Do behavioral nudges interact with prevailing economic incentives? Pairing experimental and quasi-experimental evidence from water consumption*

Daniel A. Brent[1] and Casey J. Wichman[2]

[1]Pennsylvania State University
[2]University of Chicago, Resources for the Future

January 7, 2020

Abstract

Social comparisons are a popular behavioral nudge to change behavior, partially because raising prices can be politically difficult. In many settings, nudges may interact with prevailing prices, potentially crowding out intrinsic motivation to conserve or by increasing the salience of prices. We investigate the interaction of prices and nudges for water conservation in two experiments in neighboring utilities. First, we layer randomized behavioral treatments on top of variation in price driven by arbitrary lot-size thresholds that assign marginal prices to customers exogenously. Second, we explore whether behavioral treatments affect consumers' price sensitivity. We find no consistent evidence that social comparisons are more effective at inducing conservation at higher prices or that they increase consumers' price sensitivity. Ultimately, we find little empirical support that consumers respond to behavioral treatments due to private economic benefits.

Key Words: behavioral interventions, social norms, field experiments, water conservation, price sensitivity, water demand
JEL Codes: D12, C93, H42, L95, Q21, Q25

*Brent: Department of Agricultural Economics, Sociology, & Education, Pennsylvania State University, dab320@psu.edu. Wichman: Energy & Environment Lab, University of Chicago; Energy Policy Institute at the University of Chicago; and Resources for the Future; wichman@uchicago.edu. D.A.B. and C.J.W. jointly conceived the idea, performed the analysis, and wrote the paper. The authors thank Michael Price, Nathan Chan, and Koichiro Ito for helpful comments, as well as conference and workshop participants at the 2017 AERE summer conference, Resources for the Future, University of Southern California, Pennsylvania State University, University of Chicago, and the 2019 ASSA meetings.

1 Introduction

Behavioral interventions are widespread policy options for altering consumption choices. Governments, policymakers, and industry around the world now look to behavioral economics to manage private and social costs. Behavioral economics has inspired policies targeting a wide range of outcomes including tax evasion (Hallsworth et al., 2017), charitable donations (Croson and Shang, 2008; Shang and Croson, 2009), education (Levitt et al., 2016), healthy eating (Hanks et al., 2012; List and Samek, 2015), and exercise (Royer et al., 2015). These interventions have motivated, and in some cases are the output, of government-run "nudge units" such as the United Kingdom's Behavioural Insights Team.

Nowhere have behavioral nudges been more pervasive than for managing energy and water consumption (e.g., Allcott, 2011; Ferraro et al., 2011; Allcott and Rogers, 2014; Brent et al., 2015; Ito et al., 2018; Byrne et al., 2018). Regulated industries, such as electricity or water and sewer service, are limited by how much they can use price as a tool of conservation. In the state of California, for example, water utilities cannot charge a price greater than cost of service, effectively rendering scarcity pricing illegal.[1] As a result, utilities often rely on nonprice demand-management tools to encourage conservation. Researchers have shown that social comparisons can be effective nonprice policies for conservation, reducing household energy and water consumption between two and five percent (Allcott, 2011; Ferraro and Price, 2013; Brent et al., 2015). At scale, these small reductions can generate substantial benefits for the service provider at relatively low cost, potentially delaying or avoiding investment in costly new power plants or water sources.

A notable feature of this literature on social comparisons is that the treatment effect estimates are typically causal, arising from the randomized nature of program designs implemented by companies such as OPower and WaterSmart Software. Reconciling these estimates with models of consumer behavior, however, is less transparent. Some have claimed that norm-based information treatments apply a *moral tax* to consumption of externality-producing goods (Levitt and List, 2007; Ferraro and Price, 2013). Others, however, have claimed that information treatments reduce the distortion in consumer's perceptions of price and quantity consumed, thus reducing informational "internalities"—or optimization mistakes—borne by consumers (Allcott and Taubinsky, 2015; Wichman, 2017). As such, there are competing views of whether behavioral interventions affect an individual's intrinsic motivation to conserve, provide direct economic benefits to the consumer, or both.

In line with understanding the behavioral mechanism of underlying conservation behavior, we posit that behavioral policies may interact with prevailing market mechanisms in an ambiguous way. Because behavioral interventions in electricity and water demand

[1] See, e.g., http://www.latimes.com/local/orangecounty/la-me-rates-decision-20150421-story.html.

managment are always overlaid on top of contemporaneous pricing structures, there is latent potential for economic incentives and nudges to interact in an ambiguous way. We show theoretically that existing interpretations of nudge treatment effects may confound behavioral and economic explanations if there is an economically significant interaction between prices and nudges. Within the current literature, there is virtually no evidence of whether this interaction is meaningful. Our paper fills this gap.

We explore the impacts of a social messaging experiment and large changes in marginal prices on water conservation behavior. Our analysis produces causal effects by design: first, we evaluate the effects of two independent, randomized messaging experiments implemented by WaterSmart Software at different points in time for neighboring water utilities in Southern California. Second, we exploit two sources of variation that introduce price changes at the household level. One source of price variation comes from arbitrary lot-size thresholds within nonlinear water rate structures that we exploit in a regression discontinuity design. The second source of price variation arises from the utilities' rate-setting practices, included in an instrumental variables framework. Our methodology cleanly identifies the separate impact of the social comparison treatment and price on consumer behavior, as well as their joint effect.

Within our unique empirical approach, we answer two questions. First, do customers facing different price levels respond more strongly to norm-based conservation campaigns? We refer to this as the *price-level effect*. We identify the price-level effect from comparing responsiveness to behavioral treatments for otherwise identical households on either side of a price discontinuity introduced by arbitrary lot-size thresholds within a utility's rate structure. Second, do norm-based conservation campaigns increase customers' price sensitivity? We refer to this as the *price-sensitivity effect*. We identify the price-sensitivity effect by estimating demand equations and observing whether our randomized behavioral treatment significantly alters our estimate of the price elasticity. Both effects are theoretically plausible ways in which nudges and prices interact.

Our results show no consistent evidence that social comparisons generate more conservation for households facing an exogenously larger marginal price of water. Higher prices cause small and insignificant decreases in the magnitude of the treatment effect from peer comparisons. Additionally, we find similarly weak and inconsistent evidence of a price sensitivity effect. Treatment induces small increases in the magnitude of the price elasticity in some specifications, although these effects disappear in alternative specifications.

Because norm-based policies are implemented broadly for water and electricity, the policy implications of this research are vast. Allcott and Rogers (2014) and Brent et al. (2015) both show that behavioral nudges interact with prevailing conservation policies. Recent research shows that the mechanisms through which consumers respond to behavioral nudges has important welfare implications (Allcott and Kessler, 2019; Taylor et al., 2018). Nudges generate unambiguous welfare gains if consumers conserve due to correcting internalities. However, if consumers respond due to a moral tax on consumption then welfare only increases if the price of the resource is below its marginal social cost. Strong interactions between nudges and prices would indicate that consumers are at least in part responding to increases in the private benefits from conservation, given the extensive evidence that consumers in these settings do not have full information about prices (Sexton, 2015; Wichman, 2017; Brent and Ward, 2019) or are not responding according to standard neoclassical theory (Sallee, 2014; Allcott and Wozny, 2014; Jacobsen, 2015). Therefore, although it is difficult to directly measure the welfare benefits of behavioral interventions, we find evidence that supports the fact that social comparisons operate as a moral tax on consumption.

There is, however, a growing body of evidence that focuses on comparing the effects of moral and neoclassical incentives on energy and water consumption. Ito et al. (2018) explore the effectiveness of a standard moral suasion nudge relative to dynamic electricity pricing treatments. They find that moral suasion induces sizable effects in the short-run that dissipate quickly relative to dynamic prices that exhibit longer-run effects. Our project is different in that we seek to understand how the moral suasion treatment interacts with *underlying* economic incentives. Additionally, Brandon et al. (2018) implement a randomized OPower experiment in which personalized energy reports were sent to electricity customers that targeted aggregate savings or peak-load savings, and measured the response of these treatments during peak-load and non-peak load events. They find that a combination of treatments induced a larger effect than the joint effect of each treatment in isolation or, in other words, that treatments were complimentary. This result suggests an important role for exploring other policy complimentarities, particularly with respect to interactions with economic incentives because nudges can highlight the private economic benefits of conservation. Additionally, West et al. (2019) show that strong fines for violating outdoor water restrictions do not impact the behavioral response from peer comparisons. Finally, in another project, List et al. (2017) show that economic incentives (via a rewards program) can better target electricity consumption reductions from low-use, low-variance households, who are typically less responsive to nudges. Importantly, electricity and water are often priced using nonlinear increasing-block rate structures where the economic benefits from conservation are positively correlated with consumption. Thus it feasible that low-use, low-variance consumers respond to nudges differently because of different private economic returns from conservation. This latter effect is precisely what we seek to estimate in this paper.

Overall, we find little evidence that moral nudges interact with underlying economic incentives. This is an important result because nearly all behavioral public policy has the potential to interact with existing neoclassical incentives. Placed alongside the previously mentioned literature, our study provides a clearer view of the mechanisms underlying responses to behavioral treatments. Behavioral nudges can be criticized for providing too many types of information to isolate the relevant mechanism for consumer behavior, but we fail to find convincing evidence that making economic incentives more salient is a relevant factor for behavioral interventions. This finding sharpens our view of past and future conservation policies because nearly all behavioral nudges for electricity and water are layered on top of prevailing rate structures. Additionally, in situations where resource use goes unpriced, our results suggest that behavioral treatments can still be a useful policy instrument to govern consumer behavior in a socially advantageous way.

2 Conceptual framework

To show how nudges and incentives interact conceptually, we begin with the general framework of Allcott and Kessler (2019). Consider a consumer with income y who gains utility from the consumption of water w and numeraire good x. w generates consumption utility of $f(w;\alpha)$, where α captures consumer tastes as a demand shifter. We include an internality parameter $\gamma > 0$ that affects choice but not experienced utility, such as imperfect information, mistakes in evaluating private benefits of water consumption, or some other behavioral bias. For our purposes, it is useful to think of γ as inattention to water consumption. Consumers thus have perceived utility $\hat{f}(w;\alpha,\gamma)$, which we assume takes the form $\gamma^{-1}f(w;\alpha)$. Thus, utility is expanded for $\gamma > 1$ and contracted for $0 < \gamma < 1$.

Following Levitt and List (2007) and Ferraro and Price (2013), we include a "moral utility" term, $M = m - \mu w$, which captures nonpecuinary impacts associated with consumption of w. We define $\mu \geq 0$ as a marginal "moral tax" on consumption of w.

We summarize individual-specific parameters in the vector $\theta = \{y, \alpha, \gamma, m, \mu\}$ so that the consumer maximizes

$$\max_{x,w} \hat{U}(\theta) = x + \gamma^{-1} f(w;\alpha) + m - \mu w \tag{1}$$

subject to her budget constraint

$$y = x + pw \tag{2}$$

where $p \geq 0$ is the marginal price for water consumption. The consumer allocates all non-water expenditures to the numeraire, thus satisfying her budget constraint with equality. We focus only on interior solutions.

Standard first-order conditions govern the consumer's choice of water consumption, $\tilde{w}$:

$$f'(\tilde{w};\alpha) = \gamma(\mu + p). \tag{3}$$

Eq. 3 states that consumers will choose consumption of $\tilde{w}$ to equalize their marginal experienced utility with the sum of perceived monetary and moral costs. Because γ introduces a wedge between experienced marginal utility and a consumer's true marginal utility, choice of $\tilde{w}$ is not required to be individually optimal. The framework so far is consistent with stylized formulations in Sexton (2015) and Wichman (2017) who model price (and quantity) misperceptions. The only difference is the inclusion of the Ferraro and Price (2013) moral cost parameter.

We can express changes in consumption by totally differentiating Eq. 3:

$$f''(\tilde{w};\alpha)\mathrm{d}\tilde{w} = \mu\mathrm{d}\gamma + \gamma\mathrm{d}\mu + p\mathrm{d}\gamma + \gamma\mathrm{d}p. \tag{4}$$

Now, let the nudge be represented by changes in attention to water use ($\mathrm{d}\gamma$) and changes in the moral cost of consumption ($\mathrm{d}\mu$).[2] Because the nudge does not affect the market price of water (i.e., $\mathrm{d}p = 0$), we can express the demand effect of a nudge as

$$\mathrm{d}\tilde{w} = \frac{1}{f''(\cdot)}\left[(\mu + p)\,\mathrm{d}\gamma + \gamma\mathrm{d}\mu\right]. \tag{5}$$

Under standard assumptions of demand (i.e., diminishing marginal utility), f'' is weakly negative, which implies that the nudge will (weakly) reduce water demand in equilibrium for $\gamma < 1$.[3] Eq. 5 shows that the total effect of the nudge depends on how changes in perceptions interact with moral and explicit prices as well as how changes in moral costs interact with perception. The vast majority of research to date assumes implicitly that the increased salience of private economic benefits of conservation are negligible; in other words, these studies interpret the effect of the nudge as if $p\mathrm{d}\gamma = 0$. That is, the majority of experiments focused on exploring the effects of salience or moral suasion ignore their underlying interaction with prices. This omission is a potentially important oversight because behavioral interventions for water and energy use are implemented on top of prevailing prices, which are often nonlinear and thus different for different consumers. Furthermore, many nudges aimed at water and energy conservation, including the one analyzed in this paper, explicitly communicate the private financial benefits of conservation.

