Title
Learning to live within your (water) budget: Evidence from allocation-based rates

Journal
Resource and Energy Economics, 57

ISSN
0928-7655

Authors
Pérez-Urdiales, María
Baerenklau, Kenneth A

Publication Date
2019-08-01

DOI
10.1016/j.reseneeco.2019.06.002

Peer reviewed

Learning to live within your (water) budget: evidence from allocation-based rates

María Pérez-Urdiales and Kenneth A. Baerenklau

School of Public Policy, University of California Riverside

Abstract

As allocation-based rates (ABR) become more widely adopted in the residential water sector, it is important to determine whether consumers interpret the water budget as a target for water use, or if they simply respond to the increasing block design. This paper exploits a natural experiment created when ABR were adopted by a California water district, that allows us to identify households with high water use as a treatment group. We find evidence of heterogeneous customer responses, and policy-induced behavioral changes that imply customers adjust their target level of water use based on the information signal provided by the water budget.

JEL-Classification: Q20; Q25; Q28.

Keywords: Water Demand; Allocation-based Rates; Convergence; Regression Discontinuity Design, Difference-in-differences; Latent Class Models

1 Introduction

In the current context of increasing water scarcity, both price and non-price strategies are commonly implemented to promote more sustainable water use. Among the former, approaches include seasonal rates, time of use pricing and drought surcharges, but increasing block rate (IBR) tariffs have been the most widely adopted price structure in recent decades (Zetland and Gasson, 2013). As noted by Monteiro and Roseta-Palma (2011), this type of tariff potentially provides consumers with a clear signal of water scarcity while also supplying low cost water for basic needs. However, Arbués and Barberán (2012) argue that equity problems may arise associated with the size of the household. In particular, they argue that because larger households tend to use more water even for basic needs, they will tend to pay more per unit for that water than smaller households (Arbués and Barberán, 2012). Moreover, IBR tariffs may not send an adequate scarcity signal to smaller households. That is, even if a small household is using water excessively, the higher rate may not apply to its level of consumption.

Non-price strategies have become increasingly common in the past years, mostly due to the fact that water demand is generally price inelastic (Sebri, 2014), and therefore, price policies can have limited water conservation potential. Notably, social comparisons have been used as a non-price tool to promote conservation in the electricity and water sectors. Several studies show that social comparisons can significantly decrease consumption (Allcott, 2011; Brent et al., 2015). However, Ferraro and Price (2013) find that the effect fades over time.

In recent years, allocation-based rates (ABR)[1] have been implemented in the residential water sector across the US[2] and in particular in California where they were codified in the State Water Code in 2008. These rates emerged to address the equity issues associated with IBR described above, as ABR is a type of increasing block rate structure in which the block sizes are defined based on the quantity of water that is deemed efficient for a particular consumer's indoor and outdoor needs. The applicable definition of water use efficiency is that provided by Gleick and Cain (2004): "the theoretical maximum water use efficiency occurs

[1] ABR also is referred to as "IBR water budgets" or "budget-based rates".

[2] See Mayer et al. (2008) for a list of water utilities using ABR in California, Utah, Nevada, Colorado, North Carolina and Florida.

when society actually uses the minimum amount of water necessary to do something". While this measure of water use efficiency could be determined using Frontier Analysis methods that compare the level of water use across households, in practice efficiency is determined by the water agency. That is, the definition of efficiency used here is an engineering rather than economic use of the word.

Under ABR, the block sizes are household-specific and are updated every billing period according to weather conditions (Baerenklau et al., 2014). Thus larger households (typically defined as having more residents and more irrigated area) are allocated more low-cost water than smaller households, with higher rates impinging on both households if and when consumption goes above the efficient level of use. This efficient level is labeled the "water budget" in customer billing statements, and provides a signal to each customer about their relative water use efficiency. The existence of the water budget thus potentially allows consumers to learn about their changing water needs and adapt their water use accordingly. That is, ABR can be considered a combination of both price and non-price strategies. However, rather that relying on social comparison information as above, this tariff informs consumers about their own level of water use efficiency, i.e., a measured tailored to their specific characteristics. Therefore, ABR is thought to be a fair and effective option for areas with a pressing need to encourage water conservation (Mayer et al., 2008).

However, several studies in the water economics literature raise concerns about the conservation potential of this price structure due to its complexity. For instance, Mayer et al. (2008) indicate that this tariff differs significantly from the typical rate structures most consumers are familiar with. García-Rubio et al. (2015) state that water utilities aiming to promote water conservation could implement standard IBR (with fixed block sizes) rather than ABR structures that may be difficult for consumers to understand. Pinto and Marques (2015) indicate that in order to achieve the desired outcome of a tariff structure, the consumer needs to receive a price signal or an understanding of their water consumption cost. As a consequence, they recommend avoiding complicated tariffs such as ABR as they are difficult for consumers to understand, thus potentially undermining their effectiveness. Moreover, the authors also state that ABR is costly to implement due to its information requirements, and the need for the water agency to define the efficient level of water use.

Despite these concerns, no study has attempted to test the effectiveness of ABR as a signal of efficient water use, and in particular whether the existence of the "water budget" helps inefficient customers learn to use water more efficiently. In other words, none has sought evidence that consumers are responding to the particular features of the rate structure that are designed to promote efficient use but that also make it relatively more complex than other rate structures.

