Price Perceptions in Water Demand

Daniel A. Brent*[1] and Michael B. Ward[2]

[1]Department of Agricultural Economics, Sociology, and Education, Pennsylvania State University
[2]Department of Economics, Monash University

July 2019

Abstract

Economists advocate for using the price mechanism to manage water scarcity, but complex nonlinear rate structures prevalent in markets for municipal water obscure price signals. We conduct a randomized field experiment that jointly elicits knowledge about the cost of water and examines the impact of improved information on demand by linking a survey to water billing records. Half of our sample of 30,000 single family homeowners are randomly sent an invitation to a survey that asks questions about the water bill and the costs of water-use activities (e.g. the cost of taking a shower), and subsequently provides personalized accurate information. Results show that consumers have poor information about the marginal price of water and overestimate the costs of using water. Respondents are relatively better informed about their total bill and water consumption. In aggregate, respondents increase water use in response to the survey, potentially due to learning that water is cheaper than they previously thought. Increased consumption is concentrated among low users who are more likely to over-estimate the costs of using water.

Keywords: water demand; demand elasticity; economics of information; experimental economics; behavioral economics

JEL Classification Numbers: Q21, Q25, Q54.

---

*Corresponding Author:
Daniel Brent, *email:* dab320@psu.edu. The authors are grateful for feedback provided by Katrina Jessoe, Kurt Schwabe, Dan Kaffine, and seminar participants at the JEEM Inaugural Conference, the San Diego Energy and Water Workshop, the AERE summer conference, and the Strata Environmental Policy Workshop. Casey Wichman and Nathan Chan assisted in the development of the conceptual framework. The authors are thankful to Janet Wade and Brett Mathesien from Yarra Valley Water for providing data and insightful comments throughout. We acknowledge financial support from the Australian Research Council grant: DP120103846.

1 Introduction

A fundamental principle in microeconomic analysis is that consumers respond to higher prices by reducing the quantity demanded. However, the development of new pricing structures and payment mechanisms, as well as growing interconnectedness of goods and services, complicate the price-quantity decision sphere.[1] Urban water represents a prime example where the link between price and quantity is obscured. Utilities often employ complicated rate structures that combine a fixed charge with a volumetric charge, and the marginal price often depends on the level of consumption. One of the most common rate structures is the increasing block rate (IBR) that charges low marginal prices for the initial units and higher marginal prices for subsequent units. These rate structures proliferated as a way to ensure a basic level of access to energy and water, goods that are deemed essential for modern life, while at the same time discouraging wasteful consumption. While the focus of this article is on perceptions and the role of information in municipal water demand, similar features exist in residential electricity markets.

We designed a survey to elicit perceptions of the cost of using water for consumers in Melbourne, Australia who face an increasing block rate structure for residential water demand. We randomly sent the survey to half of a sample of 30,000 customers for which we have billing data. The linked billing data allow us to personalize the survey with historical consumption data and estimate the effect of taking the survey on subsequent water consumption. This research design allows us to address two objectives. First, we assess the baseline customer knowledge of water prices and other features of the bill and estimate what type of consumers acquire accurate price information. This builds off the work of Attari (2014) that elicits information about water quantity, but not prices or costs. Additionally, we define and elicit the costs of specific activities (we term this cost-per-use), such as flushing the toilet, that combine both information on water prices and the quantity intensity of different activities. Second, we estimate the impact of providing updated information on water consumption.

The ability of simple price and quantity information to change consumer behavior in energy and water markets serves as evidence that consumers do not have accurate perceptions about water prices and the quantity they consume. However, most studies infer inaccurate or incomplete information from the behavioral response to information, as opposed to directly collecting consumers' price perceptions.[2] In contrast, our results *directly* show that consumers' price knowledge is very poor and that providing correct

[1] Grubb and Osborne (2014) and Grubb (2015) show how consumer inattention and "bill shock" has important implications in purchasing cell phone plans.

[2] One is example in energy is Allcott and Wozny (2014) and Wichman (2017) is an example in water.

information changes behavior. Consumers generally know their total bill, but not the marginal price of water. Many consumers do not know how much water they use, but they have better information on the proportion of the bill that is due to volumetric charges. These results indicate that total water consumption and the marginal price are not the most relevant metrics for consumer decisions in residential water markets. This motivates our cost-per-use (CPU) elicitation for four common water activities: irrigation, toilets, washing machines, and showers. CPU combines the marginal cost of water with the water quantity requirements for common end uses. Consumers have even worse information on the cost of common water-using activities. In general consumers overestimate the marginal price of water, as well as the cost of common water activities.

We also predict what type of consumers have accurate price information. Consumers who use more water have more accurate estimates of their total bill, marginal prices, and the percentage of their bill that is due to volumetric charges. This is consistent with endogenous information acquisition where households with high water bills have a larger incentive to acquire accurate price information.[3] Water consumption does not have an effect on the accuracy of the CPU estimates. We also elicit the respondents' confidence in each their estimates, which is a reasonably strong predictor of the accuracy of their estimates indicating that water costs are known unknowns.

In aggregate consumers increase water use after the survey. We estimate local average treatment effects (LATE) for taking the survey on water consumption where we use the randomized letter as an instrument. To generate more precise estimates we also use a genetic matching algorithm developed by Diamond and Sekhon (2013) to estimate average treatment effects on the treated (ATT) and embed the matched sample in a regression framework (Ho et al., 2007). Our estimates show that consumers increase consumption by roughly 6-10%.

We develop a conceptual framework that describes three primary mechanisms through which the survey may affect consumption: increased salience of moral costs, correcting price mis-perceptions, and alternative behavioral biases. We describe the alternative biases as internalities - defined as biases that prevent consumers from accounting for all the *private* benefits and costs of a particular good (Allcott and Sunstein, 2015; Allcott et al., 2014). Examples of internalities outside of imperfect price perceptions in our setting are inaccurate quantity information on specific water activities or mistakes in understanding the bill structure, such as the percentage of the bill due to volumetric charges.

On average, respondents overestimate the costs of using water and learn that water is

[3] A related concept is rational inattention (Sallee, 2014) where consumers do not invest in acquiring information that will not affect there choices.

cheaper than they previously thought. Additionally, there is evidence that the treatment effect is concentrated among lowest users who are more likely to over-estimate the costs of using water. Therefore, one interpretation of the results is that the survey serves as a price decrease, which causes an increase in consumption. Alternatively, consumers could be responding to the CPU information, which helps them re-optimize by learning the marginal cost of specific activities. The CPU information best fits into the general framework of internalities as opposed to price perceptions.

While other information such as the total bill and the proportion of the bill due to volumetric charges may also drive behavior change, consumers were initially better informed along these dimensions. We posit that pure inattention or moral costs are less plausible mechanisms explaining the demand response to the survey. If we assume that the survey only addresses price perceptions, a stylized model estimates that correcting price mis-perceptions increases consumer surplus by approximately 1% relative to average bills. Therefore, the costs of mis-information to individual consumers are low relative to the aggregate costs of water consumption.

We make several contributions to the literature: first, we document consumers' price perceptions; second, we explore who acquires accurate information on the cost of water; and third, we estimate the effect of accurate information on subsequent consumption. Both eliciting price perceptions and observing the behavioral response to updated information help us interpret the causal mechanisms through which information affects behavior. Collecting data on price perceptions also highlights the importance of understanding biases in consumer behavior. The existence of an energy efficiency gap implies that there are win-win opportunities that generate private benefits to consumers along with public benefits through reduced internalities.[4] However, we find that water consumers are over-consuming relative to the private optimum, which would lead to tradeoffs between private and public benefits if there are unpriced externalities in water.

Additionally, we contribute to the evaluation of informational programs using both experimental and non-experimental methods by highlighting the bias-variance tradeoff in experimental versus non-experimental econometric methods. We extend the work by Wichman and Ferraro (2017) and Ferraro and Miranda (2017) that attempt to recover treatment effects from randomized experiments using non-experimental methods in two ways. First, we apply the methods to randomized encouragement designs, and second, we apply more recent advanced matching estimators (Diamond and Sekhon, 2013). The ability to leverage non-experimental estimators that may generate more precise estimates is particularly important for randomized encouragement designs that have low enrollment

[4]A summary about the existence of an energy efficiency gap is provided in Jaffe and Stavins (1994), Allcott and Greenstone (2012), Palmer and Walls (2015), and Gerarden et al. (2015), among others.

rates and limited statistical power.

Our results are relevant for research testing the appropriate price signal for consumers facing non-linear price structure. The survey results document that consumers do not have good knowledge of either marginal prices or total consumption, which makes it unlikely that consumers respond directly to either marginal or average volumetric prices. Since consumers generally know their total bill, they may exhibit non-standard behavior such as reducing consumption by a discrete amount when a water bill exceeds some bandwidth (sometimes described as "bill shock"). This fits into the mental accounting framework of Thaler (1985), and empirical evidence from other markets such as mobile phones (Grubb and Osborne, 2014; Grubb, 2015). The lack of accurate price information is related to other research examining incomplete information or imperfectly optimizing agents in residential energy and water consumers markets (Allcott and Wozny, 2014; Sallee, 2014; Houde, 2018; Brent and Ward, 2018). Documenting price perceptions has implications on the literature examining how consumers respond to nonlinear pricing (Ito, 2014; Wichman, 2014) by collecting data on what type of price information consumers actually possess.[5]

We also show that improved price information increases water use. This finding is relevant to research studying the impact of improved price information on energy and water demand is mixed. Jessoe and Rapson (2014) show that households with easy access to real time price information via in-home displays are much more responsive to temporary energy price increases than uninformed households. Monteiro et al. (2018) find that informed water consumers are relatively more price elastic. Several other studies show that the timing and mode of billing impacts consumption. Wichman (2017) shows that increased billing frequency, which improves price information consumption, leads to higher water consumption, while Sexton (2015) finds that automatic bill payments, which reduces price information, also increases consumption. Gilbert and Zivin (2014) shows that within a billing period consumers decreases consumption immediately after receiving a bill. An important distinction is that previous research infers consumers' price information by the behavioral response to changes in the available information. We go beyond this by both estimating the demand response to information and collecting baseline information.

This research is also related to how information nudges affect energy and water demand.[6] Several studies analyze how price information affects residential electricity de-

[5]The research on average vs. marginal price assumes that it is more difficult to understand or find marginal price information (Shin, 1985), but this has not been tested empirically. Additionally, distinctions between average total price, which include the fixed cost, and average volumetric price, which does not, are not given proper attention in the literature. Implicitly, consumers must have some information about both prices and quantities to respond to average prices.

[6]Research investigating how social comparisons affect energy and water demand include, but are not

mand. Kahn and Wolak (2013) show that an energy use tutorial that informs consumers on energy costs decreases consumption, but there are heterogeneous effects based on the rate structure. Pellerano et al. (2015) finds that increasing nonlinear price salience affects energy consumption near the kink point. McRae and Meeks (2016) elicit energy price knowledge and show how consumers with different price information reacted to historical changes in the electricity rate structure. Byrne et al. (2018) show that peer information increases electricity consumption for households who overestimated their energy use, while it decreased consumption for households who underestimated energy use. Similar to Kahn and Wolak (2013), Stojanovski et al. (2018) show that in response to a randomized field experiment that improves price information, electricity users facing a high marginal price decrease consumption, while households facing a low marginal price do not. Our study is different by studying water demand, a well as collecting and providing different types of information such as cost-per-use.

The next section develops a conceptual framework to map potential mechanisms through which information can affect water demand. Section 3 describes the experimental design and institutional setting. Section 4 shows the results of the price and cost perceptions elicitation and analyzes what type of consumers acquire accurate price information. Section 5 describes the empirical methods and results of the field experiment by estimating the effect of updated information on subsequent water use. Section 6 interprets the empirical results within the conceptual model. We discuss the implications of the results and avenues for future research in Section 7.

2 Conceptual framework

We begin with a model of water demand in the presence moral costs, imperfect price perceptions, and general internalities. This is based on the theoretical framework of Brent and Wichman (2018), which uses the notation of Allcott and Kessler (2019). Consider a consumer with income y who gains consumption utility from water w via $f(w;\alpha)$ and the numeraire good x, where α is an individual taste parameter. An internality parameter $\gamma > 0$ affects choice utility but not experienced utility, and $\gamma \neq 1$ implies mistakes in evaluating the private benefits and costs of water consumption, or some other behavioral bias. Consumers have perceived utility $\hat{f}(w;\alpha,\gamma)$, which we assume takes the form $\gamma^{-1}f(w;\alpha)$. Perceived utility is higher than consumption utility when $\gamma < 1$, resulting in too much water consumption relative to the private optimum. Additionally, we allow the perceived price $\tilde{p}$ to differ from actual price p. Lastly, similar to Levitt

limited to Allcott (2011b,a); Ferraro et al. (2011); Ferraro and Miranda (2013); Bolsen et al. (2014); Brent et al. (2015); Jessoe et al. (2018).

and List (2007) and Ferraro and Price (2013), we also include a "moral utility" term, $M = m - \mu w$, which captures nonpecuinary impacts associated with consumption of w.[7] 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

$$y = x + \tilde{p} w \tag{2}$$

Th first order condition for choosing $\tilde{w}$ to maximize decision utility is given by:

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

Equation 3 states that consumers choose consumption of $\tilde{w}$ to equalize their marginal perceived utility with the sum of perceived monetary and moral costs.[8] Because γ introduces a wedge between perceived marginal utility and a consumer's true marginal utility, the choice of $\tilde{w}$ is not required to be individually optimal. The framework is consistent with stylized formulations in Sexton (2015) and Wichman (2017) who model price (and quantity) mis-perceptions.

Totally differentiating equation 3 and solving for the change in water consumption, $\mathrm{d}\tilde{w}$, relates changes in consumption to changes in perceived prices, internalities, and moral costs.

$$\mathrm{d}\tilde{w} = \frac{1}{f''(\cdot)} \left[(\mu + \tilde{p})\, \mathrm{d}\gamma + \gamma \left(\mathrm{d}\mu + \mathrm{d}\tilde{p} \right) \right]. \tag{4}$$

Our survey could conceivably operate through some combination of the channels: $\mathrm{d}\gamma$, $\mathrm{d}\mu$ or $\mathrm{d}\tilde{p}$. We assume that the survey will move the consumers towards more accurate information and correcting internalities.Under standard concavity assumptions of demand (i.e., diminishing marginal utility), f'' is weakly negative, which allows us to consider how changes in consumption link to mechanisms through which the survey affects decision-making. We consider three cases that each isolate an individual mechanism of interest.