This simplified representation of demand translates directly to our first empirical hypothesis: the existence of an economically important interaction between behavioral treatments and conventional pricing mechanisms. We define this effect as the **price-level effect (PLE)**, which measures the magnitude of $\mathrm{d}\tilde{w}$ in response to the nudge that is driven by differences in price levels. Our null price-level hypothesis posits that $p\mathrm{d}\gamma = 0$. Evidence of a nonzero price-level effect would lend support to the notion that consumers change consumption in part due to changes salience of private economic benefits from conservation. We test this by comparing the effect of randomized nudges for households who face exoge-

[2]For clarity, we assume the nudge is corrective in that it reduces information distortions, or $\mathrm{d}\gamma \implies \gamma \to 1$.

[3]For $\gamma > 1$, the nudge could increase consumption if, e.g., consumers had been initially over-perceiving the costs of consumption. This stylized result is captured empirically in Wichman (2017).

nously different marginal prices. We describe our empirical identification of this effect in the subsequent section.

Additionally, we explore a second, complementary approach to investigate whether nudges affect consumer demand through neoclassical price mechanisms. Consider a change in price, dp. Using Eq. 3, we can define the resulting price elasticity,

$$\hat{\varepsilon}^p = \frac{\gamma}{f''(\cdot)} \frac{p}{\tilde{w}} = \gamma \varepsilon^p \tag{6}$$

where ε^p is the neoclassical price elasticity of demand and the hat indicates "perceived" price elasticities. This formulation leads directly into our second hypothesis. We define the **price-sensitivity effect (PSE)** as the degree to which nudges affect price sensitivity. Because social comparisons operate through both channels of μ and γ, our null price-sensitivity hypothesis is $\partial\hat{\varepsilon}_p/\partial\gamma = 0$. Evidence of a nonzero price-sensitivity effect would suggest that consumers' sensitivity to price is affected by the nudge, thus providing support for the idea that consumers respond to nudges, at least in part, because of the private economic benefits that arise from internality correction.

3 Empirical setting and strategy

3.1 Data

The data we use in the analysis are household-level water consumption records for two utilities in Southern California. We obtained these data through partnership with WaterSmart Software. We refer to the larger utility in our sample as "Large Utility" and, correspondingly, the smaller utility is "Small Utility."[4] These two utilities share a geographic border and their residents form a common labor market along with other nearby municipalities. Both utilities combine water and sewer services and also serve as the electric utility. Figure 1 shows the geographic distribution of households in the treatment and control groups in each utility.

Large and Small Utility have different pricing structures, and the water rates have changed over time. Large Utility has "budget-based" increasing-block rates in which consumption thresholds for the marginal price blocks vary with geographic region and lot size. This means different households will be assigned to higher marginal prices at different levels of consumption. There are three geographic zones: low, medium, and high. The geographic zones refer to the water requirements for irrigation based on temperature conditions; the low zone has the most moderate weather and the high zone has hotter weather. There are five lot size thresholds leading to fifteen unique sets of consumption tiers that

[4] As part of the confidentiality agreement we cannot disclose the names of these utilities.

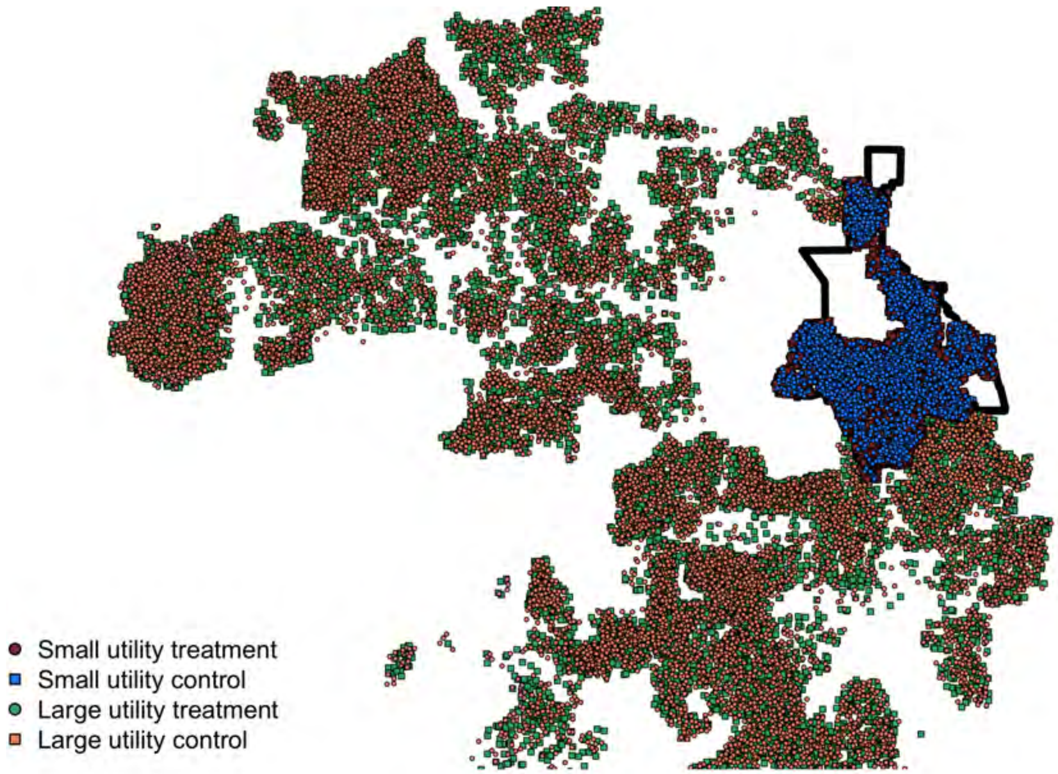


Figure 1: Water utility service area (partial), sample boundaries, and households by treatment status

Note: The small utility is contained within the black border and the large utility is outside the border. Household locations are scrambled by .001 decimal degrees to preserve anonymity.

determine marginal prices. Small Utility has a standard increasing-block rate structure. Figure 2 displays the full water rate structure over time for both utilities.

For each household in our sample, we have consumption for the given billing period and the relevant prices for consumption. Households receive water bills every two months. To protect anonymity, geographic coordinates for each household were scrambled within 0.001 degrees (a maximum of approximately 365 feet), which permits us to identify the neighborhood of the household, but not its exact address. Each account in our sample was randomized into a treatment or control group by WaterSmart Software. All households in each utility begin receiving HWRs at the same time (details on randomization and treatment are described below). Households in both utilities are billed bimonthly leading to six billing periods each year. Treatment began during the sixth billing period of 2014 in Small Utility and during the second billing period of 2015 in Large Utility.

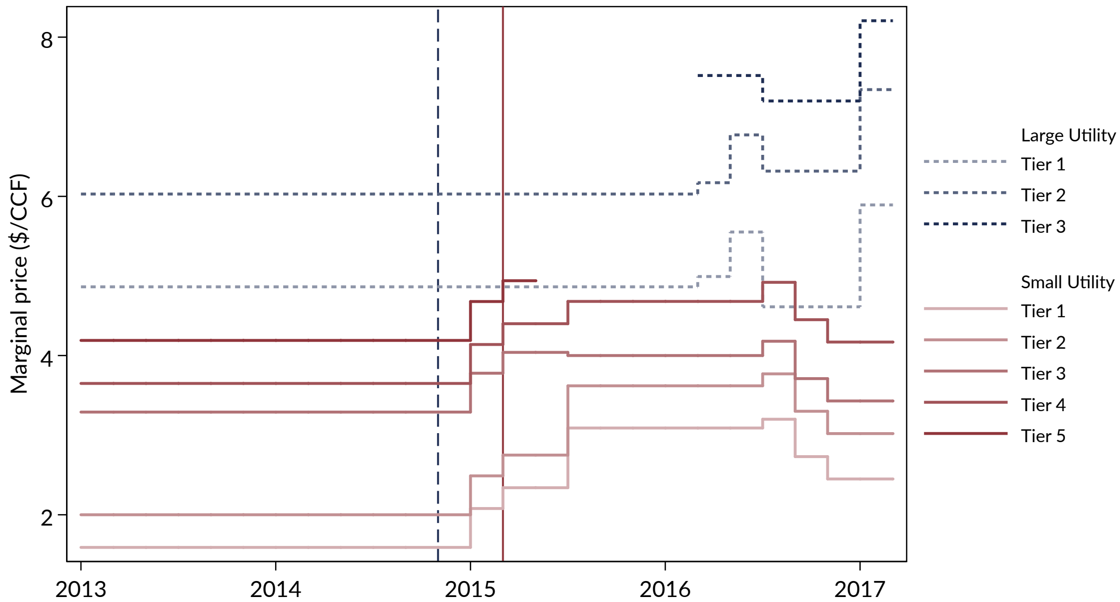


Figure 2: Marginal Prices Over Time

Note: The colors depict the marginal price for different consumption tiers. The dashed lines show prices for the Large Utility and the solid lines show marginal prices for the Small Utility. The vertical dashed and solid lines depict the treatment start date for the Large and Small Utilities respectively.

3.2 Experimental design

WaterSmart Software (henceforth WaterSmart) is a smart-technology company that contracts with water utilities to help them manage demand.[5] In addition to providing analytical support to utilities, WaterSmart primarily focuses on helping utilities reduce water consumption by providing consumers with additional information through customized Home Water Reports (HWRs) (Figure 3) and an online customer account portal. WaterSmart bears resemblance to the model of OPower for electricity customers analyzed in Allcott (2011). For many utilities WaterSmart randomizes the assignment of households who receive HWRs in order to evaluate the causal impact on water consumption (see, e.g., Brent et al. (2015)). Because customers opt-in to viewing their online account, we focus here on the treatment effect for households receiving a HWR (i.e., intent-to-treat effects).

The one-page HWR as tested has three components. The main component (in the upper left of the figure) is a social comparison. WaterSmart estimates the household's total water consumption over the prior two months from utility billing records and compared that to the consumption of "average neighbors" and "efficient neighbors." "Neighbors" are defined as households that have the same number of occupants and similar irrigable area across the utility, such that the general water requirements within a peer group are comparable.

[5]More information is available on their website: http://www.watersmartsoftware.com/.

"Efficient neighbors" were peers with consumption in the bottom 20%. Households with consumption above the median of their peer group receive a "Red" normative message (shown in Figure 3), those with consumption between the median and 20^{th} percentile receive a "Yellow" message, and those below the 20^{th} percentile receive a "Green" message. (Home Water Reports showing the latter two categories are provided in the Appendix).

The second component (across the bottom of Figure 3) is a list of three personalized recommendations for strategies to save water. Recommendations include installing low-flow toilets and switching to native plants. Based on data available from the utility (described more below) or on results from a baseline household survey with limited responses, WaterSmart personalized these recommendations to the extent possible. For example, if a household had no outdoor area it was not given a recommendation regarding irrigation. The personalized recommendations provide estimates of the water savings in gallons and in dollars, and the dollar estimates rely on the highest marginal price the household faced last month. The third component (in the upper right of Figure 3) cycles between a variety of messages about water conservation and utility programs.

To show that the randomization was conducted properly we graph average water use over time across treatment groups and perform a variety of balance tests. Figure 4 shows the average water consumption for the treatment and control groups in both utilities. The treatment and control groups have similar consumption prior to the intervention and after treatment the treatment groups use less water. Table 1 shows that treatment and control groups are well balanced on a range of observables based on a variety of parametric and non-parametric tests. Out of the 42 tests performed only three (7%) have a p-value less than 0.05 and only four (9.5%) have a p-value less than 0.1.

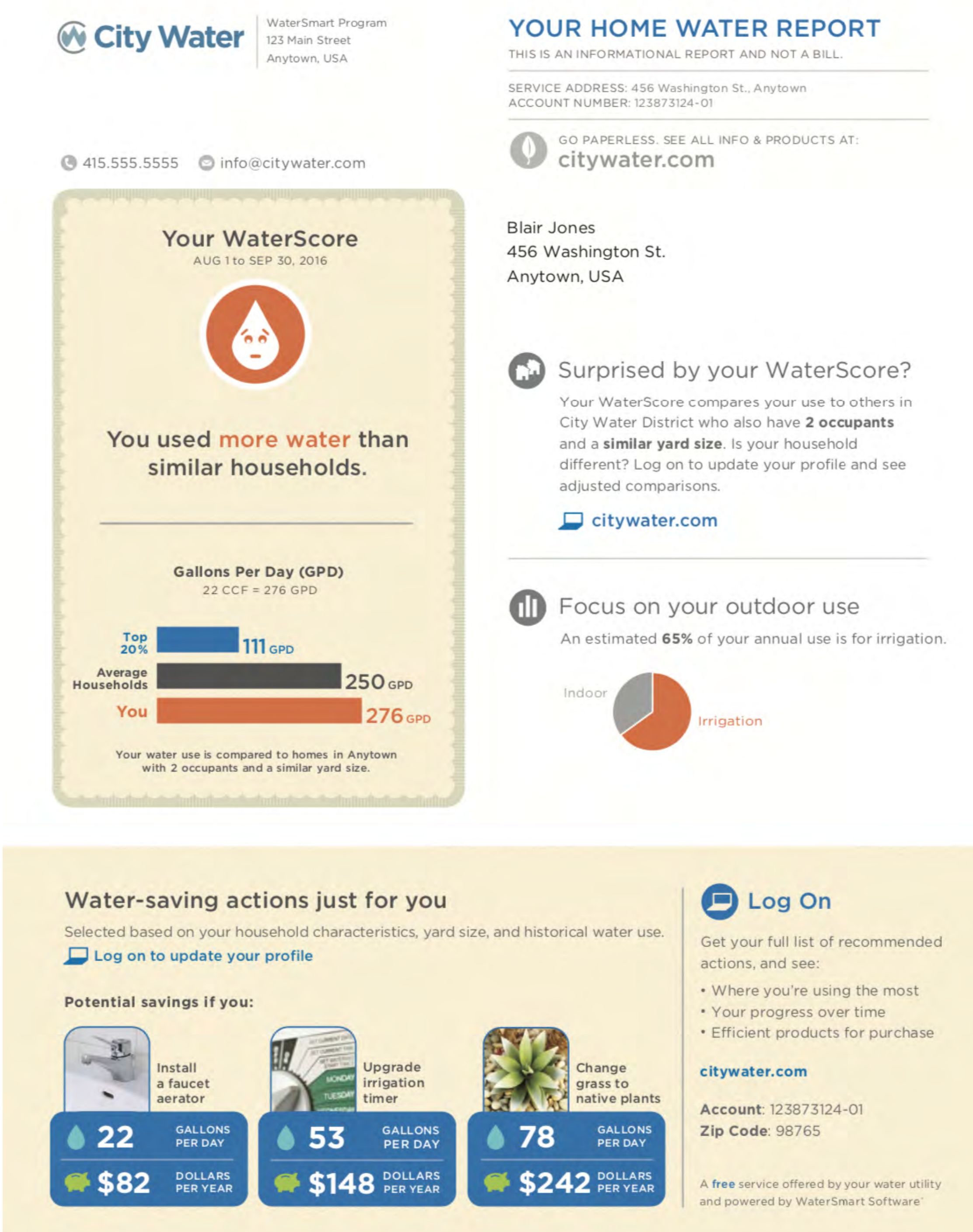


Figure 3: Home Water Report

Note: This is an example of a generic "Red" Home Water Report (HWR). These households used more than the median of their peer group.