Our goal is thus to investigate the effect of ABR on inefficient water consumers, and in particular to examine whether these consumers are converging to their efficient levels of water use through time. To do so, we use household-level panel data from the Eastern Municipal Water District (EMWD) in southern California for the period April 2006-2012. EMWD changed its tariff from a uniform rate to ABR in 2009. This change allows us to use quasi-experimental methods (specifically, a Regression Discontinuity Design (RDD)) to isolate the effect of ABR on inefficient consumers and determine whether they converge to efficient use in response to the implementation of this tariff. The idea of convergence in this context refers to a reduction in the distance between actual water use and efficient water use, over time. Therefore, we adopt an empirical strategy similar to the time series approach used to study *convergence* in the economic growth literature, along with a "difference-in-differences" (DID) methodology to estimate reductions in the distance between actual and efficient water use. Unlike typical DID methods that study the impact of a one-time change in policy (Nataraj and Hanemann, 2011), our DID approach explicitly considers a gradual learning or adaptation to the new tariff. Moreover, we expect heterogeneity in the response to the new tariff because households may differ in their ability or willingness to respond, as noted by the studies cited above. In order to address this issue, we follow the approach by Deb and Vargas (2016) and combine the DID strategy with a Latent Class Model.

In our baseline homogeneous model, we find that inefficient households experience a significant improvement in water use efficiency due to the implementation of ABR, i.e., their target level of water use decreases to just below the water budget. Moreover, the convergence rate towards the target level accelerates after the tariff change. In a heterogeneous model, we find that the effects of the tariff change are noticeably different across five latent classes of inefficient households. Finally, we show that after the tariff change, inefficient households are converging to levels of water use just below their efficient levels, implying that the efficiency signal provided by the water budget is having a detectable effect on consumer behavior.

2 Allocation-based rates: the case of the Eastern Municipal Water District

Eastern Municipal Water District is a member agency of the Metropolitan Water District of Southern California. EMWD serves an area of western Riverside County, comprising around 795,000 people over 555 square-miles. Cities served in the EMWD service area include Hemet, Menifee, Murrieta, Perris, San Jacinto and Temecula.

During the period of study, this area experienced increasing pressure on water resources that was mainly caused by two factors. First, Riverside County experienced a large increase in population, largely due to migration from Los Angeles County (U.S. Bureau of the Census, 2014). Second, California had three consecutive dry years during the period 2007-2009. This drought was the first ever for which a statewide proclamation of drought emergency was implemented (CDWR, 2010), on February 26, 2009. As a consequence, the Metropolitan Water District (EMWD's wholesaler) imposed mandatory water restrictions and rate increases on its member agencies in 2009 for the first time in 18 years.

Moreover, in 2008, Water Code section 370 et seq. was adopted by the California State Legislature. This Water Code section states that "it is in the best interest of the people of California to encourage public entities to voluntarily use allocation-based conservation water pricing, tailored to local needs and conditions, as a means of increasing efficient uses of water, and further discouraging wasteful or unreasonable use of water". (Ca. Water Code §§370, 374 (b), ch. 610 (A.B. 2882))

Faced with these challenges, and noting the recent change to the State Water Code, EMWD responded by changing its residential water tariff from a uniform rate to ABR in April 2009. There are four consumption blocks in EMWD's allocation-based tariff. The "water budget" is the sum of the first block, which represents efficient indoor water use, and the second, which takes into account efficient outdoor water use (where efficiency is defined by EMWD based on state-level standards and measured evapotranspiration needs). The remaining blocks, the third and the fourth, are "penalty" blocks that charge substantially higher prices for inefficient use. The block sizes are computed for each household as:

$$\begin{aligned}
&\text{Block 1. } \mathrm{w}_1 = \text{HHS x PPA x DF + IV} \\
&\text{Block 2. } \mathrm{w}_2 = \text{ (ET x CF x IA + OV)x DF} \\
&\text{Block 3. } \mathrm{w}_3 = 0.5 \text{ x}(\mathrm{w}_2 + \mathrm{w}_1) \\
&\text{Block 4. } \mathrm{w}_4 = \text{water use in excess of } \mathrm{w}_3
\end{aligned} \tag{1}$$

where HHS is the household size measured as number of residents in the house, PPA is the per-person water allowance, DF is a drought factor that is equal to 1 for the period of analysis,[3] ET is the sum of observed evapotranspiration values for each period measured in inches, CF a conservation factor which is equal to 1 for all the households in the sample, IA is the irrigated area measured in square feet. IV and OV are indoor and outdoor variances that can be negotiated for households with unusual indoor or outdoor circumstances.

[3] While the drought factor considered in this type of ABR directly affects the size of the blocks, an alternative would be to use a drought factor affecting the rates.

Examples of indoor variances include medical needs and in-home daycare, and examples of outdoor variances include large animals (such as horses) and turf grass establishment. Indoor variances are not affected by the drought factor whereas outdoor variances are.

The new tariff also introduced new marginal price levels. Table 1 shows the price households were paying before the change in the tariff and the marginal price paid in each of the blocks after the implementation of the new tariff. The price paid under the prior uniform rate was higher than the price associated with the first block under ABR, but slightly lower than the price paid for the water consumed within the second block.