Case 1: The survey increases salience of the moral cost of water.
We designed the survey to focus on the financial costs of water consumption. However, the process of estimating water consumption and then learning about actual consumption

[7]Levitt and List (2007) describe moral utility within a model where "utility maximization is influenced not only by wealth maximization, but also by an individual's desire to "do the right thing" or make the "moral" choice." In our context consumers feel bad about excess water consumption.

[8]Since both water and the numeraire are necessary for human survival, we are comfortable ignoring corner solutions.

might cause respondents to update the moral cost of consumption. If this is the only mechanism then $\mathrm{d}\gamma = \mathrm{d}\tilde{p} = 0$ and the change in consumption is $\mathrm{d}\tilde{w} = \frac{1}{f''(\cdot)}\gamma \mathrm{d}\mu$. The survey will increase consumption if moral costs are adjusted downward and decrease consumption if moral costs increase.

Case 2: The survey corrects price mis-perceptions.
The survey elicits price perceptions and then we provide the actual price of water. If correcting price mis-perception is the only mechanism then $\mathrm{d}\mu = \mathrm{d}\gamma = 0$ and the change in consumption is $\mathrm{d}\tilde{w} = \frac{1}{f''(\cdot)}\gamma \mathrm{d}\tilde{p}$. The expected change in consumption depends on the change in the perceived price $\mathrm{d}\tilde{p}$, which in turn depends on initial price perceptions. By assumption the new perceived price moves towards the actual price. If consumers were initially overestimating prices ($\tilde{p} > p; \mathrm{d}\tilde{p} < 0$) then water consumption increases, and consumption decreases if consumers were initially underestimating prices ($\tilde{p} < p; \mathrm{d}\tilde{p} > 0$).

Case 3: The survey corrects internalities not related to the price of water.
The survey provides information on the non-price features of the bill (e.g. % volumetric charges), and the marginal costs of specific activities. Therefore, the survey may address internalities such as incorrect information on the water intensity of specific activities (Attari, 2014), or mis-perceptions of the bill structure. If consumers respond to a general behavioral bias not related to moral costs or incorrect price perceptions, then $\mathrm{d}\mu = \mathrm{d}\tilde{p} = 0$ and the change in consumption is $\mathrm{d}\tilde{w} = \frac{1}{f''(\cdot)}\left(\mu + \tilde{p}\right)\mathrm{d}\gamma$. The expected change in consumption depends on the initial level of γ, assuming strictly positive perceived prices and moral costs ($\mu + \tilde{p}$). If perceived utility is less than actual utility ($\gamma > 1$) then initial consumption is too low ($w^* > \tilde{w}$) and the survey will increase consumption ($\mathrm{d}\gamma < 0; \mathrm{d}\tilde{w} > 0$). If perceived utility is greater than actual utility ($\gamma < 1$) then initial consumption is too high ($w^* < \tilde{w}$) and the survey will decrease consumption ($\mathrm{d}\gamma > 0; \mathrm{d}\tilde{w} < 0$).

These cases represent extreme scenarios where the survey only affects one mechanism, but realistically the survey may impact multiple mechanisms simultaneously. The design of our survey does not allow for formal predictions between these mechanisms because some mechanisms have similar empirical signatures. However, we use the model as an organizing framework to link our empirical results to potential mechanisms.

3 Background and Design

3.1 Institutional Setting

We collaborated with Yarra Valley Water (YVW), the largest water company serving greater Melbourne with over 1.7 million retail customers.[9] Customers are billed four times a year at approximately 90 day intervals. Quarterly billing is common in Australia, though it is less frequent than many water utilities in the United States that bill customers every 30 or 60 days. There are three primary charges from YVW on a customer's water bill: usage charges, a fixed water supply charge, and a fixed sewage system charge.[10] There is an additional fixed waterways and drainage charge that is paid on the YVW bill but is set by the wholesaler, Melbourne Water. In Victoria, Melbourne Water manages the water supply system including reservoirs and desalination plants, and sells water to major retail companies, including YVW, that in turn sell directly to end-users. The usage charge is comprised of both volumetric water and sewage charges.

The volumetric water charge is set by an increasing block rate tariff as shown in Figure 1. The prices within each block have changed over time but the basic rate structure has remained the same. The thresholds for the higher price tiers are set at 40 kiloliters (kL) and 80 kL.[11] For the whole year approximately 46% of customers consume in the first tier, 42% in the second tier, and 12% in the highest tier. During the summer period, when water consumption is higher due to outdoor water use the distribution shifts towards the higher price tiers; 39% of customers consume in the first tier, 43% in the second tier, and 18% in the highest tier.

There is only one volumetric sewer price, however, sewage volume must be estimated since it is not directly metered. The estimate for sewage volume is calculated by multiplying metered water use by a seasonal factor and a discharge factor. The seasonal and discharge factors approximate how much water is used outside, and hence should not be subject to the sewage charge. The seasonal factor varies by season but not by household, whereas the discharge factor depends on how much water a household uses. Below 125 kL per quarter the discharge factor is 0.9 and above 125 it decreases at a per-kL basis since water above this threshold is likely for outdoor use. The discharge factor is capped from below at 0.45 when consumption reaches 250 kL.

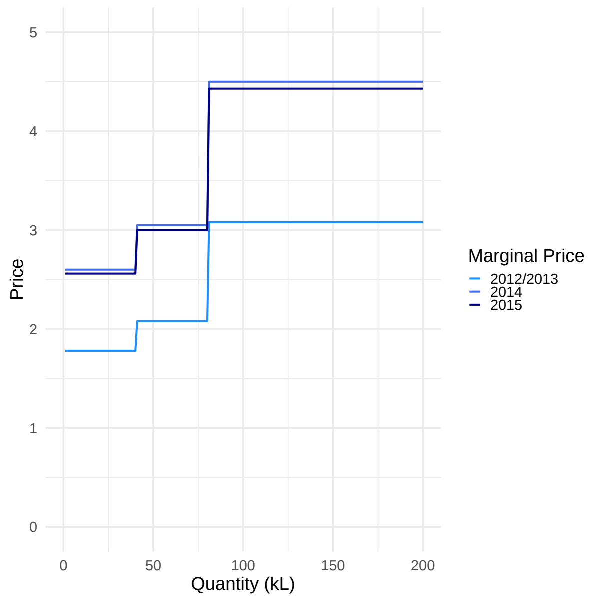
There is a substantial amount of information on the bill structure, which shows how

[9] Basic information about Yarra Valley Water can be found at https://www.yvw.com.au/Home/Aboutus/Ourorganisation/index.htm.

[10] An example water bills is presented in Figure A.1 in the Appendix.

[11] These are actually set by water consumption per day to account for differences in when the meter is read, but this works out to 40 kL and 80 kL for the median billing period.

Figure 1: Water Rates



Notes: The graph show the water rates for Yarra Valley Water during the study period and how the marginal price changes over time.

complicated water bills can be and helps put the consumer response into perspective. There are many choices the analyst must make when determining how to define prices. Should water and sewage charges be combined or treated separately? Are fixed charges incorporated into average prices, and if so are third party charges included? These are not not arbitrary issues, and the survey is an attempt to uncover what customers actually know about their relatively complex water bills.

3.2 Survey

We designed our survey with two research questions in mind. First, we elicit consumers' perceptions and compare them to the true cost of water. Second, we inform the consumers about the true cost of water and examine how the updated information affects future water use. The survey is divided into three sections: (1) demographics and structural features of the house, (2) questions eliciting information about the costs of using water, and (3) information about the cost of using water. We provide the true answers based on historical data for each question about water costs that we elicit from the respondents. The primary goals from the survey are to estimate consumers' knowledge of water use costs from section (2) and examine if updated information provided in section (3) changes water use behavior. The survey results are linked to customers' account information that pre-loads individualized data on water use into the survey. Therefore, we know each household's historical water use, typical pricing tier, and average bill. Prior to eliciting information from respondents we inform them about the basic features of their water bill (see Figure A.2). We emphasize that they should not look at their bill when answering the survey. An example of the updated price information is shown in Figure A.3.

List 1, in the Appendix, presents the survey questions with the information that we provide in italics. The 10 questions eliciting information about the consumer's perceptions of the cost of using water are divided into three categories: (1) questions 1 and 2 are about average usage and the total bill; (2) questions 3-6 relate to features of the rate structure including marginal prices; (3) questions 7-10 describe the marginal cost of activities that comprise the majority of water use. After asking the respondent how much water they use and their average bill (questions 1-2), we provide the correct answers to the respondent. This is to present ballpark figures for general water use and costs for the subsequent questions. We were concerned that without some baseline information on costs and quantities of water the respondents would not be able to answer the subsequent questions of the survey in a meaningful way. We expect that this should improve the answers in questions 3-10, and thus the accuracy of those questions should be considered upper bounds. We phrase the cost-per-use questions (7-10) as the "net bill impact" to represent the marginal cost of these activities, which includes both volumetric water and sewage charges. After each question we ask the respondent how confident they are in their answer.

3.3 Field Experiment

In order to estimate the effect of the updated information from taking the survey on subsequent water consumption we randomize the households whom we invite to take the survey. Our target period is the summer of 2015 since most discretionary water use occurs when households use water outside the home for irrigation and pools. Our experimental sample is the set of single family customers who own their home and had their meter read from January 7th through February 2nd 2014. This captures a group of customers that have relatively homogeneous weather shocks while still allowing us to stagger the invitation letters to accommodate printing and mailing constraints. We drop customers who recently moved and whose consumption has changed by over 100kL during any summer consumption quarter in the pre-treatment data.[12] These customers have erratic consumption behavior and are not representative of the typical household; including them would increase the variance of our sample data. We use a block randomization design

[12] 100kL corresponds to roughly the 90th percentile of consumption so we are just excluding customers who move form the highest to the lowest consumption during our sample.

where we first sort households into 100 quantiles based on historical consumption and then randomly assign households to either treatment or control. Our overall sample consisted of 30,825 households, of which 14,755 received an invitation to the survey. We received 1,630 responses, corresponding to an 11% response rate for the full sample. As shown in Figure 4 and Table 5 the survey respondents are not a random sample of the population, and we must consider the selection process when interpreting the results on price and cost perceptions. Respondents use less water than non-respondents, which may affect the responses to the survey in two ways. First, lower users may be inherently more concerned about water use and therefore have more accurate information. However, among our survey sample households with higher bills are better informed across some dimensions so the fact that our sample consists of more low users could lead to our sample being less informed about water costs than the general population.

4 Price and Cost Perceptions

4.1 Do consumers know how much water costs?

Our first research objective is to document consumers' knowledge about the costs of water. We elicit two types of information about the cost of using water. First, we ask if consumers' know about relevant features of the bill structure: including their total bill, total water use, marginal price, and the percentage of their bill due to volumetric charges. The survey was sent out during the summer and all of the questions are framed in terms of the average bills during the summer season when there is significant outdoor water use. Histograms of the accuracy of consumers' knowledge of the bill structure are presented in panel (a) of Figure 2. Since the questions are in different units we convert the answers to the percentage difference from the correct answer. For example, when interpreting the graph for the marginal price, a value of zero means the consumer knew the exact marginal price, whereas values of 50 and -50 represent consumers who overestimated or underestimated the price by 50%, respectively. Consumers have reasonably accurate knowledge of the total bill and the proportion of the bill that is volumetric, as evidenced by the histograms that are both relatively tightly centered around zero. Conversely, the more diffuse distribution for the marginal price shows that consumers' had very poor knowledge about the marginal price of water. Due to the increasing block pricing structure the marginal price depends on water consumption, so we also ask consumers how much water they historically used during the summer quarter.[13] Many consumers

[13] As seen in Figure A.1 consumers see the marginal prices on their bill so they could know the marginal price without knowing their quantity

knew how much water they typically use, but there are many respondents who either overestimated or underestimated their use as seen by the spike near -100 and the fat right tail.

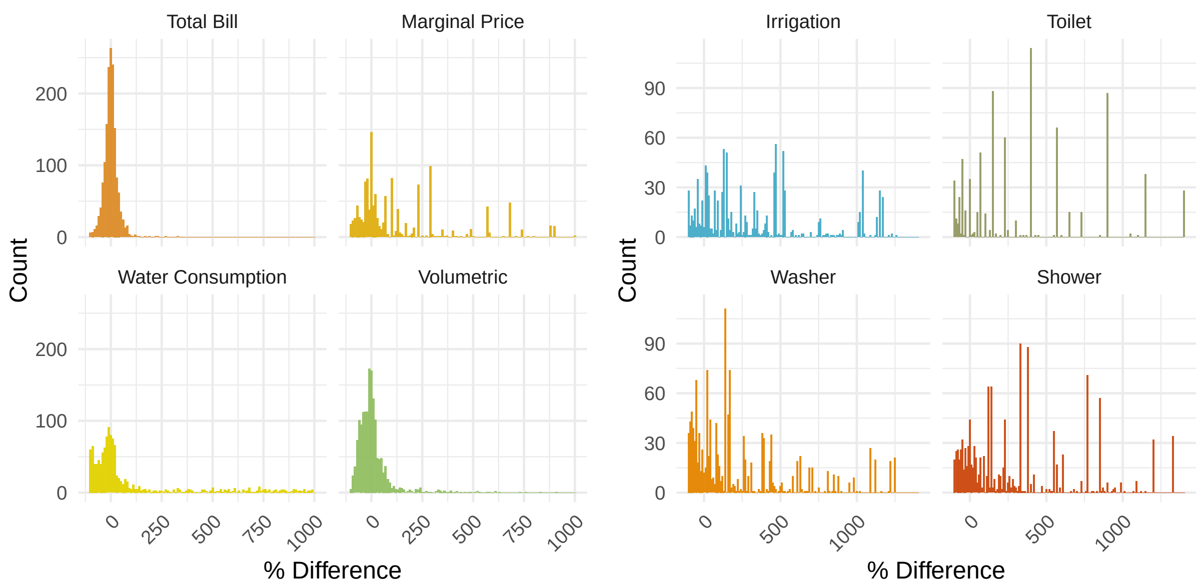
While the features of the bill are the basis for making economic decisions about water use, the marginal price and percentage of the bill that is volumetric may be opaque metrics of the cost of water to a consumer. This is because the quantity of water is not salient for many consumers; most consumers do not know what actions represent a one-kL reduction in water use. Therefore, we define the costs-per-use (CPU) of common activities that use water as an input. We focus on irrigation, showers, toilets, and washing machines since these represent over 70% of total water use in the study area (Gan and Redhead, 2013). We phrase the questions as the "net bill impact of using an additional unit of water", which accounts for both the volumetric water and sewage charges. We calculate the CPU by multiplying the average water used for the various activities by the typical marginal price that each consumer faced in the summer period. The average water use for the activities we measure is calculated by Gan and Redhead (2013) through a combination of engineering estimates and field samples in the Yarra Valley Water service area. For washing machines and toilets the CPU is the cost-per-load and cost-per-flush, respectively. For irrigation and showers we multiply the average flow rate by the average duration in the service area, which is 7 minutes for showers and 20 minutes for irrigation. We communicate that these are rough estimates and that the actual CPU will depend on the specific appliances and water-use behavior of the household. It should also be noted that since our metrics of correct CPU are based on estimates our measures of the difference between consumers answers and the truth contains some measurement error.