Figure 4: Water Consumption by Treatment Status and Utility

Note: The graph displays the average consumption of the treatment and control groups in each utility for every billing period in the sample. The solid lines represent the control groups and the dotted lines represent the treatment groups. The vertical dashed lines designate the start of the treatment period for each utility. The line colors designate utilities.

Table 1: Summary statistics and balance on observables

Sample	Variable	Treat	Control	Difference	KS	MW	T
Small Utility	Pre-Treatment Water	443.8	440.1	3.7	0.71	0.63	0.32
Small Utility	Pre-Treatment Water (Summer)	506.1	502.3	3.8	0.42	0.96	0.38
Small Utility	Pre-Treatment Water (Winter)	404.4	401.3	3.1	0.19	0.51	0.36
Small Utility	Lot Size	9955.6	9641.5	314.1	0.18	0.31	0.04
Small Utility	Sq. Ft.	1953.5	1954.3	-0.9	0.37	0.45	0.95
Small Utility	Beds	3.0	3.0	-0.0	0.34	0.99	0.78
Small Utility	Baths	2.2	2.2	0.0	0.94	0.44	0.58
Large Utility	Pre-Treatment Water	609.2	609.0	0.2	0.37	0.97	0.93
Large Utility	Pre-Treatment Water (Summer)	722.6	722.4	0.2	0.48	0.72	0.95
Large Utility	Pre-Treatment Water (Winter)	552.9	552.8	0.1	0.31	0.76	0.96
Large Utility	Lot Size	10345.6	10313.8	31.8	0.70	0.82	0.70
Large Utility	Sq. Ft.	2146.9	2173.5	-26.6	0.00	0.07	0.01
Large Utility	Beds	3.5	3.5	-0.0	0.42	0.97	0.68
Large Utility	Baths	2.5	2.5	-0.0	0.38	0.32	0.30

Note: The table shows the average values for a variety of households characteristics for the treatment and control groups in each utility. All the pre-treatment water variables are measured in gallons-per-day. Lot size and sqft (indoor living space) are measured in square feet. Beds and baths are the number of bedrooms and bathrooms. The last three columns present the p-values from test statistics. KS is the non-parametric Kolmogorov-Smirnov equality of distributions test, MW is the non-parametric rank-order test, and T is the two-sided t-test for difference in means.

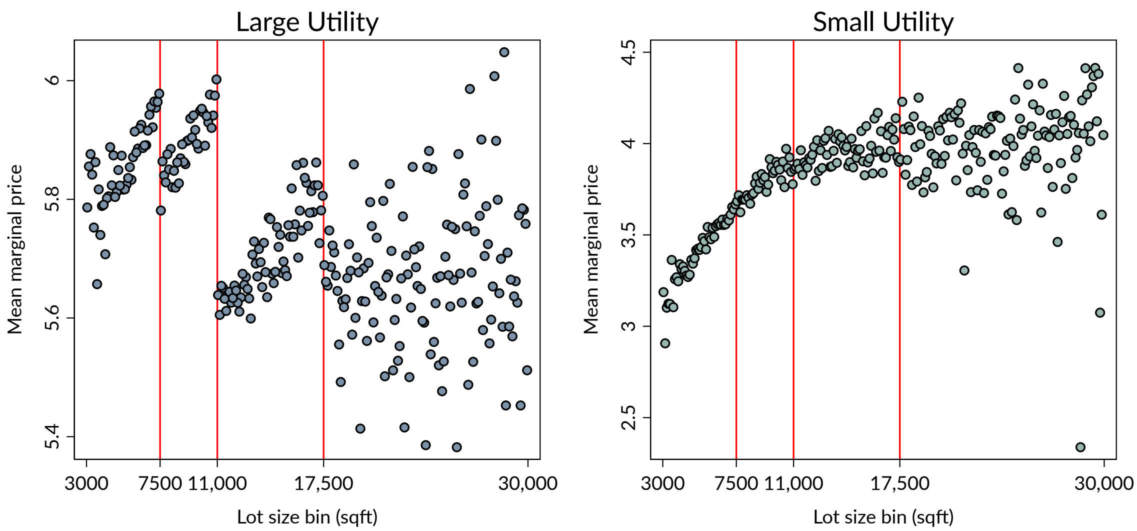
3.3 Quasi-experimental design

In order to identify the differential impact of HWRs for households facing different prices we estimate a difference-in-discontinuity model (DD) that exploits a discontinuity in the rate structure of the Large Utility. The Large Utility uses a budget-based increasing-block rate structure in which the tier thresholds depend on the climate zone and lot size (in square feet). There are five lot-size tiers (0-7499, 7500-10,999, 11,000-17,499, 17,500-43,559, $\geq$ 43,560), and households with smaller lot sizes are allocated less water before moving to a higher pricing tier. Therefore, households that are just below the lot-size tier threshold (e.g., 7499 sqft) on average face higher prices than households just above a tier threshold (e.g., 7500 sqft).[6] We exploit this threshold by restricting the analysis to various lot-size bandwidths such that the households are relatively close to the lot-size thresholds.

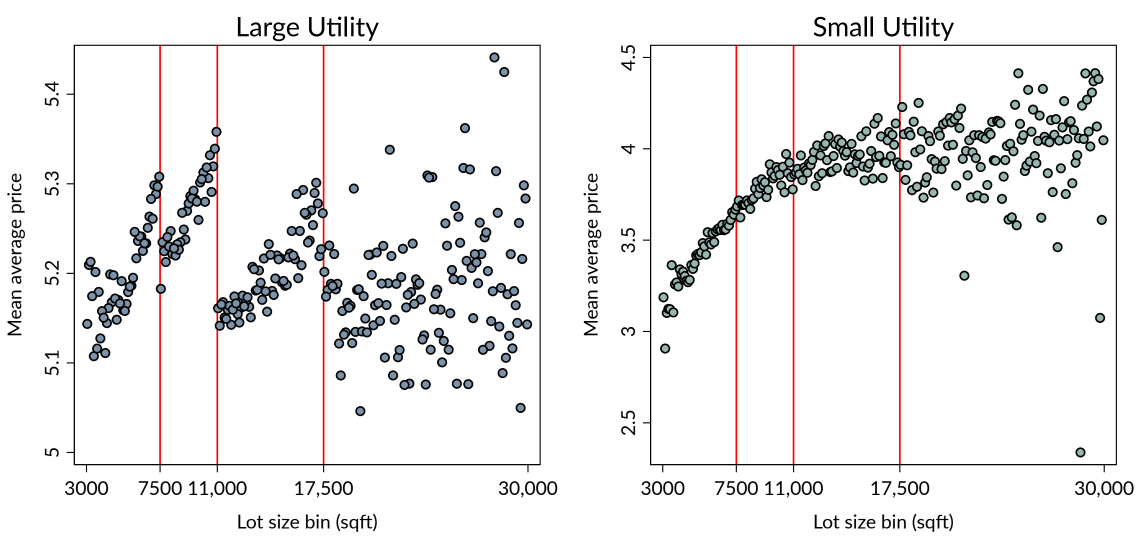
In Figure 5, we show how the lot-size threshold introduces a discontinuous effect on the likelihood of facing a higher marginal prices in each utility. These figures present mean marginal (average) prices relative to lot size in 100 sqft bins. There is a distinct jump in the expected marginal price for households just below the lot-size threshold in the Large Utility (panel (a)), but not in the Small Utility (panel (b)). To highlight the difference in typical marginal prices induced by the lot-size threshold we define "low" households as those less than 1000 feet below a lot-size threshold (e.g. 6499-7499 sqft). We define "high" households as those less than 1000 feet above a lot-size threshold (e.g., 7500-8500 sqft). The raw data show that in the Small Utility the average marginal price for low and high lot-size households is \$3.65 and \$3.77 respectively—so high lot-size households on average pay more for water because they are typically larger water consumers. In the Large Utility the average marginal price for low and high lot-size households is \$5.98 and \$5.81 respectively—low lot-size households pay more for water *despite* the fact that they are lower users on average. Note that although the average price difference in the Large Utility from the lot-size threshold is only \$0.20, the marginal price increase from moving to the higher tier is more than \$1. The average marginal prices reflect both the change in marginal prices and the probability that a household moves into the higher consumption tier. Therefore, some households will face significant marginal price increases due to the lot-size threshold discontinuity. Because the lot-size threshold introduces a discontinuity in the probability that a household faces a higher marginal price, we evaluate this price change in a fuzzy regression discontinuity design.

In Figure 6, we present the price variation that we are exploiting in a different way. Here we show the different rate structures for households in three different lot-size groups. Each

[6]Because this budget-based billing only occurs in the Large Utility, we also use a third difference as a robustness check (in a difference-in-difference-in-discontinuity design, or DDD) to compare similar households above and below the lot-size threshold across utility boundaries. These households below the lot-size discontinuity will only face higher prices in the Large Utility.



(a) RD treatment in marginal prices at lot-size thresholds for both utilities



(b) RD treatment in average prices at lot-size thresholds for both utilities

Figure 5: Discontinuities in average and marginal price driven by lot-size

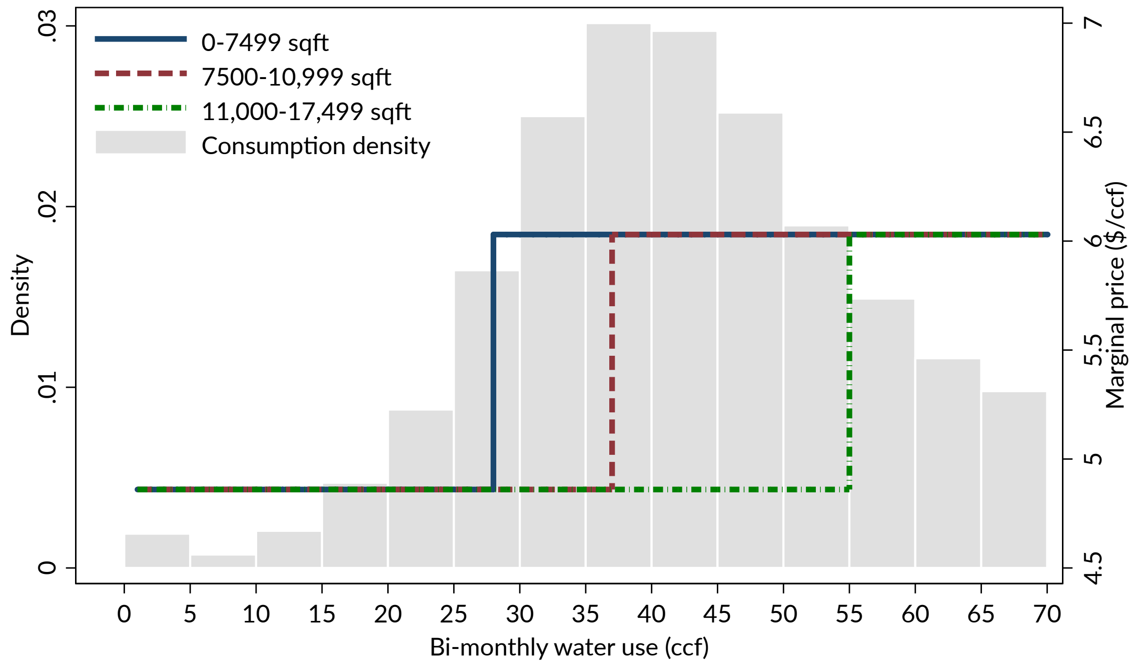
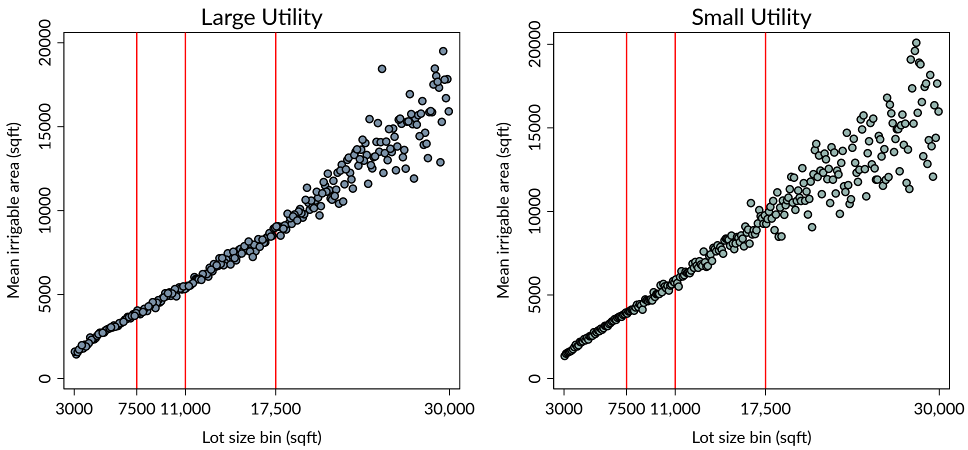