	Price	Block 1	Block 2	Block 3	Block 4
April 2006-07	Uniform rate	1.73			
April 2007-08	Uniform rate	1.79			
April 2008-09	Uniform rate	1.88			
April 2009-10	Marginal price	1.32	2.41	4.33	7.91
April 2010-11	Marginal price	1.43	2.61	4.68	8.57
April 2011-12	Marginal price	1.38	2.53	4.53	8.29

Prices are expressed in 2010 $/ccf

Table 1: Evolution in prices for the period April 2006- April 2012

3 Regression discontinuity design

Our hypothesis is that, in addition to the effect of inclining prices, the water budget aspect of ABR acts as a kind of “soft constraint” that signals to consumers the efficient level of water use. In other words, it acts like a target with a built-in penalty for exceedance (i.e. a higher marginal price). However, the penalty is delayed since it arrives with the subsequent water bill approximately one month later. Moreover, both the water budget and the penalty are somewhat detached from water use habits because few, if any, consumers know how much water they are using in real-time. Thus, consumers have an opportunity to react to this signal and adjust consumption habits in the following month. Such behavior would tend to produce convergence to efficient use through time as consumers learn how their habits affect their water use, and how their actual water use compares to an indicator of efficient use for their specific household size and environmental conditions.

In order to investigate whether ABR appears to promote this kind of learning behavior, we use a regression discontinuity design (RDD) and borrow the identification strategy used by Nataraj and Hanemann (2011). Specifically, we identify two types of consumers: inefficient households as the treatment group and efficient households as the control group. As in Nataraj and Hanemann (2011), these groups are identified prior to the tariff change, and thus prior to receiving any ongoing feedback about their water use efficiency, so we must construct an indicator of efficiency when EMWD was still using uniform rates. To do this, we calculate "shadow water budgets" using EMWD's ABR formula (described in Section 2), customer data from EMWD, and location-specific historical evapotranspiration data computed by Baerenklau et al. (2014). Treatment households are those consuming on average above their shadow water budgets during the year prior to receiving the first bill under ABR[4]. If these households continue with their consumption pattern, their ABR water bill will penalize them and signal that they are consuming inefficiently. Control households are those consuming on average within their shadow water budgets during the same period. If they maintain their consumption pattern, their ABR water bills will indicate that they are consuming efficiently.[5] That is, as in Nataraj and Hanemann (2011), we assume that households that consumed above the threshold in the past, are likely to continue consuming above the threshold in the future, and thus, receive the information signal. Moreover, by assigning households to treatment and control groups based on their water use efficiency during the year prior to the tariff change, we alleviate concerns about mean reversion. This is because an idiosyncratic

[4]For convenience, we refer to this period in time as the year prior to the change in tariff for the remainder of the paper.

[5]Our approach differs from the one used by Nataraj and Hanemann (2011) in that we define the treatment and control groups considering households' average consumption throughout the previous year rather than during summer months only. We do this because, unlike the IBR analyzed in Nataraj and Hanemann (2011), ABR adapt to weather conditions throughout the year. Therefore, we believe it is preferable to consider the entire year prior to the implementation of ABR.

consumption shock in one billing period has a much smaller impact on the assignment when considering average consumption for the entire year.

As in Nataraj and Hanemann (2011), our treatment is not randomly assigned because it depends on each household's water use with respect to its shadow water budget. However, the treatment can be considered "as good as randomized" (Lee and Lemieux, 2010) if households have imprecise control over the treatment assignment. As noted by Lee and Lemieux (2010), when households have imprecise control, observable and unobservable household characteristics prior to the treatment have statistically indistinguishable distributions on either side of the treatment threshold[6]. We can expect this situation to be more likely to occur as we examine smaller and smaller neighborhoods around the threshold, making the treatment "locally randomized". To check for this condition, we first follow the approach by Calonico et al. (2015) and perform a graphical exploratory analysis to check the continuity of observable household characteristics around the threshold. We consider the number of people living in the household (*HHS*), the size of the irrigated area (*IA*), household income (*Income*)[7], and education level[8] (*Education*). In Figure 1, we plot the value of each characteristic against the difference between the logs of water use and water budget. In particular, the black dots in each panel in Figure 1 represent the sample means of each covariate computed by partioning the support of the difference between the logs of water use and budget into 250 disjoint bins separately for treatment and control households. Moreover, in each panel the two solid gray lines show a fourth-order polynomial approximation of the conditional expectation of the observable covariate and the difference between the log of water use and budget, computed separately for treatment and control households.[9] The dashed

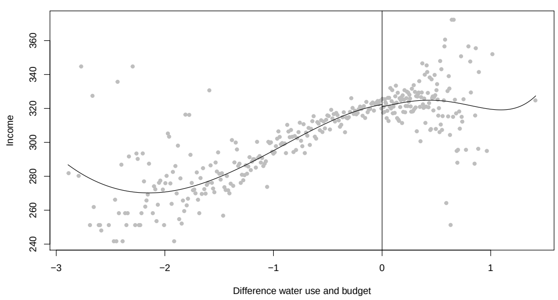
[6]The threshold is the efficient level of water use.

[7]Our income variable is derived by adjusting census income by the fraction of income usually spent on "utilities, fuels and public services" and for temporal changes in per-capita personal income for the Ontario-Riverside-San Bernardino metropolitan statistical area.

