energy conservation in California. Policy makers and utilities in the state are investigating whether water conservation programs present an opportunity to save energy as well. To date, efforts to document a 'water-energy nexus' have focused on 'embedded' energy savings from water conservation, i.e., electricity savings from the conveyance and treatment of water. We depart from this literature by contributing the first causal data point on the end-use energy impacts of a water conservation program, and show that meaningful energy savings occur from interventions aimed at water conservation.

The paper proceeds as follows. Section 2 provides details on the experimental design. Section 3 describes a conceptual framework of household water and energy use, highlighting the mechanisms through which behavioral interventions for water may impact electricity use. Section 4 describes the data and tests the quality of the randomization, and Section 5 presents our empirical specification and results. Section 6 examines the hypothesis that a purely mechanical response is driving the estimated response. Section 7 discusses cost-effectiveness, and Section 8 concludes.

2 Research Design and Setting

We partnered with a municipally owned water and electric utility and the water technology vendor WaterSmart to evaluate the potential energy and water savings of non-pecuniary water conservation instruments. Together we implemented a framed field experiment in Southern California that exposed a random sample of residential households to a behavioral intervention intended to induce water conservation.

The study occurred in the jurisdiction served by Burbank Water and Power (BWP) and spanned March 2015 to May 2016. BWP serves roughly 18,500 single family customers in the City of Burbank, an inland city in Los Angeles County that like much of Southern California is characterized by a subtropical Mediterranean climate. The timeline of the project includes the summer of 2015, a period that coincides with the worst drought in California's recorded history, and a relatively wet and cool winter.

Our sample consists of 7,341 single-family homes served by BWP. To be eligible for participation in our study, a household needed to reside in a single-family home and have six months of meter readings before the launch of the experiment. Of the roughly 18,000 eligible homes, we randomly assigned 4,559 accounts to the WaterSmart treatment and 2,782 to control.[3] All households in the study had the opportunity to access a utility maintained web portal that provided real-time information on electricity and water use. Control households did not receive notification that they were in a pilot program nor that this portal was available.

WaterSmart: Figure 1 provides an example of a WaterSmart report that treatment households received. The report includes a social comparison, water savings recommendations, and information about the report and BWP's conservation programs. The social comparison provides a household with information on its current water use and compares

[3]The households in this study represent a sub-sample of a broader experiment deployed in Burbank during this time period. The larger experiment is comprised of 17,000 households.

its use to that of similar households and an 'efficient' household.[4] An injunctive norm accompanies the comparison, conveying pro-social behavior through the display of a smiling or frowning face depending on the household's water use relative to its neighbors. The report also includes individualized recommendations on ways to use water more efficiently. The bottom panel of Figure 1 provides an example of a "hard" recommendation – changing from grass to native plants could save 78 gallons per day and $242 per year – and a "soft" recommendation – upgrading irrigation timer settings could save 53 gallons per day and $148 per year. Thus, HWRs motivate conservation by making use of interpersonal comparisons and injunctive norms, recommending changes in water use behavior, and encouraging the adoption of efficient water technologies. Importantly, no messages or water conservation tips target or mention energy conservation.

Treatment households received bi-monthly reports from May 2015 to March 2016 through the mail or a combination of email and mail. WaterSmart sent an introductory letter to all treatment households in March 2015, explaining the product and notifying them that they would be receiving HWRs over the subsequent year. WaterSmart then mailed the first home water reports to treatment households in May 2015. Our study defines the treatment period as May 15, 2015 to May 31, 2016. We exclude data spanning March 15 to May 15, 2015. We do this to account for the possibility that households may change behavior in response to the introductory letter, but before the rollout of HWRs. The medium through which households received HWRs depends on their method of bill payment to BWP. While WaterSmart physically mailed all treatment households the introduction letter and first HWR, subsequent reports were sent according to the following rule: households enrolled in electronic billing received electronic HWRs and households enrolled in paper billing received physical reports in the mail.

3 Conceptual Motivation

To understand why cross-sectoral spillovers may exist, we outline a simple conceptual framework that illustrates two channels through which a water nudge may impact energy use. Consider a household that derives utility from energy-consuming services (e.g., air conditioning), water-consuming services (e.g., a shower), and a composite good. Following Allcott

[4]WaterSmart uses proprietary software to choose a group of 'similar' households based on publicly available household characteristics (e.g., number of rooms), an assumed number of occupants, and proximity.

and Kessler (2015) and Farhi and Gabaix (2017), we allow imperfect information and behavioral biases to complicate choices about water and energy. Consumers may misperceive water and energy prices and/or be misinformed about consumption quantities because demand for electricity and water are derived (Ito, 2013, 2014; Jessoe and Rapson, 2014). Since our experiment coincided with the worst drought in California's history, consumers likely experienced 'moral utility' from their consumption of water and energy.[5]

A water conservation nudge may influence household energy demand through two channels: a mechanical (complementarities) channel, and a behavioral channel. First, water and electricity are complementary inputs in the use of many household appliances. For example, if households respond to a nudge by washing one less load of laundry per day, this leads to a reduction in electricity use along two margins: less energy is used running the washing machine, and the dryer is used less. The reduction in electricity use that occurs in response to the water nudge is a result of the Leontief nature of demand for this appliance, a channel we label "mechanical complementarities". If nudges alter the moral tax on water use, or partly correct customer misperception about water prices, a nudge will cause demand for water to change, and due to complementarities between water and electricity use, electricity use will also change.

The second channel reflects the possibility that a nudge on water use may change customer perception about electricity prices or the moral tax households place on electricity consumption. We refer to this as the behavioral channel. This will occur if misperceptions or behavioral biases about water and energy use share some common components, and the nudge influences perceptions and/or the moral tax for both water and electricity. Assume that a nudge on water use increases the moral tax on electricity use. This will lead to a decrease in electricity use above and beyond the amount attributable to the mechanical spillover. Alternatively, if the nudge decreases the moral tax on electricity use, for example through moral licensing, then this non-mechanical channel will induce households to increase electricity use.[6] Importantly, the sign of this behavioral spillover could be either positive or negative.

---

[5]Residents faced enormous conservation appeals from local and state governments, and significant social pressure to conserve water. The hashtag #droughtshaming was trending on both Twitter and Facebook, and consumers regularly posted pictures of and reported neighbors that violated outdoor watering restrictions.

[6]Moral licensing may play an important role in this setting. The psychological theory on "moral licensing" suggests that individuals may use their own "good" behavior to justify "bad" behaviors. It has been empirically corroborated in the energy setting whereby households voluntarily enrolled in "green" energy programs increase electricity consumption (Jacobsen et al., 2014; Harding and Rapson, 2014).

4 Data and Quality of Randomization

High-frequency metering data on household electricity and water use serve as the primary data for our analysis. BWP provided hourly water and electricity data for every household in our sample for the treatment period, defined from May 2015 to May 2016. Pre-treatment monthly billing data on water and electricity use were obtained for all households spanning the baseline period March 2014 to February 2015. We also collected county assessor data on housing unit attributes and matched these at the address with the electricity and water use data. Our balance tests make use of the monthly pre-treatment billing data on water and electricity use, and the assessor data.

Table 1 reports summary statistics on pre-treatment electricity use, water use and housing unit attributes, and examines the credibility of the randomization. The first two columns report means for households assigned to control and treatment, respectively, and the third column reports the difference in means. A comparison of means across control and treatment shows no statistically significant differences in monthly water use, monthly electricity use or summertime electricity use. For the households for whom we observe assessor data, housing unit characteristics are also balanced across control and treatment.[7] We further investigate the quality of the randomization by comparing baseline electricity use across control and treatment in each month of the year preceding the experiment. As shown in Figure 2, there are no significant differences in electricity use across control and treatment in any calendar month, including importantly the summertime months of May, June, July, and August. These descriptive statistics provide a first layer of support for the integrity of the randomization. They also offer visual evidence that control households may use more but not significantly more electricity than treatment households, particularly in the summer months. To account for the possibility that differences in post-treatment outcomes may be attributable to pre-treatment differences in electricity use, we follow Bruhn and McKenzie (2009) and control for pre-treatment annual, summer, and winter electricity use in our preferred specifications.

While Table 1 focuses exclusively on assignment to treatment, attrition occurs in our sample for two reasons. First, households move. This leads to the termination of an account and the omission of post-move household hours from our sample. This form of attrition affects both control and treatment households and is uncorrelated with assignment to treatment. The second source of attrition arises since a small number of households assigned to

[7] Assessor data were available for 7,105 of the 7,341 homes in our sample.

the WaterSmart treatment opt out of receiving HWRs. This selective attrition leads to the presence of "never takers" in the treatment group and may compromise the experimental research design. To address this concern, we continue to monitor water and electricity use for households that opt out of HWRs and estimate intent to treat effects.

5 Empirical Approach and Results

This section details the empirical approach used to isolate the effect of HWRs on average electricity consumption, and the results that follow. We then discuss the impacts of HWRs on the time profile of electricity usage.

5.1 Average Treatment Effects

To identify the causal effect of HWRs on energy use, we begin by comparing average hourly electricity use across treatment and control households conditional on weather controls and time fixed effects, estimating the equation:

$$y_{ih\tau} = \beta_0 + \beta_1 \mathrm{WS}_i + f(T_{h\tau}; \Theta) + \theta_p P_{h\tau} + \gamma_\tau + \gamma_h + \epsilon_{ih\tau}. \tag{1}$$

The dependent variable $y_{ih\tau}$ is the level of electricity use specified in kilowatt-hours per hour (kWh/hr) for household i during hour h of day τ.[8] The indicator variable WS_i denotes assignment to the WaterSmart treatment, and equals one if customer i is assigned to treatment. Calendar date and hour of day fixed effects, denoted by γ_τ and γ_h, control for seasonal and hourly patterns in use. Weather controls include, $f(T_{h\tau}; \Theta)$, a flexible function of hourly temperature ($T_{h\tau}$) with parameters Θ, and $P_{h\tau}$, an indicator variable denoting if precipitation was recorded in hour h of day τ. We specify $f(T_{h\tau}; \Theta)$ as a series of 5°F temperature

[8] We use the level of electricity use for two reasons. First, we are primarily interested in understanding the level change in electricity use from assignment to treatment. Given that electricity consumption varies substantially across hours of the days and months of the year, a focus on the percentage change in electricity use might mask substantial differences in level changes across hours and months. Second, logarithmic transformations dampen the impact of high electricity users, and previous work has found that these users exhibit some of the largest treatment effects (Allcott, 2011). In Appendix Tables A.4 and A.5, we show that our results are robust to a log transformation of the dependent variable.

bins,

$$f(T_{h\tau};\Theta) = \begin{bmatrix} \theta_{60} \cdot \mathbf{1}(T_{h\tau} < 65°\text{F}) \\ \theta_{65} \cdot \mathbf{1}(65°\text{F} \leq T_{h\tau} < 70°\text{F}) \\ \vdots \\ \theta_{80} \cdot \mathbf{1}(80°\text{F} \leq T_{h\tau} < 85°\text{F}) \\ \theta_{85} \cdot \mathbf{1}(T_{h\tau} \geq 85°\text{F}) \end{bmatrix}$$

where $\mathbf{1}(\cdot)$ is an indicator function that equals one whenever the outdoor temperature in an hour lies within the specified range.

