Do Pecuniary Incentives Crowd Out Social Pressure?

Jeremy West[a,b]* Robert W. Fairlie[a,c] Bryan Pratt[a] Liam Rose[d]

[a]University of California at Santa Cruz
[b]The E2e Project
[c]NBER
[d]Stanford University

January 2019

Abstract

Prior research indicates that financial encouragement often displaces intrinsic motivations, but not whether pecuniary interventions also crowd out external social pressure for prosocial behavior. Studying a large randomized field experiment across a sharp change in policy regimes, we find that normative peer comparisons cause significant water conservation invariant to the strength of pecuniary incentives targeting the same behavior. Dispelling potential threats to interpretation, we confirm using a regression discontinuity design that the financial incentive binds and also reduces water consumption. These findings demonstrate that strong economic incentives do not crowd out social pressure, an actionable insight for increasingly multidimensional policies.

JEL: D12, K42, Q25

*West (corresponding author): westj@ucsc.edu. Fairlie: rfairlie@ucsc.edu. Pratt: pratt@ucsc.edu. Rose: liamrose@stanford.edu. We are eminently grateful to WaterSmart Software for implementing the randomized *Home Water Reports* and to William Holleran in particular for assistance with the associated data. This work benefited from valuable comments from Joshua Abbott, Kelly Bishop, Daniel Brent, Lopamudra Chakraborti, Jesse Cunha, Carlos Dobkin, Daniel Friedman, Brent Haddad, Nicolai Kuminoff, Natalia Lazzati, Justin Marion, Jonathan Meer, Adam Millard-Ball, Steven Puller, Jonathan Robinson, Jason Scorse, Kerry Smith, John Sorrentino, Alan Spearot, Forrest Williams, and numerous seminar participants. Any errors, opinions and conclusions in this study are our own and do not necessarily represent the position of WaterSmart Software nor any other parties. The authors have no material financial interests related to this research.

1 Introduction

Society relies heavily on both pecuniary and non-pecuniary incentives to encourage charitable giving, resource conservation, and other prosocial behavior. Recently, institutions often combine the two tactics to internalize externalities and increase voluntary provision of public goods. For example, in large-scale media campaigns targeting risky driving behaviors such as speeding and distracted driving, the U.S. National Highway Traffic Safety Administration emphasizes both the "financial penalties from violating laws" and the moral imperative not to "put everyone around you at risk."[1] One outstanding question that is critical for the efficacy of such two-fold approaches is how pecuniary incentives interact with non-pecuniary incentives targeting the same behavior.

An extensive literature finds that pecuniary incentives to contribute to a public good often have psychological or behavioral aspects that displace intrinsic motivations to contribute to the same public good. As characterized by Gneezy et al. (2011), "Monetary incentives have two kinds of effects: the standard direct price effect, which makes the incentivized behavior more attractive, and an indirect psychological effect [that] in some cases works in an opposite direction to the price effect and can crowd out the incentivized behavior." Such crowding out of prosocial behavior by financial incentives has been demonstrated for a wide range of outcomes, from blood donations and charitable giving to day care pick up timeliness and nature conservation (e.g. Titmuss, 1970; Gneezy and Rustichini, 2000; Bénabou and Tirole, 2006; Ariely et al., 2009; Bowles and Polanía-Reyes, 2012; Rode et al., 2015). Given this evidence of crowd out of intrinsic motivations, it is natural to ask: does strengthening pecuniary incentives also crowd out non-pecuniary incentives for prosocial behavior?

In this study, we examine whether the effects of external social pressure – the influence on people by their peers – are crowded out by a sharp increase in pecuniary incentives targeting the same behavior. The specific intervention we study involves peer comparisons, a form of non-pecuniary incentive used widely to encourage behaviors such as charitable giving, civic participation, resource conservation, and workplace productivity (Frey and Meier, 2004; Gerber et al., 2008; Mas and Moretti, 2009; Allcott, 2011; Ferraro and Price, 2013; Castillo et al., 2014; Smith et al., 2015). The context we study is residential water use, and our empirical research question is whether strengthening pecuniary incentives to conserve water reduces the effectiveness of normative peer comparisons that encourage water conservation. To our knowledge, our study is the first to directly test whether pecuniary incentives crowd

[1] www.nhtsa.gov/risky-driving.

out social pressure by comparing the efficacy of randomized social pressure before versus after an exogenous strengthening of financial incentives targeting the same behavior.

To empirically investigate this question, we partnered with WaterSmart Software and a Southern Californian water utility that measures hourly residential water consumption using Advanced Metering Infrastructure (AMI). Via a field experiment, WaterSmart substantially increased social pressure for randomly-selected households to conserve water by providing them with Home Water Reports (HWR) comparing their water consumption with that of their neighbors. In this baseline condition, HWR should be highly effective; prior studies testing similar treatments in Georgia and California generally find water conservation effects ranging from three to six percent (Ferraro et al., 2011; Ferraro and Price, 2013; Bernedo et al., 2014; Brent et al., 2015; Jessoe et al., 2017).[2] Moreover, the literature finds that this effectiveness is attributable primarily to the social pressure component, rather than to any technical or financial information included in HWR (Ferraro and Price, 2013; Mitchell and Chesnutt, 2013; Allcott and Rogers, 2014; Brent et al., 2015, 2017). Altogether, the existing research strongly supports the possibility for there to be significant behavioral crowding out of HWR from strengthening financial incentives for conservation.

Several months into the field experiment, having established baseline HWR effects, the utility significantly strengthened pecuniary incentives for households to conserve water. Leveraging the AMI data, the utility undertook an innovative approach to enforcing existing mandatory irrigation policies by automating detection of irrigation violations and issuing notice to offenders. These violation notices, discussed in Section 2, warned offending households of fines for continuing to violate municipal irrigation policies and made it clear that the enforcement was now fully automated through computer algorithms. As shown in Figure 1, within one week this enforcement decision increased the share of households that had ever been warned from 4.6 to 39.2 percent. This sharp change in the policy regime and associated substantial increase in pecuniary incentives for conserving water affords us a causal test of crowd out of social pressure from the randomized HWR.

Using the field experiment to identify intent-to-treat effects in administrative data on residential water consumption among the universe of utility customers, we find that the randomized HWR reduced average household water use by 76 gallons (3.2 percent) per week prior to the intensification of the pecuniary incentive. After the automation of pecuniary enforcement, we find HWR reduced average water consumption by 75 gallons per week – an

[2] These effects of peer comparisons for water conservation are somewhat stronger than those found in the literature on similar interventions related to energy use by Opower, which tend to range from one to three percent (e.g. Allcott, 2011; Ayres et al., 2013; Costa and Kahn, 2013; Allcott and Rogers, 2014).

economically and quantitatively identical treatment effect. Using difference in (randomized) differences tests to more formally quantify the change in treatment effects between neighboring weeks before and after the strengthened pecuniary incentives, we find no evidence that pecuniary incentives crowd out the effectiveness of social pressure for water conservation.

We address several potential threats to inference. The key identification assumption in our test for crowd out is that the effect of randomized HWR would have held constant counterfactually, had the water utility not strengthened the pecuniary incentive. Prior research shows that peer comparison effects tend to gradually attenuate over time, particularly when the interventions are halted (Ferraro et al., 2011; Allcott and Rogers, 2014; Bernedo et al., 2014; Ito et al., 2018). This concern seems minimal for our analysis, which focuses only on the initial few months of the intervention, entirely within a single summer water season, during which HWR continued to be sent monthly to treated households. A related consideration is that pecuniary incentives might crowd out social pressure simply due to limited scope for residential water conservation. Based on institutional details and prior research, it seems unlikely there would be significant crowd out for purely mechanical reasons, as we discuss in Section 2 (Castledine et al., 2014; Baker, 2017; Brelsford and Abbott, 2018; Pratt, 2018). Moreover, if HWR treatment effects did attenuate for any aforementioned reasons, this would bias our estimates towards instead of against finding crowd out.

Given that we find that the effects of social pressure are invariant to the strengthened pecuniary enforcement, the most serious threat to inference would be if the pecuniary incentive were non-binding. Akerlof and Dickens (1982) argue that changes to regulative enforcement cause individuals to switch their focus from "self-motivation to obey the law" towards the severity of the punishment. Although the irrigation violation notices explicitly assert they are a precursor to monetary penalties, recipients might have perceived the warnings to be cheap talk by the water utility.[3] To negate concerns that our finding of no crowd out is due to the "strengthened" pecuniary incentive being weak and ineffective, we examine the direct effects of the automated violation notices on water conservation. Using a regression discontinuity design based on a somewhat arbitrary cutoff in the noncompliance algorithm, we find that violation notices cause significant reductions in water consumption: the local average treatment effect is a reduction of 550 gallons (29 percent) per week. In addition, the violation notices further increased policy compliance by shifting some residential water consumption within the week into irrigation-permitted time periods. Together, these findings

[3] In practice, irrigation restrictions are rarely enforced because of enforcement limitations. State reports show that during the drought we study, most water agencies that restricted irrigation to two days per week never issued a single penalty (California State Water Resources Control Board, www.waterboards.ca.gov).

demonstrate by revealed preference that the change in pecuniary incentive is economically significant, supporting our interpretation of the experimental estimates that the efficacy of social pressure is not crowded out by strengthening pecuniary incentives.

Our study makes several contributions to the literature. Most broadly, we provide insights regarding the interaction of pecuniary and non-pecuniary incentives. The three most closely-related studies are by Ito et al. (2018), who estimate separately and compare the effects of dynamic electricity price increases and those of peer comparisons, Pellerano et al. (2017), who find that layering on information about potentially relevant economic incentives may diminish the impact of normative appeals to conserve energy, and List et al. (2017), who find that adding a financial rewards program to a standard home energy report treatment provides additional conservation.[4] None of these studies directly tests whether monetary incentives crowd out or partially crowd out social pressure using a comparison of social pressure effects before and after exogenously increasing financial incentives.

In addition, our study adds to a growing literature that examines social pressure as a form of policy. Governments are often limited in their use of first-best Pigouvian remedies, especially in energy and water conservation contexts. Various regulative and economic strategies for conserving water have been explored, including use restrictions and price increases; however, Grafton and Ward (2008) show that mandatory restrictions can be welfare-reducing, and numerous studies find that water demand is fairly inelastic to price changes (Olmstead et al., 2007; Ito, 2013). Moreover, utilities typically face substantial regulatory and political constraints to adjusting prices or restricting use. In response to these limitations, agencies are increasingly using non-pecuniary interventions that operate outside the regulatory framework, and the use of social pressure – such as providing peer comparisons of consumption – is steadily rising among water utilities and in other sectors.

Interventions based around social pressure are particularly attractive due to low implementation costs for the policymaking agency. However, Allcott and Kessler (2019) find that the majority of resource conservation nudge recipients are unwilling to pay the marginal social cost of the nudge, indicating significant economic welfare considerations for social pressure policies. Our evidence of additive conservation benefits from pecuniary and non-pecuniary incentives is thus especially relevant for policymakers as they incorporate a broader variety of policies and optimize across an increasingly multidimensional set of interventions.

[4]Within a broader study of various financial incentives for energy conservation, Gillan (2018) also tests a normative treatment. Given that the study finds insignificant effects of the normative message independent of the price incentive, it is unclear how to interpret the estimated interaction effect of the combined treatments.

2 Empirical setting and research design

California has a history of extreme and persistent variation in precipitation, with the Coastal Southern California hydrological region suffering multiple periods of extended drought conditions during the past decade (see Appendix Figure A1). The vast majority of annual precipitation falls during the winter months throughout the state and in Coastal Southern California in particular. With a dry 2011-2012 winter followed by an even drier 2012-2013 winter, authorities began implementing a wide range of conservation measures. This paper focuses on two of the most common interventions, which succinctly capture two major pillars of water policy: social pressure to encourage voluntary conservation and pecuniary incentives to reduce consumption. The primary intervention we study – normative peer comparisons as social pressure for voluntary conservation – is relatively contemporary in origin. The second intervention we evaluate – pecuniary incentives tied to day-of-week and time-of-day outdoor water use restrictions (hereafter, DOWR) – has a long heritage in water conservation, although automating the associated enforcement is a first of its kind.