Consumers have very poor information about the cost of various water-using activities. All of the distributions are wide and not necessarily centered around zero. Most consumers overestimate the water cost of these common activities. This is important because changing these actions, or investing in efficiency that reduces the cost of these actions, are the primary ways that consumers can reduce water consumption. In order to make informed decisions consumers need to know both the costs and benefits of these actions.

Many consumers in our sample have very poor price information, as evidenced by estimates that are over ten times above the true value. Since consumers may be employing something close to a random guess, we also ask consumers how confident they are in their answers. This is a descriptive measure of consumers' self-perceived price knowledge that we present in Figure 3. The answers to the question about confidence line up with the accuracy of the estimates, indicating that consumers' know whether they have accurate cost information - water costs are known unknowns. Consumers are most confident about

Figure 2: Price Perceptions

(a) Bill Structure (b) Cost-per-use



Notes: These graphs display the accuracy of the estimates as the percentage difference from the correct information for the each of variables shown. Estimates that were 10 times the sample maximum of the correct value were dropped.

their bill, the % volumetric, and water use, and they also are more likely to answer these questions correctly. The vast majority of respondents have little or no confidence in the CPU estimates, and generally are not confident with their estimate of the marginal price. We formally investigate if more confident respondents actually have more accurate information in section 4.2.

We summarize the information on cost perceptions in Table 1. This provides the mean of the actual values for the various questions that we asked, along with the mean of the respondents' estimates. We also show the average error and absolute value of the error in percentage terms. As seen from the table the mean error is positive for almost all of the estimates. Most consumers also exhibit relatively low confidence in their answers.

4.2 Who acquires price information?

Next, we examine which type of consumers acquire accurate information about the cost of water. These models exclusively focus the survey respondents, which includes the correct estimates of the features of the bill and CPU based on historical billing data. We run a linear probability model of the following form:

$$Accurate_{ik} = \alpha + \boldsymbol{\beta X_i'} + \epsilon_i \tag{5}$$

Figure 3: Confidence in Estimates

Note: The figure presents a histogram for confidence of the estimates, which was asked immediately after the estimates were elicited.

Our dependent variable, $Accurate_{ik}$, is a dummy variable equal to one if respondent i answered question k accurately and $\mathbf{X_i'}$ is a vector of independent variables. Since the raw estimates are very noisy we construct a categorical metric for accuracy. Our primary metric is whether the respondent's error on question k was below the sample median. We also construct a dummy variable equal to one if the respondent was within 50% of the correct answer.

We consider standard demographics, a dummy for being above the median income, a dummy for whether the respondent has a college degree or higher, household size, and a dummy if the respondent is over 65 years old. We also include the standardized average summer water consumption and certain answers to survey questions. Respond to Prices is a dummy if the respondent said that she responded to past price increases. Motivated by money is a dummy equal to one if the respondent stated that water use was motivated by saving money. Investments is the stated number of water efficiency investments. Confidence is a dummy equal to one if the respondent stated that their confidence level

Table 1: **Summary of Cost Perceptions**

Variable	Mean	Mean Estimate	Mean Error (%)	Mean Abs. Error (%)	Low Confidence (%)
Water	53.32	189.87	374	417	51
Bill	358.37	350.59	2	26	34
% Volumetric	58.18	60.02	10	51	54
Water price	3.03	7.84	172	198	68
Sewer price	2.1	5.12	144	176	78
STC	69.94	35.22	-50	56	78
Shower	0.23	0.97	322	343	85
Washing machine	0.42	1.41	243	274	83
Flush	0.02	0.11	372	396	83
Irrigation	0.89	3.87	341	357	82

Note: The first column presents the actual mean for the sample, and the second columns presents the mean of the respondents' estimates. The Mean Error shows the average percentage difference from the estimate and the correct answer, and the Mean Abs. Error is the absolute value of the percentage difference. Low Confidence shows the percentage of respondents who recorded either "None" or "Little" to the question about the confidence of their estimate. The units of Mean and Mean Estimate are as follows: Water is in kL; Bill, Water price, and Sewer price are in $ (AUD); % Volumetric and STC are percentages; Shower is in $-per-flush; Washing machine is in $-per-load; Flush is in $-per-flush; Irrigation is in $-per-20 minutes of irrigation.

for that estimate was either high or very high. The number of observations varies due to missing data for some of the estimates.

Estimates related to the bill structure are presented in Table 2 and estimates of the CPU are presented in Table 3. Most of the coefficients are insignificant, and the low R^2 shows that variation in respondents' characteristics explains very little of the variation in respondents' knowledge of the costs of using water. One exception for the determinants of bill structure knowledge, shown in Table 2, is that the households who use more water have more accurate information about the costs of using water. This is consistent with logic that consumers with large bills invest more time and effort to learn about the water rate structure. Older respondents are more likely to know the marginal price of water and wastewater. The confidence variable is positive and significant in most specifications, indicating that respondents know whether or not they have accurate information. One exception is for the percent volumetric where consumers confidence is negatively correlated with accurate information.[14]

The CPU estimates are noisier, which is reflected in less precise estimates and lower R^2 values in Table 3. Water consumption does not improve the accuracy of the CPU estimates. One apparent pattern is that respondents who said they responded to past prices have better CPU knowledge, whereas respondents stating their water use is motivated by money have worse CPU knowledge. Confidence is correlated with better estimates for the costs of showers and toilets.

The poor model fit and general lack of statistically significant coefficients highlights

[14]The number of observations in Tables 2 and 3 vary and are less than the 1,630 survey respondents for two reasons. First, some respondents refused to answer certain demographic questions such as their household income, and second some of the perceptions data (the dependent variables) was dropped to vary due to missing or invalid responses.

Table 2: Determinants of Price Information: Bill Structure

	Bill	Water	Water Price	Sewer Price	Volumetric
	(1)	(2)	(3)	(4)	(5)
Water	0.137***	0.032**	0.027*	0.028*	0.067***
	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)
Income	−0.045	−0.020	0.033	0.029	−0.009
	(0.038)	(0.040)	(0.039)	(0.039)	(0.039)
Degree	0.028	0.013	−0.013	−0.011	0.044
	(0.031)	(0.033)	(0.032)	(0.032)	(0.032)
HH Size	−0.061***	0.005	0.009	0.007	0.010
	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)
65+	0.007	−0.017	0.108***	0.101***	0.030
	(0.033)	(0.036)	(0.035)	(0.035)	(0.034)
Respond to Prices	−0.053	−0.011	0.033	0.035	0.041
	(0.042)	(0.045)	(0.044)	(0.044)	(0.043)
Motivated by Money	−0.035	−0.018	0.009	0.009	−0.025
	(0.031)	(0.033)	(0.032)	(0.032)	(0.032)
Investments	−0.021	−0.012	0.010	0.009	0.014
	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)
Confidence	−0.046	0.076*	0.303***	0.391***	−0.117**
	(0.036)	(0.042)	(0.062)	(0.077)	(0.059)
Constant	0.749***	0.512***	0.398***	0.411***	0.436***
	(0.067)	(0.071)	(0.069)	(0.068)	(0.068)
Observations	1,226	1,172	1,198	1,198	1,229
Adjusted R^2	0.077	0.0001	0.025	0.026	0.021

Note: The dependent variable is a dummy equal to one if a respondent was below the sample median in terms of the absolute value of their estimation error and zero otherwise. Respond to Prices is a dummy if the respondent said that she responded to past price increases. Motivated by money is a dummy equal to one if the respondent stated that their water use was motivated by saving money. Investments is the number of water investments that the respondent stated they had made. Confidence is a dummy equal to one if the respondent stated that their confidence level for that estimate was either high or very high. *p<0.1; **p<0.05; ***p<0.01

the difficulty in predicting consumers' knowledge of water costs. As a robustness check we replace our metric for accurate price information with a dummy equal to one if the respondent was within 50% of the true value. This is an absolute metric for accurate price information as opposed to a relative one. The results are largely consistent and are available in the Appendix (Tables A.1 & A.2).

Table 3: Determinants of Price Information: Cost-per-Use

	Irrigation	Washing Machine	Shower	Toilet
	(1)	(2)	(3)	(4)
Water	−0.007	0.024	0.019	0.007
	(0.015)	(0.015)	(0.015)	(0.015)
Income	0.054	0.045	0.020	0.054
	(0.040)	(0.039)	(0.039)	(0.040)
Degree	0.059*	0.033	0.039	0.064**
	(0.032)	(0.032)	(0.032)	(0.032)
HH Size	0.026**	0.014	0.022*	0.004
	(0.013)	(0.013)	(0.013)	(0.013)
65+	0.055	0.009	0.049	0.025
	(0.035)	(0.035)	(0.035)	(0.035)
Respond to Prices	0.052	0.094**	0.084*	0.085*
	(0.045)	(0.044)	(0.044)	(0.044)
Motivated by Money	−0.010	−0.068**	−0.056*	−0.104***
	(0.032)	(0.032)	(0.032)	(0.032)
Investments	0.0002	−0.012	−0.008	−0.005
	(0.015)	(0.015)	(0.015)	(0.015)
Confidence	−0.038	0.041	0.453***	0.347**
	(0.062)	(0.112)	(0.129)	(0.138)
Constant	0.332***	0.390***	0.363***	0.414***
	(0.070)	(0.069)	(0.069)	(0.069)
Observations	1,192	1,199	1,202	1,187
Adjusted R^2	0.006	0.021	0.024	0.028

Note: The dependent variable is a dummy equal to one if a respondent was below the sample median in terms of the absolute value of their estimation error and zero otherwise. Respond to Prices is a dummy if the respondent said that she responded to past price increases. Motivated by money is a dummy equal to one if the respondent stated that their water use was motivated by saving money. Investments is the number of water investments that the respondent stated they had made. Confidence is a dummy equal to one if the respondent stated that their confidence level for that estimate was either high or very high. *p<0.1; **p<0.05; ***p<0.01

Next, we also estimate equation 5 where we replace the dependent variable with a indicator variable for whether the respondent overestimated the variable in question. For example, when considering the total bill the dependent variable would be equal to one if the respondent answered that their water bill was higher than it actually was. The results are reported in Table 4. Higher water use are less likely to overestimate almost all components of the costs of using water. This means that low water users are more likely to overestimate the costs of water, which is informative when analyzing heterogeneity in the demand response to the field experiment.

Table 4: **Determinants of Overestimating Costs of Water**

	Bill	Water	Water Price	Sewer Price	Volumetric
	(1)	(2)	(3)	(4)	(5)
Water	-0.212^{***}	-0.118^{***}	-0.055^{**}	-0.038	-0.170^{***}
	(0.022)	(0.025)	(0.024)	(0.027)	(0.022)
Income	0.039	0.040	0.014	0.017	0.033
	(0.038)	(0.042)	(0.039)	(0.045)	(0.037)
Degree	-0.032	-0.028	-0.019	-0.038	-0.069^{**}
	(0.031)	(0.034)	(0.032)	(0.035)	(0.030)
HH Size	0.057^{***}	0.008	-0.006	-0.031^{**}	-0.008
	(0.012)	(0.014)	(0.013)	(0.014)	(0.012)
65+	-0.026	0.065^{*}	-0.042	-0.094^{**}	-0.013
	(0.033)	(0.037)	(0.034)	(0.039)	(0.033)
Respond to Prices	0.047	0.053	0.029	-0.012	-0.042
	(0.042)	(0.047)	(0.044)	(0.049)	(0.041)
Motivated by Money	0.022	0.025	0.001	0.013	0.027
	(0.031)	(0.034)	(0.032)	(0.036)	(0.030)
Investments	0.019	0.022	0.020	0.012	-0.005
	(0.014)	(0.016)	(0.015)	(0.016)	(0.014)
Confidence	0.038	0.089^{**}	-0.263^{***}	-0.409^{***}	0.072
	(0.036)	(0.043)	(0.059)	(0.080)	(0.056)
Constant	0.488^{***}	0.519^{***}	0.750^{***}	0.715^{***}	0.618^{***}
	(0.065)	(0.071)	(0.067)	(0.076)	(0.064)
Observations	1,226	1,052	1,048	960	1,230
Adjusted R^2	0.076	0.022	0.019	0.032	0.054

Note: The is a dummy equal to one if a respondent overestimated the variable in question relative to the true value and zero otherwise. *p<0.1; **p<0.05; ***p<0.01

5 Field Experiment and Water Demand

5.1 Experimental Methods

Recall that we provided updated information to respondents for every question in List 1 after they completed the survey.[15] We use the Rubin potential outcomes framework (Rubin, 1974) to model the impact of this updated information on water consumption.[16] We only observe post-treatment water consumption for the billing period immediately after the completion of the survey. The most basic model is the average treatment effect (ATE) that estimates the effect of treatment on the population of interest. In this model Y_{it}^1 is the outcome variable given that respondent i received the treatment at time t,

[15]We provided this information immediately for total bill and water consumption to help the respondents answer the other questions. For the rest of the questions (3-10 in List 1) we provide the information at the end of the survey. See Figure A.3 for an example.