Our coefficient of interest, β_1, should be interpreted as the intent to treat effect and represents the average change in hourly electricity use from assignment to treatment. Standard errors are clustered at the household. In subsequent specifications, we follow Allcott and Rogers (2014) and Bruhn and McKenzie (2009) and build on this comparison of means by including baseline summer, winter and annual electricity use in the year preceding treatment as control variables.

Table 2 reports results for the effect of assignment to WaterSmart on average hourly electricity use.[9] To translate these level effects into percentages, this table also reports mean hourly use for control households. Column (1) displays our results from estimating equation (1). Column (2) includes pre-treatment summer, winter, and annual electricity use to control for possible differences in baseline electricity use across households. In columns (3) and (4), we restrict the sample to the summer months of 2015 (May 15, 2015 to August 31, 2015) and replicate the specifications in columns (1) and (2). We focus on summer months for two reasons. First, system-wide electricity loads are highest and water demand peaks in the summer. Second, this period corresponds to the first 75 days of our pilot when HWRs may be most salient. In columns (5) and (6), we further restrict the sample to the peak electricity use hours – 3 PM to 8 PM – in the summer.

Our interpretation of Table 2 focuses on the estimates that condition on pre-treatment electricity use. Our choice to concentrate on these results is guided by earlier work on evaluation in randomized controlled trials, and a comparison of the results presented in even and odd columns (Bruhn and McKenzie, 2009). The latter shows that the estimated treatment effects reduce in magnitude after controlling for pre-treatment electricity use. This attenuation in treatment effects reflects our finding in Figure 2 that households assigned to

[9]We exclude AMI meter reads above 20 kWh/hr as they are likely errors. This restriction reduces the sample by 1,115 observations or 0.01%. Results are similar if we exclude AMI meter reads above 10 kWh/hr or readings above the 99th percentile.

treatment use less (but not significantly less) electricity in the pre-treatment period. Moving forward, we focus on the specifications that control for baseline electricity consumption and use the results reported from the estimation of equation (1) as an upper bound.

Assignment to treatment reduced hourly electricity use by approximately 0.017 kWh/hour, or 1.35 percent, during the summer months. This amounts to every household terminating the use of one 15W light bulb over the summer. Alternatively, this is equivalent to every household reducing its use of an Energy Star dishwasher by one load every three days during the summer months. The estimated impact falls within the range of treatment effects reported from the deployment of home *energy* reports focused exclusively on energy consumption and conservation (Allcott, 2011). This comparison makes clear the economic importance of the estimated electricity spillover. Columns (5) and (6) show that the treatment effects are most pronounced during peak hours in the summer when HWRs induce a 1.5 to 2.5 percent reduction in electricity use. This provides the first piece of empirical evidence that, in addition to a conservation effect, the reports may yield savings via a reduction in peak electricity demand.

As shown in columns (1) and (2), over the duration of the year WaterSmart does not induce a change in electricity use. To illustrate the time profile of treatment effects over the duration of the experiment, we interact assignment to treatment with month of year and estimate monthly treatment effects conditional on baseline electricity use, weather controls, calendar date and hour of day fixed effects. Figure 3 plots the estimated treatment effect for each month spanning the period April 2015 to May 2016 relative to March 2015.[10] The blue line graphs estimates using all hours of the day and the red line illustrates the estimates from a regression that includes only peak hours. The figure makes clear that significant reductions in electricity use occur, but are confined to June, July, and August. After the summertime, HWRs induce no change in electricity use. The difference between the summertime and annual response may occur because earlier reports are more salient, there are more levers to reduce electricity use in the summer, or electricity conservation is more front of mind in the summer.

5.2 Time Profile of Treatment Effects

Hourly interval data allow us to estimate treatment effects across hours of the day and provide further insight into the load management and environmental impacts of home water

[10] We include March and April 2015, months in which households had received a mailer introducing them to the WaterSmart product but had yet to receive a HWR.

reports, as well as the levers by which households respond to treatment. To decompose the treatment effect, we augment equation (1) and estimate

$$y_{ih\tau} = \sum_{j=1}^{23} \alpha_h \mathbf{1}[h = j] + \sum_{j=0}^{23} \beta_h \left(\mathbf{1}[h = j] \times \mathrm{WS}_i\right) + f(T_{h\tau}; \Theta) + \theta_p P_{h\tau} + \Gamma \mathbf{X}_i + \gamma_\tau + \epsilon_{ih\tau}, \tag{2}$$

where $\mathbf{1}(\cdot)$ is an indicator function that equals one when hour h equals hour-of-day j. The coefficients α_h reflect the conditional average hourly electricity consumption for control households relative to the omitted hour, 12 AM. Controls for monthly pre-treatment average electricity use in the twelve months, summer months, and winter months preceding treatment are denoted by $\mathbf{X}_i$. The coefficients of interest, β_h, capture the average effect of assignment to treatment for each hour of the day. Given our findings in Table 2 and Figure 3, we restrict the sample to the summer of 2015.

Disaggregating the impact of HWRs by hour of day reveals substantial heterogeneity in the time profile of our average treatment effect. This can be seen in Figure 4(a) which plots the effect of assignment to treatment on electricity use for each hour of the day. From 12 AM to 5 AM we observe visual but not statistically significant evidence that treatment households use less electricity. No discernible differences occur between 5 AM and 11 AM. Around 11 AM, treatment households begin to reduce electricity use relative to control households, with this effect growing in magnitude and significance over the hours of the late morning and early afternoon. Significant reductions start to occur at 3 PM. The treatment effects are largest from 3 PM to 7 PM, peaking at 4 PM and 5 PM. They persist until 11 PM, with the effect slowly declining after 7 PM. Heterogeneity in the timing of the treatment effects has direct implications for the greenhouse gas and local pollutant reductions attributable to HWRs and perhaps energy conservation programs more generally. This is because the marginal source of electricity generation, and the greenhouse gases and local pollutants generated from this source, vary hourly. In Section 7, we weigh in on the importance of this heterogeneity by quantifying the local pollution and greenhouse gas impacts of HWRs under a uniform hourly average treatment effect assumption and a variation on the hour of day treatment effects presented in Figure 4(a).

The largest estimated treatment effects coincide with peak demand hours for both the households in our study and the electric grid as a whole.[11] This suggests that in addition

[11] A recent analysis by the California ISO classifies 'peak net load hours on the California electricity

to a conservation effect the report leads to additional savings by reducing peak load. The efficiency gain occurs because of the disparity between the marginal cost to supply electricity and the retail price of electricity. The absence of large-scale economically feasible electricity storage options means that at each moment in time, the electricity supply needs to meet demand. This leads to wholesale electricity prices that vary in near real-time as the cost of the marginal unit differs widely across generation sources. These fluctuating marginal costs, however, are not passed along to residential customers in BWP. Most customers pay a flat retail price irrespective of when they use electricity. Our finding that customers reduce electricity during the hours when it is most expensive to supply provides further private value to the utility. In Section 7, we monetize the incremental savings attributable to heterogeneity in wholesale costs and the response to treatment.

6 Mechanisms: Empirical Evidence

Our finding that the water conservation instrument induces electricity conservation leads to several questions on the mechanisms underpinning the result. One leading explanation for cross-sectoral spillovers is mechanical – reductions in energy use are due to reductions in water-consuming activities that also use electricity. An alternative hypothesis is that treatment alters consumer choice about resource use more generally, encouraging households to reduce electricity use. We now propose and implement three empirical tests and an engineering simulation to gauge the plausibility of these two channels.

6.1 Timing of Water and Energy Conservation

In a first empirical test, we hypothesize that if a mechanical relationship explains the cross-sectoral spillover, then reductions in electricity use should be accompanied by reductions in water use. One indication that energy conservation is driven exclusively by reductions in water use is if the observed timing of water and electricity conservation coincide. To examine this, we compare the estimated hour of day treatment effects from Figure 4(a) with the corresponding hour of day treatment effects for water use.

As a first step, we present results for the mean hourly effect of assignment to Water Smart on water use. Table 3 replicates the specifications from Table 2 using hourly water use as the

grid as 4 PM to 9 PM, with ‘super peak occurring at the same hours in July and August each year. See http://www.caiso.com/Documents/CaliforniaISO_Time_UsePeriodAnalysis.pdf.

dependent variable. The results make clear that HWRs induced large water savings. The intent to treat effects range from 0.45 to 0.625 gallons per hour. Our results correspond to a 4.4 and 2.9 percent reduction in water use over the whole year and in the summertime, respectively, and are equivalent to reducing the amount of time spent showering by 6 minutes per day.[12] Unlike our findings in Table 2, the annual treatment effects are larger in magnitude than the summertime effects. When we restrict the sample to the peak summer hours of 3 PM to 8 PM, we find that the estimated water reductions increase in levels and percentages.

Panel (b) of Figure 4 plots the treatment effects from estimating equation (2) using hourly water use as the dependent variable. The water treatment effects are concentrated in both magnitude and significance in the early morning hours of 5 AM to 7 AM, and at 7 PM. These hours coincide with the sunrise and sunset, suggesting that the treatment effects are likely driven by reductions in outdoor watering. We observe no significant reductions in water use in any other hours of the day.

Comparing panels (a) and (b) in Figure 4 provides our first piece of empirical evidence that a mechanical relationship between actions that use both water and electricity is not the sole driver of our observed cross-sectoral spillover. The time profile of electricity reductions runs counter to the time profile of water savings during some hours. While moderate reductions in electricity use occur from 3 PM to 11 PM, there are no statistically significant reductions in water usage over this time interval except for a large decrease at 7 PM. Under the mechanical hypothesis, if electricity reductions occur, water reductions should also occur. From 3 PM to 7 PM and 8 PM to 10 PM this is not the case, suggesting that HWRs affect at least some electricity use through a non-mechanical channel.

6.2 Response to Temperature

An ideal exercise to tease out the importance of the behavioral channel in explaining the cross-sectoral spillover would test whether electricity choices that do not require water as an input are affected by assignment to treatment. If treatment households make systematically different electricity consumption decisions, this suggests that assignment to treatment impacts electricity use through a behavioral channel. While the data needed to examine this hypothesis directly are not available, we propose and implement an indirect empirical test to investigate its plausibility.

[12] The treatment effects are similar to those found in other studies of WaterSmarts HWRs, providing further support for the integrity of the randomization (Brent et al., 2015).

We examine the behavioral hypothesis along an important and salient dimension of household electricity use: cooling. To do this, we first test whether electricity use for treated households responds differentially to increases in outdoor temperature relative to control households. We then test whether water conservation from assignment to treatment also increases in temperature. The rationale guiding this test is that cooling comprises a significant share of household summertime electricity use, its use increases in outdoor temperature, and it requires electricity but not water as an input. A differential response to high temperatures for electricity but not for water suggests that households exposed to HWRs alter their consumption of air conditioning.