In partnership with WaterSmart Software and Burbank Water and Power (BWP), we use a field experiment to study the effects of normative peer comparisons.[5] Nearly 17,000 single-family households in Burbank, California are included in the randomized control experiment we study. Treated households were provided with WaterSmart HWR by mail or email, with the timing of initial treatment rolled out over the monthly water billing cycle during late April through mid May 2015. Notably, households in the "Experimental" treatment arm began to receive HWR at least six weeks prior to the heightened enforcement of irrigation restrictions in early July. Very few treated households chose to opt-out, and nearly all continued to receive monthly HWR for at least several months after their initial treatment and throughout the time period we include in our empirical analyses.

A HWR has a few components (see Appendix Figure A2 for an example report), with the core component being the normative comparison of the treated household's water consumption with that of a peer group formed from neighboring households with the same number of occupants and similar irrigable area. A household that used more water than an "average" comparison household would receive a frowning face on a red water drop and be informed that it "used more water than most of [its] neighbors." If a household used less water than an "efficient" household it would receive a smiling green water drop. Yellow drops were given to

[5]Operating in a parallel industry to Opower/Oracle Utilities for energy utilities, WaterSmart Software provides assistance to water utilities in California and around the world through analyzing and interpreting data on water use and through providing information to customers.

households using between the efficient level and the median level of water use.[6] In all three cases, households were provided with their water use in gallons per day (GPD), the median water use in GPD, and the efficient level of water use in GPD for comparison "neighbor" households.

Although a HWR includes some technical or financial information, Ferraro et al. (2011) experimentally tested the effectiveness of this information alone in HWR and found it ineffective in the absence of the peer comparison. The broader literature also supports that the normative social comparison is by far the most effective component of HWR, including research experimentally testing WaterSmart HWR nearly identical to the treatments examined in our study (Ferraro and Price, 2013; Mitchell and Chesnutt, 2013; Allcott and Rogers, 2014; Brent et al., 2015, 2017). Moreover, as these "social comparisons impose a moral cost on consumption," we join the existing literature in interpreting the HWR treatment primarily as a form of social pressure for voluntary resource conservation (Brent et al., 2017).

In contrast to the contemporary novelty of HWR, DOWR are a common and longstanding conservation policy for water utilities. Although the benefits from water conservation do not vary across hours of the week, there are several institutional, horticultural, and behavioral reasons to impose restrictions by time of day and day of the week. First, DOWR generally prohibit irrigation between a few hours after sunrise and a few hours before sunset, which minimizes water lost to evaporation (Christiansen, 1942). Furthermore, spacing out the days on which irrigation is allowed ensures water can be spread efficiently for the benefit of the plants. In addition, there are enforcement justifications for prohibiting outdoor uses on certain days. Before the introduction of AMI, "smart meters" which record high-frequency consumption data, the only method of detecting irrigation violators was visual inspection; in other words, either a water agency employee or an informant neighbor of the violator must visually observe the illegal irrigation. Prohibiting outdoor water use on specific days of the week – rather than, say, restricting total irrigation volume – facilitates enforcement, since an extensive margin of outdoor use on a given day is much more easily observed. Finally, outdoor water use comprises a large share of residential water use and provides the potential to conserve water without generating negative health or safety consequences.[7]

During the drought, BWP joined many other water utilities in implementing irrigation

[6] The HWR does not clarify the specific thresholds, but the "efficient" neighbor benchmark is based on the 20th percentile of within-neighbor-group consumption, and the "average" neighbor benchmark is actually the 55th percentile of within-neighbor-group consumption, ensuring that most homes are considered better than "average."

[7] cf. California Department of Water Resources' *California Water Plan Update* (2013), Volume 3, Chapter 3: Urban Water Use Efficiency.

restrictions. On May 14, 2015, the Burbank City Council approved the implementation of stricter Stage 3 restrictions, which included limiting outdoor water use to Tuesdays and Saturdays. Notably, BWP initially enforced these DOWR using only the traditional method of visual inspection, resulting in only a very small number of households determined to be in violation. As seasonal dryness accumulated, the violation count increased slightly, but enforcement remained extremely low through the end of June. Then, after six weeks of the post-treatment period in our field experiment, BWP additionally undertook the unprecedented approach of using AMI data to algorithmically detect potential violators. BWP automatically sent more than one-third of single-family residential accounts in the district a violation notice during the first week of July 2015, as shown in Figure 1. These violation notices (shown in Appendix Figure A3) warned that customers would be fined $100 if they continued to irrigate more than two days a week. Fines increased to $200 for a second violation and $500 for a third violation. The notices also clearly indicate that detection of violations was now algorithmic in nature. As the remainder of the content included in the violation notices was already being publicized widely throughout the community, these factors afford this treatment an interpretation as being a shock to household beliefs about detection and enforcement probabilities for a pecuniary threat. Because of potential spillovers across households, both directly and through media coverage of the heightened enforcement in outlets such as the *Los Angeles Times*, we interpret the treatment broadly as a substantial strengthening of pecuniary incentives.[8]

Both of these policies were introduced into a landscape filled with media coverage and appeals from the State to conserve water, and BWP was facing threats of State-mandated pecuniary penalties tied to conservation goals.[9] Furthermore, recent evidence suggests that intense media coverage can directly influence residential water consumption (Quesnel and Ajami, 2017). As such, one potential concern for our study is that the strengthened pecuniary incentives could crowd out the effects of social pressure for purely mechanical reasons of limited scope to further reduce water consumption. We do not view this as a significant concern for several reasons. For one, the policies we study targeted primarily outdoor water use, for which there is ample scope for additional conservation (Castledine et al., 2014; Baker, 2017; Brelsford and Abbott, 2018; Pratt, 2018). In addition, the households treated by the automated violation notices are (unsurprisingly) almost exclusively included in the "above

[8]For example, www.latimes.com/tn-blr-bwp-to-step-up-water-saving-efforts-20150705-story.html.

[9]California passed and ultimately enforced regulations providing for fines of $10,000 per day for water agencies that did not generate sufficient mandated conservation, as noted by local media such as www.mercurynews.com/2015/04/28/water-wasting-fines-of-10000-proposed-by-gov-jerry-brown/.

average" WaterSmart tier. As these tiers are defined within household size and irrigable area, above average consumption primarily reflects choices of the household rather than need. In any case, mechanical crowd out for reasons of scope would attenuate estimates of the HWR treatment effect and bias our estimates towards finding crowd out, whereas we find no evidence of crowd out.

Our research designs allow us to examine the impacts of social pressure and pecuniary incentives within a given period across randomly and quasi-randomly assigned groups. One consideration for this particular setting is the external validity of conclusions drawn at the height of severe drought in a drought-prone region. However, this utility is representative of water utilities in Southern California, and studying policy during such an event is critical for understanding the effects of related policies in the contexts in which they are invoked.

Fresh water availability remains one of the most pressing environmental and economic concerns in many regions around the world. The United Nations forecasts that two-thirds of the world's population will live with water stressed conditions by 2025, and that the outlook will only worsen under existing climate change scenarios.[10] At a municipal level, there are also growing concerns about cities sinking as they over-extract local water resources.[11] As weather patterns are becoming more erratic, severe droughts such as the one experienced during the past several years in California, other Southwestern states, and many other regions around the world are becoming increasingly common.

3 Data and balance checks

Our study leverages the fairly recent introduction of AMI for residential water service, which records high-frequency data on water consumption. Unlike with smart meters for electricity, AMI measurement is relatively rare for household water use, although adoption of the technology exhibits a steep upward time trend. In 2015, only about seven million smart meters for water had been installed in the United States, compared to about 68 million smart electricity meters.[12] Our partner utility, Burbank Water and Power (BWP), had installed smart water meters for eligible households throughout their service area approximately one year prior to the field experiment we evaluate.

Hourly AMI data offer several significant advantages over traditional monthly water

[10] www.un.org/waterforlifedecade/scarcity.shtml.

[11] For example, www.nytimes.com/interactive/2017/02/17/world/americas/mexico-city-sinking.html and www.nytimes.com/interactive/2017/12/21/world/asia/jakarta-sinking-climate.html.

[12] cf. www.westmonroepartners.com/Insights/White-Papers/State-of-Advanced-Metering-Infrastructure

consumption data. For the researcher, the availability of hourly consumption data avoids the measurement error that is typically present when trying to map metered water use to the actual timing of consumption. In addition, AMI data enable us to analyze not only broad consumption patterns but also the distribution of water consumption within a week; a portion of our empirical study takes advantage of this hourly disaggregation to identify patterns of within-week intertemporal substitution. For the utility, one benefit of AMI is the possibility for algorithmic detection of water leaks or, less commonly, irrigation. As single-family residential accounts typically do not have separate meters for landscape irrigation, utilities are generally unable to identify irrigation disaggregated from total household water use. However, the flow rate of irrigation controllers is so high that consumption during an hour with irrigation far exceeds regular household consumption during any hour of the week without irrigation.[13] Thus, AMI technology allows for automating the enforcement of irrigation policies, and, as water utilities increasingly install smart meters throughout their jurisdictions, the scope for applying AMI technology is steadily growing.

In conducting the field experiment, WaterSmart excluded commercial, industrial, and multi-family residential accounts, leaving an experimental sample of 17,289 single-family residential water accounts. For our analyses, we drop an additional 653 households with missing water consumption data during the pre-treatment period or during the analysis period of May-October 2015.[14] Our final analysis sample includes 16,636 households in Burbank, California. WaterSmart intentionally weights their experiments to have significantly more treated participants, so 13,717 (82 percent of the total) of these households are assigned to the "Experimental" treatment arm, while 2,919 households form the untreated control.

Summary statistics and experimental balance tests for these households are presented in Table 1. The top panel shows that provision of the HWR is virtually 100 percent among households assigned to treatment; no households in the control group were treated by WaterSmart. Furthermore, the use of administrative data on water consumption for all households rules out concerns about differential attrition or loss in follow-up.

The bottom panel of Table 1 shows averages and balance t-tests for time-invariant and pre-treatment household covariates. The administrative data used in our study include several residential attributes, showing overall a high degree of balance across groups in the characteristics of their homes and landscapes. Households across groups are also equally likely to have received prior irrigation violation notices. Most importantly, the groups are

[13] cf. www.wsscwater.com/customer-service/rates/water-usage.html.

[14] We verified that these sample restrictions are orthogonal to the randomization we use for identification.

balanced in their water consumption during the year preceding our study (May 2014 - April 2015). Figure 2 confirms that the distribution – not just the average level – of pre-treatment water consumption is very similar across the two groups. Overall, as we should expect given randomization among a large set of households, the two arms are very balanced.

4 Results

This section presents our empirical findings. First, we assess the impact of social pressure on water conservation using the randomized WaterSmart field experiment. We estimate treatment effects of the normative peer comparisons over the full sample period and then evaluate crowd out by comparing estimates between a baseline period and a period with strengthened pecuniary incentives for water conservation. Second, to validate that the heightened pecuniary incentives are binding, we estimate the direct effects of the violation notices using a regression discontinuity design based on the computer algorithm used to automate enforcement.

4.1 Estimated effects of the randomized social pressure

We start by examining whether social pressure affects water conservation in both the absence and presence of strong pecuniary incentives. Our identification uses a field experiment in which randomly-selected households were provided HWR including normative social comparisons of water use.