[16]We cannot fully separate the causal effect of different parts of the survey because it was randomized, so the treatment effects are based on taking the survey.

whereas Y_{it}^0 is the outcome conditional on not being treated. In our setting the outcome variable (Y_{it}) is the natural log of average daily water consumption and the treatment is completing the survey. The ATE is the expectation of the difference of these potential outcomes over the population of interest.

$$ATE = E[Y_{it}^1 - Y_{it}^0] \tag{6}$$

In our setting we randomize sending consumers invitation letters and cannot coerce recipients to complete the survey. Since many consumers do not respond to the invitation letter we cannot estimate the ATE, and instead we estimate several other treatment effects. First, we estimate the impact of sending an invitation letter on the population, the intent to treat effect (ITT). If we denote receiving a letter as a binary variable Z_{it} that takes on the value of one if a consumer receives the letter and zero otherwise we can write the ITT as:

$$ITT = E[Y_{it}^1 | Z_{it} = 1] - E[Y_{it}^0 | Z_{it} = 0] \tag{7}$$

Our primary goal is to estimate the effect of responding to the survey, therefore, we also estimate the local average treatment effect (LATE) of responding to the survey where we use the randomized receipt of a letter as our instrument (Imbens and Angrist, 1994). Completing the survey is denoted as a binary variable D_{it} that takes on one if the consumer completes the survey, which results in the LATE model:

$$LATE = \frac{E[Y_{it}^1 | Z_{it} = 1] - E[Y_{it}^0 | Z_{it} = 0]}{E[D_{it} = 1 | Z_{it} = 1]} \tag{8}$$

The LATE model scales the ITT by the probability that the instrument induces treatment and is estimated using two stage least squares (2SLS). Since our setting is a randomized trial with one-sided noncompliance our estimates of from the LATE model are equivalent to the average treatment effect on the treated (ATT).

$$ATT = E[Y_{it}^1 - Y_{it}^0 | D_{it} = 1] \tag{9}$$

5.2 Non-experimental Methods

Above we describe the simple estimators that rely on the experimental variation that produce unbiased estimates. However, our experiment is under-powered to calculate estimates purely using experimental methods. The mean of pre-treatment consumption in the billing period of interest was 55 kL with a standard deviation of 37. We included every household that met our inclusion criteria subject to our budget constraint as described in Section 3. This resulted in roughly 30,000 households and using conventional power parameters (power= 0.8; significance = 0.05) we can identify an ITT effect size of roughly 2%.[17] This is similar in magnitude to the effect of social comparisons in energy, but a key distinction is that our letter will primarily only affect those who select into the survey. The LATE estimate is scaled by the response rate to the invitation, so a 10% response rate means that our experiment is powered to identify a 20% change in consumption due to taking the survey. This is much larger than is typically found in the literature on behavioral interventions in water and energy demand. Therefore, in addition to estimating the ATT from the LATE model, we also use matching estimators to generate more precise estimates of the ATT. Ho et al. (2007) describe how matching can in fact increase the precision despite decreasing the sample size. The primary intuition is that both the conditional variance, and the dependence between the treatment variable and covariates, will often decrease after matching. Since many other settings also suffer from relatively small effect sizes and low response rates we consider the comparison of the experimental and non-experimental estimates to be of interest from a methodological perspective.

Our matching approach follows a wide literature that estimates $E[Y_{it}^0|X_{it}]$ by conditioning on observable variables X_{it} that are not affected by treatment.

$$ATT = E[Y_{it}^1|X_{it}] - E[Y_{it}^0|X_{it}] \tag{10}$$

where we assume that D_{it} is as good as randomly assigned after conditioning on household characteristics X_{it} We employ the genetic matching algorithm developed by Diamond and Sekhon (2013) to condition on observables. The genetic algorithm iteratively generates matched samples in order to balance the treatment and control samples along the full distribution of covariates using several non-parametric tests for balance.[18] We use one-to-one matching with replacement, and a caliper of 0.25.[19] We match based on average pre-treatment water consumption, and the postcode average income, education, owner-occupied status, and percentage of single family homes.

The matching algorithm generates a set of weights that can be used to re-weight the sample in regression analysis. We estimate several variations of the treatment effects based on the genetic matching algorithm. Our first approach embeds the weighted matched sample in a regression model in the spirit of using matching to pre-process the data using matching (Ho et al., 2007). This allows us to control for covariates after matching to increase the precision of the estimates. Next, we exploit the panel structure

[17] Our sample size is slightly smaller due to some households dropping out due to moving or not having their meters read on time for the post-treatment period.

[18] The matching procedure is implemented using the Matching package in R (Sekhon, 2011).

[19] The caliper of 0.25 ensures that each matched observation is within 0.25 standard deviations of the treated observation for all covariates used in matching.

of the data by estimating a panel fixed effects difference-in-difference (DID) model on the matched sample. Ferraro and Miranda (2017) show that combining matching with a FE panel model can closely replicate results from a randomized field experiment in a non-experimental setting. Lastly, we estimate the ATT directly by taking the difference between treatment and matched control observations as shown in Diamond and Sekhon (2013). The direct matching estimator of the ATT is $1/N_1 \sum_i (Y_i^1 - \hat{Y_i^0})$, where $\hat{Y_i^0}$ represents the matched control observation to treated observation i.[20] The standard errors are based on Abadie and Imbens (2006) which account for the (asymptotic) variance induced by the matching procedure itself.

In order to show the problem, and our solution to, selection, Figure 4 shows the distributions of household water use for (a) the respondents vs. the entire control group and (b) the respondents vs. the matched control.[21] We also show balance tables for historical average water use, historical water use for the same period of out dependent variable (second quarter of the year), and demographics at the postcode level from the Australian Bureau of Statistics. Since we are concerned with heterogeneity across the distribution of consumption we show not only p-values for t-statistics based on the difference in means but also non-parametric Mann-Whitney and Kolmogorov-Smirnov tests. Table 5 shows the balance on observables before and after pre-processing the data using the genetic matching algorithm of Diamond and Sekhon (2013). As seen from Table 5, matching greatly increases balance on observables; the lowest p-value for any of the statistical tests is 0.297.

5.3 Regression Models

We estimate our both our experimental and non-experimental estimators of the treatment effects in a regression framework. Our dependent variable is the natural log of average daily water consumption. We obtain average daily water use by dividing the water use by the number of billing days in the quarter.[22] Our primary estimating equations:

$$w_i = \alpha + \delta Letter_i + \theta \bar{w}_i + \epsilon_i \tag{11}$$

$$w_i = \alpha + \gamma \widehat{Survey}_i + \theta \bar{w}_i + \epsilon_i \tag{12}$$

[20]Within one-to-one matching there is one matched control observation for each treated observation. When multiple control observation are equally suitable matches they are averaged together. Treated observations without suitable controls are dropped, but in our setting we do not drop any treated observations.

[21]Figure A.4 shows densities of the entire control group and the matched control for summer pre-treatment consumption corresponding to the post-treatment period.

[22]Based on the meter reading schedule household have different numbers of days in each quarter, and the median number of days is 92.

Figure 4: Balance Densities

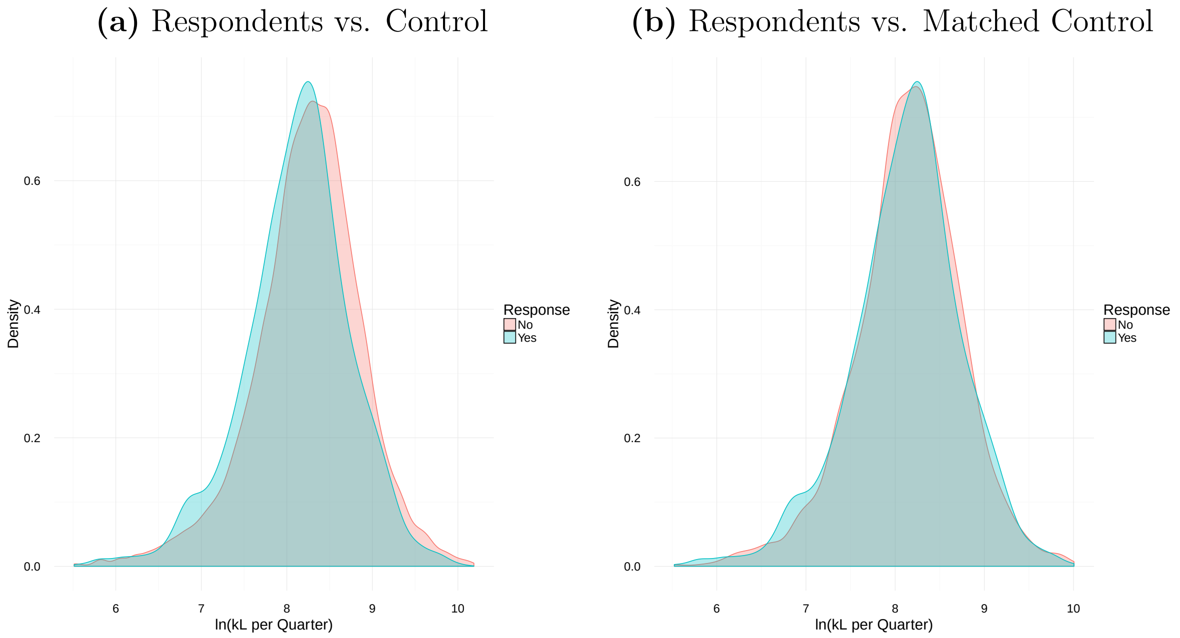


Table 5: Balance in Full and Matched Samples

(a) Full Sample

Variable	Respond	Control	Difference	KS	MW	T
Water	44.817	50.496	-5.679	0.000	0.000	0.000
Water (Q2)	50.933	55.856	-4.923	0.000	0.000	0.000
Income	83796.431	81221.752	2574.679	0.000	0.000	0.000
Education	0.045	0.052	-0.007	0.000	0.000	0.000
Owner	0.749	0.751	-0.002	0.000	0.006	0.361
SFH	0.795	0.803	-0.008	0.000	0.000	0.001

(b) Nonparametric Genetic Matching

Variable	Respond	Control	Difference	KS	MW	T
Water	44.701	45.315	-0.614	0.297	0.319	0.542
Water (Q2)	50.796	50.058	0.738	0.773	0.919	0.540
Income	83794.391	83789.299	5.093	1.000	0.985	0.993
Education	0.046	0.046	-0.000	1.000	0.960	0.908
Owner	0.749	0.749	-0.000	1.000	0.926	0.957
SFH	0.795	0.795	-0.001	1.000	0.934	0.889

Notes: The columns show the average covariate values in among respondents and either the full (a) or matched control (b), as well as the the difference in means and the p-values for t-tests (T), Mann-Whitney tests (MW) and Kolmogorov-Smirnov (KS) tests.

$$w_i = \alpha + \beta Survey_i + \theta \bar{w}_i + \epsilon_i \tag{13}$$

$$w_{it} = \alpha_i + \tilde{\beta} Post_t \times Survey_i + \phi Post_t + \tau_t + \epsilon_{it} \tag{14}$$

In these equations w is the log of daily water use use by household i and ϵ is an idiosyncratic error term. In the cross sectional models we include average water consumption prior to the intervention ($\bar{w}_i$) as a control term to improve the precision of the estimates.[23] Equation 11 is estimated on the whole sample in the period following the survey (the treatment period) and δ represents the ITT. Equation 12 is also estimated on the whole sample in the period following the survey but estimates the effect of taking the survey, which is instrumented with receiving a letter, and γ is the LATE. Equation 13 is estimated on the matched sample using the weights from the genetic matching algorithm in the period following the survey and β is the ATT. Equation 14 is estimated on the matched sample using all time periods in a DID model with time (τ_t) and individual (α_i) fixed effects and $\tilde{\beta}$ is alternative estimator of the ATT. Examining the results of equations 12, 13, and 14 allow the comparison of the ATT using a combination of experimental and non-experimental methods. We also compare our estimates of the ATT using matching weights in the regression model to conventional one-to-one matching estimates obtained from Diamond and Sekhon (2013).[24]

We investigate heterogeneity based on quartiles of pre-treatment consumption. We also estimate exploratory regressions examining heterogeneity due to survey responses on the treated sample in a fixed effect panel model. These regressions exploit within-household variation and do not have a control group, so the results should not be interpreted as causal.

5.4 How does the survey affect water consumption?

Table 6 shows estimates of the treatment effects using different estimators. The column labels describe the treatment effect estimator and the use of pre-processing matching is indicated in the bottom portion of the table. The first two columns show estimates of the ITT and LATE using experimental variation. The last three columns use non-experimental methods, showing both naive estimates and regression coupled with weights from genetic matching.

The first column shows that the ITT - the effect of receiving an invitation to the survey - increases consumption by approximately 1%. The estimate is relatively noisy and significant only at the 10% level. Next, column (2) reports the LATE estimate using 2SLS where the randomized invitation letter serves as an instrument for completing the survey.[25] The LATE estimate is roughly 10% and significant at the 10% level. The experimental results show that respondents increase water use after the survey, but the

[23] We also drop outliers above the 99th percentile of average summer consumption, which have an outsized effect on the variance of water consumption.

[24] The regression models using matching condition on pre-treatment water use to reduce the conditional variance and therefore may vary slightly from the direct matching estimates of Diamond and Sekhon (2013).

[25] The LATE model estimates the ATT in our setting.

estimates are imprecise.

Next we move to the non-experimental methods in columns (3)-(5). Column (3) estimates a naive version of the ATT to highlight the problem of selection where we regress water use on the survey indicator, dropping all non-respondents. This model compares the respondents to all households in the control group who never received an invitation. The ATT without controlling for selection is negative and statistically significant, however, this is primarily a product of selection; respondents used less water compared to non-respondents. Columns (4) and (5) show the matching estimates for the ATT in cross section and panel settings. Column (5) uses the genetic matching sample with probability weights specified by the matching algorithm. Similar to columns (1) and (2), the model in column (5) focuses on the sole post-treatment period. The results are similar in magnitude to the LATE model, but are much more precisely estimated. The standard errors in the matching method are roughly three times smaller than in the LATE model. Estimates of the ATT using the matched sample show that treatment increased water use by 8%, which is significant at the 1% level. The results are quite similar when using the matched sample to estimate a panel DID model with individual and time fixed effects. We also report the matching estimates from one-to-one matching with replacement in Table 7 - this is the same setup as Abadie and Imbens (2006) (including the calculation of standard errors) except the propensity score is generated from the genetic matching algorithm of Diamond and Sekhon (2013). The matching estimate for the ATT, shown in the first row of Table 7, is qualitatively similar is statistically significant at the 1% level, though the magnitude is slightly smaller at roughly 6%.