We implement this empirical test by augmenting equation (1) and estimating

$$y_{ih\tau} = f(T_{h\tau};\Theta) + g(\mathrm{WS}_i \times T_{h\tau};\Lambda) + \theta_p P_{h\tau} + \Gamma\mathbf{X}_i + \gamma_\tau + \epsilon_{ih\tau}, \tag{3}$$

where all variables are defined as before. We specify $g(\mathrm{WS}_i \times T_{h\tau};\Lambda)$ as the interaction of each of the temperature bins in $f(T_{h\tau};\Theta)$ with the indicator variable for assignment to treatment. As before, we limit our sample to summer 2015.

Our empirical strategy identifies the differential impact of a given temperature bin on electricity (and water) use across treatment and control households. The coefficients θ_j estimate the conditional average electricity (water) use for temperatures in bin j relative to the excluded bin, temperatures less than 65°F. The coefficients λ_j estimate the difference in average electricity and water use when temperatures fall in bin j between treatment and control households. If treatment households respond to HWRs by reducing their air conditioning use, λ_j will increase in temperature for electricity but not for water use.

We face an important challenge in using this empirical strategy – the water consumption response to treatment may increase in temperature as well. The literature on the relationship between ambient air temperature and residential water demand highlights that residential water demand increases linearly in temperature, and/or may exhibit a threshold response above specific temperatures (Balling and Gober, 2006; Gato et al., 2007). This work points to the possibility that the water response to treatment may increase in temperature, though the mechanism through which this would occur is not obvious ex-ante. Households may respond to treatment by shifting outdoor irrigation from the afternoon to the morning, or by reducing outdoor watering altogether on hot days. A unique feature of our setting allows us to imperfectly control for the relationship between temperature and outdoor irrigation. Due to the ongoing drought, BWP imposed outdoor watering restrictions and households were only permitted to water on two exogenously determined days of the week. As such, we split our sample into utility-wide watering and non-watering days. We hypothesize that on watering days the water response to WaterSmart is likely increasing in temperature, while on non-watering days the relationship between temperature and response to treatment should be substantially attenuated. If our cooling hypothesis holds, then (i) the electricity response to treatment should be increasing in temperature but should not vary substantially across watering and non-watering days, and (ii) the water response to treatment should be relatively constant across temperatures on non-watering days.

Our first set of results highlights that differences in electricity use across treatment and control households are increasing in temperature, a first step in providing evidence consistent with a cooling hypothesis. Columns (1) and (2) of Table 4 report the results from the estimation of equation (3) on non-watering and watering days, respectively. On non-watering days the magnitude of the treatment effect jumps from 0 to 0.027 kwh per hour as temperature increases from below 65 F to above 85 F, and on watering days the response increases from around 0 to 0.036 kwh per hour. The findings highlight that the relationship between electricity use and treatment is increasing in temperature and is nearly identical across watering and non-watering days, suggesting that it is unaffected by watering restrictions.

In contrast to the electricity results, the water response to treatment does not increase in temperature on watering days or non-watering days. We report these results in columns (3) and (4) of Table 4. On both watering and non-watering days, the largest treatment effects occur during the relatively mild temperature hours of 70F or cooler. This highlights that the water use response to treatment exhibits no temperature gradient, suggesting that factors uncorrelated with temperature are driving water conservation.

Our findings provide indirect evidence that households exposed to HWRs alter electricity use through decreased cooling. A comparison of the electricity and water response to treatment at different temperatures makes explicit that while the electricity response to treatment is increasing in temperature, the water response to treatment is not. We find that the water use response to treatment is not correlated with temperature increases, and exhibits the largest response in both magnitude and significance during hours with relatively mild temperatures.

6.3 Household Survey of Electricity Use Behavior

To explore the plausibility of our hypothesis that HWRs affect customer choices about electricity use, we designed and conducted a post-treatment household survey on electricity and water use decisions for a subset of households. Comparing responses across control and treatment households will provide empirical insight into whether assignment to treatment is correlated with electricity conservation actions.

We mailed our survey to 2,400 households (1,600 treatment and 800 control) in August 2016. The survey asked respondents whether they undertook a variety of actions to reduce their water and electricity use over the period June 2015 to June 2016. All surveys included a $2 bill to increase completion rates, and households received two follow-up reminders. In total, 38% of households that received a survey completed it, with compliance rates equivalent across the control and treatment groups.

Two empirical challenges arise in drawing inference from the survey data. First, within each control and treatment group, the households that complete the survey may systematically differ from the households that do not. Second, conditional on survey completion, treatment and control households may be systematically different. Both issues may pose empirical challenges since unobservable factors that influence a household's decision to complete the survey may impact their response to treatment. We test for each of these issues in Table 5. In columns (1) and (2), we regress an indicator variable for whether a household completed the survey on pre-treatment water and electricity use for treatment households and control households, respectively. Treatment households that do and do not complete survey are balanced on pre-treatment water and electricity use but control households are not. Larger summertime electricity users are less likely to complete the survey. To study whether survey attrition may be asymmetric across treatment and control households, we restrict the sample to households that completed the survey and formally test for differences in observables across control and treatment households. Results presented in column (3) suggest that treatment and control households differ in baseline electricity use. These imbalances inform our analysis of the survey data, motivating a comparison of means across energy conservation actions conditional on baseline electricity and water use.

For every energy conservation question asked in the survey, households assigned to treatment are (weakly) more likely to respond that they participated in that action. We also find that even with our small sample, households assigned to the WaterSmart treatment are significantly more likely to report turning off the lights or turning off the TV. These results are shown in Table 6; results are reported from the estimation of a linear probability regressing participation in an energy conservation activity on assignment to treatment, controlling for households pre-treatment electricity and water use. Comparing self-reported participation in various energy consuming actions across treatment and control households highlights that the largest differences in response, both in magnitude and statistical precision, are for actions that affect electricity but not water use: (i) turning off the lights or the TV; (ii) air drying clothes; and (iii) installing a programmable thermostat. These comparisons provide another line of evidence consistent with our behavioral hypothesis.

6.4 Simulation to Bound Mechanical Spillovers

We develop an engineering model of household water and energy use to understand the potential magnitude of mechanical complementarities in explaining the cross-sectoral electricity spillover. The model simulates expected changes in yearly electricity use assuming all electricity reductions come from reductions in end uses that use both water and electricity as inputs. The model makes several restrictive assumptions, and thus should be viewed as an approximation of the expected electricity savings from the deployment of WaterSmart if the behavioral channel was "shut off."

We simulate the expected change in yearly electricity use from the WaterSmart intervention assuming that HWRs only impact electricity use through mechanical complementarities. We use individual response data from the 2009 California Statewide Residential Appliance Saturation Study (RASS) to parameterize an appliance-level model of a representative household's water and electricity use. The RASS data include detailed information on appliance ownership and estimates of yearly, appliance-level electricity use for over 25,000 homes in California. We restrict our sample to around 8,500 single-family homes in Southern California to ensure the RASS respondents closely resemble the households in our study. From the list of appliances covered in the RASS, we identify the following end uses as directly or indirectly consuming water and energy: water heating, clothes washing, clothes drying, dishwashing, groundwater pumping, evaporative cooling, and pool/spa operation. We use the average ownership rates, the distribution of annual energy use for each appliance, and the average water treatment effect from column (4) of Table 3 to simulate the distribution of average electricity use reductions under four scenarios.[13]

Table 7 reports our results. In our baseline scenario, we assume that water use reductions are distributed equally across all appliances. Our additional scenarios take extreme stances on appliance use and ownership patterns. Our intention is to gauge under each of these

[13]Appendix A.1 discusses the RASS data and simulation procedure in more detail.

implausible assumptions the extent to which mechanical complementarities could explain the estimated electricity reductions, and make obvious that mechanical complementarities are unlikely the sole driver behind the cross-sectoral spillover.

In our preferred scenario – in which water use reductions are distributed equally across appliances – household electricity consumption decreases by approximately 39 kWh per year. For comparison, the estimated treatment effect from column (4) of Table 2 is 149 kWh per year. In this scenario, we can attribute about 25 percent of the observed electricity savings from the WaterSmart intervention to mechanical complementarities.

The second scenario assumes that all water reductions come from decreases in the use of indoor appliances that require both water and electricity as inputs. Specifically, we assume that the quantity of water savings estimated in column (4) of Table 3 occurs from a reduction in the use of clothes washers, electric dryers, and dishwashers only. This implies that on average dishwasher and clothes washer use would need to decrease by nearly 40 percent in response to treatment.[14] Under this implausible assumption, household electricity use decreases by 140 kWh per year and can explain 90% of the cross-sectoral spillover. Framed differently, if households respond to WaterSmart by reducing the use of clothes washers, electric dryers, and dishwashers by 20% this could only explain 45% of the estimated reduction in electricity use.

Our third scenario assumes that water use reductions are distributed equally across appliances and that every household uses an electric water heater. While ownership of electric water heaters is well below 10 percent in our area of study, this scenario allows us to gauge the importance of water heating in explaining mechanical complementarities. Even under this extreme assumption, the simulated electricity savings explain only 80 percent of the estimated electricity spillover. If we assume that 10% of households have electric water heaters and all the estimated water savings occur within these households, then these households would need to reduce water use by almost 45% for the electricity savings to be explained by mechanical complementarities.

Our last scenario casts light on one potential lever that could explain the cross-sectoral spillover: pools and spas. If we assume that all households with pools and spas reduce their use of pool filters and spas by 100 percent, mechanical complementarities lead to electricity

[14] According to the RASS data, households run approximately 256 clothes washer loads per year, 228 dryer loads per year, and 133 dishwasher loads per year. If we assume clothes washers use 35 gals/load and dishwashers use 8 gals/load, total average water use from the appliances is 10,024 gallons per year compared to our estimated treatment effect of 3,925 gallons per year.

savings far exceeding our estimated treatment effect. Phrased differently, if around twenty percent of the households in our sample that own pools and spas discontinued their use of pools and spas entirely, mechanical complementarities could explain our estimated cross-sectoral spillover. We can assess the extent to which pool users drive our results directly. To do this, we replicate our results from Table 2 using only the sample of households that do not own pools. Results reported in Table 8 make clear that households with pools do not drive the estimated electricity spillover. The robustness of the electricity savings to the exclusion of these households rules out the hypothesis that mechanical complementarities in the operation of pools and spas explain the cross-sectoral spillover.

7 Valuing Electricity Conservation Spillovers

Our finding that social norms messaging about water consumption induces electricity conservation implies that existing cost-effectiveness estimates of HWRs, and perhaps more generally social norms messaging, are understated. In this section, we set forth a cost-benefit framework to quantify the net benefits from the deployment of home water reports. This analysis departs from existing work on the cost-effectiveness of social norms messaging in the residential water and energy settings (Allcott, 2011; Brent et al., 2015; Ferraro and Price, 2013). Cost-effectiveness estimates in earlier studies provide a ratio of program costs to the energy or water savings, yielding a dollar per kWh or dollar per gallon metric. Our framework uses dollars as the common unit to aggregate water and energy savings, and incorporates savings from reductions in local and global pollutants. In addition to reporting average savings, our framework takes advantage of granular data on electricity and water use to explicitly account for when these reductions occur. The timing of electricity conservation affects both wholesale electricity costs and pollution emissions, and will impact cost-benefit estimates.