Figure 3 displays average post-treatment water consumption by week for both the treatment group that received social comparisons and the control group that did not. The time range shown in the figure spans from late May through December 2015; treated households each had been sent exactly one HWR at the start of this time period and then continued to receive monthly reports throughout. The AMI data enable us to see treatment effects immediately following the initial HWR, and weekly water use is lower for the treatment group than the control group for every week over the experimental period.[15] The magnitude of treatment-control differences also appears to be fairly consistent across weeks. Visually, the evidence strongly supports that providing social comparisons to households causes sub-

[15] The immediacy of the treatment effect is consistent with Reiss and White (2008), who find that consumers in the electricity sector respond promptly to both price changes and normative appeals to conserve. These findings support our use of May-June consumption as a counterfactual for the magnitude of the effect of WaterSmart HWR in subsequent months.

stantial water conservation.

Of importance for exploring the crowd out hypothesis, there is no discernible break over time in the magnitude of treatment-control differences in average weekly water use, despite the substantial increase in pecuniary enforcement discussed earlier and shown in Figure 1. The post-treatment sample period can be divided into three ranges. During the six weeks from late May through June, the statutory summer watering season under irrigation restrictions was in effect, but it was prior to automation of the associated pecuniary enforcement. Then, during July through October the statutory summer watering season under irrigation restrictions remained in effect, and automated violation notices were issued. Finally, Figure 3 displays water use through November and December, part of the statutory and technical winter water season, when precipitation seasonally increases and legal irrigation was further restricted to only on Saturdays.[16] Visually examining the treatment-control differences across these three periods, there does not appear to be any clear break in the magnitude of the treatment effect. In fact, there does not appear to be a break in the magnitude of treatment-control differences at any point over the entire sample period.

We investigate these patterns of water use more formally by estimating several regression specifications. Because of the significant change in seasonal rainfall patterns in California beginning in November (i.e. the winter rainy season), we focus our quantitative analyses on the summer watering season from late May to October. The starting regression equation is straightforward in the context of the random experiment:

$$\text{consumption}_{it} = \beta' X_i + \gamma \cdot \mathrm{I}\{\mathrm{HWR}\}_i + u_{it} \tag{1}$$

In Equation (1), consumption_{it} is weekly water use for household i, X_i is a vector of baseline controls, $\mathrm{I}\{\mathrm{HWR}\}_i$ is the randomly assigned WaterSmart Home Water Reports treatment indicator, and u_{it} is the econometric error term. The baseline controls include administrative data on residential lot size, irrigable area, and the home's square footage, year of construction, number of floors, number of bedrooms, and number of bathrooms. The effect of receiving a HWR – the "intent-to-treat" estimate of social pressure – is captured by γ, which can be estimated separately under different regimes of pecuniary incentives. All

[16] Following plans to the tariff structure set years in advance, water tariffs changed once near the beginning of our study period on June 2, 2015. The price change was a relatively small increase of 5.2 cents per hundred cubic feet (about 748 gallons) for the first consumption tier, with slightly larger increases on higher tiers of consumption. The median May water bill of 8,550 gallons would have been increased by only $1.59, inclusive of a $1.00 increase to the fixed service charge. There were no additional changes to tariffs until July 2016, over a year later and after the end of our study period.

specifications are estimated using OLS and robust standard errors are reported, two-way clustered by household and week.[17]

Table 2 reports regression estimates for the effects of randomized HWR on post-treatment water consumption. Each column presents an estimate of the average intent-to-treat effect of the HWR for weekly water consumption during the months of 2015 indicated by the column titles. We first discuss the results for the late May-June period, which is covered by summer irrigation restrictions but prior to automated pecuniary enforcement. Columns (1) and (2) report estimates without controls and with controls, respectively. As expected with randomization of households into treatment, the inclusion of controls has no effect on the point estimate but does slightly improve precision of the estimate. Prior to heightened pecuniary incentives, we find an intent-to-treat estimate from the social pressure treatment of 76 gallons per week reduction in water use per treated household. The reduction in water consumption represents 3.2 percent of average weekly water consumption. These findings closely align with those in the existing literature on HWR.

We now turn to the results for the July-October period. This represents the period when pecuniary incentives were heightened and the remainder of the statutory summer watering season under irrigation restrictions. From what we gather based on newspaper articles and institutional sources, there is evidence of abundant general public awareness of the monetary fines and associated automated enforcement for wrong day-of-week irrigation. Columns (3) and (4) report estimates without controls and with controls, respectively. Again, the inclusion of controls has no effect on the point estimate but does slightly improve precision. During this period, we find an estimate from the social pressure treatment of a 75 gallons per week reduction in water use per treated household.

The comparison of treatment estimates from before to after heightened pecuniary incentives reveals that they are virtually identical, suggesting no crowd out of treatment effects by the pecuniary incentives. To formally investigate this comparison, we build on Equation (1) by including the full late May-October sample period and estimating the following equation:

$$\text{consumption}_{it} = \beta' X_i + \gamma \cdot \mathrm{I}\{\text{HWR}\}_i + \lambda \cdot \mathrm{I}\{\text{Post July}\}_t + \delta \cdot \mathrm{I}\{\text{HWR}\}_i \cdot \mathrm{I}\{\text{Post July}\}_t + u_{it} \quad (2)$$

In Equation (2), $\mathrm{I}\{\text{Post July}\}_t$ is an indicator variable that equals zero during May-June and one in the post-pecuniary incentive period (i.e. July-October). In this specification, δ captures the difference between the two periods' average treatment effects and is the

[17]As our study period includes a fairly small number of weeks, we verified that standard errors are very similar when clustering only by household.

coefficient of interest because it provides a direct estimate of the crowd out effect. This empirical test is in the spirit of a difference in differences estimator, with one of the differences determined by a randomly assigned treatment. The other difference is a comparison across closely neighboring weeks. The identifying assumption for this test is thus much less stringent than that for typical difference in differences strategies based on heterogeneous policy changes across treatment units and time periods (often years).

Column (5) of Table 2 reports estimates of Equation (2). As expected given the previous results, we find a point estimate of crowd out that is essentially zero. The standard error is small enough to also rule out even moderate amounts of crowd out: under the null hypothesis of positive crowd out, at conventional significance levels we can rule out anything larger than 25 gallons per week, one third of the average treatment effect of the social pressure. Finally, in Column (6) of Table 2 we estimate a specification that includes household fixed effects. The inclusion of these fixed effects – which control for all unobserved time-invariant differences across households – does not change the results. The estimate of δ is very similar to that reported in Column (5). From this model that absorbs even more heterogeneity, we continue to find no evidence of crowd out of social pressure by the heightened pecuniary incentives.

The interpretation of these findings is made stronger given that several factors potentially bias our estimates towards finding crowd out. As discussed in Section 1, prior research shows that the effects of peer comparisons tend to attenuate over time, although generally over a much longer time span than that included in our analysis. As discussed in Section 2, there is also some minor possibility for mechanical crowd out due to limited scope for additional conservation. Finally, the intensity of the direct pecuniary treatment might be slightly stronger for the HWR control group. Initial HWR were sent more than a month prior to the automated violation notices and they induced conservation among treated households. Because a household's total water consumption factors into the algorithm for detecting irrigation violations, HWR-treated households are about two percentage points less likely to have been sent an automated irrigation violation notice than those in the control group (36.9 vs. 39.0 percent). This minor difference in relative pecuniary treatment intensity across groups does not affect causal inference about the randomized HWR treatment, but it does potentially bias the difference in differences estimates towards showing crowd out.[18] In light of these

[18]The degree of bias from this heterogeneity in intensity of pecuniary incentives across HWR arms is well within our estimated confidence intervals; if anything, at face value this implies our estimated null for δ in Equation (2) supports there is some crowd-*in*. A more nuanced consideration would be if the HWR treatment effect were concentrated among a very small share of households, but this conjecture runs counter to the entire body of evidence in the literature on home energy reports and HWR, including our own findings.

factors, the weight of empirical evidence is strengthened against there being crowd out of social pressure.

We conduct several additional robustness checks of these findings in Table 3, which presents estimates for six variants of Equation (2). Columns (1) and (2) reproduce the difference in differences tests from Columns (5) and (6) of Table 2. In Columns (3) and (4) we add fixed effects for the week of the sample, more flexibly controlling for any heterogeneity in consumption across time periods. Again, we find no evidence of crowd out. In Columns (5) and (6) of Table 3, we narrow the analysis window to include only late May through August, more tightly surrounding the July policy change. Using this shorter post-treatment time period also yields no evidence of crowd out.

In Table 4, we explore heterogeneity in the treatment effect and assess any crowd out. Again, Columns (1) and (2) reproduce the difference in differences tests from Columns (5) and (6) of Table 2. Columns (3) and (4) and Columns (5) and (6) estimate these specifications for the subsets of households whose pre-treatment consumption was, respectively, below or above the median. Perhaps not surprisingly, we find that low-volume water users show smaller and statistically insignificant effects of HWR, whereas large-volume consumers exhibit large and significant conservation effects. Neither subset shows any evidence of crowding out between the two interventions. Especially when considering that nearly two-thirds of the high-volume group was treated with violation notices, the estimates in Table 4 serve as a further strong robustness check of the primary finding of our study. In sum, these estimates show consistent water conservation effects from social pressure. The effectiveness of social pressure does not change with respect to the strength of pecuniary incentives targeting the same behaviors.

4.2 Regression discontinuity estimates for the pecuniary incentive

Given that we find that the effects of social pressure are invariant to altered pecuniary incentives, the most serious potential threat to inference would be if the "stronger" pecuniary incentive were non-binding. To negate concerns that our finding of no crowd out is attributable to the pecuniary incentive being weak and ineffective, we next examine the direct revealed preference effects of the automated violation notices on water conservation.

In doing so, we employ a regression discontinuity design based on a cutoff in the noncompliance algorithm the water agency initially used to determine irrigation violations (discussed above in Section 2). Specifically, BWP conservatively estimated the number of days per week that each household was irrigating by counting the number of days during one week of June 2015 on which the household consumed more than 125 gallons in any individual hour of the day. This arbitrary cutoff of 125 gallons for the daily peak consumption hour forms the basis of our regression discontinuity design. For internal institutional reasons, in automating the violation notices the agency decided to allow for comparatively more detected irrigation days per week for (both HWR Treatment and Control) accounts with "average" or "efficient" consumption, per the WaterSmart tier categorizations discussed above in Section 2. Household HWR arm assignment is not a factor in the algorithm and households in both the HWR and Control arm were sent violation notices. However, because we imperfectly observe households' historical tier statuses (particularly among the WaterSmart control group), we assign all households to the running variable based on the peak consumption hour of the day that would have determined a violation under the strictest allowance. For this reason, the RDD is "fuzzy" and treated households nearly-exclusively fall into the "above average" water consumption tier.[19]

Before turning to the estimates, we conduct some standard exercises to support the validity of our RDD. First, we test for manipulation along the running variable, which measures the distance to the irrigation violation cutoff. Given that the automation of detecting irrigation violations was unprecedented and unannounced, a priori there should be no concern. As shown in Figure A4, there is some measurement lumpiness from the underlying meter technology, but there is no evidence of any sorting of households around the threshold, which visually confirms the results of our statistical implementation of McCrary's (2008) test for manipulation. Further supporting the identification strategy, Figure A5 demonstrates that there is also smoothness across the threshold in pre-treatment water consumption along the running variable. Finally, in Appendix Table A1 we replicate the balance tests from Table 1 for only the subset of residences within an 80 gallon bandwidth of the RD cutoff to show that the RDD is not disproportionately capturing behavior of only one of the randomly-assigned HWR treatment arms. On the whole, the RDD appears to be very strongly supported and well-positioned to provide credible causal inference.[20]

Next, we check compliance. Figure 4 displays the share of households receiving initial automated violation notices along the running variable. For visual clarity, the running variable uses bins of 10 gallons in Figures 4-5, and the size of the markers corresponds to the number of households included in the local average. As evidenced graphically, zero

[19] While the vast majority of households detected using this scheme were likely violating the city's restrictions on outdoor water use, a small fraction of residences that received notices claimed to be using the water for other purposes or hand-watering, which was permitted under the policy.