Figure 6: Changes in marginal prices driven by lot-size groups

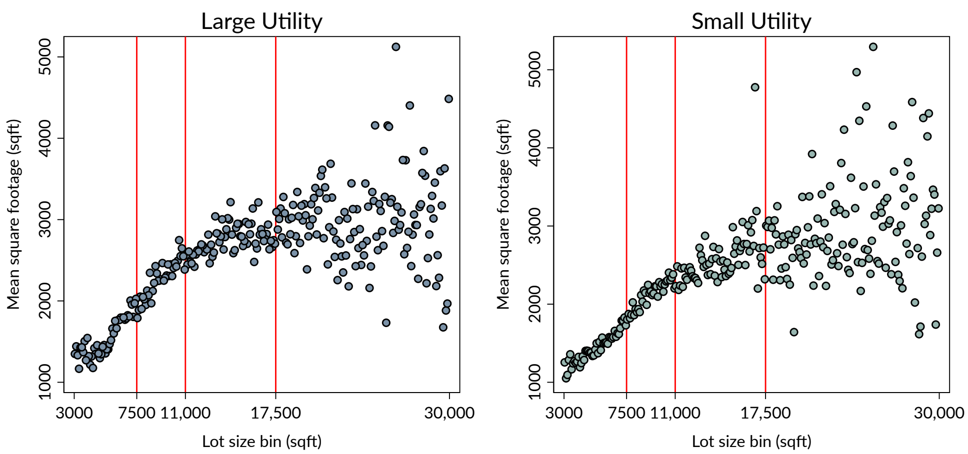
Note: Bi-monthly water distribution is truncated at 70 ccf. Rate structure presented is for the Large Utility prior to the introduction of a third price tier later in the sample.

lot-size group faces the same set of marginal prices, but larger lots are allocated a larger proportion of bi-monthly consumption at the lower marginal price. As shown, a household with a 7400 sqft lot moves into the second price tier at 28 ccf, whereas a household with a 7500 sqft lot moves into the second price tier at 37 ccf. Further, a household with an 11,000 sqft lot does not enter the second price tier until 55 ccf. Moreover, these inframarginal price differences are not trivial: households with lots smaller than 7500 sqft face a marginal price increase of 19.4% nine units of consumption sooner than do households with slightly larger lot sizes. At the 11,000 sqft threshold, this inframarginal price difference is sustained for 18 ccf every two months.

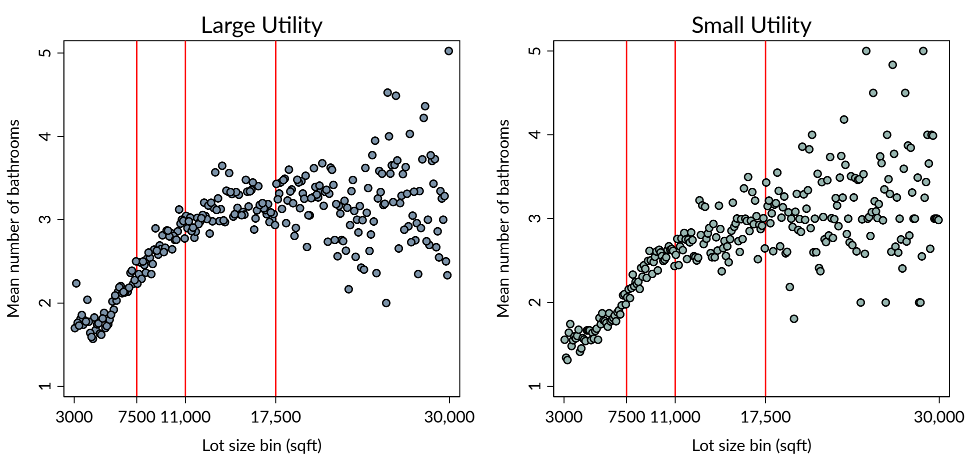
The standard identifying assumptions in RD frameworks are that: (a) other covariates move smoothly through the discontinuity induced by the running variable, and (b) the running variable cannot be manipulated. The latter assumption is satisfied by noting that lot sizes are fixed over time and recorded by county surveyors. In our setting, if other variables associated with water consumption changed discontinuously then we would worry that (a) is not satisfied. As a visual test of this assumption, we present in Figure 7 three relevant variables for water consumption across our RD threshold: irrigable area of lot, indoor square footage of the home, and number of bathrooms. Notably, irrigable area, which we



(a) Mean irrigable area



(b) Mean square footage



(c) Mean number of bathrooms

Figure 7: Covariate distributions across lot-size thresholds for both utilities

anticipate to be highly correlated with lot size, moves nearly linearly through the lot-size discontinuities, which provides convincing support for the RD assumptions. We observe no obvious discontinuity in square footage and number of bathrooms at the discontinuities either. We present the same graphical analysis for additional covariates in the appendix (See Figure A.3).

3.4 Estimating baseline treatment effects of home water reports

Our primary regression framework is a panel difference-in-difference design to estimate the effect of the randomly assigned HWR treatment on average household water use. We calculate normalized water use by dividing each household's water use in gallons-per-day (GPD) by the average consumption of the control group in the post-treatment period within the same utility. This specification maintains the interpretation of coefficients as percentage changes in water consumption, but unlike the logarithmic transformation does not dampen the effect of high users. This is important in the context of social comparisons because prior research shows that most of the savings are concentrated among high users (Allcott, 2011; Brent et al., 2015). We include household fixed effects to control for all static household heterogeneity. Although we are not concerned about traditional forms of endogeneity due to random assignment of treatment we prefer the specification with household fixed effects to focus on how any price effects from the lot-size discontinuity change once a household starts receiving HWRs.

To estimate our primary treatment effects, we specify the following equation:

$$\tilde{w}_{it} = \alpha_i + \gamma_1 \text{Treat}_{it} + \gamma_2(\text{Treat}_{it} \times \text{Large}_i) + \beta X_{it} + \tau_{it} + \varepsilon_{it}, \tag{7}$$

where $\tilde{w}_{it}$ is normalized average daily water consumption for household i during billing period t. Treat_{it} is an indicator if household i was in the randomized treatment group in a treated time period. We interact the treatment indicator with Large_i, an indicator for whether the household is in the Large Utility, to account for potential heterogeneity in treatment effects across the utilities. Lastly, X_{it} is a vector of weather controls, α_i is a household fixed effect, τ_{it} is a period-by-utility fixed effect, and ε_{it} is the residual error term. We cluster all standard errors at the household level. $\hat{\gamma}_1$ and $\hat{\gamma}_1 + \hat{\gamma}_2$ are average treatment effect estimates of the HWRs for the Small and Large utilities, respectively.

Additionally, to ensure that we are comparing similar populations who face similar temporal shocks in both utilities, we examine the treatment effect model restricted to households within 10 kilometers (km) from the shared utility border.

3.5 Estimating the price-level effect

The price-level effect (PLE) is the differential responsiveness to HWR treatment driven by different economic incentives that households face. We exploit two sources of variation to identify the PLE in a difference-in-discontinuity (DD) design. The first source of variation is the random assignment of treatment status and the second source is identified by the lot-size discontinuity in the rate structure.

Formally, our approach interacts the variables in equation 7 with an indicator for whether the household is below the lot-size threshold:

$$\tilde{w}_{it} = \alpha_i + \gamma_1 \text{Treat}_{it} + \gamma_2(\text{Treat}_{it} \times \text{Low}_i) + \gamma_3(\text{Treat}_{it} \times \text{Lot}_i) + \beta X_{it} + \tau_{it} + \varepsilon_{it} \tag{8}$$

where all variables are the same as in equation 7, except we add an interaction of treatment with a new indicator (Low_i), which signifies that a household is below any of the lot-size thresholds in the Large Utility (and, thus, more likely to face an exogenously higher marginal price). Because the lot-size discontinuity exists only for the Large Utility, we identify the PLE using Large Utility customers only. We vary the bandwidth of lot-size from $+/-1000$ sqft, $+/-750$ sqft, $+/-500$ sqft, and $+/-250$ sqft of lot-size thresholds, as well as an optimally chosen bandwidth following Calonico et al. (2014). We include an interaction with treatment and the continuous lot size (Lot_i) to control for differential treatment effects based on lot size.[7] The base effect of the lot-size discontinuity is absorbed by the household fixed effects.

We include several variants of this specification. We estimate a model with differential treatment-lot size interactions on either side of the lot-size threshold. We also estimate this model for the three primary lot-size discontinuities (at 7000 sqft, 11,000 sqft, and 17,500 sqft) individually.

The specification in equation 8 is the reduced form of a difference-in-discontinuity design. The discontinuity at the lot-size thresholds (captured by Low_i) changes the probability that a given household will face the higher marginal price in that billing period, although Low_i does not assign higher marginal prices to customers perfectly. Thus, we are operating with a fuzzy discontinuity. To estimate a local average treatment effect of exogenous marginal price assignment, we instrument for a whether a customer faces the higher marginal price (High Price_i) with being below the lot-size threshold (Low_i). Thus, we estimate:

$$\tilde{w}_{it} = \alpha_i + \gamma_1 \text{Treat}_{it} + \gamma_2(\text{Treat}_{it} \times \widehat{\text{High Price}}_i) + \gamma_3(\text{Treat}_{it} \times \text{Lot}_i) + \beta X_{it} + \tau_{it} + \varepsilon_{it} \tag{9}$$

[7]Lot size is correlated with water use, which is an important driver in treatment heterogeneity in peer comparisons for water conservation (Ferraro and Price, 2013; Brent et al., 2015).

where $\widehat{\text{High Price}}_i$ is predicted by Low_i in the first stage.

The *price-level effect*, the amount of the HWR treatment effect that is driven by exogenous differences in marginal price levels, is given by γ_2 in equations 8 and 9. These models allows us to test the hypothesis that $\gamma_2 = 0$. The regressions include both household and weather controls (X_{it}), household fixed effects (α_i), and billing period-by-utility (τ_{it}) fixed effects.

We also estimate the analogs of equations 8 and 9 for baseline effects on water consumption. That is, we re-estimate equations 8 and 9 without our randomized treatment indicator. This framework allows us to assess how price variation from the lot-size discontinuities affects demand directly.

As a robustness check, we include a third source of variation across utilities. The primary motivation for this robustness check is that our discontinuity depends on lot size, which in turn is correlated with water consumption. Because many studies find that high-use households are more responsive to social comparisons, we estimate a double-difference-in-discontinuities (DDD) model that nets out any primary effect of the lot-size threshold. We add the additional difference across utilities in the following framework:

$$\begin{aligned} \tilde{w}_{it} = \alpha_i &+ \gamma_1 \text{Treat}_{it} \\ &+ \gamma_2(\text{Treat}_{it} \times \text{Large}_i) + \gamma_3(\text{Treat}_{it} \times \text{Low}_i) + \gamma_4(\text{Treat}_{it} \times \text{Large}_i \times \text{Low}_i) \\ &+ \gamma_5(\text{Treat}_{it} \times \text{Lot}_i) + \gamma_6(\text{Treat}_{it} \times \text{Lot}_i \times \text{Large}_i) \\ &+ \beta X_{it} + \tau_{it} + \varepsilon_{it}, \end{aligned} \tag{10}$$

In this setup, γ_4 is our estimate of the PLE. This specification exploits the fact that while the Large Utility has a lot-size discontinuity in the rate structure the Small Utility does not.[8] This specification addresses the potential confounding of treatment heterogeneity associated with lot size around the threshold that is not accounted for by the linear lot-size interaction with treatment.[9]

3.6 Estimating the price-sensitivity effect

Next, we estimate the *price-sensitivity effect* (PSE), which we defined to be how treatment induces differential responses to price changes. We exploit price changes over time and across the utilities in order to estimate a demand equation and then interact the price variable with our randomized HWR treatment variables. Our demand regressions take the

[8] Because the Small Utility does not have a corresponding "High Price" at the discontinuity, we do not estimate equation 10 in a fuzzy RD framework.

[9] One might wonder why we did not consider the utility boundary as a spatial regression discontinuity similar to Ito (2014). In our setting, water utility boundaries also serve as political boundaries that induce numerous other changes in tax rates, city regulations, and so forth, thus we did not believe the abrupt change in prices at utility borders would provide a viable identification strategy.

following form:

$$\ln(w_{it}) = \alpha_i + \beta_1 \ln(\hat{p}_{it}) + \beta_2(\ln(\hat{p}_{it}) \times \text{Treat}_{it}) + \gamma_1 \text{Treat}_{it} + \tau_t + \varepsilon_{it} \tag{11}$$

where $\hat{p}_{it}$ is our endogenous price variable, τ_t are billing-period fixed effects, and all other variables are defined the same as in equations 8 and 9. We estimate equation 11 on our full sample including both utilities. Additionally, we run an additional specification limiting our sample to households within 10km of a common district boundary. Similar to the price-level effect models, we also estimate a baseline demand model without any treatment interactions for comparison.