An ideal valuation exercise would move beyond a benefit-cost framework and contribute to a growing literature that quantifies the welfare effects of nudges (Handel, 2011; Bernheim et al., 2015). Of particular relevance is recent work that conducts a follow-on experiment to elicit consumer willingness to pay for home energy reports, and provides a first estimate of the welfare impacts of nudges in the residential electricity sector. Despite the importance of evaluating the welfare effects of nudges, our research setting does not provide an opportunity to do this effectively. In contrast to Allcott and Kessler (2015), we did not elicit willingness to pay for nudges, rendering it challenging if not impossible to trace out a demand curve for HWRs. One option would be to use average WTP values from Allcott and Kessler (2015) as a proxy for consumer welfare in our setting. We do not pursue this for two reasons. First, Allcott and Kessler (2015) find substantial heterogeneity in WTP across consumers, even within a single service territory in upstate New York. Second, we question the external validity of transferring estimates obtained in upstate New York to Southern California. Instead, we outline a cost-benefit framework that provides a transparent and practical metric to evaluate the relative importance of electricity spillovers.

Our cost-benefit exercise aggregates the water savings, electricity savings, and savings from reductions in local and global pollutants into an annual, dollars per household metric. We measure the average annual net benefits per household from the deployment of HWRs as:

$$\text{Net Benefit} = P^{\text{Ret,w}}\Delta w + P^{\text{Ret,e}}\Delta e + \phi^{e}\Delta e + \Delta\Pi. \tag{4}$$

The first two terms estimate the average reduction in water and electricity bills from the intervention, where $P^{\text{Ret,w}}$ and $P^{\text{Ret,e}}$ denote the water and electricity retail prices, respectively. Households in BWP's territory face an increasing block rate pricing structure for both water and electricity, complicating our determination of marginal prices.[15] To approximate the marginal price, we calculate each household's cumulative water and electricity use for each calendar month, and assign each household the marginal price associated with the final unit of consumption.[16]

The term $\phi^{e}\Delta e$ accounts for social benefits from electricity conservation. These occur because a reduction in upstream pollution generation reduces global and local pollutants. Our cost-benefit framework focuses on the benefits from changes in the local pollutants SO_2, NO_x, and $PM_{2.5}$, and CO_2 emissions. To value the avoided pollution from the intervention, we use estimates from Holland et al. (2016) on hourly marginal damages, broken out separately for CO_2 and local air pollutants.[17]

[15] Households pay \$0.1153/kWh for the first 300 kWh, and \$0.1672/kWh for all consumption above 300 kWh. Similarly, households pay \$1.291/HCF for the first 15 hundred cubic feet (HCF) of water consumption, \$1.59/HCF for the next 15 HCF, and \$2.001/HCF for all consumption above 30 HCF. One hundred cubic feet is approximately 750 gallons.

[16] This is an imperfect measure of marginal price since monthly electricity and water use are based on billing cycles not calendar months.

[17] CO_2 damages are valued at the EPA's social cost of carbon, \$41/ton, and marginal local pollutant damages are estimated with the AP2 model, an extension of the model developed by Muller and Mendelsohn (2009), using an \$8.1 million value of a statistical life.

Our net benefits measure also includes the net change in utility revenue, $\Delta\Pi$, from the deployment of HWRs. HWRs affect utility revenue along three dimensions: a reduction in water and electricity sales to consumers; a decrease in wholesale water and electricity acquisition costs; and an increase in costs to pay for HWRs. We calculate the change in utility revenue as,

$$\Delta\Pi = (P^{\mathrm{Wh,e}} - P^{\mathrm{Ret,e}})\Delta e + (P^{\mathrm{Wh,w}} - P^{\mathrm{Ret,w}})\Delta w - C,$$

where $P^{\mathrm{Wh,j}}$ are wholesale costs for $j = e, w$; C is the annual cost per household to supply bi-monthly HWRs; and all other terms are as defined in equation (4). Based on utility audit reports, we assume the price incurred by the utility to purchase a marginal unit of water is \$2.072/HCF. We use hourly, day-ahead wholesale electricity prices for a large aggregation point near our utility in Southern California as our measure of wholesale electricity costs, and an average annual cost per household of \$10 to supply bi-monthly HWRs.[18]

If we substitute the measure of net utility revenue ($\Delta\Pi$) into equation (4), we show that customer bill savings are fully offset by utility revenue losses,

$$\text{Net Benefit} = P^{\mathrm{Wh,e}}\Delta e + P^{\mathrm{Wh,w}}\Delta w + \phi^{e}\Delta e - C. \tag{5}$$

We choose to measure the savings and losses in retail expenditures because their inclusion makes explicit how our calculation differs from a welfare analysis. In a welfare analysis, we would replace household bill savings with the net change in consumer surplus attributable to HWRs (Allcott and Kessler, 2015). The implication of this is that the revenue losses incurred by the utility from a reduction in sales would no longer fall out of the net benefits calculation.

Panel A of Table 9 presents our baseline results, itemizing each term from equation (4). In this panel, we assume that the reductions in water and electricity use are uniform across all hours and months. All columns report results from the estimation of a modified version of equation (1) in which (i) we condition on baseline electricity or water use and (ii) the dependent variable is the outcome of interest listed in each row.[19] In the first column we restrict our attention to the water savings and revenue losses from the deployment of HWRs. Column (2) focuses exclusively on the net benefits from energy conservation, including the

[18]We use day-ahead wholesale electricity prices from the default load aggregation point (DLAP) SCE-APND node.

[19]Tables A.2 and A.3 in the Appendix present the regression results.

social benefits from reductions in local pollutants and greenhouse gas emissions. The final column of the table aggregates these savings and losses.

The table makes clear the importance of incorporating spillovers into a benefit-cost analysis. The net benefits increase by almost 40% from \$2.91 per household to \$4.04 when we account for private electricity savings and the social benefits of electricity conservation. On average, households save \$22 and \$3 on their water and electricity bills from the deployment of HWRs, respectively, though lost utility revenues perfectly offset these savings. Net benefits accrue from a reduction in the cost to purchase water and electricity. Wholesale water savings amount to \$13 per household per year, and electricity savings total at \$0.86 per household per year. The value of the CO_2 reductions from the HWRs is \$0.16 per household per year, and the value from reductions in local pollutants amounts to \$0.11 per household per year. While the net savings from water conservation alone are positive, our analysis suggests that the internalization of spillovers into a cost-benefit framework increases the net benefits of HWRs substantially.

Our valuation of the electricity spillover in Panel A does not allow for temporal heterogeneity in when these treatment effects occur. Accounting for heterogeneity in electricity use reductions both in the hour of day and month of year may substantially affect the value of spillovers. This is because the wholesale price of electricity varies considerably across months of the year and hours of the day, and the marginal source of generation differs in emissions rates. To incorporate this form of heterogeneity into our net benefits framework, we follow Novan and Smith (2017) and Boomhower and Davis (2016) and estimate treatment effects for each month of the year and hour of the day.

As shown in Panel B of Table 9, accounting for variation in the timing of when electricity reductions occur further increases the net benefits of HWRs by 16% from 4.04 to 4.70. In our preferred specification, the estimated savings from electricity reductions now amount to \$1.79 per household. This increase in savings aligns with our expectations because treated households reduced electricity use during the hours of the day when electricity is most expensive to supply, and the marginal source of generation is dirtier. In this more flexible and more realistic framework, the incorporation of cross-sectoral spillovers increases the net benefits from the deployment of HWRs by 62%.

We view these cost-effectiveness estimates as a starting point to quantify the importance of electricity spillovers, and caution against interpreting them as the welfare impacts from the deployment of HWRs. For a number of reasons, the net benefits calculated using our cost-benefit framework likely exceed those that would be calculated using a welfare analysis. First, our analysis does not account for potential costs incurred from efforts to conserve water and electricity, or an increase in the moral costs attached to water or electricity consumption. Instead, we assume that the change in benefits equals a customer's bill savings. Previous work finds that consumer willingness to pay for home energy reports is roughly 55 percent of the energy cost savings, implying that our use of bill savings overstates the consumer surplus from HWRs (Allcott and Kessler, 2015). Second, the disparity between marginal prices for electricity and water, and the social marginal costs to supply these goods, indicates that there may be reductions in consumer surplus from conservation. In our setting, the marginal price for residential electricity exceeds (on average) the marginal cost to supply it. These inefficient prices lead to underconsumption of electricity relative to the social optimum, and suggest that additional conservation will lead to further decreases in consumer surplus. A welfare analysis would account for differences between marginal prices and the social marginal cost of production, something that our cost-benefit framework does not do. Even with these caveats in mind, our benefit-cost framework provides a relevant and valuable launchpad with which to incorporate spillovers into policy evaluation.

8 Conclusions

This paper evaluates whether behavioral interventions spill over into unintended sectors through the lens of urban water conservation instruments and energy use. To do this, we designed and implemented a field experiment that allows us to measure the effect of HWRs on residential electricity use. We find that water conservation instruments induce conservation beyond the water sector, leading to a 1.3 to 2.2% reduction in summertime electricity use, where the magnitude of the response is comparable to that reported from the deployment of home energy reports. High-frequency data allow us to explore heterogeneity in the timing of the treatment effect, and reveal that electricity conservation is most pronounced during peak hours in the summer when electricity is most expensive to provide and marginal emissions from generation are higher.

We formalize two channels through which HWRs could affect electricity use: mechanical complementarities and behavioral choices. Under the former, electricity reductions are explained by actions that use both water and energy, such as doing a load of laundry. While mechanical complementarities account for some of the savings, empirical tests, household survey data, and simulation results all point to the likelihood that HWRs also alter consumer choices about electricity consumption. Treatment households reduce electricity consumption during hours of the day when no water consrevation occurs; increase electricity conservation as ambient temperatures increase; and report a significantly higher frequency of engagement in energy conservation actions. Results from an engineering simulation imply that only 25% of the electricity savings can be explained by mechanical complementarities.

To date, economists have primarily evaluated behavioral interventions using a partial equilibrium framework. The presence of cross-sectoral spillovers points to the importance of broadening the focus in the direction of a general equilibrium framework, both when considering the net benefits of these programs as well as their welfare impacts. In our setting, reductions in electricity use augment the net benefits of the water conservation instruments by 62% from a benefit to cost ratio of 2.91 to 4.7. We focus exclusively on one possible cross-sectoral spillover, but spillovers from this intervention may extend beyond energy. Moving forward, we should improve our understanding of the conditions and circumstances under which one would expect cross-sectoral spillovers to occur.

References

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

Allcott, H. and J. Kessler (2015). The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons. Working Paper.

Allcott, H. and S. Mullainathan (2010). Behavior and Energy Policy. *Science 327*(5970), 1204–1205.

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.