[20] In principle, we could test for crowd out by evaluating the difference in discontinuities across the HWR Treatment and Control arms. In practice, such an exercise is statically underpowered as there are too few households in the neighborhood of the RD cutoff (particularly among the much smaller Control group).

households below the threshold received a violation notice. At the treatment threshold, we find a discontinuous jump of roughly 25 percentage points for receipt of an automated violation notice in the first week of July, a strong first stage.

In other words, households with peak hourly water consumption of amounts just above the 125 gallon threshold are 25 percentage points more likely to have received a violation notice from the water utility than households just below the threshold. In addition, because households that have higher peak water consumption are more likely to be "above average" and thus treated using the stricter threshold, the treatment propensity increases with the running variable. At the right end of the range displayed in Figure 4, we find that 40 percent of households received a notice, an increasing slope that continues past the displayed range. Because of the heterogeneous treatment across the different consumption tiers, as discussed earlier, the computer algorithm used by BWP resulted in perfect compliance below the threshold but not above the cutoff. Thus, because the running variable represents a necessary but not sufficient condition for a household to be sent an automated violation notice, the first stage supports a fuzzy regression discontinuity design.

Having established the validity of our first-stage, we examine the effects of the pecuniary violation notices on water use. We start by presenting figures plotting local averages of post-treatment water consumption measures against the running variable. Figure 5(a) plots the reduced-form relationship for water consumption during targeted periods of the week, when irrigation was not allowed (all hours excluding Tuesday and Saturday before 9:00 a.m. and after 6:00 p.m.). Figure 5(b) shows the reduced-form relationship for water consumption during the entire week. Average water consumption in these figures is pooled over July-October 2015, the four-month period immediately following the automated violation notices treatment and continuing through the end of the statutory local summer water season. This period corresponds with the "Post July" period used earlier in our RCT analysis.

We find a discontinuous drop in water consumption at the threshold for water use during irrigation-restricted periods of the week. The discontinuity in reduced-form is roughly 200 gallons per household per week, or about 7.5 percent of the pre-treatment average weekly water consumption. For water consumption during the entire week, in Figure 5(b) we also find a large drop at the threshold, although comparatively smaller, consistent with possible intertemporal substitution in response to the enhanced enforcement of an asymmetric restriction. Overall, water consumption both during the targeted hours of the week and across the entire week decreases across the violation notice threshold.

We investigate these patterns more formally by estimating nonparametric local linear regressions (Imbens and Kalyanaraman, 2012). Table 5 reports RD estimates of the effects of irrigation violation notices. Panel [A] presents reduced-form estimates and Panel [B] presents the local average treatment effects, which essentially rescale the reduced-form estimates by the estimated magnitude of the first stage. Each cell in the table presents a nonparametric RD estimate at the cutoff for automated violation notices. Following Lee and Card (2008), standard errors are clustered along the running variable, which is discrete in gallons. The bandwidth for each specification is selected independently and nonparametrically using a triangular kernel (Lee and Lemieux, 2010).[21]

In Panel [A], we present the reduced-form estimates that correspond to Figure 5. Column (2) reports estimates for water consumption during irrigation-restricted times of the week. We find a statistically significant drop of 194.4 gallons per week at the threshold. Some of this decrease in water consumption, however, is offset by an increase in water consumption during non-restricted days and hours of the week. Column (3) shows that water consumption increased during the irrigation-allowed hours on Tuesdays and Saturdays before 9:00 a.m. or after 6:00 p.m.: specifically, water use during these periods increases at the threshold by 47.7 gallons per week. Thus, the violation notices partly shifted water consumption from irrigation-restricted times to irrigation-allowed times, showing intertemporal substitution in response to a policy with (intentionally) partial coverage. Focusing on total weekly water conservation in Column (4), we also find significant reduced-form effects at the threshold, with total water consumption decreasing by 139.7 gallons per week.

As displayed in Figure 4, household receipt of violation notices increases substantially at the threshold. Confirming this visible discontinuous jump, nonparametric RD estimates indicate an increase of 0.2555 at the cutoff for automated violation notices. The estimate is reported in the first column of Panel [A] in Table 5. Given that only one out of four "barely-eligible" households received the violation notice, it is useful to rescale the RD estimates so that they can be interpreted as the effect of receiving a violation notice instead of as a reduced-form estimate of the effects of crossing the arbitrary threshold.

Panel [B] of Table 5 reports RD estimates for local average treatment effects of receiving a violation notice. As expected, the LATE estimates are roughly four times larger than the reduced-form estimates. The effect of receiving a violation notice is to reduce post-notice water consumption by 765.6 gallons per week during irrigation-restricted times of the week. In contrast, water use during irrigation-allowed portions of the week increases by 189.8 gallons

[21] Estimates are quantitatively and qualitatively similar when forcing a common bandwidth, such as 40 gallons, to be used across all outcomes.

per week. Finally, total weekly water use decreases by 550.5 gallons per week on average for households sent a violation notice. To place this into perspective, average household water use per week during the pre-treatment period is about 2700 gallons (Table 1), implying that these are economically significant effects on the order of more than 20 percent of total water consumption – and nearly 30 percent relative to the barely-untreated side of the RD treatment cutoff.[22]

We draw three main inferences from these estimates. First, strengthened pecuniary incentives directly cause significant reductions in water consumption, in addition to any spillovers or other indirect conservation effects from enhanced enforcement. Second, the overall effectiveness of automating violation notices indicates that we can view the "Post July" period as one with now-much-stronger pecuniary incentives to conserve water. Third, the results from our difference in (randomized) differences tests in Section 4.1 are not simply due to the heightened pecuniary incentive being weak and ineffective, supporting the hypothesis that there is no crowd out from interacting the two policy instruments.

5 Conclusions

The research literature includes ample evidence that pecuniary incentives can crowd out intrinsic motivations for voluntary contributions to public goods across a wide range of settings, but these studies do not address whether monetary incentives also reduce the efficacy of external social pressure interventions to behave prosocially. Focusing on water conservation, we examine whether the effects of social pressure are crowded out by strengthened pecuniary incentives. An important feature of the setting that we analyze is that the heightened pecuniary policy is well-enforced and automated because of a new technology monitoring precise hourly water use.

We test whether social pressure yields economically significant conservation in the face of binding, technology-enforced pecuniary incentives for water conservation. Using a large field experiment among water customers, we find that randomly-provided normative peer comparisons cause substantial water conservation over time periods in which strengthened pecuniary incentives were in place as well as when they were not in place. The evidence is compelling: in every post-treatment week we find lower water use among treated households

[22] In percentage terms, our estimated conservation effects for automating landscape irrigation violation notices are on par with those from subsidizing 'Water Smart Landscape' conversion (e.g. Baker, 2017; Brelsford and Abbott, 2018). Although estimating the effects of enforcing irrigation restrictions is not the primary focus of our study, this is an interesting and policy-relevant finding in its own right.

that received normative social comparisons. More formally quantifying crowd out using difference in differences tests, in which one difference is randomly generated through the experiment and the other difference compares across neighboring weeks, we find no evidence that the effects of social pressure are crowded out by heightened pecuniary incentives; rather, the evidence supports that the conservation benefits could be fully additive.

We also rule out a serious potential threat to the inference of no crowd out, which is that the "strengthened" pecuniary incentives might have been weak and non-binding. If this were the case, it would not be surprising to find that the effects of the experimentally generated social pressure to conserve water did not change across neighboring weeks. To rule out this explanation, we employ a regression discontinuity design based on a somewhat arbitrary cutoff in the noncompliance algorithm used by the water utility. These estimates indicate that the heightened pecuniary incentives also cause significant water conservation, confirming that we are comparing a baseline period to a period with substantially stronger pecuniary incentives in our test of the crowd out hypothesis.

Governments rely on three major types of policies to encourage resource conservation and public goods provision: regulations, economic incentives, and social pressure. Because of limited ability to implement first-best corrective policies, often due to legal or political constraints, policymakers are increasingly drawing on normative appeals through social pressure as a means to change behavior. Our findings provide evidence that these relatively new policies based around non-pecuniary incentives can strengthen other-regarding preferences to serve as an effective complement to binding pecuniary incentives. Demonstrating that social pressure remains effective even when interacted with economic incentives is especially relevant as government and non-government institutions increasingly utilize combinations of normative interventions and technology-driven enforcement of pecuniary policies.

References

George A. Akerlof and William T. Dickens. The economic consequences of cognitive dissonance. *The American Economic Review*, 72(3):307–319, June 1982.

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

Hunt Allcott and Judd B. Kessler. The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, 11(1):236–276, January 2019.

Hunt Allcott and Todd Rogers. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *The American Economic Review*, 104(10):3003–3037, October 2014.

Dan Ariely, Anat Bracha, and Stephan Meier. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *The American Economic Review*, 99(1): 544–555, March 2009.

Ian Ayres, Sophie Raseman, and Alice Shih. Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *The Journal of Law, Economics, and Organization*, 29(5):992–1022, October 2013.

Jonathan E. Baker. Subsidies for succulents: Evaluating the Las Vegas Cash-for-Grass rebate program. Harvard Environmental Economics Program Discussion Paper 17-74, June 2017.

Roland Bénabou and Jean Tirole. Incentives and prosocial behavior. *The American Economic Review*, 96(5):1652–1678, December 2006.

Maria Bernedo, Paul J. Ferraro, and Michael K. Price. The persistent impacts of norm-based messaging and their implications for water conservation. *Journal of Consumer Policy*, 37 (3):437–452, September 2014.

Samuel Bowles and Sandra Polanía-Reyes. Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, 50(2):368–425, June 2012.

Christa Brelsford and Joshua K. Abbott. How smart are "Water Smart Landscapes"? arXiv.org Working Paper, March 2018.

Daniel A. Brent, Joseph H. Cook, and Skylar Olsen. Social comparisons, household water use, and participation in utility conservation programs: Evidence from three randomized trials. *Journal of the Association of Environmental and Resource Economists*, 2(4):597–627, December 2015.

Daniel A. Brent, Corey Lott, Michael Taylor, Joseph H. Cook, Kim Rollins, and Shawn Stoddard. Are normative appeals moral taxes? Evidence from a field experiment on water conservation. LSU Working Paper 2017-07, June 2017.

Marco Castillo, Ragan Petrie, and Clarence Wardell. Fundraising through online social networks: A field experiment on peer-to-peer solicitation. *Journal of Public Economics*, 114:29–35, June 2014.

Anita Castledine, Klaus Moeltner, Michael K. Price, and Shawn Stoddard. Free to choose: Promoting conservation by relaxing outdoor watering restrictions. *Journal of Economic Behavior and Organization*, 107(A):324–343, November 2014.

Jerald Emmet Christiansen. *Irrigation by Sprinkling*, volume 670. University of California, College of Agriculture: Agricultural Experiment Station, Berkeley, California, 1942.

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

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

Paul J. Ferraro, Juan José Miranda, and Michael K. Price. The persistence of treatment effects with norm-based policy instruments: Evidence from a randomized environmental policy experiment. *The American Economic Review: Papers & Proceedings*, 101(3):318–322, May 2011.

Bruno S. Frey and Stephan Meier. Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *The American Economic Review*, 94(5): 1717–1722, December 2004.

Alan S. Gerber, Donald P. Green, and Christopher W. Larimer. Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, 102(1):33–48, February 2008.

James M. Gillan. Dynamic pricing, attention, and automation: Evidence from a field experiment in electricity consumption. Energy Institute at Haas WP284, March 2018.

Uri Gneezy and Aldo Rustichini. A fine is a price. *The Journal of Legal Studies*, 29(1):1–17, January 2000.

Uri Gneezy, Stephan Meier, and Pedro Rey-Biel. When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, 25(4):191–210, Fall 2011.