The presence of increasing block rates makes price endogenous because the marginal price the consumers faces depends on the quantity consumed. Thus, we estimate equation 11 using two-stage-least squares (2SLS) where price and the associated interactions are endogenous variables. Following the framework in Olmstead (2009) and Wichman et al. (2016), we instrument for the actual price the consumer faces (either marginal or average) with the full set of marginal prices from the rate structure. All price instruments are transformed by natural logarithms. Therefore, our identification comes from variation in water rates set by the utility as opposed to changes in the households' consumption. There is an ongoing debate whether marginal or average price is the relevant price signal for decision-making when consumers face increasing block rates (Nataraj and Hanemann, 2011; Ito, 2014; Wichman, 2014), so we model price as both average and marginal price.[10]

The estimated coefficient $\hat{\beta}_2$ is a direct estimate of our price-sensitivity effect. That is, the degree to which randomized HWR treatments affect consumers' price sensitivity. Our framework allows for a direct test of the price-sensitivity hypothesis, i.e., that $\hat{\beta}_2 = 0$. Evidence of a nonzero PSE would suggest that behavioral treatments interact with structural parameters of demand. On the other hand, evidence of a PSE equal to zero would provide support for the notion that our randomized nudges act solely through behavioral channels.

4 Results and discussion

We first summarize the baseline results for the field experiment and the natural experiment. These results set the stage for understanding the interactions between prices and social pressure when presenting the results for the price-level effect and the price-sensitivity effets.

[10]We define average price as the volumetric proportion of the bill divided by quantity consumed that month.

4.1 Baseline treatment effects

In Table 2, we present treatment effects for the randomized home water reports (HWRs).[11] In the first two columns, our treatment effects for both the Large and Small utilities are -3.7% and -3.8% reductions in water consumption due to randomized HWRs. Both treatment effects are significant at the $p < 0.01$ level. In the third column, we pool both utilities, but allow for different treatment responses by including an interaction between our treatment variable and and an indicator for Large Utility. In the final column, we restrict the sample of the Large Utility to households within 10km of the Small Utility's border to ensure common support across both utilities. Overall, we find consistent evidence in line with previous research that HWRs reduce water consumption by $3-5\%$ (Ferraro and Price, 2013; Brent et al., 2015) and we observe nearly identical treatment effects across utilities.

Table 2: Baseline treatment effects

	(1) Large	(2) Small	(3) Both	(4) 10km
Treat	-0.037***	-0.038***	-0.038***	-0.038***
	(0.004)	(0.008)	(0.008)	(0.008)
Treat*Large			0.001	0.001
			(0.009)	(0.010)
Observations	602,415	453,624	1,056,039	606,876
Households	26,729	19,395	46,124	26,174
Household FEs	Yes	Yes	Yes	Yes
Sample	Full	Full	Full	10km
Period-by-utility FEs	Yes	Yes	Yes	Yes

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

4.2 Baseline quasi-experimental estimates

We next present the baseline effects of how inframarginal price changes driven by lot-size discontinuities affect consumer demand. In Figure 8, there is no obvious graphical evidence of a discontinuous change in consumption that matches the discontinuity in prices shown in Figure 6. Additionally, we present results from the reduced form of the local linear discontinuity model in panel (a) of Table 3. The coefficients on Low—the indicator that a household is below the lot-size threshold, thus face a higher expected marginal price—suggest a small negative effect at the largest bandwidth that shrinks and becomes statistically insignificant at smaller bandwidths. Because the discontinuity only increases

[11] Because we do not observe whether households actually read the HWRs, these treatment effects should be interpreted as intent-to-treat effects.

the probability that consumers face a higher marginal price we also estimate a local linear fuzzy regression discontinuity model and present results in panel (b) of Table 3. The results are similar to the reduced form, where there is a significant negative effect of facing higher marginal prices due to the lot size discontinuity at larger bandwidths, but no effect at smaller bandwidths. The coefficients from fuzzy discontinuity models are larger, however, because the estimated first-stage coefficient is approximately 0.2.

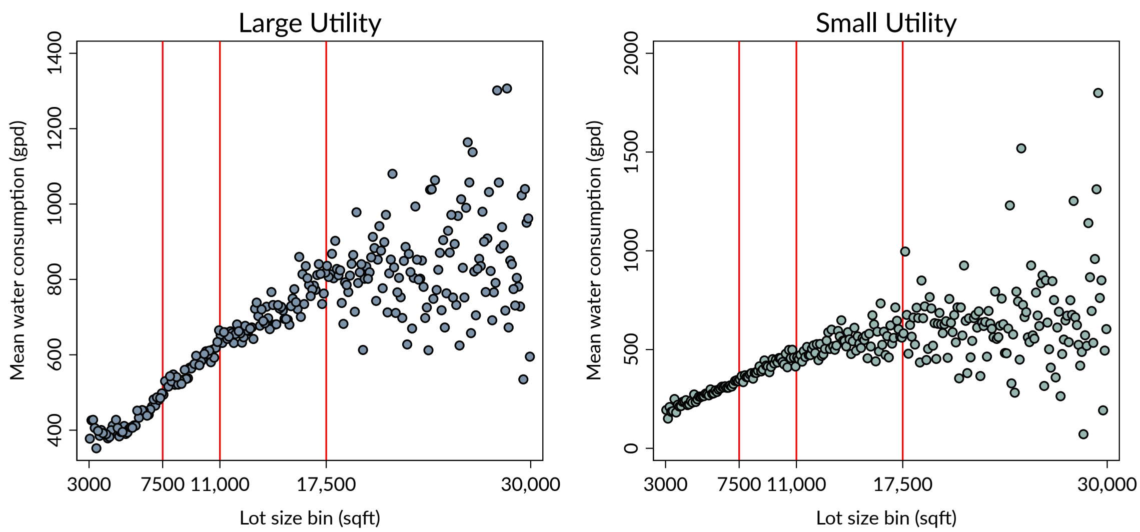


Figure 8: Effect of discontinuity on consumption at lot-size thresholds

These results provide evidence that there is no statistically significant first-order response of the lot-size threshold on consumption. Coefficients in smaller bandwiths, although insignificant, are positive. This result is surprising considering that both Low and High Price indicate that the customer faces a higher price. There are several explanations for why consumers do not appear to respond to this price differential as anticipated. First, consumers may respond to average as opposed to marginal prices, and although the effect of the discontinuity on average prices is still present it is not as large as the impact on marginal prices. A slightly different interpretation is that customers do not respond to *inframarginal* prices, and that they learn about their consumption at the end of the billing period. Alternatively, the difference in prices may not be sufficiently large to warrant a demand response. However, as noted above, the change in marginal prices at the lot-size threshold is roughly 20%.

Overall, our initial analysis produces estimates of average treatment effects for HWRs that are squarely within the results of previous studies. This consistency provides us with confidence that the experiments were conducted accurately. We must consider the lack of a primary demand effect in the natural experiment using lot-size discontinuities to identify

Table 3: Baseline effects of lot-size discontinuities on household water consumption

(a) Reduced form

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Low	-0.023**	-0.010	-0.010	0.012	0.017
	(0.012)	(0.013)	(0.015)	(0.021)	(0.026)
Sq.ft.	0.833***	1.303***	1.425***	3.071***	1.951
	(0.135)	(0.201)	(0.361)	(1.037)	(2.418)
Low*Sq.ft.	-0.027	-0.071	-0.113**	-0.065	-0.062
	(0.045)	(0.050)	(0.051)	(0.061)	(0.071)
Observations	160,209	125,142	90,021	54,849	66,530
Households	12,615	9,849	7,082	4,309	2,934
Household FEs	No	No	No	No	No
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Bandwidth (sqft)	1000	750	500	250	128

(b) Fuzzy discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
High Price	-0.134**	-0.057	-0.055	0.062	0.058
	(0.064)	(0.065)	(0.074)	(0.078)	(0.080)
Sq.ft.	0.880***	1.344***	1.479***	2.915***	2.599***
	(0.124)	(0.167)	(0.308)	(0.730)	(0.874)
Low*Sq.ft.	-0.018	-0.069***	-0.112***	-0.067*	-0.055
	(0.024)	(0.026)	(0.031)	(0.037)	(0.039)
Observations	160,209	125,142	90,021	54,849	50,770
Households	12,615	9,849	7,082	4,309	3,990
Household FEs	No	No	No	No	No
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Bandwidth (sqft)	1000	750	500	250	212
First-Stage Coef	0.18	0.19	0.19	0.20	0.21
First-Stage SE	0.007	0.008	0.009	0.01	0.01

Note: Sample includes households from the Large Utility only. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

variation in marginal prices when interpreting the results. Despite the lack of a primary effect we believe it is important to examine the interaction between prices and nudges for two reasons. First, the lack of a response may constitute an internality that the nudge may correct. Second, the information in the HWR on financial savings is dictated by the highest marginal price the consumer faced last billing period (see bottom of Fig. 3), so the financial benefits of conservation will appear larger to households below the lot-size threshold relative to similar households above lot size thresholds.

4.3 Estimates of the price-level effect

We now turn to Table 4, in which we present our primary results of the price-level effect. Recall, the PLE in our setting is the interaction between the treatment effect of the HWR and the exogenous assignment of a higher marginal price via the lot-size discontinuity. We present results only for the Large Utility because the Small Utility does not have discontinuous changes in price due to lot-size thresholds (see Fig. 5). We focus first on the results from the reduced form equation in panel (a). The coefficient on the interaction Treat*Low is our estimate of the PLE. We vary the bandwidth (distance from the lot-size discontinuity) in each of the columns. For all bandwidths, we estimate a small positive effect, which means that the higher prices in the Low group decreases the conservation generated from the HWRs. However, the effect is not always statistically significant. In our narrowest bandwidth, the interacted coefficient is 0.013 with a standard error (robust to within-household correlation) of 0.013.

Next, in panel (b) of Table 4, we estimate the fuzzy RD version of the PLE. In this specification we interact an indicator for whether a household faces the higher price with our randomized treatment indicator, where the interaction variable is instrumented by the Lot*Treatment interaction. In this model there is a positive and significant effect of being treated while facing higher marginal prices in all but the specifications with the tightest two bandwidths. These results may be due to financial incentives crowding out intrinsic incentives (Pellerano et al., 2017). The first-stage coefficient in these specifications ranges from 0.10 to 0.14.

In both panels of Table 4, our primary treatment effect increases from its baseline level. This result suggests that by focusing only on households within narrow bandwidths around the lot-size threshold may change the composition of households from our primary sample. We suspect that the larger treatment effects are driven by larger water users who might also have larger lot sizes. Thus, because we pool all lot-size discontinuities in the rate structure together, our PLE estimates in Table 4 may mask important heterogeneity.

In Table 5, we estimate PLEs for each discontinuity separately, again for the Large Utility only. For the 7500 and 11,000 sqft discontinuities in panels (a) and (b), we again find precisely estimated null effects, with standard errors increasing slightly with smaller bandwidths. For these models, the baseline treatment effect is also much more similar to the baseline effects in Table 2. These results are based on larger, more representative subsamples of our data. We only present coefficients from reduced form models. Because our reduced-form coefficients are 0.007 and -0.001 in the optimal bandwidth models, the first-stage coefficients would need to be extraordinarily small for the LATEs to be economically meaningful in the fuzzy RD framework.

For the discontinuity at 17,500 sqft, however, we observe both a substantially larger base

Table 4: Price-level effect

(a) Reduced form

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.061***	-0.064***	-0.059***	-0.062***	-0.065***
	(0.006)	(0.007)	(0.008)	(0.010)	(0.011)
Treat*Low	0.012*	0.013	0.017*	0.011	0.013
	(0.007)	(0.008)	(0.009)	(0.012)	(0.013)
Observations	284,298	222,168	160,100	97,496	90,271
Households	12,615	9,849	7,082	4,309	3,990
Sample	Large only	Large only	Large only	Large only	Large only
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	Yes	Yes	Yes	Yes	Yes
Bandwidth	1000	750	500	250	212

(b) Fuzzy discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.110***	-0.112***	-0.113***	-0.096***	-0.101***
	(0.019)	(0.020)	(0.021)	(0.025)	(0.026)
High Price*Treat	0.120***	0.116***	0.135***	0.085	0.091
	(0.041)	(0.044)	(0.044)	(0.054)	(0.057)
Observations	284,298	222,168	160,100	97,496	73,472
Households	12,615	9,849	7,082	4,309	3,241
Household FEs	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Bandwidth	1000	750	500	250	157
First-Stage Coef	0.10	0.11	0.13	0.14	0.14
First-Stage SE	0.006	0.006	0.007	0.010	0.01

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Columns designate the bandwidths around the lot size thresholds in sqft Robust standard errors are clustered at the household level. *p<0.1; oi**p<0.05; ***p<0.01

treatment effect ($11 - 14\%$ reductions in average daily consumption) and a large, positive PLE. This PLE estimate, however, is estimated with large confidence intervals, due in part to the smaller number of households near this discontinuity (only 374 households are included in the optimal bandwidth subsample). Households with larger lots tend to use more water for irrigation, which is why we suspect we see larger base treatment effects near the 17,500 sqft discontinuity. Because this estimate is positive, we interpret this as suggestive evidence that higher prices crowd out conservation among high use households on larger lots. Overall, we place more confidence in our precisely estimated null effects based on