Angelucci, M. and G. D. Giorgi (2009). Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption? *American Economic Review 99*(1), 486–508.

Ashraf, N., D. Karlan, and W. Yin (2006). Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. *The Quarterly Journal of Economics 121*(2), 635–672.

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.

Balling, R. and P. Gober (2006). Climate Variability and Residential Water Use in the City of Phoenix, Arizona. *Journal of Applied Meterology and Climatology 46*, 1130–1138.

Bernheim, B. D., A. Fradkin, and I. Popov (2015). The Welfare Economics of Default Options in 401(k) Plans. *American Economic Review 105*(9), 2798–2837.

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.

Boomhower, J. (2016). Peer Effects in Energy Efficiency Program Participation. Working Paper.

Boomhower, J. and L. Davis (2016). Does Energy Efficiency Deliver at the Right Time? Energy Institute at Haas Working Paper 271.

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.

Brown, J., T. Hossain, and J. Morgan (2010). Shrouded Attributes and Information Suppression: Evidence from the Field. *Quarterly Journal of Economics 125*(2), 859–876.

Bruhn, M. and D. McKenzie (2009). In Pursuit of Balance: Randomization in Practice in Development Economics. *American Economic Journal: Applied Economics 1*(4), 200–232.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and Taxation: Theory and Evidence. *American Economic Review 99*(4), 145–1177.

Costa, D. and M. Kahn (2013). Energy Conservation "Nudges" and Environmental Ideology: Evidence from a Randomized Residential Electricity Field Experiment. *Journal of the European Economic Association 1*(3), 680–702.

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

DellaVigna, S. and J. Pollet (2009). Investor Inattention and Friday Earnings Announcements. *The Journal of Finance 64*(2), 709–749.

Dolan, P. and R. Metcalfe (2015). Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation. Working Paper.

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.

Farhi, E. and X. Gabaix (2017). Optimal Taxation with Behavioral Agents. Working Paper.

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.

Fletcher, J. and R. Marksteiner (2017). Causal Spousal Health Spillover Effects and Implications for Program Evaluation. *American Economic Journal: Economic Policy 9*(4), 144–166.

Fowlie, M., C. Wolfram, C. A. Spurlock, A. Todd, P. Baylis, and P. Cappers (2017). Default Effects, Follow-on Behavior and Welfare in Residential Electricity Pricing Programs. Energy Institute at Haas Working Paper 280.

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.

Gato, S., N. Jayasuriya, and P. Roberts (2007). Temperature and Rainfall Thresholds for Base Use Urban Water Demand Modelling. *Journal of Hydrology 337*, 364–376.

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.

Handel, B. (2011). Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review 103*(7), 2643–2682.

Harding, M. and D. Rapson (2014). Do Voluntary Carbon Offsets Induce Energy Rebound? A Conservationist's Dilemma. Working Paper.

Holland, S., E. Mansur, N. Muller, and A. Yates (2016). Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors. *American Economic Review 106*(12), 3700–3729.

Hutton, B. and D. McNeill (1981). The Value of Incentives in Stimulating Energy Conservation. *Journal of Consumer Research 8*(3), 291–298.

Ito, K. (2013). How Do Consumers Respond to Nonlinear Pricing? Evidence from Household Water Demand. Working Paper.

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

Jacobsen, G., M. Kotchen, and M. Vandenbergh (2014). The Behavioral Response to Voluntary Provision of an Environmental Public Good: Evidence form Residential Electricity Demand. *European Economic Review 56*(1), 946–960.

Jessoe, K. and D. Rapson (2014). Knowledge is (Less) Power: Experimental Evidence from Residential Energy Usage. *American Economic Review 104*(4), 1417–1438.

Johnson, E. J. and D. G. Goldstein (2003). Do Defaults Save Lives? *Science 302*(5649), 1338–1339.

Madrian, B. and D. Shea (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *The Quarterly Journal of Economics 116*(4), 1149–1187.

Miguel, E. and M. Kremer (2003). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica 72*(1), 159–217.

Mitchell, D. and T. Chesnutt (2013). Evaluation of East Bay Municipal Utility Districts Pilot of WaterSmart Home Water Reports. Working Paper.

Muller, N. and R. Mendelsohn (2009). Efficient Pollution Regulation: Getting the Prices Right. *American Economic Review 99*(5), 1714–1739.

Nolan, J., P. W. Schultz, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius (2008). Normative Influence is Underdetected. *Personality and Social Psychology Bulletin 32*(1), 913–923.

Novan, K. and A. Smith (2017). The Incentive to Overinvest in Energy Efficiency: Evidence from Hourly Smart-Meter Data. *Working Paper*.

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.

Sexton, S. (2015). Automatic Bill Payment and Salience Effects: Evidence from Electricity Consumption. *Review of Economics and Statistics 97*(2), 709–749.

Table 1: Balance Tests

	Control	Treatment	Difference	Households (Total)
Elec Use (kWh/mth)	735.11	729.07	6.03	7,341
	(6.72)	(5.29)	(8.57)	
Summer Elec Use (kWh/mth)	1020.76	1010.81	9.95	7,341
	(9.58)	(7.59)	(12.26)	
Water Use (gals/mth)	12480.36	12333.84	146.52	7,341
	(132.39)	(109.53)	(174.23)	
Year Built	1944.91	1944.87	0.03	7,105
	(0.27)	(0.21)	(0.34)	
Bedrooms	2.91	2.89	0.02	7,105
	(0.02)	(0.01)	(0.02)	
Bathrooms	1.92	1.93	-0.01	7,105
	(0.02)	(0.01)	(0.02)	
Square feet	1621.35	1620.14	1.21	7,105
	(12.66)	(10.15)	(16.31)	
TotalValue	391,670	394,493	-2,823.40	7,105
	(4915.66)	(3938.27)	(6329.74)	
Pool (Indicator)	0.23	0.22	0.00	7,105
	(0.01)	(0.01)	(0.01)	

Notes: The first two columns report means for control and treatment households with standard deviations in parentheses below. The third column displays the difference in means between treatment and control, with the standard error reported in parentheses below. The last column displays the total number of households included in the balance test. The assessor data on housing unit attributes are missing for 236 households. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2: Electricity Intent to Treat Effects
(Dependent Variable: Electricity Use (kWh/hr))

	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.010	-0.002	-0.028**	-0.017**	-0.049**	-0.029**
	(0.011)	(0.004)	(0.014)	(0.008)	(0.023)	(0.015)
Observations	59,989,591	59,840,480	14,321,113	14,292,331	3,616,831	3,609,534
Mean Control Group Use	1.00	1.00	1.26	1.26	1.99	1.99
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Hour of Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Electricity Use	No	Yes	No	Yes	No	Yes
Hours	All	All	All	All	Peak	Peak
Sample	5/15-5/16	5/15-5/16	5/15-8/15	5/15-8/15	5/15-8/15	5/15-8/15

Notes: The table reports intent to treat results from an OLS regression of hourly electricity use on assignment to the treatment. Columns (1) and (2) include all observations from May 15, 2015 to May 31, 2016. Columns (3) and (4) restrict the sample to the summer of 2015 (May 15 to August 30). Columns (5) and (6) further limit the sample to include only peak demand hours (3 PM to 8 PM). Pre-treatment electricity use controls include mean monthly electricity use in the summer, winter and year preceding treatment. Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

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

	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.585***	-0.528***	-0.572***	-0.448***	-0.724***	-0.625***
	(0.179)	(0.110)	(0.217)	(0.133)	(0.257)	(0.234)
Observations	68,502,709	68,412,093	19,932,386	19,932,386	4,984,984	4,984,984
Mean Control Group Use	12.9	12.9	15.4	15.4	14.2	14.2
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Hour of Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Water Use	No	Yes	No	Yes	No	Yes
Hours	All	All	All	All	Peak	Peak
Sample	5/15-5/16	5/15-5/16	5/15-8/15	5/15-8/15	5/15-8/15	5/15-8/15

Notes: The table reports intent to treat results from an OLS regression of hourly water use on assignment to the treatment. Columns (1) and (2) include all observations from May 15, 2015 to May 31, 2016. Columns (3) and (4) restrict the sample to the summer of 2015 (May 15 to August 30). Columns (5) and (6) further limit the sample to include only peak demand hours (3 PM to 8 PM). Pre-treatment water use controls include mean monthly water use in the summer, winter and year preceding treatment. Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

Table 4: Treatment Effect Heterogeneity and Outdoor Temperature (Dependent Variable: Electricity and Water Use)

	Electricity Use		Water Use	
	(1)	(2)	(3)	(4)
WaterSmart (<65F)	-0.002	-0.000	-0.505**	-0.244
	(0.017)	(0.017)	(0.212)	(0.390)
WaterSmart (65F-70F)	-0.011	-0.011	-0.360*	-0.823**
	(0.009)	(0.009)	(0.187)	(0.353)
WaterSmart (70F-75F)	-0.017**	-0.020***	-0.489***	-0.473
	(0.008)	(0.008)	(0.177)	(0.313)
WaterSmart (75F-80F)	-0.019**	-0.022**	-0.385**	-0.349
	(0.009)	(0.010)	(0.172)	(0.283)
WaterSmart (80F-85F)	-0.019	-0.017	-0.395**	-0.330
	(0.012)	(0.013)	(0.193)	(0.266)
WaterSmart (85F-90F)	-0.027	-0.036*	-0.486**	-0.405
	(0.019)	(0.019)	(0.193)	(0.262)
Observations	8,340,181	5,952,150	10,449,277	7,756,130
Calendar Date FE	Yes	Yes	Yes	Yes
Day of Week	No Watering	Watering	No Watering	Watering
Sample	5/15-8/15	5/15-8/15	5/15-8/15	5/15-8/15

Notes: The table reports intent to treat effects across 5F temperature bins. Columns (1) and (3) report results for electricity use and columns (4) and (5) report results for water use. The data include the period spanning May 15, 2015 to August 31, 2015. Day of Week indicates the days included in our sample; 'No Watering' restricts the sample to days of the week when outdoor watering was banned, and 'Watering' restricts the sample to days of the week when outdoor watering was allowed. All regressions include controls for the temperature bins 65F-70F, 70F-75F, 75F-80F, 80F-85F and $\geq$ 85F, hourly outdoor precipitation, and calendar date fixed effects. Pre-treatment electricity use controls in columns 1 and 2, and water use use controls in columns 3 and 4 include mean monthly use in the summer, winter and year preceding treatment.Standard errors in parentheses are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1% level.