R. Quentin Grafton and Michael B. Ward. Prices versus rationing: Marshallian surplus and mandatory water restrictions. *Economic Record*, 84(S):S57–S65, September 2008.

Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959, July 2012.

Koichiro Ito. How do consumers respond to nonlinear pricing? Evidence from household water demand. Mimeo, April 2013.

Koichiro Ito, Takanori Ida, and Makoto Tanaka. Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, 10(1):240–267, February 2018.

Katrina Jessoe, Gabriel E. Lade, Frank Loge, and Edward Spang. Spillovers from behavioral interventions: Experimental evidence from water and energy use. E2e Working Paper 033, December 2017.

David S. Lee and David Card. Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674, February 2008.

David S. Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355, June 2010.

John A. List, Robert D. Metcalfe, Michael K. Price, and Florian Rundhammer. Harnessing policy complementarities to conserve energy: Evidence from a natural field experiment. NBER Working Paper 23355, April 2017.

Alexandre Mas and Enrico Moretti. Peers at work. *The American Economic Review*, 99(1): 112–145, March 2009.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, February 2008.

David L. Mitchell and Thomas W. Chesnutt. Evaluation of East Bay Municipal Utility District's pilot of WaterSmart Home Water Reports. Commissioned report, for the California Water Foundation and the East Bay Municipal Utility District, December 2013.

Sheila M. Olmstead, Michael Hanemann, and Robert N. Stavins. Water demand under alternative price structures. *Journal of Environmental Economics and Management*, 54 (2):181–198, September 2007.

José A. Pellerano, Michael K. Price, Steven L. Puller, and Gonzalo E. Sánchez. Do extrinsic incentives undermine social norms? Evidence from a field experiment in energy conservation. *Environmental and Resource Economics*, 67(3):413–428, November 2017.

Bryan Pratt. Symmetric responses to threshold policies: The impact of temporary drought restrictions. Mimeo, November 2018.

Kimberly J. Quesnel and Newsha K. Ajami. Changes in water consumption linked to heavy news media coverage of extreme climatic events. *Science Advances*, 3(10):1–9, October 2017.

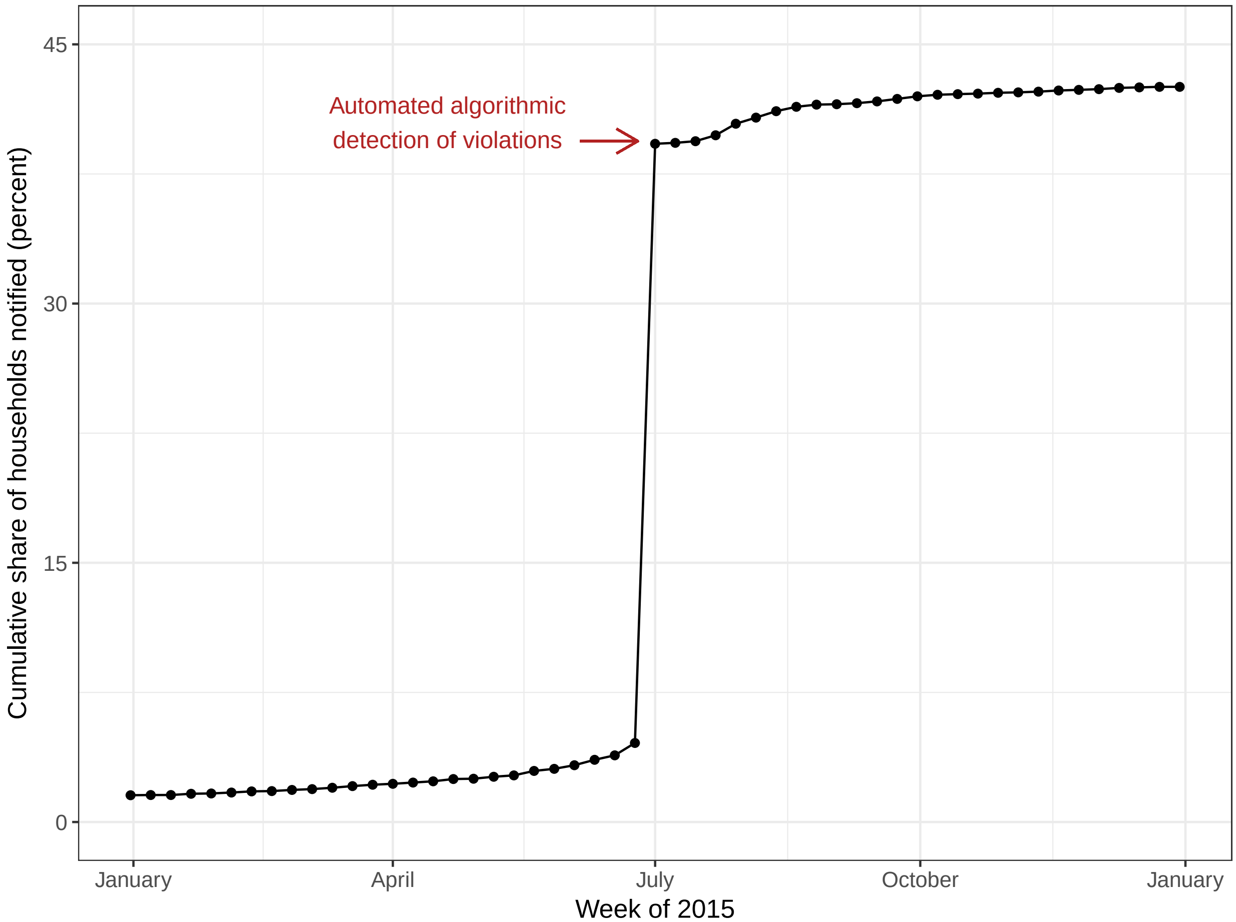
Peter C. Reiss and Matthew W. White. What changes energy consumption? Prices and public pressures. *RAND Journal of Economics*, 39(3):636–663, Autumn 2008.

Julian Rode, Erik Gómez-Baggethun, and Torsten Krause. Motivation crowding by economic incentives in conservation policy: A review of the empirical evidence. *Ecological Economics*, 117:270–282, September 2015.

Sarah Smith, Frank Windmeijer, and Edmund Wright. Peer effects in charitable giving: Evidence from the (running) field. *The Economic Journal*, 125(585):1053–1071, June 2015.

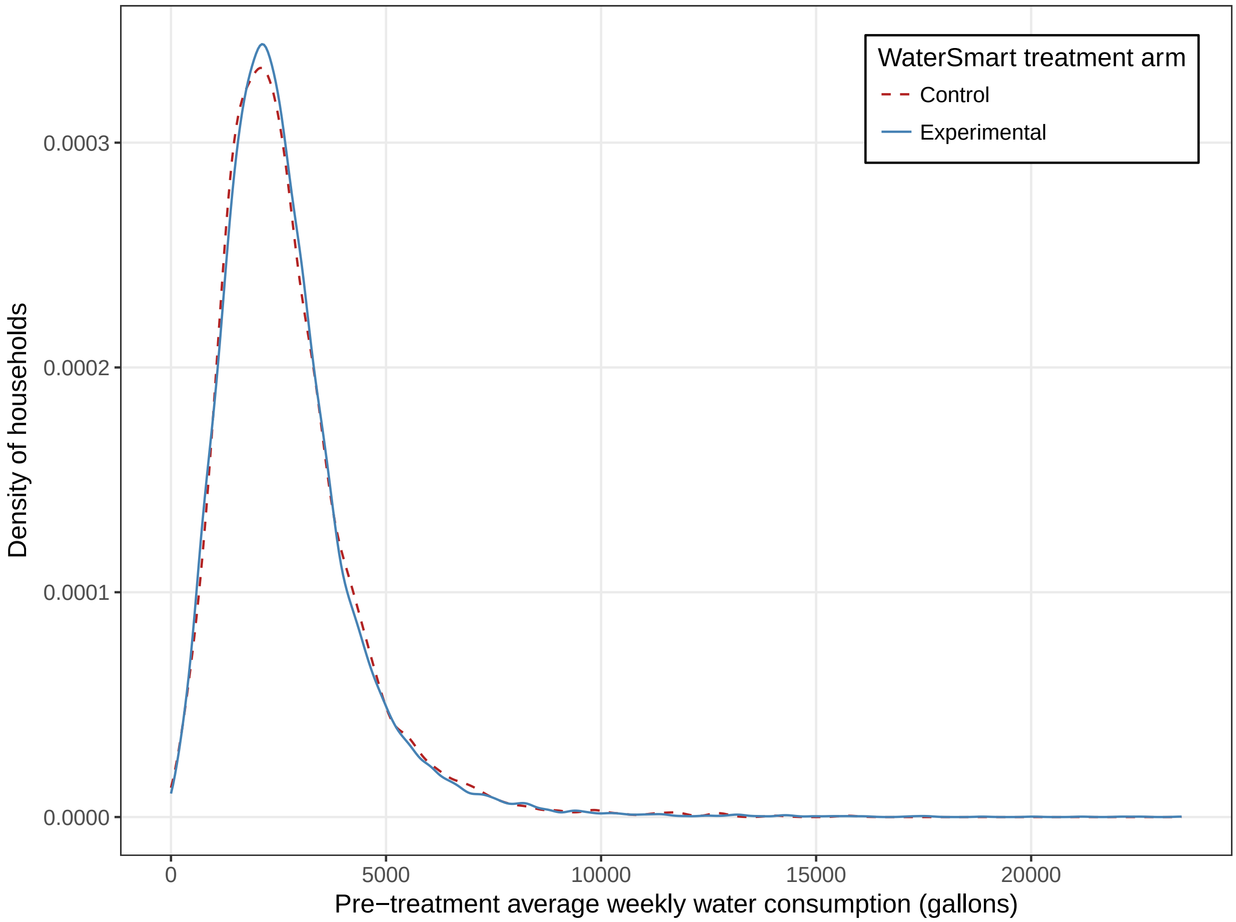
Richard Titmuss. *The Gift Relationship: From Human Blood to Social Policy*. Allen and Unwin, 1970.