Table 6: Treatment Effect Regressions

	Experimental		Non-Experimental		
	(1) ITT	(2) LATE	(3) Naive ATT	(4) ATT	(5) ATT
Letter	0.0113*				
	(0.00671)				
Survey		0.0986*	-0.0383**	0.0779***	0.0662***
		(0.0585)	(0.0173)	(0.0208)	(0.0148)
Observations	27,934	27,934	16,066	2,904	78,453
Matching	None	None	None	Genetic	Genetic
Panel w/ Household FEs	No	No	No	No	Yes
Baseline Consumption	Yes	Yes	No	Yes	No

Note: The dependent variable is the log of daily water use. Columns (1)-(4) are estimated on the cross sectional sample for the billing period directly after the survey was completed. Columns (4) and (5) construct the control group using the genetic matching algorithm of Diamond and Sekhon (2013). Column (5) nests the matched sample in a panel DID model. The models reported in columns (1), (2), and (4) control for pre-treatment water consumption. Robust standard errors are reported in parentheses except for column (5) where robust standard errors are clustered at the household level.*p<0.1; **p<0.05; ***p<0.01

Table 7: Genetic Matching Estimates

Sample	Estimate	SE	p-value	N_{Treat}	$N_{Control}$ (weighted)
Matched:All	0.0610	0.0129	0.0000	1624	1624
Matched:Q1	0.0644	0.0275	0.0194	512	512
Matched:Q2	0.0680	0.0267	0.0109	439	439
Matched:Q3	0.0506	0.0185	0.0061	369	369
Matched:Q4	0.0033	0.0270	0.9022	304	304

Note: The dependent variable is the log of daily water use. The columns designate the parameters and the rows designate the matching sample. All refers to all respondents, and Q1-Q4 represent matching conducted on subsets of the sample based on quartiles of pre-treatment water use. A different matched sample is generated for each treated sample and they each use the same covariates for matching. The standard errors are based on Abadie and Imbens (2006) which account for the (asymptotic) variance induced by the matching procedure itself. *p<0.1; **p<0.05; ***p<0.01

The demand effect of our intervention generates increases in water use between 6-10%, which is larger in magnitude than other behavioral interventions in water and energy.[26] We offer several explanations for the larger treatment effects in our setting. First, our intervention provided more information than most behavioral interventions. Similar work providing an online course in electricity by Kahn and Wolak (2013) also finds treatment effects that are larger than typical behavioral interventions in energy.[27] Second, we believe sample selection affects who is "local" in a way that increases the LATE. Respondents to our survey use less water compared to non-respondents. We also find that treatment effects are larger for lower users, who are more likely to overestimate the cost of water. Therefore, our LATE is likely larger than if we were able to randomly induce households to take the survey. Lastly, the standard errors for the experimental estimates are reasonably large. The more precise matching and matching/ panel data methods estimate treatment effects of roughly 6-8%, which is larger, but not much larger than other behavioral interventions in water.

5.5 Does the survey have a differential effect on consumption for low and high users?

Next, we examine heterogeneity by pre-treatment water consumption. The experimental methods should produce unbiased heterogeneous treatment effects because we

[26]For example, social comparisons change demand by roughly 2% in energy and 3-5% (Allcott, 2011b) in water (Ferraro and Price, 2013; Brent et al., 2015).

[27]Kahn and Wolak (2013) find LATEs of -1.712 and -12.77 kWh/day respectively for two unnamed California Utilities. They do not provide baseline consumption data so it is not possible to convert these to percentage terms. However, according to an evaluation for the California Public Utility for Opower's intervention in PG&E (Comission, 2016), the average daily kWh was 0.608 for the intervention with the largest sample.

explicitly randomized the invitation letters across the distribution of baseline water consumption. Based on the conceptual framework laid out in 2, consumers with different levels of baseline water use are more or less likely to respond to different mechanisms. For example, low users are more likely to suffer have under-consumption internalities ($\gamma > 1$) and therefore should display the largest increases in consumption. Additionally, baseline water use is also a predictor of overestimating the price of water as shown in Section 4.

We augment equations 11 and 12, and 13 by interacting the treatment variables (letter or respond) with dummies for each quartile of pre-treatment consumption.[28] Table 8 shows the heterogeneous treatment effects for the ITT, the LATE, and ATT cross sectional matching models.[29] Columns (1) and (2) show that the treatment effect is concentrated among the lowest water users. The ITT for the lowest quartiles of households is 3% and the LATE is 23%. However, Wald tests fail to reject the null of equality of coefficients across quartiles. This provides suggestive evidence that low users, who are more likely to overestimate the costs of water, are the primary respondents driving the positive treatment effects.

We also run the same model using the matched sample in column (3). The heterogeneous results for the full matched sample do not follow the same pattern as the experimental results; both the lowest and highest quartiles show large and significant increases in consumption. However, column (3) does not explicitly generate matches within the pre-treatment consumption quartiles. We will refer to the pre-treatment consumption quartiles as Q1, Q2, Q3, and Q4 where Q1 is the lowest consumption quartile and Q4 is the highest. Therefore, we divide the data into four separate samples based on quartiles of pre-treatment consumption and then perform the matching procedure within each sample. We then estimate the ATT within each of the four samples noted Q1-Q4 in columns (4)-(7).[30] While there are some differences, the separately matched estimates much better replicate the pattern of heterogeneity from the LATE estimator. The first quartile (column (4)) is the largest and statistically significant, whereas the highest quartile (column (7)) is small and insignificant. The point estimates for the second and third quartiles align well with the LATE estimates, and the matching estimator generates much more precise estimates. One lesson from the analysis of heterogeneity is that even though the genetic matching algorithm balances across the full distribution of covariates it does not generate the same pattern of heterogeneity as the experimental methods. Thus, as

[28]We also perform the same exercise after dividing the sample based median pre-treatment consumption and the results are similar.

[29]These models correspond to columns (1), (2), and (4) in Table 6.

[30]The matched samples have different numbers of observations because the response rate was not homogeneous across the pre-treatment consumption distribution. Households with lower baseline water use were more likely to respond to the survey, resulting in larger samples sizes for the lower consumption quartiles.

suggested in Ho et al. (2007), it is important to generate separate matched samples for subgroups of interest.[31]

Table 8: Heterogeneity Based on Baseline Consumption

	Experimental		Matching				
	(1) ITT	(2) LATE	(3) All	(4) Q1	(5) Q2	(6) Q3	(7) Q4
Letter:Q1	0.0336**						
	(0.0155)						
Letter:Q2	0.00866						
	(0.0119)						
Letter:Q3	0.00736						
	(0.0119)						
Letter:Q4	-0.00165						
	(0.0121)						
Survey:Q1		0.230**	0.105***				
		(0.106)	(0.0395)				
Survey:Q2		0.0721	0.0408				
		(0.0991)	(0.0366)				
Survey:Q3		0.0703	0.0153				
		(0.113)	(0.0344)				
Survey:Q4		-0.0191	0.124***				
		(0.140)	(0.0405)				
Survey				0.104***	0.0678**	0.0738***	0.0273
				(0.0379)	(0.0284)	(0.0278)	(0.0311)
Observations	27,934	27,934	2,904	905	796	686	518
Wald Test p-value	0.35	0.50	0.13				
Matching	None	None	Genetic	Genetic	Genetic	Genetic	Genetic
Panel w/ Household FEs	No	No	No	No	No	No	No
Baseline Consumption	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The dependent variable is the log of daily water use. In the first two columns the letter (column 1) or survey (columns 2 & 3) variable is interacted with dummies for quartiles of pre-treatment consumption. Columns (4)-(7) estimate separate regressions based on matched samples using the genetic matching algorithm of Diamond and Sekhon (2013), where the matching takes place on subsets of the data divided by quartiles of pre-treatment consumption. All models are estimated on the cross sectional sample for the billing period directly after the survey was completed and control for pre-treatment water consumption. Robust standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

We did explore whether initial information has any differential effect on the demand response to the survey. This is challenging because we do not have data initial infor-

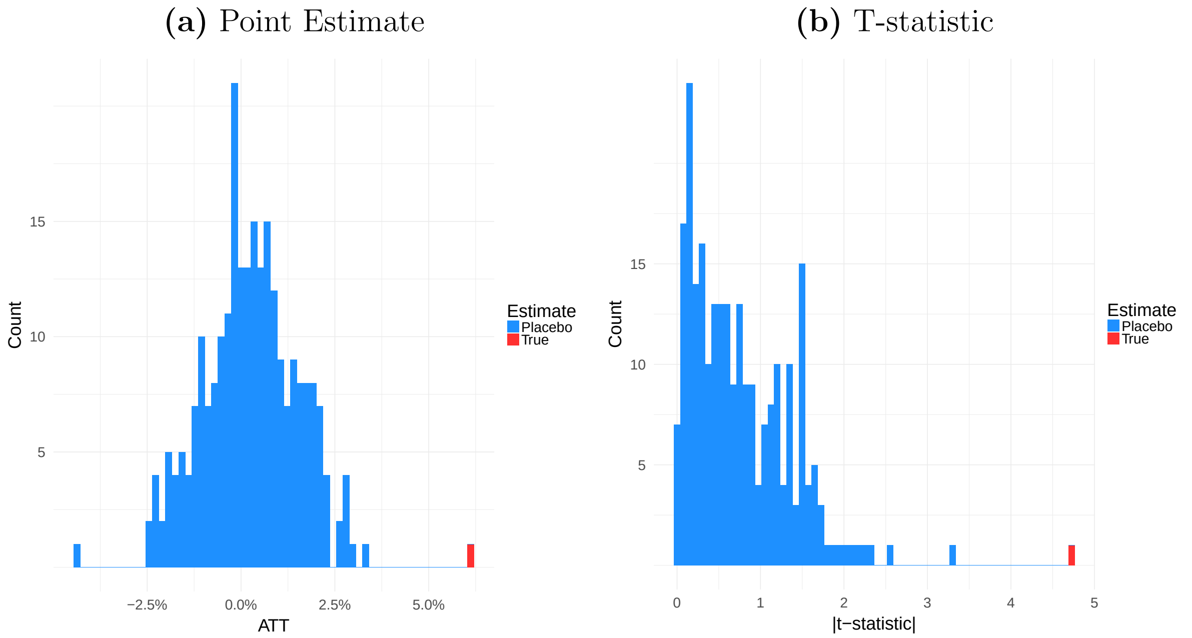
[31] The estimates of the survey on the lowest quartile are substantially different using the experimental and matching methods (columns (2) Survey:Q1 v, column (4) of Table 8). One explanation is simply that the heterogeneous LATE estimate is relatively noisy compared to the matching estimate. The second is the different information used in addressing heterogeneous selection effects. The heterogeneous LATE model uses the random invitation letter interacted with quartiles of pre-treatment consumption as instruments, and therefore assumes that a household's quartile of pre-treatment water use is the only driver of heterogeneity in the selection process for completing the survey. Conversely, the matching model in column (4) has a separate model for finding comparable control households within the first consumption quartile that uses the full distribution of consumption within this quartile as well as demographics. Therefore, the matching may generate a more flexible approach to accounting for heterogeneous selection effects.

mation for the control group. Therefore, we estimate a variety of panel data models where the sample is restricted to the survey respondents, and we estimate the change in consumption after the survey. These models cannot be interpreted as purely causal and for this reason we only briefly discuss the results although interested readers can view the results in the Appendix. There are not strong patterns of heterogeneity based on the survey data including whether respondents overestimated the cost of water. Similarly, other features of the survey such as whether the respondent indicated the information useful or surprising, and the confidence in the estimates explain heterogeneity in the demand response. Again, we caution the lack of any effects as definitive due the analysis investigating correlations as opposed to causal effects.

5.6 Robustness

In order to test the robustness of our matching estimates we generate a falsification test using randomly selected non-respondents as a placebo treatment group. First, we draw a binomial random variable for each of the non-respondents where the probability of assignment to the placebo treatment is equal to the actual response rate. After generating the placebo treatment sample from the set of non-respondents we generate a matched control sample using the same genetic matching model used to construct the actual matched sample. Next, we use the matching estimator to estimate an ATT and save the results. We repeat the this process 250 times, producing 250 placebo estimates of the ATT, which should be equal to zero since the placebo treatment group did not actually complete the survey. Our primary results are presented in Figure 5. Panel (a) shows the distribution of placebo point estimates for the ATT, which is centered at zero and our true estimate is at the far right tail. The mean of the distribution is 0.2%; more than 23 times smaller than our actual estimate. Panel (b) plots the distribution of the absolute values of the t-statistics. None of the 250 placebo samples generates an estimate of the ATT with a t-statistic as large as our actual estimate. Only 3% of the matched samples produced t-statistics for the ATT above an absolute value of 1.96; which is even lower than would be expected by pure chance. The falsification test strengthens the validity of our matching estimate and initial randomization by producing a null result for consumers that did not complete the survey.

Figure 5: Placebo Tests



Note: The figure shows the results of 250 placebo tests along with our preferred point estimate of the ATT using the genetic matching algorithm to select a matched sample from the control group for random subset of non-respondents. Panel (a) shows the matching estimates of the ATT and panel (b) shows the absolute value of t-statistics constructed from Abadie and Imbens (2006) standard errors. The preferred estimate using the the true set of respondents is also shown for comparison.