Table 5: Price-level effect at individual discontinuities

(a) 7500 sqft discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.040***	-0.039***	-0.043***	-0.034***	-0.036***
	(0.007)	(0.007)	(0.009)	(0.011)	(0.011)
Treat*Low	0.007	0.007	0.010	0.002	0.007
	(0.007)	(0.008)	(0.010)	(0.013)	(0.013)
Observations	182,130	141,984	98,483	58,625	56,411
Households	8,152	6,350	4,393	2,611	2,512
Sample	Full	Full	Full	Full	Full
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	No	No	No	No	No
Bandwidth	1000	750	500	250	212

(b) 11,000 sqft discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.038***	-0.041***	-0.030**	-0.047***	-0.050***
	(0.011)	(0.012)	(0.014)	(0.017)	(0.017)
Treat*Low	0.009	0.005	0.010	0.011	-0.001
	(0.014)	(0.017)	(0.019)	(0.024)	(0.027)
Observations	74,331	57,842	45,132	28,856	24,999
Households	3,249	2,526	1,969	1,259	1,089
Sample	Full	Full	Full	Full	Full
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	No	No	No	No	No
Bandwidth	1000	750	500	250	212

(c) 17,500 sqft discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.119***	-0.143***	-0.112***	-0.121***	-0.121***
	(0.025)	(0.027)	(0.032)	(0.041)	(0.044)
Treat*Low	0.051*	0.065*	0.066*	0.074	0.086*
	(0.030)	(0.034)	(0.037)	(0.047)	(0.051)
Observations	26,948	21,599	15,903	9,623	8,514
Households	1,176	941	695	422	374
Sample	Full	Full	Full	Full	Full
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	No	No	No	No	No
Bandwidth	1000	750	500	250	212

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption within 1,000 sqft of the lot-size discontinuity. All specifications control for evapotranspiration and precipitation. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

larger and more representative samples nearby the 7000 and 11,000 sqft thresholds.[12]

We suspect that large effects at the 17,500 sqft threshold are driven by relatively few unrepsentative households, and may be biasing the our primary PLE estimates. We replicate our primary models presented in Table 4 excluding households nearby the 17,500 sqft threshold. These results, presented in Table A.1 for the reduced form and Table A.2 for the fuzzy RD, support our intuition. Removing the 17,500 threshold results in PLE estimates closer to zero with smaller standard errors. In our optimally chosen bandwidth models, our reduced-form PLE estimate is 0.003 (0.011) and our fuzzy RD PLE estimate is 0.025 (0.047).

Because high water-use households are more responsive to social comparisons, it is important to evaluate how the interactions may differ across the distribution of pre-treatment consumption. Figure 9 shows that treatment heterogeneity is similar across quartiles of lot size and baseline water use. Households with large lots and high-use households possess the largest treatment effects. The average consumption levels using the 1000 sqft bandwidth for the 7500, 11,000, and 17,500 sqft thresholds roughly correspond to the 50th, 75th, and 90th percentiles of the full consumption distribution. Therefore, the results presented in Table 5 also provide insight on the treatment heterogeneity, and we see no significant interactions between prices and nudges across the distribution of consumption.[13]

4.3.1 Robustness of price-level effects

We include several additional analyses to support our results. First, we add a third difference to our difference-in-discontinuity design because it is possible that responsiveness to HWRs is greater for households with higher consumption levels, which is correlated positively with lot size (our running variable in the regression discontinuity). To implement the third difference, we estimate Equation 11 on a sample including both utilities. We present these results in Table 6. In these specifications, the reduced-form PLE is identified by the coefficient on Treat*Low*Large, or, the marginal change in the treatment effect due to facing an exogenously higher marginal price by being just below the lot-size threshold relative to similar households in the small utility who face no price discontinuity. In these specifications, we find no statistical evidence of a PLE, which is a precisely estimated zero

[12] We explore this result further in Figure A.4 in the appendix, in which we plot the base treatment effect coefficients interacted with 250 sqft lot-size bins near the lot-size thresholds. We do so for both utilities. Evidence of a nonzero PLE would be revealed by a discontinuous jump in treatment effect estimates at the lot-size thresholds. Specifically, in the presence of a PLE that increases conservation from HWRs we expect the treatment effect immediately to the left of the threshold to be larger in magnitude than the treatment effect immediately to the right of the threshold. For the 7,500 sqft discontinuity, we observe the treatment effect move smoothly through the discontinuity for both utilities. All estimates are statistically similar, shown by overlapping 95% confidence intervals. The results for the 11,000 and 17,500 sqft thresholds are noisier, but confidence intervals also overlap for all estimates within a utility.

[13] For reference, using the 1,000 sqft bandwidth the average consumption around the 7,500 sqft threshold is 492 GPD, 662 GPD around the 11,000 sqft threshold, and 891 GPD around the 17,500 sqft threshold.

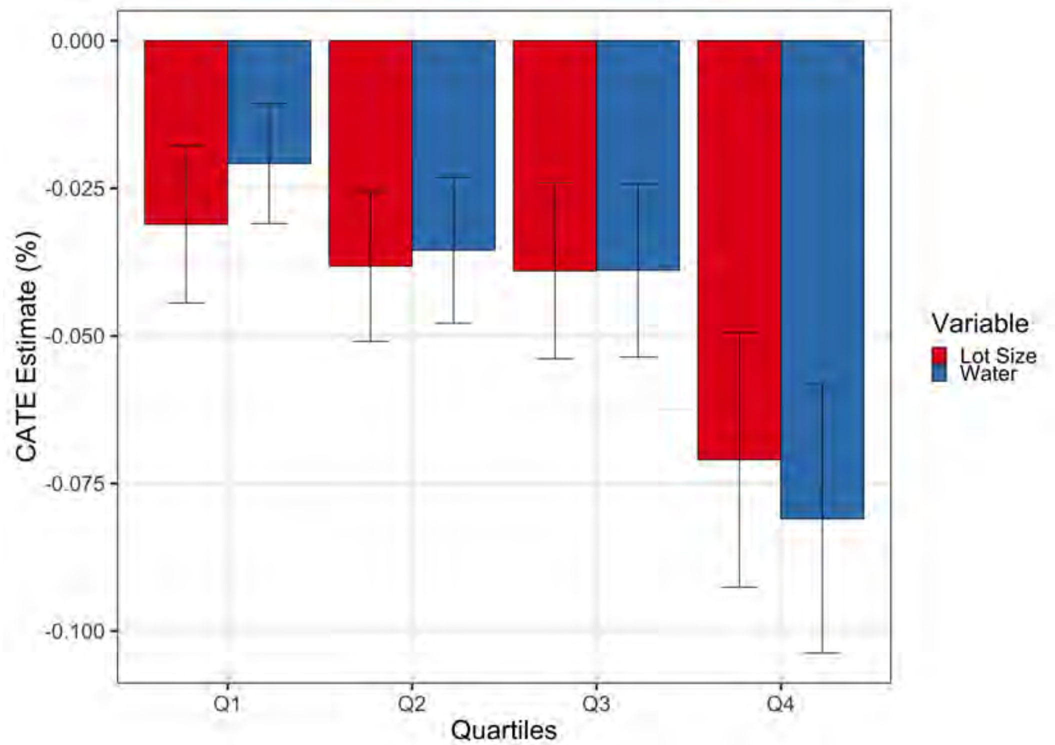


Figure 9: Treatment heterogeneity by baseline consumption and lot size quartiles

for bandwidths of 1000 sqft and 750 sqft. At the smaller bandwidths the effect becomes negative but is still small and insignificant. When we replicate these models excluding households nearby the 17,500 threshold, we find PLEs that are closer to zero with smaller standard errors. Results are presented in Table A.3 in the appendix.

We include several other robustness checks. First, we include interaction terms with lot-size on both sides of the threshold, as is typical in local linear RD designs. The results, in Table A.4, are virtually unchanged: we find small positive but insignificant effects for the PLE. Additionally, we perform a falsification test in the Large Utility at false discontinuities of 9000 sqft and 13,000 sqft We choose these thresholds because they are near our true thresholds without overlapping at the largest bandwidths (1000 sqft). These falsification tests examine whether our lot size thresholds would partially pick up the smaller treatment effects (in absolute value), associated with smaller lots that use less water. If the true PLE is negative (households who face higher prices are more responsive to HWRs) the small lot size effect will bias our estimates of the PLE towards zero. These results are presented in Table A.5. Here again, we find statistical zeros, and the point estimates switch between positive and negative values.[14]

[14]There is a larger and noisy negative PLE at the false 13,000 sqft discontinuity, but that is likely due to noisy estimates in a small sample ($\approx$ 400 households).

Table 6: Price-level effect: Difference-in-difference-in-discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.068***	-0.067***	-0.045*	-0.072**	-0.079**
	(0.016)	(0.019)	(0.023)	(0.033)	(0.036)
Treat*Large	0.007	0.002	-0.015	0.010	0.013
	(0.017)	(0.020)	(0.025)	(0.035)	(0.038)
Treat*Low	0.013	0.009	0.007	0.026	0.042**
	(0.010)	(0.012)	(0.015)	(0.019)	(0.020)
Treat*Low*Large	-0.000	0.005	0.011	-0.014	-0.029
	(0.013)	(0.015)	(0.017)	(0.023)	(0.024)
Observations	414,260	318,956	226,140	134,954	123,053
Households	18,176	13,983	9,904	5,907	5,389
Sample	Both Utilities	Both Utilities	Both Utilities	Both Utilities	Both Utilities
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	Yes	Yes	Yes	Yes	No
Bandwidth	1000	750	500	250	212

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Columns designate the bandwidths around the lot size thresholds in sqft Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

Lastly, we consider the possibility for dynamic adjustment in the PLE. Ito et al. (2018) shows different levels of persistence for financial incentives compared to moral suasion in energy and Brent et al. (2017) show that financial nudges are more persistent than nudges using moral suasion for water conservation. Therefore, the PLE may not be present until the conservation effect of social pressure begins to wane. We estimate price-level effects interacted with an indicator for the year after treatment begins to show differences in the persistence of the PLE. We present the regression results in Table A.6. There is no evidence of a PLE in the year of or year after treatment. We also explore seasonal effects by interacting the PLE with a dummy variable for summer in Table A.7. These results suggest that negative PLEs are observed in the summer months (around $2-3\%$ reductions), when conservation signals and prices might be more salient, and we observe positive PLEs of similar magnitudes in non-summer months. In our tightest, optimally chosen bandwith, both of these effects are statistically similar to zero.

To recap, we find little evidence that suggests that exogenously assigned differences in marginal prices increase the effectiveness of HWRs. In fact, we uncover some evidence that higher prices may slightly decrease conservation effects from behavioral interventions among high-use households. This result is somewhat surprising because the HWRs make the private economic benefits of water conservation more salient (e.g., bottom panel in Figure 3). HWRs provide cost-savings information that consumers might expect from changing behavior or technology. Consumers just above/below the price discontinuities we use for identification would thus face nontrivial differences in expected cost-savings despite being otherwise similar types of households, but we observe no statistical difference in their response to HWRs. Our analysis thus far suggests that the primary mechanism for the HWR operates through channels of increasing (the salience of) the moral costs of water consumption.

4.4 Price-sensitivity effects

We now turn to our results of the price-sensitivity effects. We first present our initial demand specifications in Table 7. In the first two columns, we present naïve models using endogenous marginal and average price variables. As expected with increasing block-rate structures, we observe positive price elasticities. Our IV approach, in columns (3) and (4), performs comparatively better, providing sensible demand elasticities (-0.25 for MP and -0.17 for AP) which are within the range of previous estimates for both marginal and average prices arising from reduced-form and structural models of water demand (Dalhuisen et al., 2003; Olmstead, 2009; Nataraj and Hanemann, 2011; Wichman, 2014). In the present analysis, we do not take a stand on whether average or marginal price responsiveness is the correct specification, rather we model them side-by-side. In columns (5) and (6), we restrict the sample to households within 10km of the shared border, and our elasticity estimates are similar to the full sample.

Table 7: Baseline demand models

	(1) MP	(2) AP	(3) MP	(4) AP	(5) MP	(6) AP
ln(MP)	0.549***		-0.246***		-0.278***	
	(0.007)		(0.012)		(0.017)	
ln(AP)		0.564***		-0.169***		-0.187***
		(0.008)		(0.011)		(0.016)
Observations	939,775	929,842	928,032	921,537	572,528	570,981
Households	43,133	43,132	43,124	43,120	25,990	25,987
Household FEs	Y	Y	Y	Y	Y	Y
Period FEs	Y	Y	Y	Y	Y	Y
IV	–	–	Y	Y	Y	Y
Sample	Full	Full	Full	Full	10km	10km

Note: Dependent variable is the natural log of average daily water consumption. All specifications control for evapotranspiration and precipitation. Prices are instrumented with full set of marginal prices from the utility rate schedule and associated interactions with exogenous variables. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

We present results for the PSE in Table 8. Recall, the PSE is the degree to which HWRs increase consumers' price sensitivity, e.g., by making the costs of consumption more salient. Identification of this effect is straightforward: we estimate price elasticities of water demand equation as in Equation 11 and interact our price variables with the randomized HWR treatment. The resulting coefficient on that interaction is the PSE.