Figure 1: Time series for issued irrigation violation notices



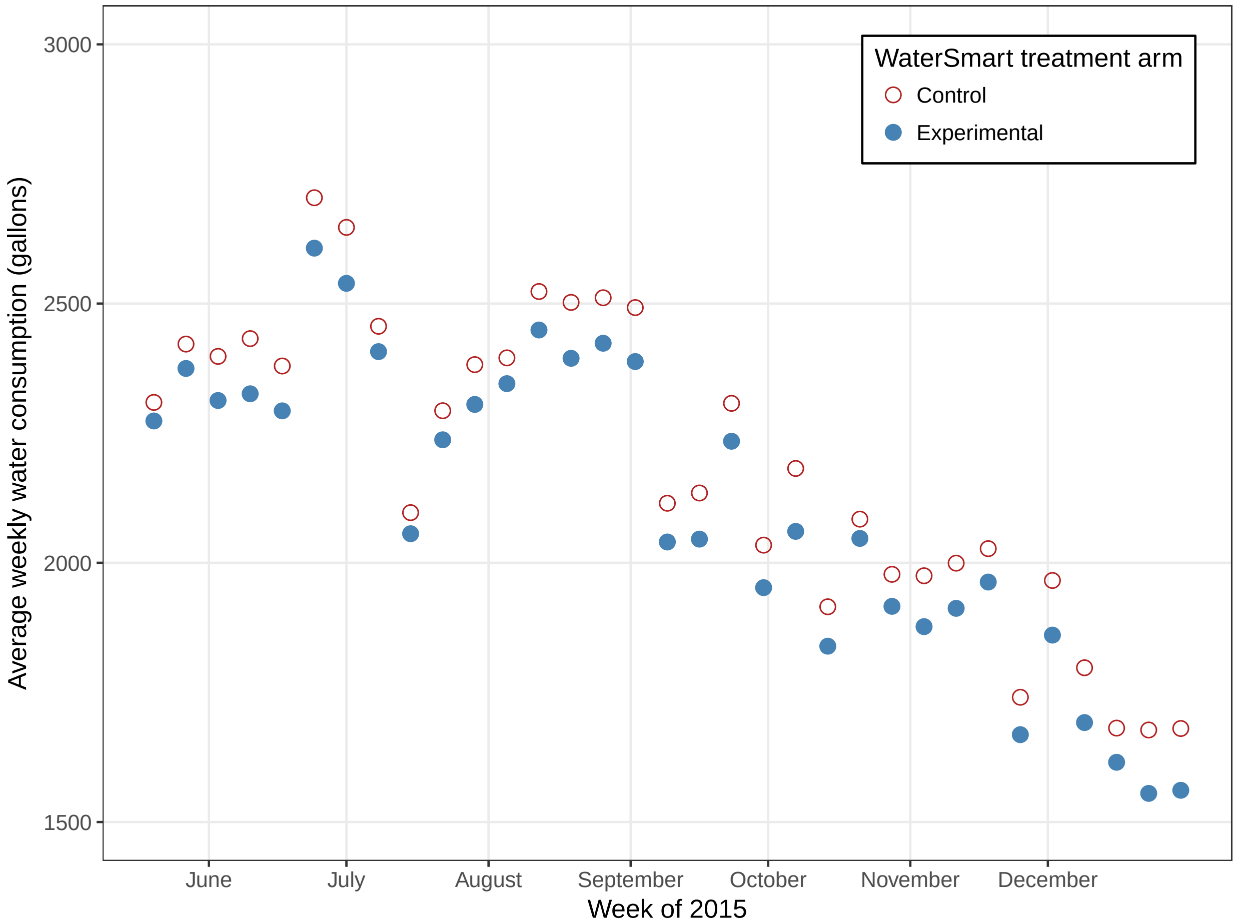
Notes: Figure 1 plots the cumulative share of in-sample households that had ever received an irrigation violation notice by week. Throughout this period, violations were determined when either a municipal employee or a neighbor of the offender reported unlawful irrigation to the city. As indicated by the annotation, the city also implemented an automated algorithmic detection of violations in early July.

Figure 2: Balance test of distributions of pre-treatment water consumption



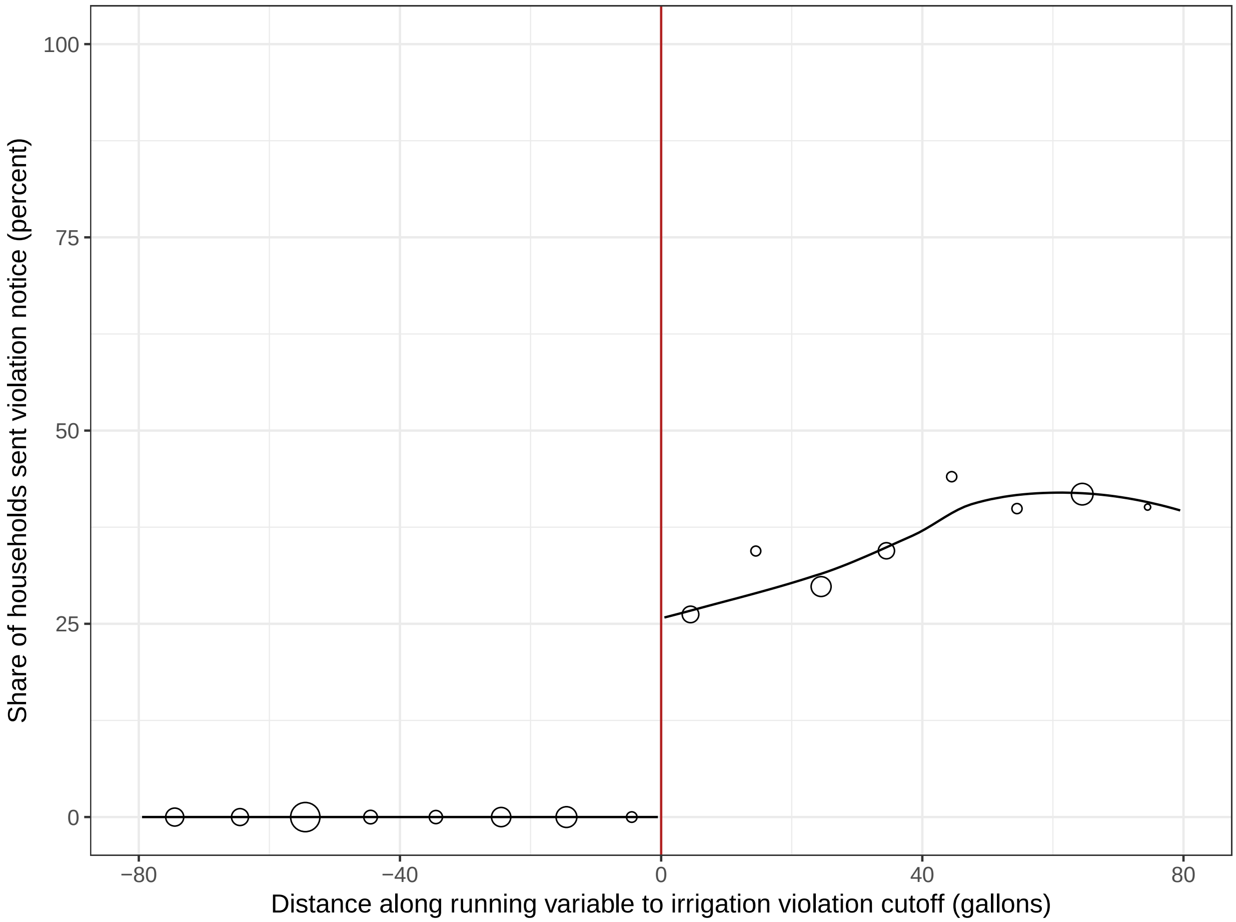
Notes: Figure 2 plots the distributions of average weekly water consumption for in-sample households during May 2014 through April 2015, the full year prior to both the social comparison and pecuniary treatments. The solid line shows the distribution for households assigned to the WaterSmart Home Water Reports (HWR) treatment arm, and the dashed line includes the untreated control group.

Figure 3: Post-treatment water consumption by treatment arm



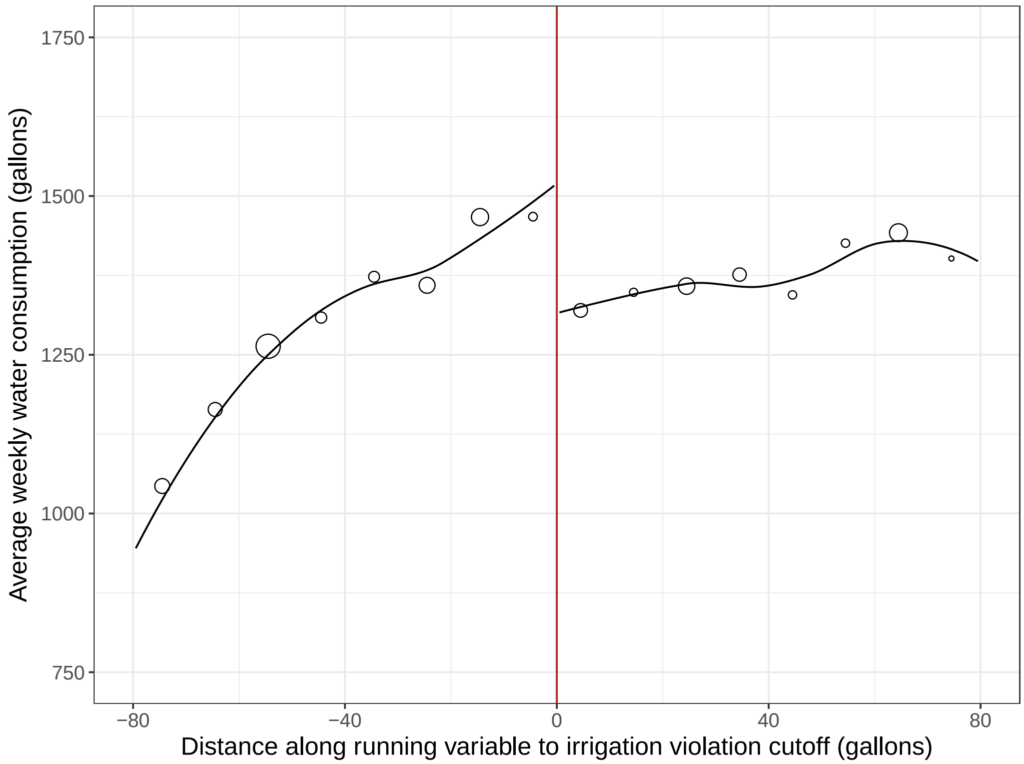
Notes: Figure 3 plots average water consumption by week for each WaterSmart treatment arm during late May through December 2015. "Experimental" households each had been sent one Home Water Report (HWR) as of the start of this time period and monthly reports continued to be sent to treated households throughout this time period.

Figure 4: First stage for automated violation notices in the regression discontinuity design

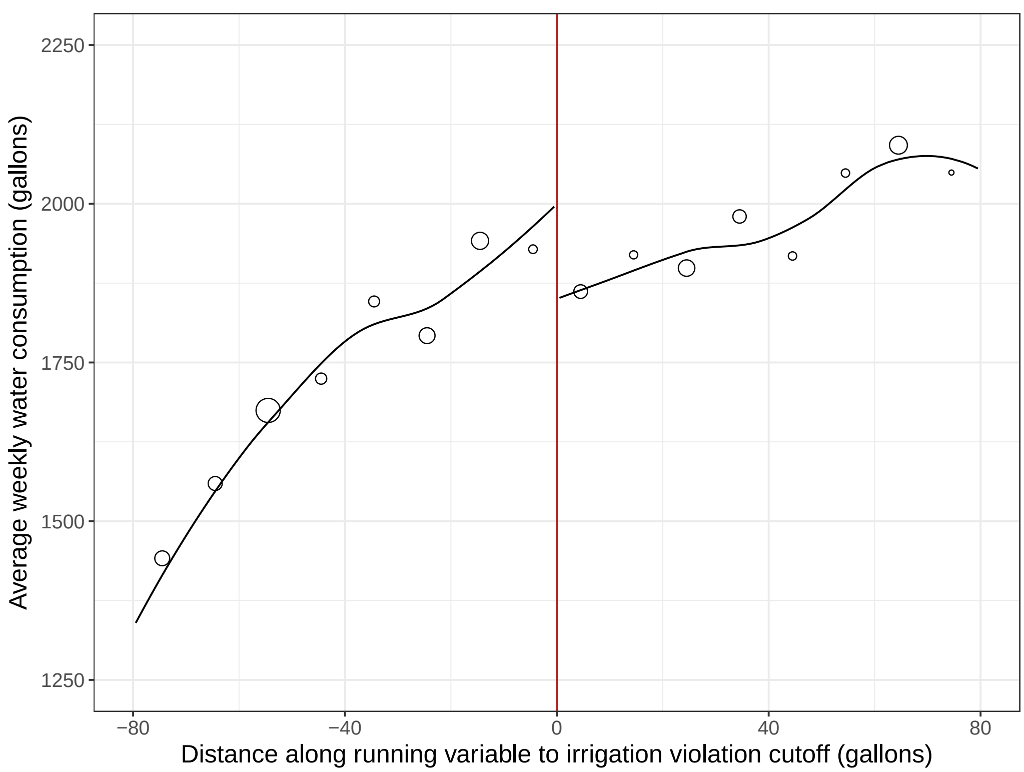


Notes: Figure 4 plots local averages for the first stage outcome of whether a household received an automated irrigation violation notice during the first week of July 2015. For clarity, the running variable uses 10 gallon bins. The size of the markers corresponds to the number of households included in the local averages. The LOESS curves shown are fit to the underlying microdata separately on each side of the threshold. Because the running variable represents a necessary but not sufficient condition for a household to be sent an automated violation notice, the first stage supports a “fuzzy” regression discontinuity design.