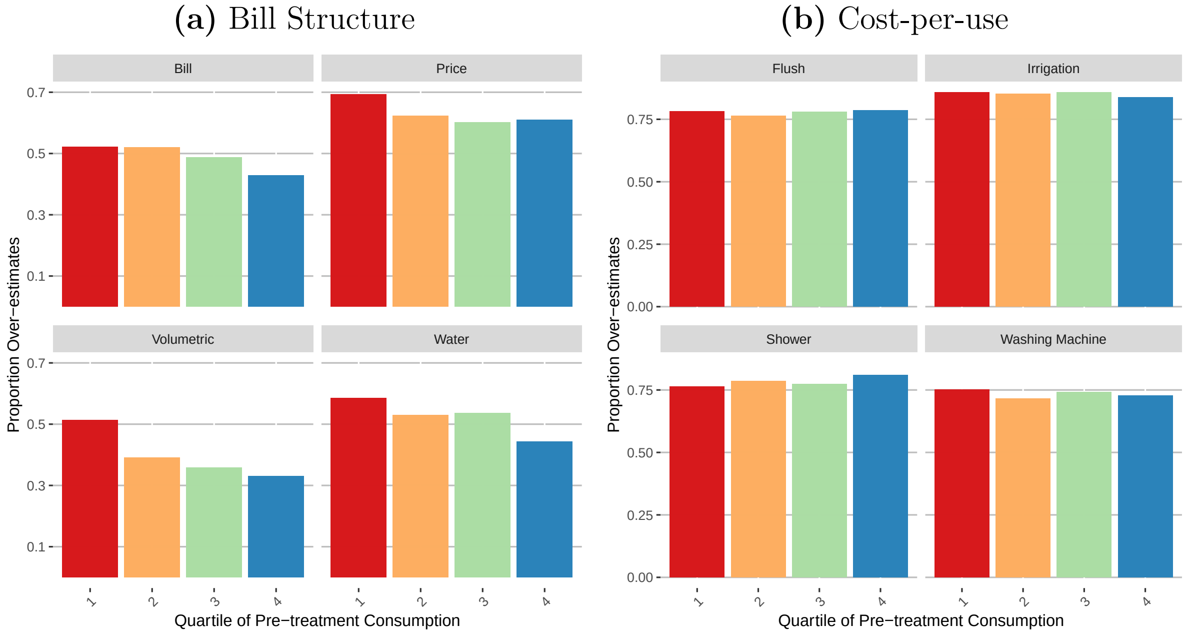
6 Interpreting Results within Conceptual Framework

In order to help interpret the empirical results, we consider how the empirical evidence fits into our conceptual framework laid out in Section 2. We will primarily rely on the base effects and the heterogeneous effects based on pre-treatment water use, combined with the summary statistics from the survey.[32] To assist in interpreting the heterogeneous results based on quartiles of baseline baseline water use with the summary statistics on cost perceptions, we graph the percentage of respondents that overestimated each variable by quartile of baseline water use in Figure 6. The graph shows that low users are more likely to overestimate most bill components relative to high users. There is little variation in overestimating the CPU questions across the consumption distribution.

Case 1: The survey increases salience of the moral cost of water.
Most nudges in water demand are designed to reduce water consumption, The most popular nudge, social comparisons, is successful at reducing consumption by raising moral costs and reducing consumption (Ferraro and Price, 2013; Brent et al., 2015, 2017; Brent and Wichman, 2018). Most interventions do not raise consumption, however Byrne et al. (2018) elicited both estimates of electricity consumption and provided peer comparisons.

[32]In addition to not generating causal estimates, the heterogeneity based on survey pattern did not produce clear intuitive patters.

Figure 6: Cost Perceptions by Quartile of Pre-treatment Consumption



Note: The figure shows the percentage of respondents that overestimated the variable in question by quartile of baseline water use.

Households who overestimated their consumption used more electricity and households who underestimated their consumption decreased electricity use. Our results are somewhat consistent with Byrne et al. (2018) since low users who are likely to overestimate consumption increase their water use after the survey. However, unlike Byrne et al. (2018) and other social comparison studies, we do not observe any decreases in consumption even among high users. Additionally, we do not provide any peer information that would highlight moral costs of consumption. Therefore, we think that it is plausible, but unlikely, that the primary mechanism is increased salience of moral costs.

Case 2: The survey corrects price mis-perceptions.
Another potential mechanism is that consumers were correctly optimizing with respect to the wrong price. This explanation fits the empirical results because on average households thought the price of water was higher than the true price, and the lowest quartile were most likely to over-estimate the price. Therefore, it is reasonable that consumers updated their price perceptions downward and reacted to the survey as a price decrease.

Case 3: The survey corrects internalities not related to the marginal price of water.
It is possible that original consumption is sub-optimally low or high prior to the survey, and the survey corrects these internalities. For example, households may overestimate the water necessary to take a shower, or they have challenges understanding the components of their rate structure. This is plausible given that households widely overestimate the CPU of all activities. Therefore, it is plausible that the survey helps households re-

optimize water consumption. If this is the case, the respondents were initially under-consuming water relative to the optimal amount. The CPU results do not explain the heterogeneous treatment effects, although alternative information such as over-estimating the percentage of the bill due to volumetric charges maps to the heterogeneous treatment effects.

Lastly, within our conceptual model internalities include inattention to prices, which has been shown to have affects on electricity consumption (Gilbert and Zivin, 2014; Sexton, 2015). However, as shown in Sexton (2015) pure inattention to water prices should result in an initial over-consumption ($\gamma < 1$), and correcting inattention should decrease water consumption. Pure price inattention is inconsistent with our results where households increase consumption after the survey. Additionally, we test for a pure salience mechanism by interacting the treatment variables (letter and survey) with a variable for the date sent, which is a strong predictor for the days between the last bill received and the completion of the survey. The results, reported in Table A.7, show the interaction term is small and insignificant and changes sign between the experimental and matching models.

Approximate Welfare Calculations

Recent research by Allcott and Kessler (2019) shows that welfare effects from behavioral nudges depend on the behavioral mechanisms through which they operate. As shown above the information contained in the survey may operate through several behavioral mechanisms, and our design does not let us completely rule out any of the three primary mechanisms. We believe the pattern of heterogeneity is most consistent with updating prices mis-perceptions or correcting some other type of internality. Importantly, both of these mechanisms should improve welfare by shifting consumption towards the optimal value ($\tilde{w} \Rightarrow w^*$). We generate an approximation of the welfare change by employing the method from Wichman (2017) to estimate welfare effects from summary statistics on the impact of improved information on water demand. We parameterize the consumption benefit of water as $f(w, \alpha) = \alpha W^{1/\sigma+1}$. We use Case 2 as a motivating example where the change in consumption is completely due to changes in price perceptions. We model perceived price as, $\tilde{p} = \theta p$ where θ is the degree of mis-perception. If consumers perfectly know water prices $\theta = 1$ and consumers can overestimate ($\theta > 1$) or underestimate ($\theta < 1$) the true price. If we assume that (1) improved price information cannot harm consumers and (2) the survey moves price perceptions move closer to the true price, then the change in consumer surplus can be calculated by integrating the demand function from the initial perceived price, $\tilde{p_0}$ to the new perceived price, $\tilde{p_1}$.[33]

$$\Delta CS = \int_{\tilde{p_1}}^{\tilde{p_0}} w(\tilde{p})d\tilde{p} \cong -\frac{1}{2}\Delta\tilde{p}\frac{\partial W}{\partial Survey} \tag{15}$$

Wichman (2017) use elasticities from the literature to back out the perceived price change, but since we elicit perceptions of price and update consumers on the true price we can calculate the sample average change in price perceptions. We use the matching estimate of 7.79% scaled by average consumption of 53 kL to estimate $\frac{\partial W}{\partial Survey} = 4.13$ kL, and our sample average of the difference between marginal price perceptions and true marginal prices to estimate $\Delta\tilde{p} = -\$1.97$. Using these estimates we calculate that the survey increased consumer surplus by \$4.07 per quarter or roughly 1.1% of the quarterly bill. This is on the high end of the percentage change in consumer surplus due to increased billing frequency estimated from Wichman (2017).

Our estimates of $\frac{\partial W}{\partial Survey}$ and $\Delta\tilde{p}$ correspond to a demand elasticity of -0.13, which is on the low end of existing estimates of demand elasticity. Since we also provide multiple types of information on the cost of water using water the change in marginal prices may not fully capture the change in consumers' price perceptions. Therefore, we also infer$\Delta\tilde{p}$ from $\frac{\partial W}{\partial Survey}$ and common demand elasticity parameters. Using elasticity of -0.2, -0.3, and -0.4 generates estimates of increases in consumer surplus of 0.6%, 0.4%, and 0.3% respectively. Since we also inform consumers about their total bill and water consumption, the treatment effect could be due to a change in quantity perceptions, but Wichman (2017) show that the summary statistics for the welfare estimates are also sufficient for a model of quantity mis-perceptions. As stated above the true mechanisms causing a change in consumption may be some combination of increased salience, correcting non-price internalities, and price mis-perceptions. However, using price perceptions likely provides a rough estimate of the welfare effects, and our back of the envelope calculations show that the welfare effects are quite small.

7 Conclusion

Economists have long argued for using prices to manage scarce water resources. Due to political pressures many water utilities have adopted complicated, non-linear, two-part pricing structures that attempt to charge low prices for water used to meet basic human needs such as drinking and sanitation while charging higher prices for discretionary uses such as irrigation. The proliferation of these rates structures has made water pricing

[33]Since we do not elicit a full demand function from the survey, the welfare generated from the treatment effect estimates ($\frac{\partial W}{\partial Survey}$) and perceived price change ($\Delta\tilde{p}$) are approximate.

more complicated and at times obfuscated the price signal.

We design a survey nested within a randomized field experiment to learn both consumers' baseline understanding of the cost of water and how improved information on the costs of water changes their consumption decisions. The results of our survey indicate that consumers have reasonable estimates of their total bill, but have very poor information on specific features of their bill such as the marginal price. Consumers also have very little knowledge of the cost of water-using activities such as flushing a toilet; not knowing the marginal costs of water-using activities inhibits consumer optimization in the municipal water sector. On average, consumers overestimate the cost of water, and in aggregate, learning the true cost of water increases consumption. The increase in consumption could be due to a combination of increased moral salience, correcting internalities, or updating perceived prices. We argue that the empirical results are most consistent with consumers updating price mis-perceptions or correcting some other form of internality. A stylized model that assumes the demand response is completely due to price mis-perceptions shows that improved price information increases consumer surplus by approximately 1%.

The findings are consistent with studies that show consumers do not respond to the marginal price of water, since they do not even know the marginal price they face (Ito, 2014; Wichman, 2014). However, consumers do not actually know how much water they use, suggesting they may not actually respond to average prices either. Rather, consumer behavior may lie outside the standard optimization framework such as mental accounting where consumers respond only when their total bill moves outside of some predefined range (Thaler, 1985). This is consistent with consumers knowing their total bill but not the marginal price of water, and that learning that water is cheaper than they anticipated affects subsequent water use. The research has implications for water rate design. Most discussion of designing water rates has focused on tradeoffs between equity and efficiency while generating enough revenue to cover costs. If the mental accounting model is correct, municipal water demand utilities may be able to raise the fixed cost and generate reductions in consumption. This is important because higher fixed costs will reduce the variation in revenue and bring revenue generation closer in line with costs, which in the water sector are primarily fixed. Additionally, simplifying bill structures will help send simple price signals that consumers can understand. This research suggests that utilities should consider the way that consumers perceive their water bills during rate design.

Another implication of the research is the importance of documenting the sources and direction of behavioral biases. In our setting consumers are likely consuming less than the private optimum. Conservation policies attempting to reduce demand may still be justified from a social welfare perspective if consumption externalities are sufficiently high. However, these policies will move consumption further away from the private optimum and generate lower welfare benefits relative to a setting without internalities. This is in contrast to proponents of the energy efficiency gap where conservation and/or behavioral policies are argued to improve welfare by reducing both internalities and externalities. It is worthwhile to test assumptions about the direction of internalities when designing behavioral policies.

While we attempt to disentangle competing mechanisms through which the survey affects water demand, we are not able to definitely isolate specific mechanisms. Future research can improve on identifying the specific causal mechanisms by explicitly randomizing the type of information that consumers receive. Examining the persistence of the results can provide additional insights on potential mechanisms.

References

Abadie, Alberto and Guido W Imbens, "Large sample properties of matching estimators for average treatment effects," *Econometrica*, 2006, *74* (1), 235–267.

Allcott, Hunt, "Consumers' Perceptions and Misperceptions of Energy Costs," *American Economic Review*, 2011, *101* (3), 98–104.

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

_ **and Cass R Sunstein**, "Regulating internalities," *Journal of Policy Analysis and Management*, 2015, *34* (3), 698–705.

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

_ **and Michael Greenstone**, "Is There an Energy Efficiency Gap?," *Journal of Economic Perspectives*, 2012, *26* (1), 3–28.

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

_ **, Sendhil Mullainathan, and Dmitry Taubinsky**, "Energy policy with externalities and internalities," *Journal of Public Economics*, apr 2014, *112*, 72–88.

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

Bolsen, Toby, Paul J. Ferraro, and Juan Jose Miranda, "Are Voters More Likely to Contribute to Other Public Goods? Evidence from a Large-Scale Randomized Policy Experiment," *American Journal of Political Science*, January 2014, *58* (1), 17–30.

Brent, Daniel A. and Casey J. Wichman, "Do behavioral nudges interact with prevailing economic incentives? Pairing experimental and quasi-experimental evidence from water consumption," 2018.

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

_ **, Corey Lott, Michael Taylor, Joseph Cook, Kim Rollins, and Shawn Stoddard**, "Are Normative Appeals Moral Taxes? Evidence from a Field Experiment on Water Conservation," 2017.

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

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.

Comission, California Public Utilities, "Review and Validation of 2014 Pacific Gas and Electric Home Energy Reports Program Impacts (Final Report)," Technical Report, California Public Utilities Commission 2016.

Diamond, Alexis and Jasjeet S Sekhon, "Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies," *Review of Economics and Statistics*, 2013, *95* (3), 932–945.

Ferraro, Paul J. and Juan José Miranda, "Heterogeneous Treatment Effects and Mechanisms in Information-Based Environmental Policies: Evidence from a Large-Scale Field Experiment," *Resource and Energy Economics*, April 2013, *Forthcomin.*

Ferraro, Paul J and Juan José Miranda, "Panel data designs and estimators as substitutes for randomized controlled trials in the evaluation of public programs," *Journal of the Association of Environmental and Resource Economists*, 2017, *4* (1), 281–317.

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

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

Gan, Kein and Michael Redhead, "Melbourne residential water use studies," *Smart Water Fund*, 2013.

Gerarden, By Todd, Richard G Newell, and Robert N Stavins, "Deconstructing the Energy-Efficiency Gap : Conceptual Frameworks and Evidence," *American Economic Review*, 2015, *105* (5), 183–186.