In columns (1) and (2) of Table 8, we report PSE estimates for our pooled sample. We find statistically significant evidence that HWRs increase price sensitivity under the

Table 8: Price-sensitivity effect

	(1) MP	(2) AP	(3) MP	(4) AP
Treat	-0.035***	-0.042***	-0.041***	-0.044***
	(0.003)	(0.003)	(0.005)	(0.005)
ln(MP)	-0.221***		-0.236***	
	(0.011)		(0.015)	
ln(MP)*Treat	-0.030**		-0.012	
	(0.014)		(0.020)	
ln(AP)		-0.191***		-0.189***
		(0.010)		(0.014)
ln(AP)*Treat		-0.002		0.002
		(0.012)		(0.016)
Observations	928,032	921,537	572,528	570,981
Households	43,124	43,120	25,990	25,987
Household FEs	Y	Y	Y	Y
Period FEs	Y	Y	Y	Y
IV	Y	Y	Y	Y
Sample	Full	Full	10km	10km

Note: Dependent variable is the natural log of average daily water consumption. All specifications control for evapotranspiration and precipitation. Prices are instrumented with full set of marginal prices from the utility rate schedule and associated interactions with exogenous variables. Interactions with indicators for treatment periods are included but coefficients are not reported for clarity. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

marginal price, but not the average price, demand specification. The PSE coefficient in column (1) is -0.03, which increases price sensitivity by approximately 13% for the MP specification. In column (2), this coefficient is -0.002, which suggests that there is no PSE for specifications that include average prices. We might be concerned that in addition to price variation, there is significant variation in unobservables across utilities that could impact price elasticity. We control for these cross-border differences by restricting the sample to households that are less than 10km away from the Small Utility's boundary. Columns (3) and (4) present the PSE estimates in the restricted sample, and we find no evidence of a statistically or economically significant PSE in either the MP or AP specification. Overall, these results are slightly less decisive than our PLE results, but still do not find consistent evidence that social comparisons have meaningful interactions with prevailing economic incentives.

5 Concluding remarks

In this paper, we explore the interaction of prices and behavioral nudges. We compare water consumer responses to a randomized behavioral messaging campaign for households who face differential exogenously assigned marginal prices. The results from our analysis suggest that households who have a greater economic incentive to conserve respond to the social comparison similarly to households with less economic incentive to conserve. This result suggests that the private economic benefits of conservation may be inconsequential for behavioral treatments to be effective. Additionally, we estimate the degree to which the behavioral treatment affects price sensitivity, finding limited evidence that there is an economically meaningful relationship between prices and nudges. Overall, our results suggest that although there is theoretical justification for why behavioral nudges and economic incentives should interact, we find little empirical support that this interaction is meaningful.

Behavioral nudges do not exist in a vacuum. Although the randomized deployment of many behavioral nudges provides strong internal validity for the estimation of causal effects, as Allcott (2015) shows, the treatment effects from any given location may be a function of the underlying characteristics of the specific population. In order to ensure that the estimates from any one location are externally valid, it is critical to identify the sources of heterogeneity and adjust the magnitude based on the characteristics of the target population. This is challenging when the entire experimental sample faces the same set of existing policies. We focus on how variation in prevailing water prices affects consumer responsiveness to behavioral nudges intended to encourage water conservation that include social comparisons. We do not find any evidence that the response to this prevalent behavioral nudge has any meaningful interactions with underlying water rates. This is an important result for many settings in which resource prices are low or zero, or when scarcity pricing may not be politically feasible, we show that behavioral nudges may still be effective to mechanisms to encourage conservation.

In addition to external validity, our results have implications for the behavioral mechanisms through which nudges operate. Finding no evidence of heterogeneity due to different private benefits of conservation leads us to conclude that consumers are primarily responding to social comparisons due to increased salience of the moral cost of water consumption. This finding has important implications for the welfare effect of nudges as shown by Allcott and Kessler (2019). If nudges operate as a moral tax, which is consistent with our findings in this paper, they will only be welfare enhancing if the social cost of energy/water exceeds the current private costs. This suggests a re-thinking of behavioral policies that specifically target welfare improvements as opposed to simply changing behavior. Given substantial evidence of behavioral biases in energy and water markets (Allcott and Wozny, 2014; Sexton, 2015; Wichman, 2017; Brent and Ward, 2018, 2019) it is worthwhile to find ways to promote pro-social behavior that also improves private decisions.

References

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

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

_ **and Dmitry Taubinsky**, "Evaluating Behaviorally-Motivated Policy: Experimental Evidence from the Lightbulb Market," *American Economic Review*, 2015, *105* (8), 2501–2538.

_ **and Judd B. Kessler**, "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons," *American Economic Journal: Applied Economics*, 2019, *11* (1).

_ **and Nathan Wozny**, "Gasoline prices, fuel economy, and the energy paradox," *Review of Economics and Statistics*, 2014, *96* (5), 779–795.

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

Brandon, Alec, John A. List, Robert D. Metcalfe, Michael K. Price, and Florian Rundhammer, "Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity," *Proceedings of the National Academy of Sciences*, 2018, p. 201802874.

Brent, Daniel A. and Michael B. Ward, "Energy Efficiency and Financial Literacy," *Journal of Environmental Economics and Management*, 2018, *90* (181–216).

_ **and** _ , "Price Perceptions in Water Demand," *Journal of Environmental Economics and Management*, 2019, *98.*

_ **, Corey Lott, Michael Taylor, Joseph Cook, Kim Rollins, and Shawn Stoddard**, "Are Normative Appeals Moral Taxes? Evidence from a Field Experiment on Water Conservation," Technical Report, Department of Economics, Louisiana State University 2017.

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

Byrne, David P., Andrea La Nauze, and Leslie A. Martin, "Tell Me Something I Don't Already Know: Informedness and the Impact of Information Programs," *Review of Economics and Statistics*, 2018, *100* (3), 510–527.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

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

Dalhuisen, Jasper M., Raymond J.G.M. Florax, Henri L.F. De Groot, and Peter Nijkamp, "Price and income elasticities of residential water demand: a meta-analysis," *Land Economics*, 2003, *79* (2), 292–308.

Ferraro, Paul J. and Michael K. Price, "Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment," *The Review of Economics and Statistics*, 2013, *95* (1), 247–264.

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

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev, "The behavioralist as tax collector: Using natural field experiments to enhance tax compliance," *Journal of Public Economics*, 2017, *148*, 14–31.

Hanks, Andrew S., David R. Just, Laura E. Smith, and Brian Wansink, "Healthy convenience: nudging students toward healthier choices in the lunchroom," *Journal of Public Health*, 2012, *34* (3), 370–376.

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

_ , Takanori Ida, and Makoto Tanaka, "Moral suasion and economic incentives: Field experimental evidence from energy demand," *American Economic Journal: Economic Policy*, 2018, *10* (1), 240–67.

Jacobsen, Grant D., "Do energy prices influence investment in energy efficiency? Evidence from energy star appliances," *Journal of Environmental Economics and Management*, 2015, *74*, 94–106.

Levitt, Steven D. and John A. List, "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?," *Journal of Economic Perspectives*, June 2007, *21* (2), 153–174.

_ , _ , Susanne Neckermann, and Sally Sadoff, "The behavioralist goes to school: Leveraging behavioral economics to improve educational performance," *American Economic Journal: Economic Policy*, 2016, *8* (4), 183–219.

List, John A. and Anya Savikhin Samek, "The behavioralist as nutritionist: leveraging behavioral economics to improve child food choice and consumption," *Journal of Health Economics*, 2015, *39*, 135–146.

— , **Robert D. Metcalfe, Michael K. Price, and Florian Rundhammer**, "Harnessing Policy Complementarities to Conserve Energy: Evidence from a Natural Field Experiment," Technical Report, National Bureau of Economic Research 2017.

Nataraj, Shanthi and W. Michael Hanemann, "Does marginal price matter? A regression discontinuity approach to estimating water demand," *Journal of Environmental Economics and Management*, March 2011, *61* (2), 198–212.

Olmstead, Sheila M., "Reduced-form versus structural models of water demand under nonlinear prices," *Journal of Business & Economic Statistics*, 2009, *27* (1), 84–94.

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

Royer, Heather, Mark Stehr, and Justin Sydnor, "Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company," *American Economic Journal: Applied Economics*, 2015, *7* (3), 51–84.

Sallee, James M., "Rational Inattention and Energy Efficiency," *Journal of Law and Economics*, 2014, *57* (3), 781–820.

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

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

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

West, Jeremy, Robert W. Fairlie, Bryan Pratt, and Liam Rose, "Do Regulatory Policies Crowd Out Social Pressure for Resource Conservation?," 2019.

Wichman, Casey J., "Perceived price in residential water demand: Evidence from a natural experiment," *Journal of Economic Behavior & Organization*, 2014, *107*, 308–323.

— , "Information provision and consumer behavior: A natural experiment in billing frequency," *Journal of Public Economics*, 2017, *152*, 13–33.

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

Online Appendix: Additional Results

Figure A.1: Home Water Report

Note: This is an example of a generic "Yellow" Home Water Report (HWR). Households receiving this report used less water than their peer group average.

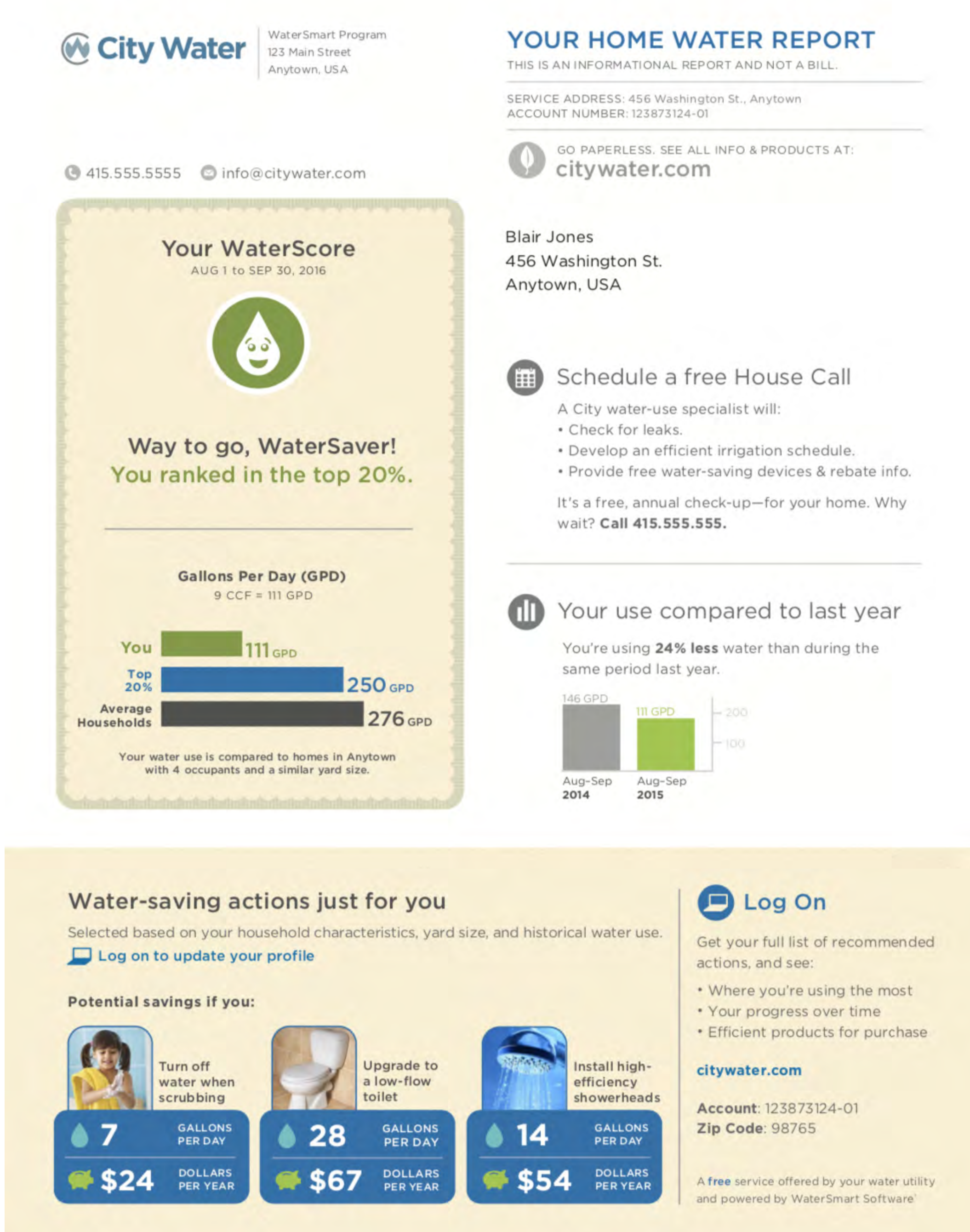


Figure A.2: Home Water Report

Note: This is an example of a generic "Green" Home Water Report (HWR). Households receiving this report were in the bottom 20% of their peer group.