Table 5: Survey Participation: Attrition and Balance Tests

Dependent Variable	(1) Completed Survey	(2) Completed Survey	(3) WS Treatment
Annual Water Use (1,000 gals.)	-0.009	0.015	-0.021
	(0.011)	(0.015)	(0.015)
Summer Water Use (1,000 gals.)	0.008	-0.007	0.012
	(0.006)	(0.008)	(0.009)
Winter Water Use (1,000 gals.)	-0.003	-0.007	0.005
	(0.006)	(0.008)	(0.006)
Annual Elec Use (1,000 kWh)	-0.347	0.564	-0.753*
	(0.306)	(0.454)	(0.446)
Summer Elec Use (1,000 kWh)	0.145	-0.328*	0.347*
	(0.136)	(0.197)	(0.190)
Winter Elec Use (1,000 kWh)	0.026	-0.283	0.330
	(0.164)	(0.253)	(0.257)
Observations	1504	746	867
Sample	WS	Control	Complete Survey
Response Rate	37.80%	38.23%	38.09%

Notes: Columns (1) and (2) report results from a linear probability model regressing survey completion on pre-treatment water and electricity use. The sample in column (1) is comprised only of households assigned to the WaterSmart treatment, and the sample in column (2) is comprised of households assigned to control. Column (3) reports results from a linear probability model regressing treatment on pre-treatment water and electricity use for all households. Standard errors are robust to heteroskedasticity. *, **, *** denote significance at the 10%, 5%, and 1% level.

Table 6: Survey Responses Across Control and Treatment
(Dependent Variable: Survey Response)

	(1)	(2)	(3)	(4)	(5)	(6)
	Turned Off Lights, TV, etc.	Adjusted Thermostat	Installed Programmable Thermostat	Air Dried Clothes	Reduced Dishwasher Use	Reduced Laundry Use
Water Smart	0.049*	0.003	0.031	0.023	0.022	0.008
	(0.028)	(0.033)	(0.032)	(0.032)	(0.035)	(0.034)
Observations	867	867	867	867	867	867
Control Response	0.80	0.71	0.26	0.26	0.54	0.67
WS Response	0.85	0.71	0.29	0.29	0.56	0.67

Notes: The table reports results from a linear probability model regressing the response to each survey question on assignment to the treatment. Controls include mean monthly water and energy use in the summer, winter and year preceding treatment. The bottom two rows report the mean survey response for control and treatment households, respectively. Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

Table 7: Household Energy-Water Use Appliance Model:
Electricity Savings (kWh/year)

Scenario	Mean	5th Percentile	95th Percentile	% TE Explained
All appliances	39.09	37.85	40.36	26.23
Clothes washer, dishwasher, and dryers only	139.73	135.86	143.68	93.78
All appliances: Electric water heater	119.63	117.95	121.33	81.43
Pool and spa use only	766.12	731.96	800.74	514.17

Notes: The table presents the mean, 5th percentile, and 95th percentile of the estimated electricity savings for each scenario in kWh/year. The last column reports the percent of the estimated electricity treatment effect (TE) explained by the mean electricity savings in each scenario. The estimated electricity TE is 149 kWh/year on average for each household based on our estimate in column (4) of Table 2.

Table 8: Electricity ITT Effects: No Pools
(Dependent Variable: Electricity Use (kWh/hr))

	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.014	-0.016	-0.032**	-0.043***	-0.048*	-0.068***
	(0.013)	(0.013)	(0.015)	(0.015)	(0.026)	(0.025)
Observations	45,068,294	44,974,419	10,810,160	10,792,459	2,728,578	2,724,094
Mean Control Group Use	0.91	0.90	1.17	1.17	1.87	1.87
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Hour of Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Electricity Use	Yes	Yes	Yes	Yes	Yes	Yes
Hours	All	All	All	All	Peak	Peak
Sample	5/15-5/16	5/15-5/16	5/15-8/15	5/15-8/15	5/15-8/15	5/15-8/15

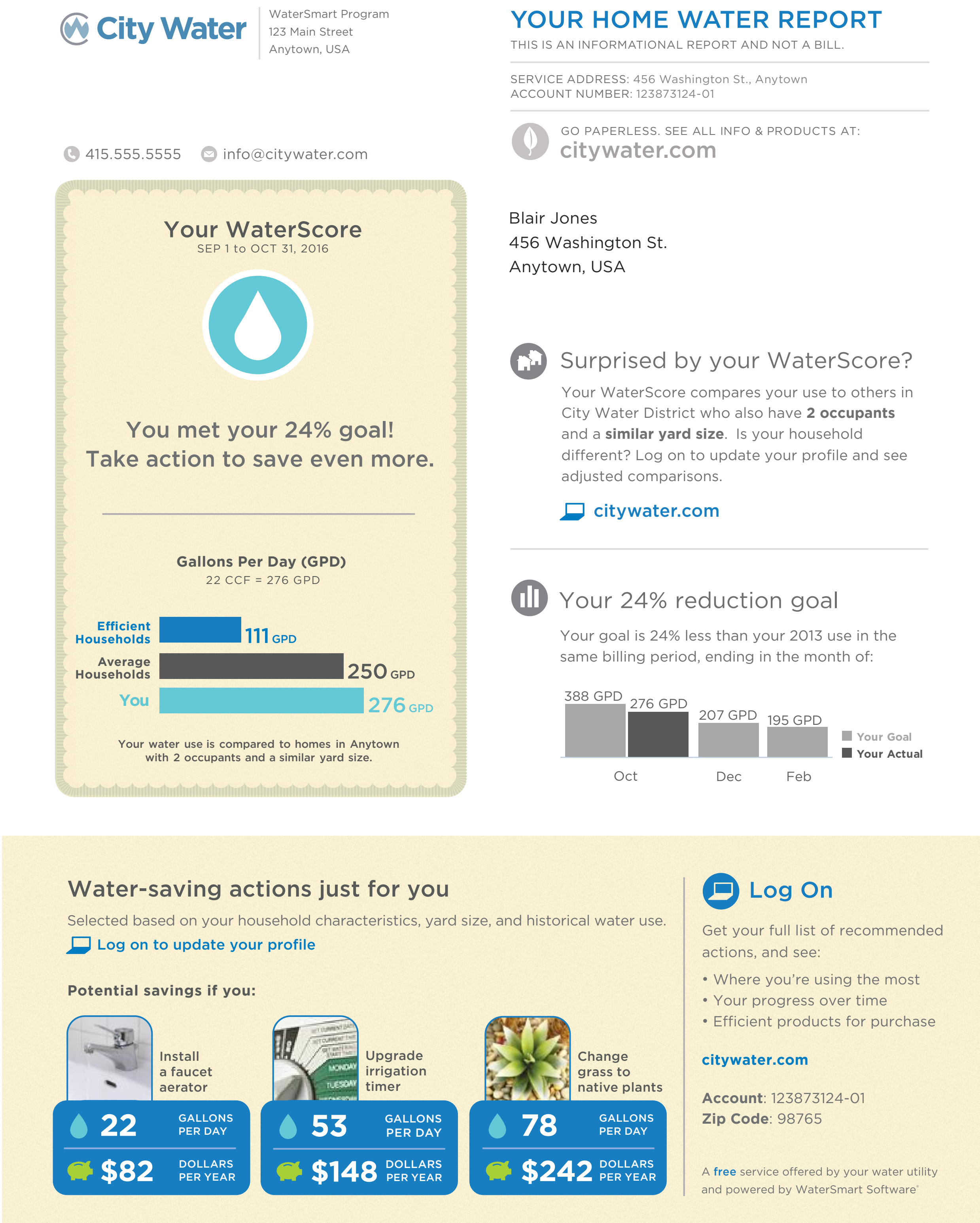
Notes: The table reports intent to treat results from an OLS regression of hourly electricity use on assignment to the treatment. The sample is comprised exclusively of households without pools. Columns 1 and 2 include all observations from May 15, 2015 to May 31, 2016. Columns 3 and 4 restrict the sample to the summer of 2015 (May 15 to August 30). Columns 5 and 6 further limit the sample to include only peak demand hours (3 PM to 8 PM). Pre-treatment electricity use controls include mean monthly electricity use in the summer, winter and year preceding treatment. Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

Table 9: Net Benefits of Home Water Reports ($/Household)

	Water (1)	Electricity (2)	Total (3)
Panel A: Year Average Treatment Effects			
(+) Customer Bill Savings	21.65	3.15	24.80
(-) Utility Revenue Loss	21.65	3.15	24.80
(+) Utility Wholesale Expenditure Savings	12.91	0.86	13.77
(+) Local Externalities Benefits	–	0.11	0.11
(+) Global Externalities Benefits	–	0.16	0.16
(-) HWR Cost	10	10	10
(=) Private Net Benefits	2.91	0.86	3.77
(=) Private & Public Net Benefits	2.91	1.13	4.04
Panel B: Monthly-by-Hour Average Treatment Effects			
(+) Customer Bill Savings	21.65	4.53	26.18
(-) Utility Revenue Loss	21.65	4.53	26.18
(+) Utility Wholesale Expenditure Savings	12.91	1.26	14.17
(+) Local Externalities Benefits	–	0.26	0.26
(+) Global Externalities Benefits	–	0.27	0.27
(-) HWR Cost	10	10	10
(=) Private Net Benefits	2.91	1.26	4.17
(=) Private & Public Net Benefits	2.91	1.79	4.70

Notes: The table reports our benefit-cost calculations on an annual, per household basis.

Figure 1: Generic Home Water Report



City Water

WaterSmart Program
123 Main Street
Anytown, USA

415.555.5555 info@citywater.com

YOUR HOME WATER REPORT

THIS IS AN INFORMATIONAL REPORT AND NOT A BILL.

SERVICE ADDRESS: 456 Washington St., Anytown
ACCOUNT NUMBER: 123873124-01

GO PAPERLESS. SEE ALL INFO & PRODUCTS AT:
citywater.com

Blair Jones
456 Washington St.
Anytown, USA

Your WaterScore
SEP 1 to OCT 31, 2016

You met your 24% goal!
Take action to save even more.

Gallons Per Day (GPD)
22 CCF = 276 GPD

Efficient Households 111 GPD
Average Households 250 GPD
You 276 GPD

Your water use is compared to homes in Anytown with 2 occupants and a similar yard size.

Surprised by your WaterScore?

Your WaterScore compares your use to others in City Water District who also have **2 occupants** and a **similar yard size**. Is your household different? Log on to update your profile and see adjusted comparisons.

citywater.com

Your 24% reduction goal

Your goal is 24% less than your 2013 use in the same billing period, ending in the month of:

388 GPD
276 GPD
207 GPD
195 GPD
Oct
Dec
Feb
Your Goal
Your Actual

Water-saving actions just for you

Selected based on your household characteristics, yard size, and historical water use.