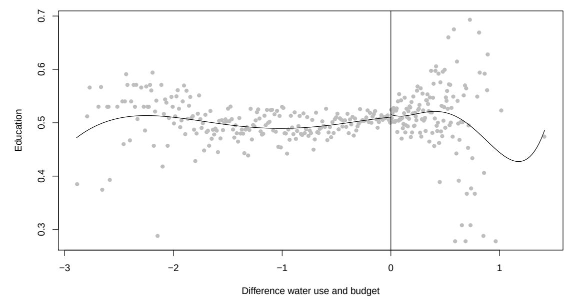
[8]Our education variable is defined as the fraction of the census tract reporting at least "some college'.'

[9]The plots in Figure 1 have been generated using the package "rdrobust" developed by Calonico et al. (2018) in R Core Team (2013).

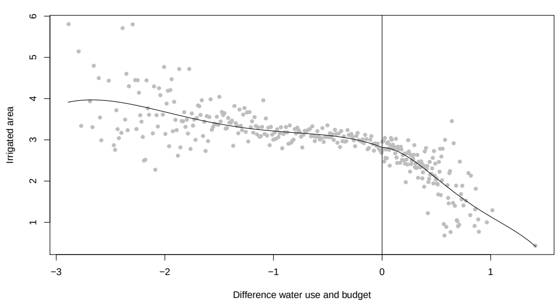
vertical line represents the threshold. We observe no significant visual evident discontinuity at the threshold in any of the panels. Following Lee and Lemieux (2010), Table 2 shows the RD estimates and Standard Errors for each of the observable covariates. Confirming the graphical results, we find no evidence of significant discontinuity.



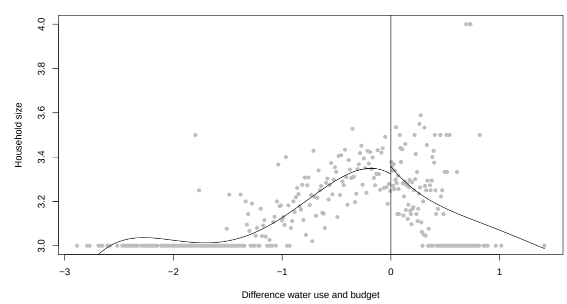
(a) *Income* vs Difference between water use and budget



(b) *Education* vs Difference between water use and budget



(c) *IA* vs Difference between water use and budget



(d) *HHS* vs Difference between water use and budget

Figure 1: Demographic and socioeconomic variables by the ratio of water use to water budget

	Coefficient	Std. Error
Income	-2.214	(2.084)
Education	0.009	(0.009)
Irrigated Area	0.077	(0.071)
Household size	0.061	(0.051)

Table 2: RDD estimates for observable covariates

Moreover, we also test for differences in means of these variables for the treatment and control groups. As in Nataraj and Hanemann (2011), we first compare all efficient households against all inefficient households,[10] and we find that treatment and control groups are significantly different in terms of their socioeconomic variables. However, we also consider households with an average water use in the year before the implementation of ABR within the $\pm 50\%$, $\pm 25\%$ and $\pm 10\%$ neighborhoods of the threshold, and we observe that the differences between groups decrease with the size of the neighborhood. Treatment and control groups are not significantly different when we consider households within 0.10 units around the threshold.[11]

Next, although it is not possible to directly test for differences in unobservable characteristics, we follow Lee and Lemieux (2010) and Oldenburg et al. (2016) and examine whether the density of the assignment variable (the relative household water use efficiency) is continuous. If households exert control over the assignment, there should be a discontinuity in this density at the threshold. Figure 2 shows a histogram of the relative water use efficiency. This variable is partitioned into equally spaced bins of 0.05, and the dashed vertical line represents the threshold. The figure shows no visual evidence of discontinuity at the threshold.

Once we define the treatment and control groups and check the validity of the assignment, we plot in Figure 3 the evolution of the mean water use (black line) and mean efficient use (gray line) for each group during the period of study. The vertical dashed line in each panel denotes the tariff change. In general, we observe that the control group's mean efficient water use is higher than the treatment group's. This is consistent with the descriptive statistics reported in Table A1 (in Appendix A), where we observe that both the *IA* and *HHS* are larger for the control group. The differences in these two variables are not significant when

[10]Households with a pre-ABR water use extremely close to the efficient level (shadow water budget) may have been unsure about whether they would be inefficient after the tariff change. Therefore, including these households in our analysis may confound the results. Following Nataraj and Hanemann (2011), we exclude households extremely close to the threshold. In our case, we do not consider those households consuming in the $\pm 2\%$ of the threshold.

[11]A summary of these tests is provided in Table A1 in Appendix A.