Gilbert, Ben and Joshua Graff Zivin, "Dynamic Salience with Intermittent Billing: Evidence fromSmart Electricity Meters," *Journal of Economic Behavior & Organization*, March 2014, pp. 1–15.

Grubb, Michael D, "Consumer inattention and bill-shock regulation," *The Review of Economic Studies*, 2015, *82* (1), 219.

_ and Matthew Osborne, "Biased Beliefs, Learning, and Bill Shock," *The American Economic Review*, 2014, *105* (1), 234–271.

Ho, Daniel E, Kosuke Imai, Gary King, and Elizabeth A Stuart, "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference," *Political analysis*, 2007, *15* (3), 199–236.

Houde, Sébastien, "How consumers respond to product certification and the value of energy information," *The RAND Journal of Economics*, 2018, *49* (2), 453–477.

Imbens, Guido W and Joshua D Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, pp. 467–475.

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

Jaffe, Adam B. and Robert N. Stavins, "The energy-efficiency gap What does it mean?," *Energy Policy*, 1994, *22* (10), 804–810.

Jessoe, Katrina and David Rapson, "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use," *American Economic Review*, April 2014, *104* (4), 1417–1438.

—, Gabriel E Lade, Frank Loge, and Edward Spang, "Spillovers from Behavioral Interventions: Experimental Evidence from Water and Energy Use," 2018.

Kahn, Matthew E. and Frank A Wolak, "Using Information to Improve the Effectiveness of Nonlinear Pricing: Evidence from a Field Experiment," 2013.

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

McRae, Shaun and Robyn Meeks, "Price perception and electricity demand with nonlinear tariffs," 2016.

Monteiro, Henrique, Rita Martins, Esmeralda A. Ramalho, and Joaquim J.S. Ramalho, "Are ill-informed residential water consumers less price-responsive," 2018.

Palmer, Karen and Margaret Walls, "Limited Attention and the Residential Energy Efficiency Gap [†]," *American Economic Review*, 2015, *105* (5), 192–195.

Pellerano, José A, Michael K Price, Steven L Puller, and Gonzalo E Sánchez, "Price salience and social comparisons as policy instruments: Evidence from a field experiment in energy usage," 2015.

Rubin, Donald B, "Estimating causal effects of treatments in randomized and nonrandomized studies.," *Journal of educational Psychology*, 1974, *66* (5), 688.

Sallee, James M, "Rational inattention and energy efficiency," *The Journal of Law and Economics*, 2014, *57* (3), 781–820.

Sekhon, Jasjeet S, "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R," *Journal of Statistical Software*, 2011, *42* (i07).

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

Shin, Jeong-Shik, "Perception of Price When Price Information is Costly: Evidence from Residential Electricity Demand," *The Review of Economics and Statistics*, 1985, *67* (4), 591–598.

Stojanovski, Ognen, Gordon Leslie, Frank Wolak, Juan Enrique Huerta Wong, and Mark C Thurber, "Promoting energy efficiency in emerging economies through consumer education: Results from a field experiment in Mexico," 2018.

Thaler, Richard, "Mental accounting and consumer choice," *Marketing science*, 1985, *4* (3), 199–214.

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

Wichman, Casey J, "Information provision and consumer behavior: A natural experiment in billing frequency," *Journal of Public Economics*, 2017.

_ **and Paul J Ferraro**, "A cautionary tale on using panel data estimators to measure program impacts," *Economics Letters*, 2017, *151*, 82–90.

Online Appendix

A Additional Tables, Figures, and Survey Questions

Figure A.1: Bill Information

(a) Total Charges

ACCOUNT SUMMARY

[redacted]

Product/Service	Amount
Usage Charges	$107.29
Water Supply System Charge	$44.34
Sewerage System Charge	$86.68
Yarra Valley Water Total	**$238.31**
Other Authority Charges	
Waterways and Drainage Charge on behalf of Melbourne Water	$22.46
TOTAL (GST does not apply)	**$260.77**

See reverse for details

(b) Price and Quantity

ACCOUNT DETAILS

Water Usage from 07/08/2013 to 08/11/2013.

Meter Number	Current Reading		Last Reading		Usage
[redacted]	4,342kL	-	4,317kL	=	**25kL**

In 93 days you used 25 kilolitres, equalling 269 litres per day.
One kilolitre (kL) equals 1,000 litres.

Usage*			Price $/kL		Amount
STEP 1	25.000	x	2.5970	=	**$64.93**

*Rising step tariffs (formerly known as block) are adjusted according to the days in your meter reading period, and applied on a daily basis.

Sewage Disposal from 07/08/2013 to 08/11/2013.
For the disposal and treatment of sewage from your property. It is based on your water usage and adjusted for seasonal variations.

Usage		Seasonal Factor		Seasonal Volume		Discharge Factor		Sewage Volume
25.000kL	x	0.9005	=	22.513	x	0.900	=	20.261kL

Sewage Volume		Price $/kL		Amount
20.261	x	2.0908	=	$42.36

Total Usage Charges	$107.29

The Water Supply System Charge from 01 Oct 13 to 31 Dec 13 is a fixed charge of $44.34 per property based on a daily rate.

The Sewerage System Charge from 01 Oct 13 to 31 Dec 13 is a fixed charge of $86.68 per property based on a daily rate.

Waterways and Drainage Charge from 01 Oct 13 to 31 Dec 13.
This charge, based on a daily rate, is collected on behalf of Melbourne Water and used to manage and improve waterways, drainage and flood protection. From 1 July 2012 the charge will be billed at a flat rate for each separate occupancy on a rateable property or for vacant land. For information about the charge and projects it funds in your region, visit melbournewater.com.au/yarravalleywater.

***NAV =** Net Annual Value of your property which is capped at 1990 levels.

Notes: These are extracts from the customer water bill for Yarra Valley Water. Panel (a) shows the total cost information and panel (b) shows usage and price information.

Figure A.2: Explaining Features of the Bill

Residential Water Use

Most customers receive a water bill every quarter (3 months). Your water bill contains several charges:

-Fixed water service charge
-Fixed sewage charge
-Water usage charge
-Sewage usage charge
-Fixed waterways charge (from Melbourne Water)

The fixed charges do not depend on how much water and sewage you use, whilst the usage charges vary with the amount of water and estimated sewage use. Water and sewage use for the usage charges are measured in kilolitres (1 kL = 1000 litres).

We will not be considering the $23 quarterly waterways charge from Melbourne Water in this survey.

Now we will ask some questions about your water bill. We will focus on your upcoming bill that generally comes in March or April, based on your water use in the first quarter of the calendar year.

Please answer the questions to the best of your ability without looking at your water bill or other material online!

Back Reset Next

Powered by snap surveys

Figure A.3: Providing Correct Information - Water Price

Residential Water Use

Please hit the "Next" arrow when you finish reading the information and answer whether the information is surprising and useful.

Water is charged per kL (1000 litres), how much do you think it costs your household to purchase one additional kL?

Your answer: **$6**

Correct answer: **$2.55**

Q41 Is this new information surprising? Is it useful to you? (select all that apply)

☐ Useful
☐ Surprising

Back Reset Next

Powered by snap surveys

List 1: Questions Eliciting Consumer Perceptions

1. What is your average historical water use in this quarter in kL (1 kL = 1000 litres)?
2. What is your total bill in this quarter based on current rates and your average historical water use? (in dollars)
   Your average historical water use in this quarter is: XXX kL (1 kL = 1000 litres)
   Your total bill based on current rates and Water kL is: XXX
3. How much of your combined bill is determined by the the usage charge (water + sewage)? (from 0 to 100%)
4. Water usage is charged per kL (1000 litres); how much do you think it costs your household to increase water use by one additional kL (water usage charge)?
5. Sewage usage is charged per kL (1000 litres); how much do you think it costs your household to dispose of one additional kL of sewage?
6. Because some water is used outside sewage usage is estimated as a percentage of the water you use. What do you think the sewage use fraction is for this season? (from 0 to 100%)
   For the rest of the questions consider your net bill impact of using an additional unit of water. This includes both water and sewage usage charges and takes into account the estimated sewage use fraction.
7. What is your best estimate of the net bill impact of irrigating an average garden for 20 minutes? Please provide your answer in standard currency format (dollars then decimal point then cents).
8. What is your best estimate of the net bill impact of flushing a toilet?
9. What is your best estimate of the net bill impacts of doing a load of laundry?
10. What is your best estimate of the net bill impact of taking an average (7 minute) shower?

At the end of the survey we provided the following information based on questions 3-10:
Your answer: XXX
Correct answer: YYY
Then we asked if the new information was useful and surprising. For an example see Figure A.3

Figure A.4: Balance Densities for Summer Consumption

(a) Respondents vs. Control **(b)** Respondents vs. Matched Control

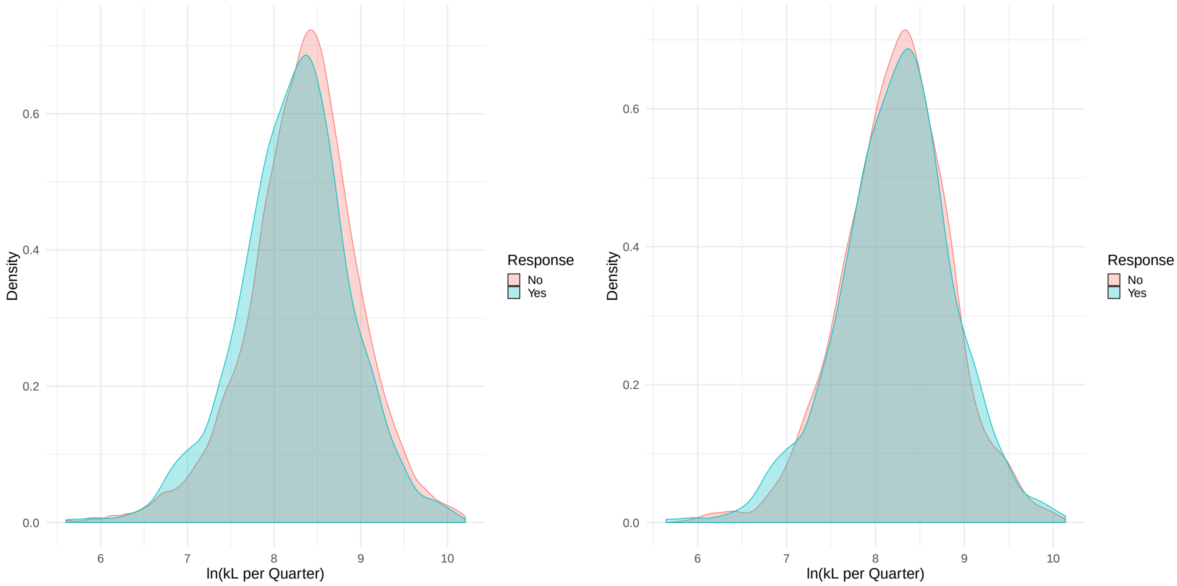


Table A.1: Determinants of Price Information: Bill Structure

	Bill	Water	Water Price	Sewer Price	Volumetric
	(1)	(2)	(3)	(4)	(5)
Water	0.018***	0.017	0.054***	0.055***	0.055***
	(0.007)	(0.015)	(0.015)	(0.015)	(0.009)
Income	−0.021	−0.002	0.020	0.015	0.001
	(0.018)	(0.040)	(0.039)	(0.039)	(0.024)
Degree	0.010	0.007	−0.001	0.002	0.021
	(0.015)	(0.033)	(0.031)	(0.031)	(0.020)
HH Size	−0.012**	0.017	0.012	0.010	0.023***
	(0.006)	(0.013)	(0.013)	(0.013)	(0.008)
65+	0.011	−0.008	0.108***	0.100***	0.00000
	(0.016)	(0.035)	(0.034)	(0.034)	(0.022)
Respond to Prices	0.008	−0.010	0.018	0.021	0.002
	(0.020)	(0.045)	(0.043)	(0.043)	(0.027)
Motivated by Money	−0.018	−0.037	0.014	0.014	−0.033*
	(0.015)	(0.032)	(0.031)	(0.031)	(0.020)
Investments	0.001	−0.014	0.003	0.002	−0.016*
	(0.007)	(0.015)	(0.015)	(0.015)	(0.009)
Confidence	0.013	0.106**	0.346***	0.447***	−0.158***
	(0.017)	(0.042)	(0.061)	(0.076)	(0.037)
Constant	0.982***	0.525***	0.304***	0.319***	0.832***
	(0.032)	(0.070)	(0.067)	(0.067)	(0.043)
Observations	1,226	1,172	1,198	1,198	1,229
Adjusted R^2	0.006	0.003	0.041	0.043	0.064

Note: The dependent variable is a dummy equal to one if a respondent was within 50% of the true estimate and zero otherwise. *p<0.1; **p<0.05; ***p<0.01

Table A.2: Determinants of Price Information: Cost-per-Use

	Irrigation	Washing Machine	Shower	Toilet
	(1)	(2)	(3)	(4)
Water	-0.021^{*}	0.003	0.010	-0.0004
	(0.012)	(0.014)	(0.013)	(0.009)
Income	0.001	0.026	-0.005	0.020
	(0.030)	(0.037)	(0.033)	(0.025)
Degree	0.034	0.027	0.041	0.018
	(0.025)	(0.030)	(0.027)	(0.020)
HH Size	0.007	0.015	0.013	0.012
	(0.010)	(0.012)	(0.011)	(0.008)
65+	-0.020	-0.024	0.038	-0.007
	(0.027)	(0.033)	(0.029)	(0.022)
Respond to Prices	0.010	0.066	0.003	0.058^{**}
	(0.034)	(0.041)	(0.037)	(0.028)
Motivated by Money	-0.034	-0.112^{***}	-0.045^{*}	-0.020
	(0.025)	(0.030)	(0.027)	(0.020)
Investments	0.013	-0.019	-0.011	0.011
	(0.012)	(0.014)	(0.013)	(0.009)
Confidence	0.020	0.065	0.589^{***}	0.359^{***}
	(0.047)	(0.105)	(0.109)	(0.086)
Constant	0.127^{**}	0.276^{***}	0.171^{***}	0.008
	(0.053)	(0.065)	(0.058)	(0.043)
Observations	1,192	1,199	1,202	1,187
Adjusted R^2	0.002	0.028	0.025	0.018

Note: The dependent variable is a dummy equal to one if a respondent was within 50% of the true estimate and zero otherwise. *p<0.1; **p<0.05; ***p<0.01

B Does initial information have a differential effect on consumption?