Log on to update your profile

Potential savings if you:

Install a faucet aerator
22 GALLONS PER DAY
$82 DOLLARS PER YEAR

Upgrade irrigation timer
53 GALLONS PER DAY
$148 DOLLARS PER YEAR

Change grass to native plants
78 GALLONS PER DAY
$242 DOLLARS PER YEAR

Log On

Get your full list of recommended actions, and see:

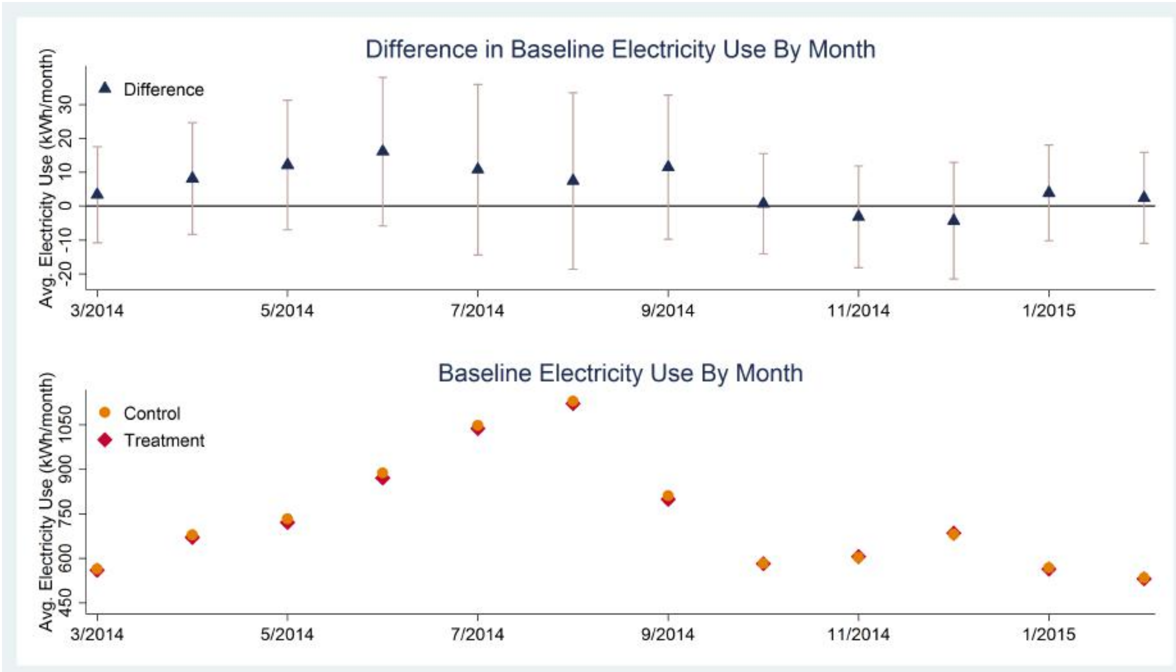
- Where you're using the most
- Your progress over time
- Efficient products for purchase

citywater.com

Account: 123873124-01
Zip Code: 98765

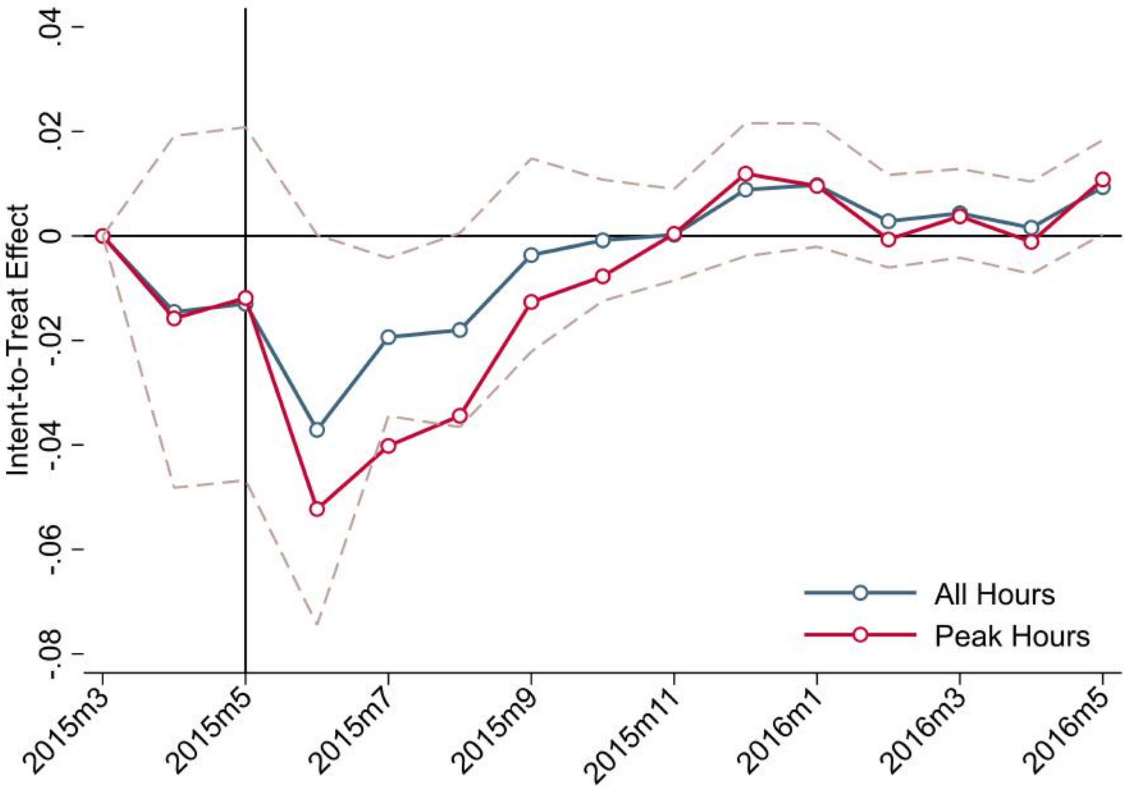
A **free** service offered by your water utility and powered by WaterSmart Software

Figure 2: Balance: Pre-treament Electricity Use by Month (kWh/mth)



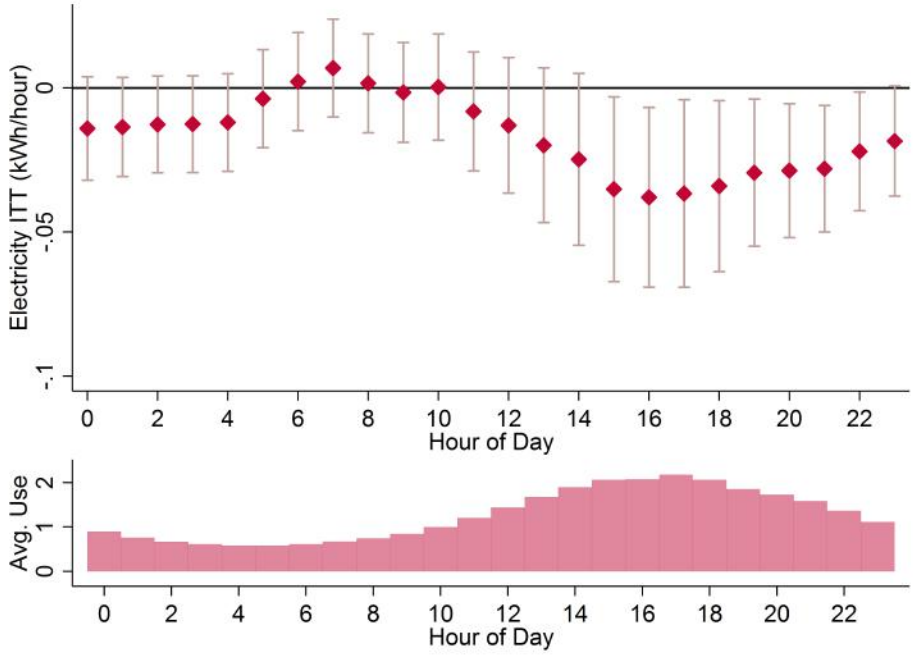
Notes: The upper portion of the figure plots the difference in mean monthly electricity use across control and treatment. The vertical lines are the 90% confidence intervals. The lower portion of the figure plots mean electricity use in a given month for control and treatment households.

Figure 3: Electricity Intent to Treat Effects Over Time

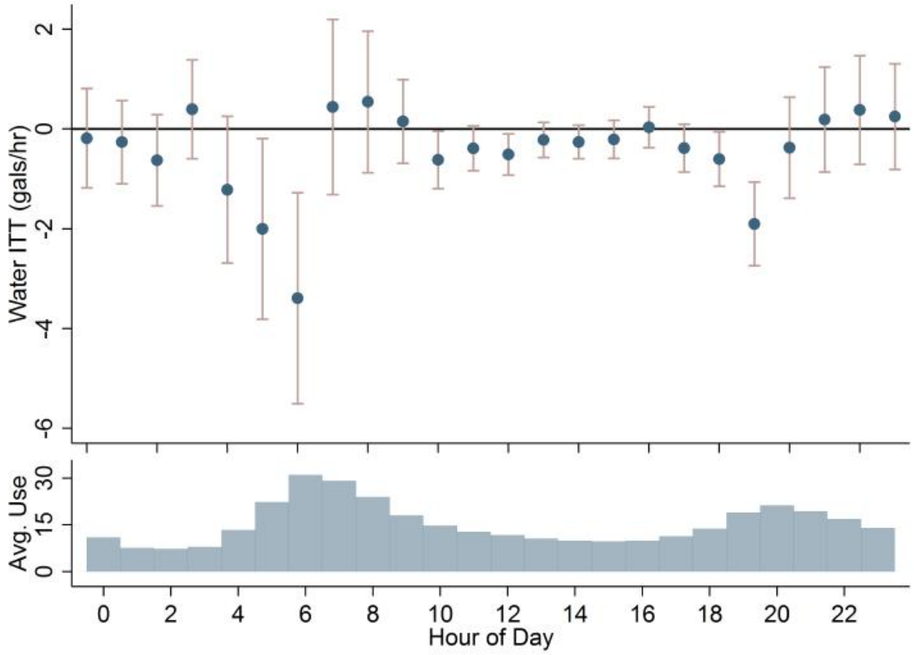


Notes: The figure plots intent to treat electricity effects (kWh/hr) for each month in the post-treatment period relative to March 2015. The blue and red lines plot estimates using all hours of the day and peak hours, respectively. The dashed lines are 90% confidence intervals for the 'all hours' treatment effects. The vertical line denotes May 2015, the first month in which HWRs were mailed to treatment households. All regressions include weather controls, hour of day fixed effects, calendar date fixed effects, and household pre-treatment electricity use controls.

Figure 4: Heterogeneity by Hour of Day



(a) Electricity Intent to Treat Effects (kWh/hr)



(b) Water Intent to Treat Effects (gals/hr)

Notes: The figure plots hourly intent to treat electricity effects in panel (a) and water effects in panel (b) from assignment to the WaterSmart treatment. The estimated treatment effects are denoted as diamonds in panel (a) and circles in panel (b). The 90% confidence intervals and mean hourly use for control households are plotted as well. All regressions include weather controls, calendar date fixed effects, and mean monthly electricity(water) use in the summer, winter and year preceding treatment.

A Appendix: For Online Publication

A.1 Simulation Model Details

We develop a model of residential households' water and electricity use to quantify the magnitude of electricity savings attributable to mechanical spillovers. The model uses individual response data from the 2009 California Statewide Residential Appliance Saturation Study (RASS).[20] The RASS reports appliance ownership statistics as well as estimates of appliance unit energy consumption (UEC) for over 25,000 households in California. UECs are household- and appliance-specific annual electricity use estimates based on households' appliance ownership portfolio and electricity billing data.[21]

The RASS data include information on households throughout the state of California. We restrict the data in two ways to ensure the sample we use is similar to those in our utility area. First, we limit the sample to single-family homes. Second, we include only households in Southern California.[22] With these restrictions in place, we are left with 8,442 household survey responses.