Figure 5: Reduced-form local averages for post-treatment water consumption



(a) Consumption during hours of the week when irrigation is not allowed



(b) Total weekly water consumption

Notes: Figure 5 plots local averages for weekly water consumption during July-October 2015, the period following the automated violation notices treatment. For clarity, the running variable uses 10 gallon bins. The size of the markers corresponds to the number of households included in the local averages. The LOESS curves shown are fit to the underlying microdata separately on each side of the threshold.

Table 1: Summary statistics and randomization balance checks

	Group means		t-tests	
	(1)	(2)	(3)	(4)
Covariate	Control	Experimental	Difference	p-value
Number of households	2919	13,717		
Sent WaterSmart HWR	0	0.9972		
Prior water violation	0.0483	0.0452	-0.0031	0.48
Lot size (SqFt)	7347	7320	-27	0.71
Irrigable area (SqFt)	3829	3794	-35	0.45
House size (SqFt)	1619	1620	1	0.95
Year built	1945	1945	0	0.82
Number of floors	1.062	1.067	0.005	0.28
Number of bedrooms	2.912	2.919	0.007	0.69
Number of bathrooms	1.93	1.939	0.009	0.61
Weekly water gallons	2693	2678	-15	0.66

Notes: Table 1 shows statistics by WaterSmart Home Water Reports (HWR) treatment arm for household-level covariates. The first two columns show means by treatment arm for all households in the randomization sample, Column (3) shows the difference in means, and Column (4) shows the p-values for t-tests of whether the difference in group means is significantly different from zero. By design, WaterSmart weighted the randomization to have about 82 percent of the sample in the "Experimental" treatment arm. Initial HWR were sent to treated households during the billing cycle ending in mid May 2015. All outcomes in the lower panel are determined prior to the randomization and prior to the automated pecuniary treatment. For pre-treatment weekly water consumption, we use each household's average weekly gallons consumed during May 2014 through April 2015, spanning one full year prior to both treatments. Figure 2 shows the density distributions of households' average weekly pre-treatment consumption by treatment arm.

Table 2: Estimated effects of randomized WaterSmart Home Water Reports

	Weekly water consumption (gallons)					
	Late May - June		July - October		Late May - October	
	(1)	(2)	(3)	(4)	(5)	(6)
I{HWR}	-76.36^{**}	-76.24^{***}	-75.51^{***}	-74.95^{***}	-76.14^{***}	
	(31.23)	(29.00)	(28.25)	(25.55)	(28.71)	
I{Post July}					-160.61^{**}	-164.47^{**}
					(73.35)	(76.16)
I{HWR} X I{Post July}					1.15	3.11
					(14.42)	(19.88)
Household controls	No	Yes	No	Yes	Yes	—
Household fixed effects	No	No	No	No	No	Yes
Num. of households	16,636	16,636	16,636	16,636	16,636	16,636
Observations	99,575	99,575	292,542	292,542	392,117	392,117

*p<0.1; **p<0.05; ***p<0.01 Columns (1) - (4) present estimates of the average intent-to-treat effect of the randomized WaterSmart HWR for weekly water consumption during 2015 for the month ranges indicated by the column titles. "Experimental" households each had been sent one HWR as of the start of this time period, and monthly reports continued to be sent throughout this time period. Automated irrigation violation notices were sent during the first week of July. The legal and technical local summer water season runs through the end of October. Columns (5) and (6) present difference in differences estimates testing whether the effect of the randomized HWR differed before versus after violation notices were sent. Standard errors in parentheses are two-way clustered by household and week. The household control terms include residential lot size, irrigable area, and the home's square footage, year of construction, number of floors, number of bedrooms, and number of bathrooms. Figure 3 shows average weekly water consumption by treatment group separately for each post-treatment week spanning from late May through December 2015.

Table 3: Robustness checks of difference in differences test for crowd out

	Weekly water consumption (gallons)					
	Late May - October				Late May - August	
	(1)	(2)	(3)	(4)	(5)	(6)
I{HWR}	-76.14^{***}		-76.15^{***}		-76.26^{***}	
	(28.71)		(28.73)		(28.76)	
I{Post July}	-160.61^{**}	-164.47^{**}				
	(73.35)	(76.16)				
I{HWR} X I{Post July}	1.15	3.11	0.68	2.28	4.44	6.34
	(14.42)	(19.88)	(14.49)	(19.95)	(12.94)	(20.58)
Household controls	Yes	—	Yes	—	Yes	—
Household fixed effects	No	Yes	No	Yes	No	Yes
Week fixed effects	No	No	Yes	Yes	Yes	Yes
Num. of households	16,636	16,636	16,636	16,636	16,636	16,636
Observations	392,117	392,117	392,117	392,117	247,463	247,463

*p<0.1; **p<0.05; ***p<0.01 All columns present difference in differences estimates testing whether the effect of the randomized Home Water Reports (HWR) differed before versus after violation notices were sent. Standard errors in parentheses are two-way clustered by household and week. Columns (1) and (2) reproduce the estimates from Columns (5) and (6) of Table 2. Columns (3) - (6) add in fixed effects for the week of sample. Columns (5) and (6) further restrict the time period to May - August, 2015. The household control terms include residential lot size, irrigable area, and the home's square footage, year of construction, number of floors, number of bedrooms, and number of bathrooms. Figure 3 shows average weekly water consumption by treatment group separately for each post-treatment week spanning from late May through December 2015.

Table 4: Heterogeneity in difference in differences test for crowd out

	Dependent variable: weekly water consumption (gallons)					
	Full sample of households		Low volume consumers		High volume consumers	
	(1)	(2)	(3)	(4)	(5)	(6)
I{HWR}	-76.14^{***}		-33.72		-116.89^{***}	
	(28.71)		(21.11)		(42.03)	
I{Post July}	-160.61^{**}	-164.47^{**}	-42.84	-44.36	-275.52^{***}	-284.16^{***}
	(73.35)	(76.16)	(53.58)	(55.58)	(95.25)	(99.62)
I{HWR} X I{Post July}	1.15	3.11	5.13	5.13	-5.12	0.49
	(14.42)	(19.88)	(14.06)	(17.18)	(26.58)	(33.01)
Household controls	Yes	—	Yes	—	Yes	—
Household fixed effects	No	Yes	No	Yes	No	Yes
Num. of households	16,636	16,636	8,317	8,317	8,319	8,319
Sent violation (percent)	37.29	37.29	12.77	12.77	61.80	61.80
Observations	392,117	392,117	196,103	196,103	196,014	196,014

*p<0.1; **p<0.05; ***p<0.01 All columns present difference in differences estimates testing whether the effect of the randomized Home Water Reports (HWR) differed before versus after violation notices were sent. Standard errors in parentheses are two-way clustered by household and week. Columns (1) and (2) reproduce the estimates from Columns (5) and (6) of Table 2, including all in-sample households. Columns (3) and (4) include only households with below-median pre-treatment water consumption. Columns (5) and (6) include only households with above-median pre-treatment water consumption. The household control terms include residential lot size, irrigable area, and the home's square footage, year of construction, number of floors, number of bedrooms, and number of bathrooms.

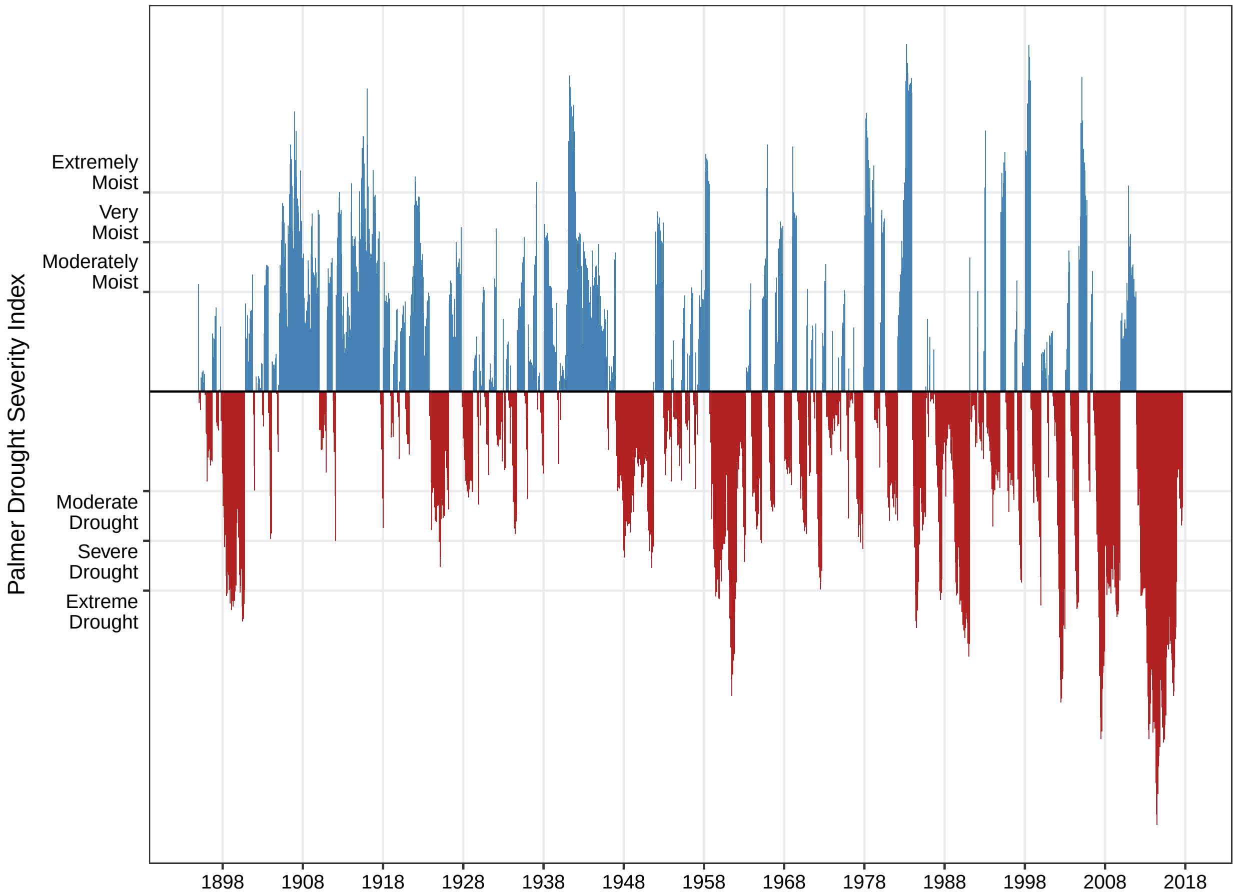
Table 5: Regression discontinuity estimates of effects of irrigation violation notice

		Weekly water consumption: July-Oct (gallons)		
	(1)	(2)	(3)	(4)
	First-stage	Non-irrigation hours	Irrig. hours	All hours
Panel [A]: Reduced-form estimates				
Discontinuity	0.2555***	-194.4***	47.66**	-139.7**
	(0.0154)	(56.14)	(21.5)	(66.41)
Panel [B]: Local average treatment effects				
Discontinuity		-765.6***	189.8**	-550.5**
		(227.9)	(84.86)	(264.3)
Bandwidth (gal)	45.17	39.16	68.84	36.37
Observations	57,259	49,316	97,861	48,509

*p<0.1; **p<0.05; ***p<0.01 Each cell presents a nonparametric regression discontinuity estimate at the cutoff for automated violation notices. Following Lee and Card (2008), standard errors in parentheses are clustered along the running variable, which is discrete in gallons. The bandwidth for each specification is selected nonparametrically using a triangular kernel (Lee and Lemieux, 2010). Column (1) provides the estimated first-stage for automated violation notices, corresponding to Figure 4. These notices were sent to households during the first week of July 2015. Columns (2) - (4) present estimates for average weekly water consumption during July through October 2015, the remainder of the legal and technical local summer water season following the violation notices. Panel [A] shows the reduced-form estimates and Panel [B] shows the estimated local average treatment effects. Column (2) includes consumption only during hours of the week when irrigation was not legally allowed, and Column (3) includes consumption only during hours irrigation was legally allowed: Tuesdays and Saturdays before 9:00 a.m. or after 6:00 p.m. Column (4) includes water consumption pooled across all hours. Figure 5 graphs the reduced-form local averages corresponding to Columns (2) and (4).