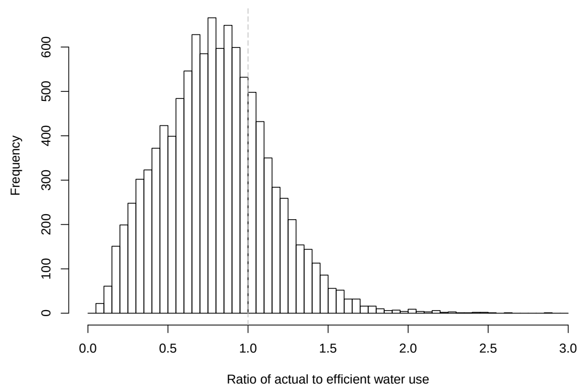


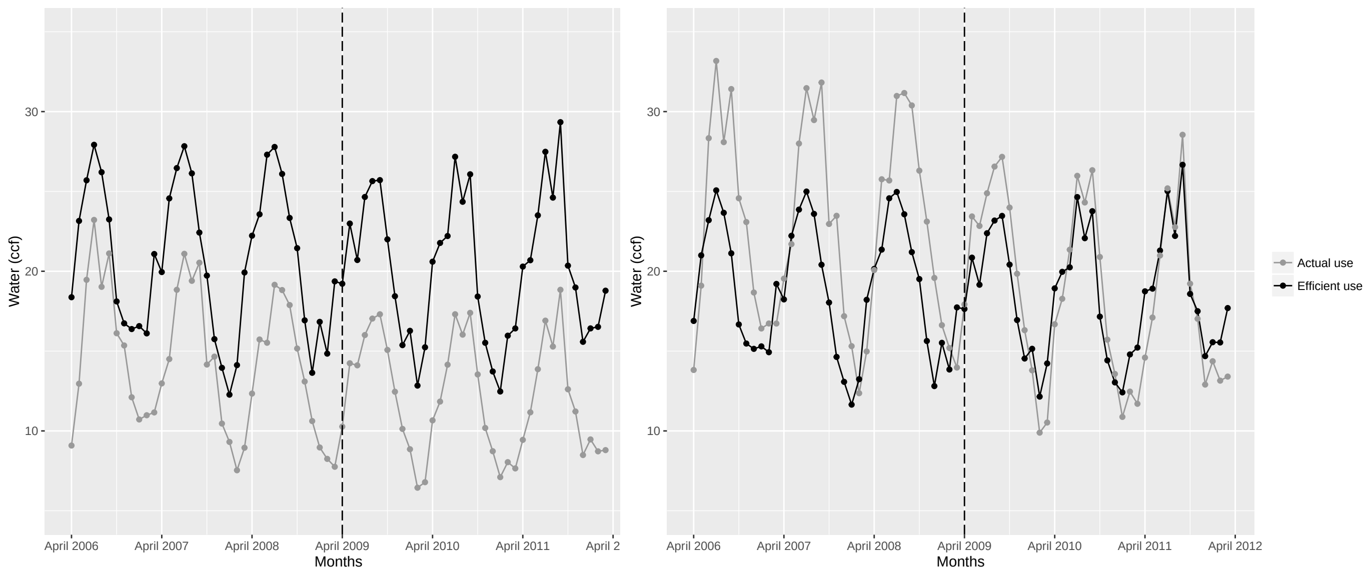
Figure 2: Histogram of Relative Water Use Efficiency in the year prior to the tariff change

we consider a small neighborhood around the treatment threshold. Prior to the tariff change, we observe that the control group's water use is consistently efficient, whereas the average water use for the treatment group is usually inefficient. However, the overall patterns are similar across groups. After the tariff change, we observe that the distance between average water use and efficient use for the control group seems to remain almost unchanged (perhaps a slight decrease), whereas this distance substantially decreases for treatment households.

4 Homogeneous model

4.1 Difference-in-differences specification

Difference-in-differences (DID) methods are commonly used to study the impact of a "treatment" on a group. In the simplest case, outcomes are observed for two groups over two time periods (Wooldridge, 2007). One group is exposed to a treatment in the second period but not in the first, whereas the second group is never exposed to the treatment. Therefore, the control group represents the evolution that would have characterized the treatment group in the absence of treatment.



(a) Control: efficient households (b) Treatment: inefficient households

Figure 3: Evolution of water use and budget

This basic model can be estimated using the following regression:

$$y_{ir} = \alpha_1 + \alpha_2 D_r + \alpha_3 D_j + \beta D_r D_j + \epsilon_{ir} \tag{2}$$

where y is the outcome of interest, r indexes the period, and j indexes the group. D_r is a dummy variable that takes value 1 for the second period (that is, when $r = 1$), and D_j is a binary indicator that takes value 1 for the treatment group and 0 for the control group. In this case, the estimated coefficient of interest is β. This coefficient applies to the interaction term $D_t D_j$ which takes value 1 for observations in the treatment group in the second period. Therefore, the DID estimate is:

$$\hat{\beta} = (\bar{y}_{11} - \bar{y}_{10}) - (\bar{y}_{01} - \bar{y}_{00}) \tag{3}$$

As noted by Slaughter (2001), most DID economic studies analyze a one-time treatment, which allows the analyst to delineate two periods: before and after the change. In this regard, the implementation of ABR is somewhat different. First, although we have a discrete treatment (i.e. the tariff change), the efficiency signal provided by the water budget arrives each month in the household water bill. Thus the effect of the treatment is repeated at regular intervals. Second, as explained in Section 2, water budgets are tailored to each household's specific characteristics including weather, and thus change every month depending on the level of evapotranspiration. Therefore, for a relatively inefficient household to modify its behavior and consume more efficiently, the household must learn to effectively predict the size of the budget each month. In practice, this is accomplished by experimenting with and changing household consumption habits and observing the effect on the subsequent water bill. In this sense, the implementation of the tariff change could be anticipated to have a gradual effect over multiple periods rather than a large one-time effect. Therefore, the analytical framework should accommodate this anticipated response.