The RASS data include ownership and UECs for twenty-seven electric and ten natural gas appliances. Of those, we identified nine that either directly or indirectly use both water and electricity. Table A.1 lists these appliances along with the average ownership (saturation) rates and average UECs. The most commonly owned appliances that use electricity and water are clothes washers and dishwashers, which have 97% and 77% saturation rates, respectively. Large electricity using appliances include electric water heaters, electric dryers, evaporative (or swamp) coolers, pool pumps, well pumps, and electrically heated spas. However, the saturation rates of these appliances are generally below 20%. The table highlights the variation in appliance UECs and ownership. For example, electric dryers have a mean UEC of 625 kWh/year. However, the standard deviation is 624 kWh/year, reflecting significant variation in the intensity of dryer use across households.

[20]More information on the RASS survey is available at http://www.energy.ca.gov/appliances/rass/. We thank Glen Sharp from the California Energy Commission for giving us access to anonymous individual response data.

[21]More information on the methods used to calculate appliance UECs are available at http://www.energy.ca.gov/reports/400-04-009/2004-08-17_400-04-009VOL2B.PDF.

[22]Specifically, we restrict the data to households located in the CEC's building climate zones 5, 6, 7, 9, 10, 14 and 16. For a map of the CEC's building climate zones, see http://www.energy.ca.gov/maps/renewable/BuildingClimateZoneMap.pdf.

Simulating the average reduction in household electricity use attributable to appliances that use both water and electricity involves dealing with several layers of uncertainty. These include (1) household appliance ownership and the intensity of appliance use; (2) the impact of WaterSmart treatment on household water use; and (3) appliance use patterns. We use the RASS saturation rates and UEC data to address the first source of uncertainty and our estimated water treatment effect to address the second. To address the last issue, how households changed their appliance use behavior, we consider four behavioral scenarios that we elaborate on below. Our simulation procedure proceeds as follows:

1. Draw 5,046 Bernoulli random appliance ownership variables based on estimated saturation rates for each appliance.[23]

2. Draw 5,046 UECs from the estimated distribution of UECs for each appliance, and multiply by ownership indicators.

3. Calculate the energy savings from water-saving scenarios 1-5. Collect average electricity savings across all households.

4. Repeat (1)-(3) 10,000 times.

Step 1 generates a vector of 5,046 zeros and ones for each appliance, where the frequency of ones is equal in expectation to the saturation rates in Table A.1. Step 2 converts the ownership indicators into annual energy use values for each appliance and household. We estimate the distribution of UECs for each appliance by fitting a beta distribution using a maximum likelihood estimator to the UECs from the RASS data. Using a beta distribution ensures that all UECs drawn from each distribution are positive. Steps 3 and 4 give the distribution of mean changes in electricity use to determine the distribution of average electricity savings under each scenario.

To implement step 3, converting water reductions to electricity reductions, we use the estimated water treatment effect from column (4) of Table 3. In all scenarios, we assume that water and energy use in appliances is proportional, i.e., a 1% reduction in water use by an appliance leads to a 1% reduction in energy use. To account for heterogeneity in the estimated treatment effect, we draw a treatment effect from a normal distribution with mean 0.034, the average treatment effect divided by the mean hourly water use, and standard deviation 0.015. The resulting electricity savings estimates are specified in kWh/year, which

[23] We use a sample of 5,046 since this is the size of our treatment group.

we compare to our estimated average treatment effect from column (3) of Table 2, which is 298 kWh/year.

We consider the following four scenarios:

Scenario 1: Equal reduction across all appliances. Households reduce water use across all appliances equally. This amounts to reducing use of all appliances on average by 3.4%.

Scenario 2: Reductions in indoor appliance water use. We assume that all water savings come from reductions in dishwasher, clothes washer, and dryer use. According to the RASS data, households in our sample on average run 256 loads of laundry and 133 dishwasher loads each year. While the RASS data do not include information on appliance water use, Home Water Works, a project sponsored by the Alliance for Water Efficiency, found that old clothes washers (dishwashers) use 40-45 (10-15) gals/load and newer, more efficient clothes washers (dishwashers) use 14-25 (<5.5) gals/load. To be conservative, we assume clothes washers use 35 gals/load and dishwashers use 8 gals/load. Under these assumptions, the average water treatment effect is 4,572 gals/year. To achieve these reductions, households would need to curtail their use of clothes washer and dishwasher use by 45.6% on average. This scenario further assumes that households would reduce electric dryer use by the same amount, and includes savings from reductions in the use of water heaters and well pumps.

Scenario 3: Electric Hot Water Heaters. Electric hot water heaters are large energy users. While the RASS data confirms information from our partner utility that ownership of electric hot water heaters is low, to understand if electric hot water heaters are driving our results, we replicate Scenario 1 assuming that all households own an electric water heater.

Scenario 4: Reductions in outdoor appliance water use. The drought saw mandates for substantial reductions in water use among residential households. This included some cities passing regulations restricting construction of new pools and banning homeowners from refilling their existing pools.[24] To explore whether reductions in pool and spa appliance use could explain our findings, we consider a scenario where all households that own a pool stop using them entirely. This amounts to a 100% reduction in pool pumps and spa use among households owning them.[25]

[24] See http://www.latimes.com/local/california/la-me-pool-drought-20150602-story.html.

[25] We also include reduction in well pump use in this scenario.

Table A.1: Household Energy-Water Use Appliance Model: Ownership and UECs

	UEC (kWh/year)	Saturation Rate
Electric Water Heater	2,936.94	0.06
	(1,431.80)	(0.23)
Clothes Washer	94.28	0.97
	(175.27)	(0.16)
Electric Dryer	625.25	0.21
	(624.05)	(0.41)
Dish Washer	76.27	0.77
	(49.50)	(0.42)
Evaporative Cooler	669.36	0.09
	(340.53)	(0.28)
Pool Pump	3,604.22	0.18
	(1,068.90)	(0.41)
Spa	300.43	0.10
	(230.47)	(0.30)
Spa Electric Heat	995.94	0.08
	(664.12)	(0.28)
Well Pump	518.01	0.03
	(214.13)	(0.17)

Notes: The table presents mean and standard deviation statistics for all household appliances included in the 2009 California RASS survey that use both water and electricity. 'Saturation Rate' is the average ownership, and UEC=unit energy consumption in kWh per year. (Source: 2009 CEC RASS)

A.2 Net Benefit Calculations: Regression Results

Table A.2: Effects of HWRs on Electricity Expenditures and Externalities

Dep. Variable	Retail Expend.		Wholesale Expend.		CO_2 Damages		Local Damages	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WaterSmart	-0.0018	-0.0004	-0.0004	-0.0001	-0.0001	-0.00002	-0.0002	-0.00001
	(0.0020)	(0.0007)	(0.0004)	(0.0001)	(0.0002)	(0.00006)	(0.0002)	(0.0001)
Yearly Savings	15.40	3.15	3.34	0.86	1.22	0.16	1.53	-0.11
Std. Error	(17.14)	(6.30)	(3.21)	(1.21)	(1.47)	(0.54)	(1.93)	(0.71)
Observations	59,989,591	59,840,480	59,982,748	59,833,656	59,989,591	59,840,480	59,989,591	59,840,480
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour of Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-treatment Electricity Use	No	Yes	No	Yes	No	Yes	No	Yes

Notes: The table reports intent to treat results from an OLS regression of either (i) retail electricity expenditures, (ii) wholesale electricity expenditures, (iii) CO_2 damages from electric generation for CA, or (iv) local damages from electric generation for CA on assignment to the treatment. Yearly savings are calculated by multiplying each ATE by 24 hours and 365 days. Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

Table A.3: Effects of HWRs on Water Expenditures

Dep. Variable	Retail Expend.		Wholesale Expend.	
	(1)	(2)	(3)	(4)
WaterSmart	-0.003***	-0.002***	-0.002***	-0.001***
	(0.001)	(0.001)	(0.000)	(0.000)
Yearly Savings	24.10	21.65	14.37	12.91
Std. Error	(7.52)	(6.70)	(4.35)	(3.86)
Observations	63,034,882	62,852,349	63,034,882	62,852,349
Weather Controls	Yes	Yes	Yes	Yes
Hour of Day FE	Yes	Yes	Yes	Yes
Calendar Date FE	Yes	Yes	Yes	Yes
Pre-treatment Electricity Use	No	Yes	No	Yes

Notes: The table reports intent to treat results from an OLS regression of either (i) retail electricity expenditures, (ii) wholesale electricity expenditures, (iii) CO_2 damages from electric generation for CA, or (iv) local damages from electric generation for CA on assignment to the treatment. Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

A.3 Logarithm Regression Results

Table A.4: Electricity Intent to Treat Effects
(Dependent Variable: Log Electricity Use)

	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.014	-0.006	-0.020*	-0.013	-0.020	-0.010
	(0.012)	(0.006)	(0.012)	(0.009)	(0.014)	(0.011)
Observations	59,890,587	59,741,506	14,287,890	14,259,108	3,608,205	3,600,908
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Hour of Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Hours	All	All	All	All	Peak	Peak
Sample	5/15-5/16	5/15-5/16	5/15-8/15	5/15-8/15	5/15-8/15	5/15-8/15

Notes: The table reports intent to treat results from an OLS regression of the log of hourly electricity use on assignment to the treatment. Columns 1 and 2 include all observations from May 15, 2015 to May 31, 2016. Columns 3 and 4 restrict the sample to the summer of 2015 (May 15 to August 30). Columns 5 and 6 further limit the sample to include only peak demand hours (3 PM to 8 PM). Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.

Table A.5: Water Intent to Treat Effects
(Dependent Variable: Log Water Use)

	(1)	(2)	(3)	(4)	(5)	(6)
WaterSmart	-0.046***	-0.047***	-0.037***	-0.037***	-0.047***	-0.047***
	(0.013)	(0.011)	(0.013)	(0.012)	(0.015)	(0.014)
Observations	63,034,882	62,651,855	18,319,342	18,205,407	4,581,513	4,553,021
Mean Control Group Use						
Weather Controls	No	Yes	No	Yes	No	Yes
Hour of Day FE	No	Yes	No	Yes	No	Yes
Calendar Date FE	No	Yes	No	Yes	No	Yes
Hours	All	All	All	All	Peak	Peak
Sample	5/15-5/16	5/15-5/16	5/15-8/15	5/15-8/15	5/15-8/15	5/15-8/15

Notes: The table reports intent to treat results from an OLS regression of the log of hourly water use on assignment to the treatment. Because there are many hours of the day with zero consumption, we use add one gallon per hour to every households' consumption. Columns 1 and 2 include all observations from May 15, 2015 to May 31, 2016. Columns 3 and 4 restrict the sample to the summer of 2015 (May 15 to August 30). Columns 5 and 6 further limit the sample to include only peak demand hours (3 PM to 8 PM). Standard errors are clustered at the household. *, **, *** denote significance at the 10%, 5%, and 1%level.