In an attempt to further disentangle the mechanism through which the survey increases water consumption, we estimate heterogeneity based on the answers to the survey. Consumers have heterogeneous baseline information about the cost of water so the survey provides differential information. Some consumers will learn that water is cheaper than they thought, while others will learn that water is more expensive than they thought. A challenge is that, unlike baseline water consumption, the survey responses are not exogenous. Additionally, we do not have survey data for households in the control group; ideally we would want to know how control households would have responded had they taken the survey. Therefore, we analyze heterogeneous responses due to information contained in the survey using a fixed effects panel data model using only the survey respondents. This model utilizes the within-household variation in consumption before and after the survey to estimate heterogeneous treatment effects. While we acknowledge that these results cannot be interpreted as causal we believe they provide additional insights into how consumers respond to different types of price information.[34]

We construct standardized measures of price information from each of the questions that elicit cost perceptions and then provide the true response. First, we take the difference between the respondent's estimate and the true value. Then we subtract the mean from this variable and divide by the standard deviation. A unit change in the variable represents a one standard deviation change in the degree of the respondent's estimation error. Positive values indicate that the respondent overestimated the relevant variable and negative values represent underestimates relative to the average respondent. Consider, for example, the variable constructed for the estimates of the total bill. A value of zero represents the average difference between the estimated bill and the true value and a value of one represents a respondent who overestimated their bill by one standard deviation above the sample average.

The results are shown in Table A.3. Column (1) in Table A.3 shows the panel estimates of the ATT for reference, and column (2) shows the results of a model that adds the standardized errors for all the questions. Most of the estimates are small and not statistically significant at conventional levels and the base effect of taking the survey barely changes. The two questions that generate statistically significant results are water consumption and irrigation CPU. Overestimates of water consumption and irrigation CPU are positive and roughly 2%. The interpretation is that respondents that overestimated

[34]Our CPU estimates also have measurement error, which will add bias into regressions that account for the accuracy of the CPU questions.

water use or irrigation costs by one standard deviation more than the average respondent experienced an even larger increase in consumption. This is consistent with consumers learning that they are more efficient in their water use than they anticipated and then decide to increase their consumption.

Since the answers to the questions are highly correlated we also use cluster analysis and principal components to determine if there are classes of respondents that have differential treatment effects. We use kmeans clustering to estimate two groups; one group overestimated their costs and one group underestimated their costs. The summary statistics for the two clusters are shown in Table A.4. Column (3) interacts the post survey indicator with a dummy for the cluster that overestimated water costs. The interaction term is very small and insignificant. Column (4) interacts the three predicted scores from the principal component analysis with the post survey indicator. Two of the three scores are insignificant and the second principal component is positive and significant at the 10% level. The principal component analysis, presented in Table A.5, shows the the second component is related to overestimates of water use, the total bill, the % volumetric and the STC, as well as underestimates of the all CPU variables. There are some patterns that suggest respondents that overestimate total water use increase their use by more, although we caution the interpretation of these results as causal.

Lastly, we analyze whether consumers' confidence, perceptions of the usefulness and novelty of the information, and stated motivations impact the demand response. Similar to the results using cost perceptions these regressions are estimated only on the sample who participated in the survey, and therefore we caution any causal interpretation of the results. The results are reported in Table A.6. Similar to the other survey data, there are not clear pattersns of heterogeneity based on these survey answers.

Table A.3: **Heterogeneity Based on Cost Perceptions**

	(1) Base	(2) Standardized Errors	(3) Clusters	(4) Principal Components
Post	0.127***	0.127***	0.167***	0.127***
	(0.0135)	(0.0135)	(0.0127)	(0.0135)
Water*Post		0.0235**		
		(0.00962)		
Bill*Post		0.000986		
		(0.00931)		
Volumetric*Post		-0.00987		
		(0.0101)		
Water Price*Post		-0.00771		
		(0.0151)		
Sewer Price*Post		-0.000336		
		(0.0134)		
STC*Post		0.0120		
		(0.0103)		
Irrigation*Post		0.0220*		
		(0.0119)		
Flush*Post		-0.0200		
		(0.0131)		
Washer*Post		0.00607		
		(0.0157)		
Shower*Post		-0.0113		
		(0.0173)		
Overestimates Cluster*Post			0.000595	
			(0.0197)	
Principal Component 1*Post				0.000140
				(0.00484)
Principal Component 2*Post				0.0122*
				(0.00688)
Principal Component 3*Post				0.000512
				(0.0104)
Observations	40,151	40,151	40,151	40,151
Household and Time FEs	Yes	Yes	Yes	Yes

Note: The dependent variable is the log of daily water use. The estimation sample is a panel DID restricted to the survey respondents. The variables Water*Post - Shower*Post in column (2) are standardized errors for the respective question multiplied by a post-survey indicator. The Overestimates Cluster*Post in column (3) is a dummy for the cluster that overestimated water costs multiplied by a post-survey indicator. Column (3) shows the three predicted scores of the principal components analyss multiplied by a post-survey indicator. Summary statistics of the clusters are presented in Table A.4 and the results of the principal component analysis are presented in Table A.5. Robust standard errors clustered at the household level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A.4: Summary Statistics of Clusters

	Cluster 1			Cluster 2		
	Observations	Mean	Std. Dev	Observations	Mean	Std. Dev
Water	635	0.26	1.14	914	-0.18	0.84
Bill	635	0.14	1.36	914	-0.10	0.62
Volumetric	635	0.19	1.44	914	-0.13	0.47
Water Price	635	0.67	1.12	914	-0.47	0.54
Sewer Price	635	0.66	0.97	914	-0.46	0.73
STC	635	0.19	1.35	914	-0.13	0.63
Irrigation	635	0.74	0.66	914	-0.52	0.86
Flush	635	0.73	0.42	914	-0.51	0.98
Washer	635	0.88	0.94	914	-0.61	0.41
Shower	635	0.90	0.72	914	-0.63	0.61

Note: The summary statistics are based on the two clusters identified via k-means clustering with Euclidean distance. Each of the variables is the percentage difference from the accurate answer for the respective question.

Figure A.5: Principal Component Scree Plot

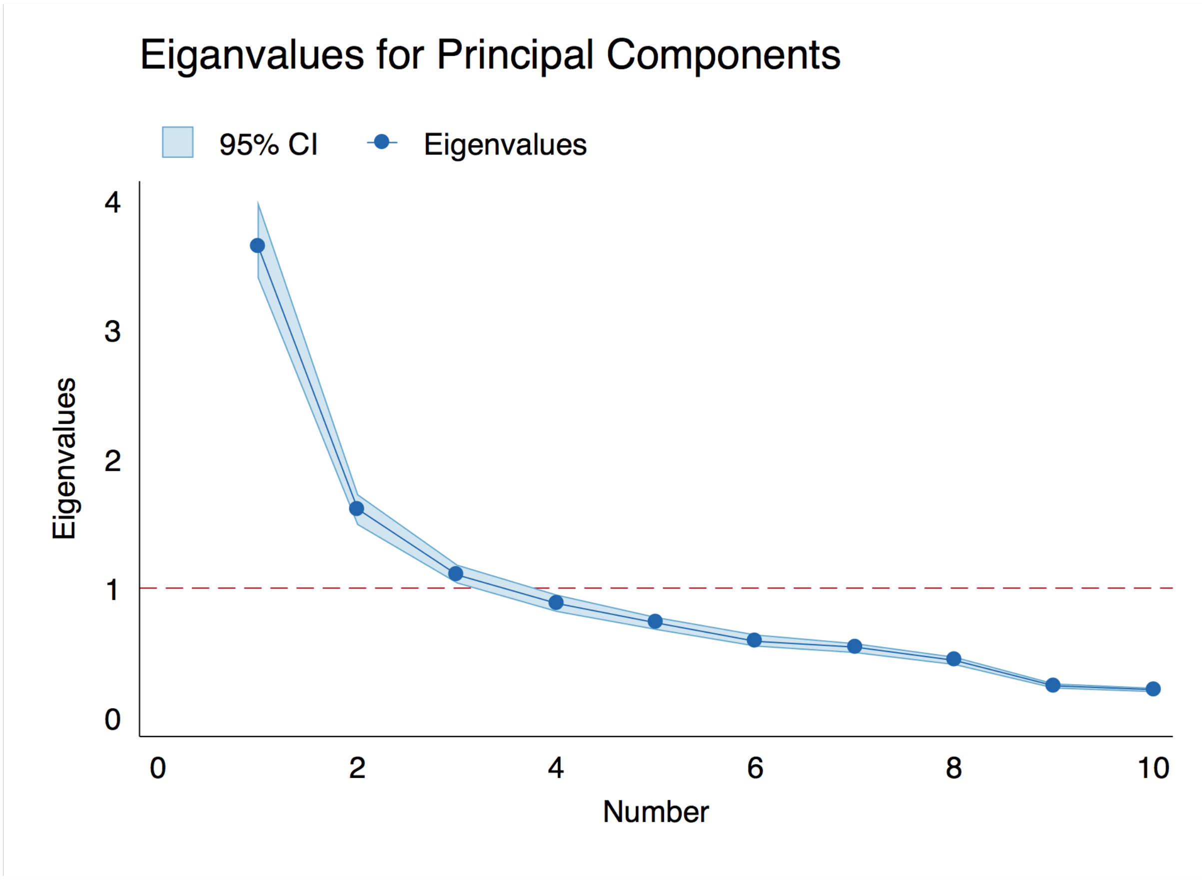


Table A.5: Principal Components

	Eigenvalues	**Component 1**	**Component 2**	**Component 3**
Component 1	3.643***			
	(0.131)			
Component 2	1.612***			
	(0.0577)			
Component 3	1.102***			
	(0.0392)			
Water		0.172***	0.354***	0.0200
		(0.0179)	(0.0351)	(0.0923)
Bill		0.126***	0.521***	0.190***
		(0.0198)	(0.0279)	(0.0618)
Volumetric		0.192***	0.486***	0.203***
		(0.0187)	(0.0275)	(0.0586)
Water Price		0.366***	0.0641	-0.582***
		(0.0129)	(0.0412)	(0.0154)
Sewer Price		0.362***	0.0121	-0.596***
		(0.0130)	(0.0421)	(0.0147)
STC		0.185***	0.425***	0.0564
		(0.0183)	(0.0306)	(0.0766)
Irrigation		0.367***	-0.168***	0.0786**
		(0.0119)	(0.0240)	(0.0357)
Flush		0.370***	-0.277***	0.201***
		(0.0128)	(0.0237)	(0.0316)
Washer		0.414***	-0.170***	0.313***
		(0.0107)	(0.0273)	(0.0239)
Shower		0.419***	-0.218***	0.285***
		(0.0109)	(0.0255)	(0.0237)
Observations	1549			

Note: These are the estimates for the first three principal components assuming that the variables are distributed multivariate normal. The first column presents the eigenvalues for the first three components, and the next three columns contains the coefficient (weights) for each of the variables for the first three components. The standard errors are presented under the coefficients in parentheses. The standard errors and p-values are approximates. *p<0.1; **p<0.05; ***p<0.01

Table A.6: Heterogeneity Based on Confidence, Perceptions of Information, and Motivations

	(1) Base	(2) Confidence	(3) Confidence	(4) Useful/Surprising	(5) Motivation	(6) All
Post	0.1267***	0.1323***	0.1332***	0.1616***	0.1465***	0.1803***
	(0.0135)	(0.0138)	(0.0137)	(0.0463)	(0.0149)	(0.0464)
High Confidence (All)*Post		-0.0157				-0.0137
		(0.0111)				(0.0110)
High Confidence (Bill)*Post			-0.0595			
			(0.0368)			
Useful (All)*Post				-0.0050		-0.0040
				(0.0062)		(0.0063)
Surprising (All)*Post				-0.0022		-0.0017
				(0.0070)		(0.0071)
Useful & Surprising (All)*Post				-0.0020		-0.0034
				(0.0080)		(0.0081)
No Price Response*Post					-0.0068	-0.0076
					(0.0325)	(0.0328)
Money Motivated*Post					-0.0585***	-0.0558***
					(0.0214)	(0.0215)
Observations	40,151	40,151	40,151	40,151	40,151	40,151
Household and Time FEs	Yes	Yes	Yes	Yes	Yes	Yes

Note: The dependent variable is the log of daily water use. The estimation sample is a panel DID with household fixed effects restricted to the survey respondents. The High Confidence (All)*Post in column (2) is equal to the total number of times the respondent reported either "High" or "Very High" confidence about her estimate multiplied by a post-survey indicator. High Confidence (Bill)*Post in column (3) is equal to an indicator equal to one if the respondent reported "High" or "Very High" confidence about her total bill multiplied by a post-survey indicator. Useful (All)*Post, Surprising (All)*Post, and Useful & Surprising*Post in column (4) are variables equal to the number of times a respondent answered that the correct information was useful and/or surprising multiplied by a post-survey indicator. No Price Response*Post is a dummy for whether the respondent did not respond to previous price increases times a post-survey indicator and Money Motivated is a dummy equal to one if the primary motivation for water conservation was due to money times a post-survey indicator. Robust standard errors clustered at the household level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A.7: Heterogeneity Based on Days from Bill to Survey Send Date

	(1) ITT	(2) LATE	(3) Matching
Letter	0.00905		
	(0.00788)		
Letter*Days	0.000907		
	(0.00242)		
Survey		0.0784	0.0932***
		(0.0712)	(0.0268)
Survey*Days		0.00779	-0.00754
		(0.0224)	(0.00755)
Observations	27,796	27,796	2,892
Matching	None	None	Genetic
Panel w/ Household FEs	No	No	No
Baseline Consumption	Yes	No	Yes

Note: The dependent variable is the log of daily water use. In the first two columns the letter (column 1) or survey (column 2 and 3) variables is interacted with the number of days from the last bill until the survey was sent. We also include the base variable for all households because we also calculated when survey would have been sent to a control household. All models are estimated on the cross sectional sample for the billing period directly after the survey was completed and control for pre-treatment water consumption. Robust standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01