In order to measure the effect of such a treatment, we base our empirical strategy on the concept of convergence used in the economic growth literature (Barro et al., 1991; Barro and Sala-i Martin, 1992). There exist several measures of convergence, but here we base our analysis on time series tests (Bernard and Durlauf, 1996; Ben-David, 1993, 1996). In the growth literature, convergence usually refers to reductions in dispersion of per-capita income or productivity levels among regions over time. However, in our case, we aim to test if households' water use is converging to an efficient level over time and, importantly, whether the average difference between actual and efficient, and the rate of convergence were affected by the tariff change. Therefore, our variable of interest is a measure of dispersion that can be written as:

$$z_{it} = \log(W_{it}) - log(E_{it}) \tag{4}$$

where W_{it} and E_{it} are the actual and the efficient water use for household i and time period t, respectively. Following Ben-David (1993), we can model convergence as:

$$z_{it} = \alpha + \phi z_{it-1} + \epsilon_{it} \tag{5}$$

where α measures the average difference between the logs of actual and efficient water use or target level, ϕ is a parameter measuring convergence to that target level and ϵ is an econometric error. As discussed by Bernard and Durlauf (1996), there exists convergence if z_{it} is a zero-mean stationary process, that is if $\alpha = 0$ and $\phi < 1$. As in Ben-David (1993), we define $\Delta z_{it} = z_{it} - z_{it-1}$ so the Dickey-Fuller form of Equation (5) can be written as:

$$\Delta z_{it} = \alpha + \beta z_{it-1} + \epsilon_{it} \tag{6}$$

Here β is the convergence coefficient, which equals $\phi - 1$. Therefore, a negative (positive) β coefficient indicates convergence (divergence).

Following Slaughter (2001), the approach by Ben-David (1993) can be extended to estimate convergence in the treatment group both before and after the treatment:

$$\Delta z_{it} = \alpha_1 + \alpha_2 D_r + \beta_1 z_{it-1} + \beta_2 z_{it-1} D_r + \epsilon_{it} \tag{7}$$

where r denotes the period before or after the tariff change. Since both periods have the same length in our data (36 months), $r = 0$ for months from $t = 1$ to $t = T/2$, and $r = 1$ for months from $t = 1+T/2$ to $t = T$, where T is 72 months. In this model, we are assuming that relative water use efficiency may improve over time as households receive repeated feedback on their water use. Both before and after the tariff change, such feedback takes multiple forms including a positive marginal cost for water consumption, messaging campaigns that educate consumers about water use efficiency, peer influence, and perhaps others. After the tariff change, feedback also arrives in the form of the water budget. Therefore, we should expect consumers to be making efficiency improvements throughout our observation window, but possibly converging at different rates and to different levels of water use before and after the tariff change. Given the nature of the treatment, we could reasonably expect convergence rates to increase and target levels to decrease towards the water budget after tariff change. Thus, Equation (7) measures behavioral change by comparing the convergence parameters α and β before and after the treatment. However, this equation may not estimate the true effect of the treatment if other factors affecting convergence behavior also changed around the time of the tariff change. Therefore, we introduce a control group that would have been affected by the same confounding factors in order to properly identify the treatment effect. Following Slaughter (2001), we can estimate the convergence parameters by combining the convergence model with the basic DID approach:

$$\Delta z_{it} = \alpha_1 + \alpha_2 D_r + \alpha_3 D_j + \alpha_4 D_r D_j + \beta_1 z_{it-1} + \beta_2 z_{it-1} D_r + \beta_3 D_j z_{it-1} + \beta_4 D_r D_j z_{it-1} + \epsilon_{it} \quad (8)$$

We note that one main difference between our approach and Slaughter (2001) is the inclusion of an intercept in our model.[12] The main consequence of this difference is that we may not be able to interpret the regression t-statistics as usual in the presence of unit root. As noted by Levin et al. (2002) "regression t-statistics for stationary panel data converge to the standard normal distribution". Therefore, before estimating equation (8), we must first establish that our dispersion measure z_{it} is a stationary process. To do so, we can use two of the most popular types of unit root tests for panel data. Both types of tests consider the same null hypothesis: the presence of a unit root for all individuals in the panel. However, the alternative hypothesis differs. The first type of test (Levin et al., 2002; Harris and Tzavalis, 1999) assumes that the water use-water budget dynamics are the same for each household, and thus the alternative hypothesis is that all household series are stationary. For the second

[12] As discussed by Ben-David (1996), the intercept equals 0 by construction when testing per capita income convergence due to a convenient transformation of variables. However, this transformation does not apply to our analysis.

type of test (Im et al., 2003; Breitung and Pesaran, 2008), the alternative hypothesis is that at least one of the series is stationary. Given our data, we also must employ unit root tests for fixed T. Therefore, we utilize Harris and Tzavalis (1999), hereafter HT, and Im et al. (2003), hereafter ITS, as our unit root tests.