A Additional figures and tables

Figure A1: Historical perspective on severity of drought in southern California



Notes: Figure A1 plots historical monthly observed drought severity on the Palmer Drought Severity Index for the hydrological region of Coastal Southern California. Our study period during 2015 lies within the most severe drought on record for the region, but lengthy periods of drought are common.

Figure A2: Example of a WaterSmart Home Water Report

Figure A3: Example of an automated irrigation violation notice

Water and Power

July 7, 2015

RE: Violation of Irrigation Requirements at St

Dear Customer,

Our system has detected that you are watering your landscape more than two days per week, for longer than 15 minutes per irrigation station, or on incorrect days, all of which are in violation of Burbank's Sustainable Water Use Ordinance.

Burbank is in Stage III of the Ordinance and landscape watering during April through October is limited to **Tuesdays and Saturdays, before 9am or after 6pm,** and **no more than 15 minutes per watering station**. Starting November 1, watering will be limited to Saturdays.

Immediately, please take the necessary steps to comply with the drought requirements, including appropriately programming your irrigation controller if you have one. If you use a gardener or landscape professional, make sure that they are aware that only Tuesdays and Saturdays are allowed for watering in Burbank, either before 9am or after 6pm.

There is no need to contact BWP. Simply correct your irrigation schedule as quickly as possible. Provided your changes are made within the week, our system will detect your corrections and discontinue sending you notices. However, if you continue watering your landscape more than two days per week, you will receive a second notice. If you remain in violation of the Ordinance thereafter you will be fined $100. Any additional fines will be $200 for the second violation and $500 for the third or more violations.

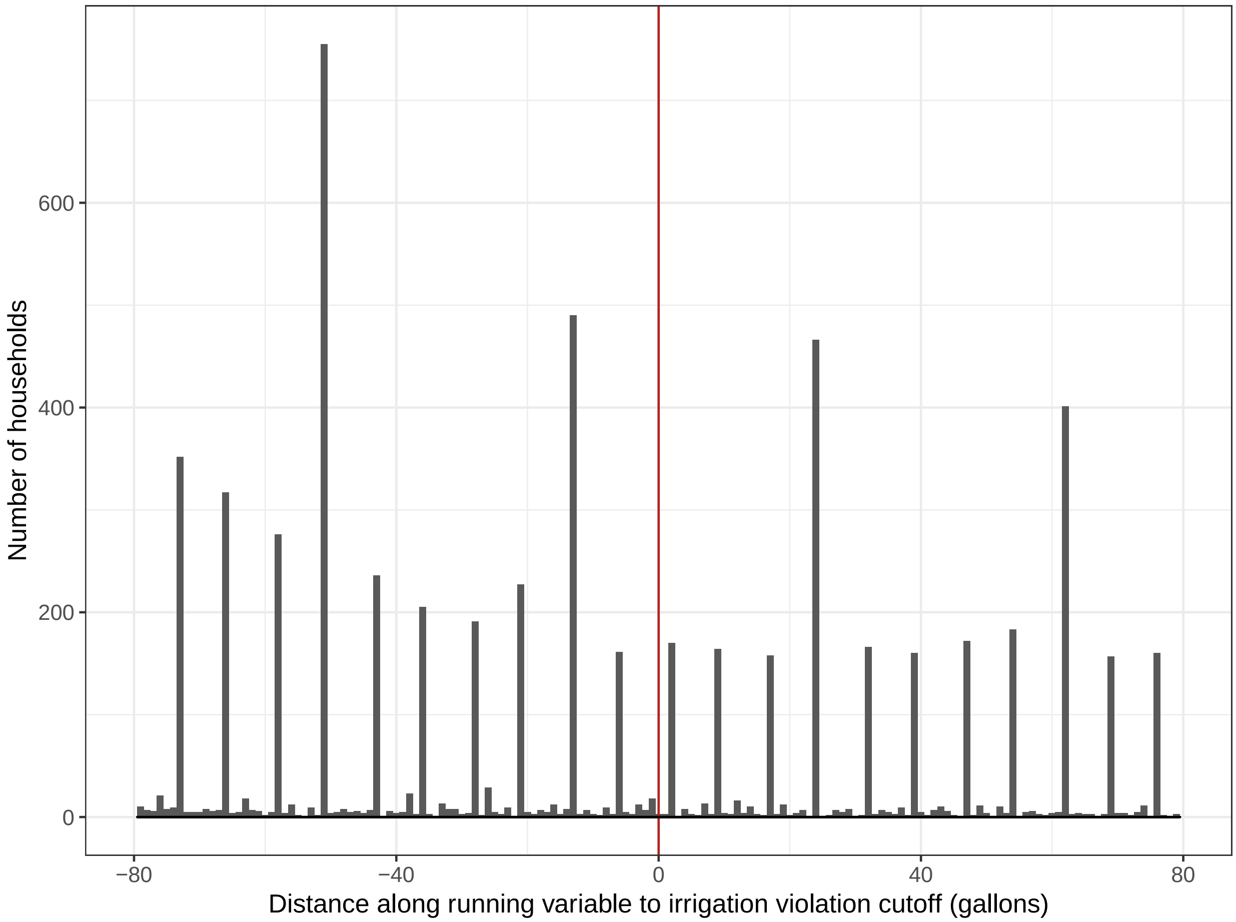
Governor Brown has mandated that Burbank reduce water usage by one billion gallons by February 2016 or face fines of $10,000 per day. We thank you in advance for doing your part to meet this mandate and preserve the water supply to get us through this historic drought.

Sincerely,

Burbank Water and Power

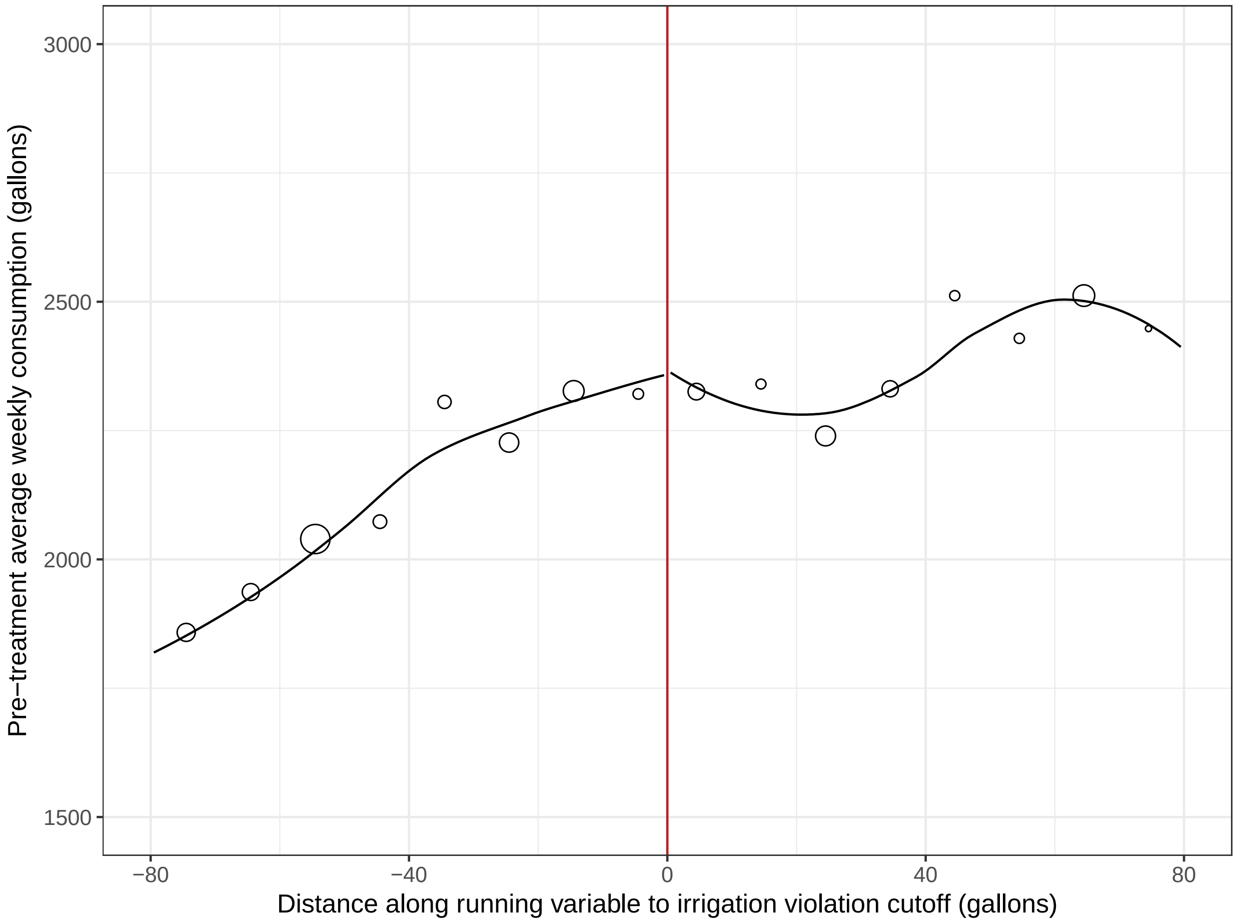
Burbank Water and Power
164 West Magnolia Blvd., P.O. Box 631, Burbank, CA 91503-0631

Figure A4: Distribution of households along the running variable



Notes: Figure A4 plots the distribution of households along the running variable used in our regression discontinuity design, providing a graphical version of the McCrary (2008) bunching test for manipulation with respect to treatment assignment. Due to heterogeneity in the granularity of measurement for included water meters, there is significantly more mass at cubic foot (7.48 gallons) and five cubic feet increments. Importantly, there is no evidence of any excess distributional mass in the region surrounding the cutoff used for determining automated irrigation violation notices.

Figure A5: Identification check for pre-treatment consumption along the running variable



Notes: Figure A5 plots local averages for weekly water consumption during May 2014 through April 2015, the full year prior to both the pecuniary and social comparison treatments. For clarity, the running variable uses 10 gallon bins. The size of the markers corresponds to the number of households included in the local averages. The LOESS curves shown are fit to the underlying microdata separately on each side of the threshold. This identification check shows that pre-treatment water consumption is smooth across the cutoff used for determining automated irrigation violation notices.

Table A1: Randomization balance checks in 80 gallon bandwidth around RD cutoff

	Group means		t-tests	
	(1)	(2)	(3)	(4)
Covariate	Control	Experimental	Difference	p-value
Number of households	1104	5222		
Sent WaterSmart HWR	0	0.9992		
Prior water violation	0.03261	0.03658	0.00397	0.5
Lot size (SqFt)	6861	6906	45	0.53
Irrigable area (SqFt)	3594	3576	-18	0.68
House size (SqFt)	1532	1555	23	0.27
Year built	1944	1944	0	0.87
Number of floors	1.057	1.055	-0.002	0.75
Number of bedrooms	2.854	2.892	0.038	0.18
Number of bathrooms	1.846	1.871	0.025	0.36
Weekly water gallons	2191	2236	45	0.25

Notes: Table A1 shows statistics by WaterSmart Home Water Reports (HWR) treatment arm for household-level covariates for the subset of households near to the RD cutoff. See notes for Table 1 for additional details.