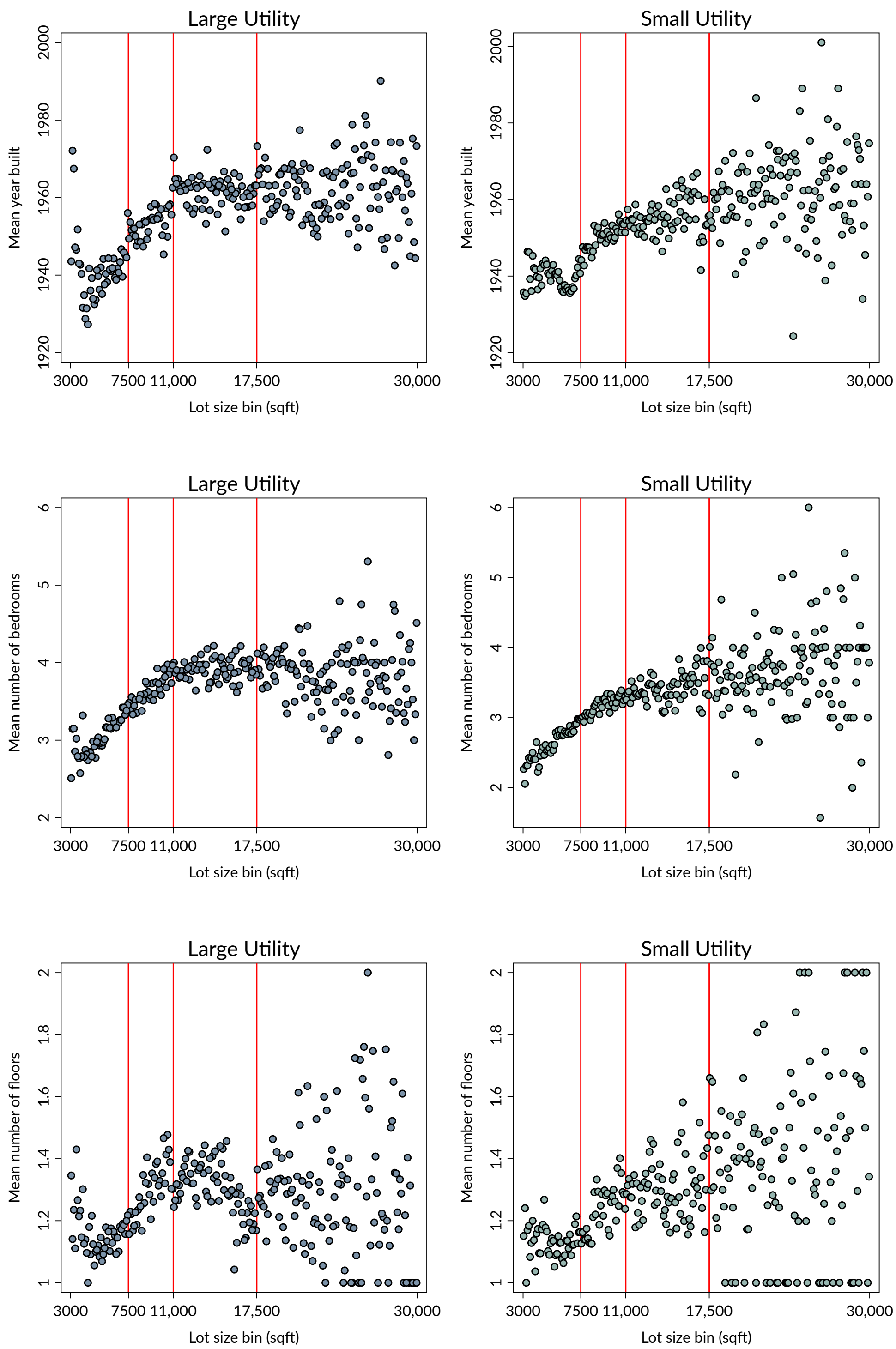


Figure A.3: Additional covariate distributions across lot-size thresholds for both utilities

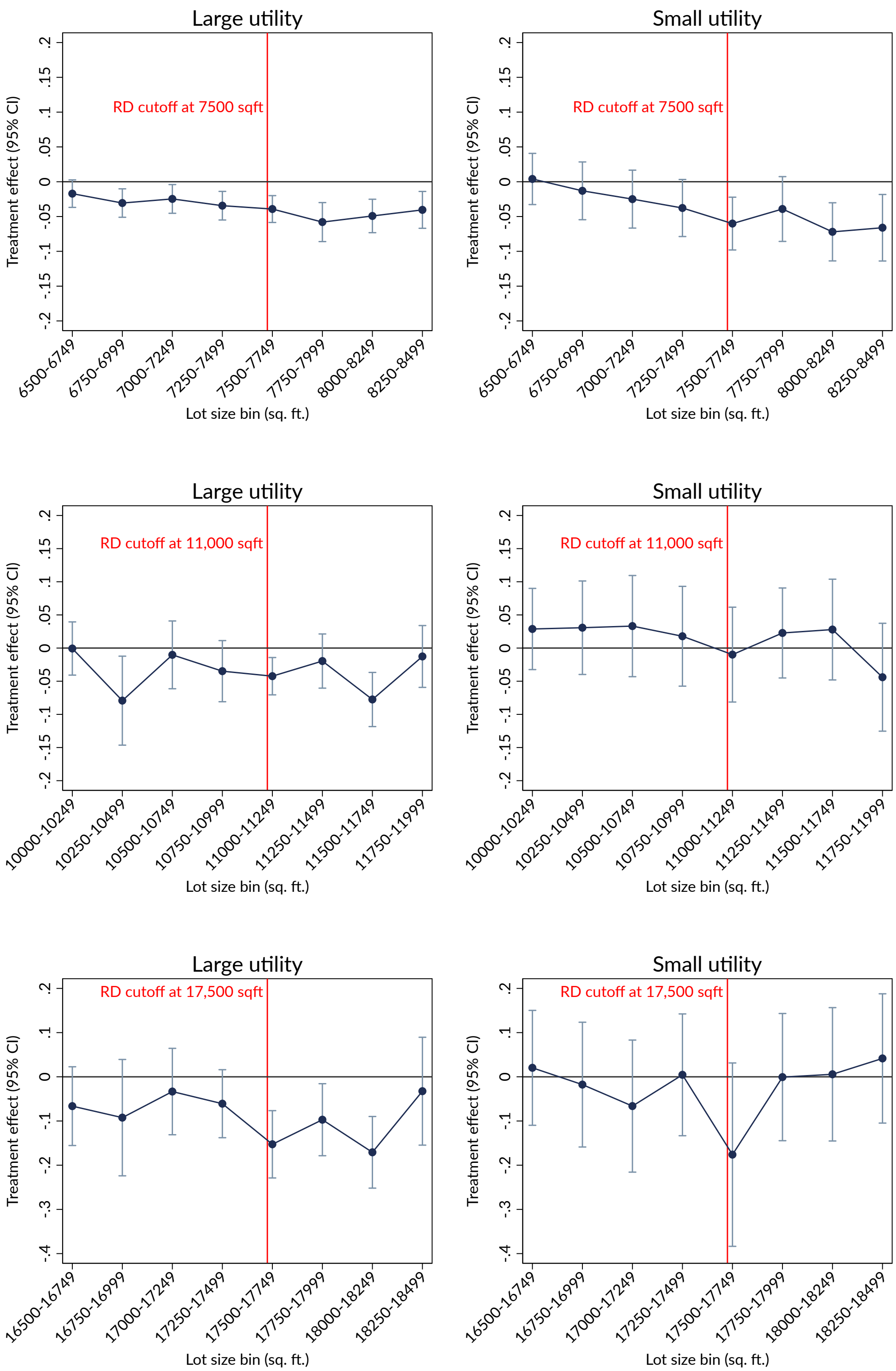


Figure A.4: Treatment effects by lot-size bin

Table A.1: Reduced-form estimates of the price-level effect without 17,500 lot-size threshold

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.062***	-0.065***	-0.058***	-0.063***	-0.055***
	(0.007)	(0.008)	(0.009)	(0.011)	(0.010)
Treat*Low	0.006	0.005	0.009	0.003	0.003
	(0.008)	(0.008)	(0.010)	(0.012)	(0.011)
Observations	256,461	199,826	143,615	87,481	105,927
Households	11,401	8,876	6,362	3,870	4,684
Sample	Large only	Large only	Large only	Large only	Large only
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	Yes	Yes	Yes	Yes	Yes
Bandwidth (sqft)	1000	750	500	250	341

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

Table A.2: Fuzzy RD estimates of the price-level effect without 17,500 lot-size threhold

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.085***	-0.083***	-0.086***	-0.070***	-0.065***
	(0.018)	(0.019)	(0.019)	(0.022)	(0.021)
High Price*Treat	0.058	0.046	0.072*	0.019	0.025
	(0.042)	(0.045)	(0.042)	(0.051)	(0.047)
Observations	256,461	199,826	143,615	87,481	105,927
Households	11,401	8,876	6,362	3,870	4,684
Household FEs	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Bandwidth	1000	750	500	250	341
First-Stage Coef	0.10	0.11	0.13	0.14	0.14
First-Stage SE	0.006	0.007	0.008	0.01	0.009

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

Table A.3: Reduced-form diff-in-diff-in-discontinuity estimates of the price-level effect without 17,500 lot-size threshold

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.106***	-0.097***	-0.078***	-0.108***	-0.088***
	(0.016)	(0.020)	(0.025)	(0.037)	(0.031)
Treat*Large	0.043**	0.031	0.019	0.044	0.032
	(0.018)	(0.021)	(0.027)	(0.038)	(0.033)
Treat*Low	0.003	0.004	0.004	0.016	0.015
	(0.010)	(0.012)	(0.014)	(0.019)	(0.016)
Treat*Low*Large	0.003	0.001	0.006	-0.013	-0.011
	(0.013)	(0.015)	(0.017)	(0.022)	(0.020)
Observations	379,115	290,848	205,660	123,161	151,534
Households	16,648	12,762	9,011	5,392	6,629
Sample	Both Utilities	Both Utilities	Both Utilities	Both Utilities	Both Utilities
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	Yes	Yes	Yes	Yes	No
Bandwidth	1000	750	500	250	341

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

Table A.4: Price-level effect: Large Utility with lot size interactions on both sides of the threshold

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.062***	-0.066***	-0.060***	-0.063***	-0.066***
	(0.006)	(0.007)	(0.008)	(0.011)	(0.011)
Treat*Low	0.019**	0.020**	0.024**	0.016	0.018
	(0.009)	(0.010)	(0.011)	(0.014)	(0.015)
Observations	284,298	222,168	160,100	97,496	90,271
Households	12,615	9,849	7,082	4,309	3,990
Sample	Large only	Large only	Large only	Large only	Large only
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	Yes	Yes	Yes	Yes	Yes
Bandwidth	1000	750	500	250	212

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Columns designate the bandwidths around the lot size thresholds in sqft Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

Table A.5: Price-level effect falsification test: Large utility only, with lot size interactions, at individual discontinuities

(a) False 9000 sqft discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.030**	-0.040***	-0.034**	-0.038**	-0.041**
	(0.012)	(0.013)	(0.015)	(0.019)	(0.020)
Treat*Low	-0.002	0.009	0.004	-0.002	0.006
	(0.012)	(0.014)	(0.017)	(0.021)	(0.023)
Observations	86,467	61,694	40,375	21,396	18,280
Households	3,845	2,742	1,793	947	806
Sample	Large only	Large only	Large only	Large only	Large only
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	No	No	No	No	No
Bandwidth	1000	750	500	250	212

(b) False 13,000 sqft discontinuity

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.096***	-0.086***	-0.095***	-0.041	-0.041
	(0.020)	(0.023)	(0.029)	(0.039)	(0.045)
Treat*Low	0.021	0.004	-0.010	-0.063	-0.069
	(0.024)	(0.027)	(0.036)	(0.051)	(0.058)
Observations	36,675	26,487	17,100	9,288	7,473
Households	1,605	1,160	748	408	328
Sample	Large only	Large only	Large only	Large only	Large only
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	No	No	No	No	No
Bandwidth	1000	750	500	250	212

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption within 1,000 sqft of the lot-size discontinuity. All specifications control for evapotranspiration and precipitation. Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

Table A.6: Persistence of the price-level effect

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.060***	-0.065***	-0.062***	-0.063***	-0.071***
	(0.006)	(0.007)	(0.008)	(0.011)	(0.012)
Treat*Year 2	-0.001	0.001	0.008	0.002	0.015
	(0.006)	(0.007)	(0.008)	(0.010)	(0.011)
Treat*Low	0.017**	0.015*	0.021**	0.008	0.011
	(0.008)	(0.008)	(0.010)	(0.013)	(0.014)
Treat*Low*Year 2	-0.012	-0.007	-0.009	0.009	0.004
	(0.007)	(0.008)	(0.010)	(0.013)	(0.014)
Observations	284,298	222,168	160,100	97,496	73,472
Households	12,615	9,849	7,082	4,309	3,241
Sample	Large only	Large only	Large only	Large only	Large only
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	Yes	Yes	Yes	Yes	Yes
Bandwidth	1000	750	500	250	157

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Columns designate the bandwidths around the lot size thresholds in sqft Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01

Table A.7: Seasonality of the price-level effect

	(1) 1000sqft	(2) 750sqft	(3) 500sqft	(4) 250sqft	(5) Optimal
Treat	-0.061***	-0.064***	-0.056***	-0.056***	-0.058***
	(0.006)	(0.007)	(0.008)	(0.011)	(0.013)
Treat*Summer	0.001	-0.001	-0.008	-0.015*	-0.017*
	(0.006)	(0.006)	(0.007)	(0.009)	(0.009)
Treat*Low	0.030***	0.030***	0.032***	0.022*	0.021
	(0.008)	(0.008)	(0.010)	(0.013)	(0.014)
Treat*Low*Summer	-0.042***	-0.043***	-0.037***	-0.026**	-0.020
	(0.007)	(0.007)	(0.009)	(0.012)	(0.012)
Observations	284,298	222,168	160,100	97,496	73,472
Households	12,615	9,849	7,082	4,309	3,241
Sample	Large only	Large only	Large only	Large only	Large only
Household FE	Yes	Yes	Yes	Yes	Yes
Period-by-utility FEs	Yes	Yes	Yes	Yes	Yes
Lot Size Int.	Yes	Yes	Yes	Yes	Yes
Bandwidth	1000	750	500	250	157

Notes: Dependent variable is average daily water consumption normalized by utility-specific control group consumption. All specifications control for evapotranspiration and precipitation. Columns designate the bandwidths around the lot size thresholds in sqft Robust standard errors are clustered at the household level. *p<0.1; **p<0.05; ***p<0.01