After performing these tests, and presuming rejection of the null, we can use Equation (8) to estimate convergence parameters for the treatment and control groups, before and after the treatment. Table 3 shows the formulas for these parameter estimates. The β coefficients indicate the convergence rate for each group and period, and the α coefficients indicate the target levels. A negative (positive) rate parameter β indicates convergence (divergence) over time, and as indicated by Ben-David (1993), the larger the β, the faster the rate. A negative (positive) intercept indicates that households are converging to a level below (above) the efficient level. Following Slaughter (2001), the effect of the implementation of the new tariff on treatment households can be obtained by computing the DID of the convergence rates in the table. The difference in convergence rates between the treatment and the control groups before the change is given by $(\beta_1 + \beta_3) - \beta_1 = \beta_3$. Moreover, the difference in convergence rates between the treatment and the control groups after the implementation of the tariff is given by $(\beta_1 + \beta_2 + \beta_3 + \beta_4) - (\beta_1 + \beta_2) = \beta_3 + \beta_4$. Therefore, β_4 identifies the effect of the tariff on the treatment group's convergence rate since the difference between the two previous DID rates is $(\beta_3 + \beta_4) - \beta_3 = \beta_4$. Similarly, α_4 indicates the change in the treatment group's target level due to the implementation of the new tariff.

Convergence to an efficient level of use exists if the difference between the log of water use and the log of efficient use is a stationary process around zero, that is, the relevant convergence parameter is negative and the relevant target level is equal to 0. We use an F test of the null hypothesis that each target in table 3 equals zero. If the null hypothesis is rejected, there exists convergence to a non-zero target: a negative (positive) intercept indicates that households are converging to a level below (above) the efficient level.

Table 3: Convergence rates and intercepts

Household group and period	Intercept	Convergence rate
Control group pre-IBR water budget	α_1	β_1
Control group post-IBR water budget	$\alpha_1 + \alpha_2$	$\beta_1 + \beta_2$
Treatment group pre-IBR water budget	$\alpha_1 + \alpha_3$	$\beta_1 + \beta_3$
Treatment group post-IBR water budget	$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4$	$\beta_1 + \beta_2 + \beta_3 + \beta_4$

4.2 Convergence to efficient water use assuming homogeneity

We begin our analysis by performing the unit roots tests discussed in Section 4.1. These tests are computed for both treatment and control groups, before and after the tariff change, using all households as well as households in the $\pm$50%, $\pm$25% and $\pm$10% neighborhoods of the threshold. Results from these tests are provided in the Appendix B. In all cases we reject the null hypothesis of the presence of unit roots for the treatment and control groups, before and after the implementation of the new tariff, and for the five neighborhoods considered. Therefore, we can conclude that the difference between the log of water use and the log of efficient use is indeed a stationary (convergent) process.

Turning to the DID analysis, we first present estimates for a one-class (homogeneous) model using Equation (8). Table 4 presents these results. The columns in the table show the results using all households and households in the $\pm$50%, $\pm$25% and $\pm$10% neighborhood of the threshold. We mainly discuss the results for the $\pm$10% neighborhood of the threshold, as it is the sample in which we do not observe significant differences between treatment and control households. However, we also report the results for larger samples for comparison purposes. The negative and significant coefficient α_1 indicates that the target level of water use for control households in the period before the tariff change was already below efficient use (as seen also in Figure 3). However the target for treatment households during the same period is above efficient use: α_3 is positive and larger (in absolute terms) than α_1. After the tariff change there is a decrease in the target level for both treatment and control groups as the negative coefficient α_2 indicates. This general decrease in the average difference between actual use and efficient use may be due to several factors, including higher marginal prices under the new tariff and increased messaging about the need for water conservation due to the statewide proclamation of drought emergency in 2009. Regarding the effect of the treatment, we see that the estimated coefficient α_4 is negative and significant, implying that the tariff change produced a significant downward shift in the target level for inefficient consumers relative to the control group. Although the sign of this effect is consistent across all sample sizes the magnitude decreases as we consider smaller neighborhoods. That is, the effect is more pronounced when considering the more inefficient households that are further away from the threshold. However, as seen in Table A1 in Appendix A, larger neighborhoods imply more dissimilar treatment and control households, which reduces the likelihood that we are observing the true causal effect of the tariff change.

In terms of the convergence rate coefficients for the sample comprising households in the $\pm 10\%$ neighborhood of the threshold, we observe that β_1 is negative while β_3 is positive but smaller (in absolute terms) than β_1. This indicates that there is convergence to the pre-ABR targets before the tariff change for both control and treatment groups. Such convergence suggests that households may be using accumulated experience to target desired levels of water services and, by extension, water use, which is consistent with household production theory. However, at least in part because they are not receiving clear signals about efficient use, these target levels may or may not be efficient, as we observe in the treatment group.[13] After the tariff change, there is a general decrease in the convergence rate, as indicated by the positive β_2. However, this decrease is less pronounced for treatment households, as β_4 is negative but smaller (in absolute terms) than β_2. To understand what may be driving

[13]As indicated by the tests shown in Table B1, the difference between actual and efficient water use is a stationary process around a positive (negative) target for the treatment (control) group before the implementation of ABR.

	All HH	HH$\pm$ 50%	HH$\pm$25%	HH$\pm$10%
Intercept (α_1)	-0.123^{***}	-0.122^{***}	-0.076^{***}	-0.039^{***}
	(0.001)	(0.001)	(0.001)	(0.002)
Post (α_2)	-0.011^{***}	-0.027^{***}	-0.047^{***}	-0.065^{***}
	(0.001)	(0.001)	(0.002)	(0.003)
Treatment (α_3)	0.184^{***}	0.173^{***}	0.107^{***}	0.046^{***}
	(0.002)	(0.002)	(0.002)	(0.003)
Post x Treatment (α_4)	-0.075^{***}	-0.058^{***}	-0.032^{***}	-0.008^{*}
	(0.002)	(0.002)	(0.003)	(0.004)
Dif$_{t-1}$ (β_1)	-0.240^{***}	-0.410^{***}	-0.471^{***}	-0.482^{***}
	(0.001)	(0.002)	(0.003)	(0.005)
Dif$_{t-1}$ x Post (β_2)	0.046^{***}	0.098^{***}	0.111^{***}	0.108^{***}
	(0.002)	(0.002)	(0.004)	(0.007)
Dif$_{t-1}$ x Treatment (β_3)	-0.160^{***}	-0.019^{***}	0.026^{***}	0.036^{***}
	(0.003)	(0.003)	(0.004)	(0.007)
Dif$_{t-1}$ x Post x Treat (β_4)	0.038^{***}	-0.0004	-0.011^{*}	-0.019^{*}
	(0.004)	(0.005)	(0.006)	(0.010)
N	694152	528120	299808	110088

Standard errors in parentheses

*p<0.1; **p<0.05; ***p<0.01

Table 4: Estimation results - Homogeneous model

this result, recall that both control and treatment households experienced a downward shift in their target levels in the period after the tariff change. In order to approach the new target, they may change habits and technologies, which may induce a learning process that requires a gradual adaptation. The more pronounced decrease in convergence rate for the control group could be because it is relatively challenging for already-efficient consumers to figure out how to be even more efficient and make progress towards a new, more demanding, target. For the larger samples, we observe similarities in terms of the signs of the effects when looking at β_1 and β_2. However, β_3 is negative, and β_4 is positive or not significantly different from 0 when we consider the two biggest samples. This could imply that there is a high level of heterogeneity when we consider observations far from the budget, and that the effect is "cleaner" closer to the threshold. However, another explanation is there may still be unobserved household heterogeneity around the threshold. We explore this possibility further in Section 5.2, after some falsification tests.

Finally, Table 5 presents an alternative way to consider our results that may facilitate more intuitive before-and-after comparisons. This table shows the intercepts and convergence rates computed using the sums in Table 3 and the estimated coefficients in Table 4. The F tests reject the null hypothesis of the sum of the coefficients being equal to 0 in all cases.

			All HH	HH +/- 0.5	HH +/-0.25	HH +/-0.10
Control	Pre-ABR	Intercept	-0.123	-0.122	-0.076	-0.039
		Convergence rate	-0.240	-0.410	-0.471	-0.482
	Post-ABR	Intercept	-0.133	-0.149	-0.123	-0.104
		Convergence rate	-0.399	-0.430	-0.445	-0.447
Treatment	Pre-ABR	Intercept	0.061	0.051	0.031	0.007
		Convergence rate	-0.193	-0.312	-0.360	-0.374
	Post-ABR	Intercept	-0.025	-0.033	-0.048	-0.066
		Convergence rate	-0.315	-0.332	-0.345	-0.357

Table 5: Convergence rates and target levels

4.3 Robustness checks: Price vs Label effects

The preceding results suggest that the implementation of ABR affected the water use efficiency of inefficient consumers. However, we cannot yet say which features of the new rate structure are driving these behavioral changes. Consider the comparison of prices paid before and after the implementation of ABR shown in Table 6.

	pre-ABR price	Post-ABR price			
		Marginal Price		Average Price	
		Control	Treatment	Control	Treatment
All HH	1.88	2.41	4.33	1.80	2.44
HH +/- 0.5				1.92	2.40
HH +/-0.25				2.03	2.33
HH +/-0.10				2.11	2.26

Table 6: Comparison of Prices before and after the implementation of ABR

Both marginal and average prices paid under ABR are higher than those paid before the tariff change (except for the control group's average price for the largest sample). Moreover, marginal and average prices for the treatment group are higher than those for the control group. And recall that the water budget coincides with second kink point in the new rate structure. Therefore, it remains unclear whether the results in Table 4 are driven by changes in marginal price or average price, or by the signal provided by the water budget.

In order to address these questions, we conduct a series of falsification tests using the two remaining kink points in the rate structure (i.e. between blocks 1 and 2, and between blocks 3 and 4). Conveniently, these kink points are not associated with the water budget. Therefore, we repeat our analysis using these two alternative treatment thresholds.

Tables 7 and 8 show comparisons of the prices paid before and after the implementation of ABR using these new thresholds. We observe that the average price decreased for both treatment and control groups when we consider the first kink point as the threshold, whereas the marginal price increased for the treatment group and decreased for the control group.

When we consider the third kink point as the threshold, we see that marginal and average prices increased for both treatment and control households, with larger increases for the treatment group.

	pre-ABR price	Post-ABR price			
		Marginal Price		Average Price	
		Control	Treatment	Control	Treatment
All HH	1.88	1.32	2.41	1.37	2.06
HH +/- 0.5				1.38	1.65
HH +/-0.25				1.41	1.57
HH +/-0.10				1.44	1.53

Table 7: Comparison of Prices before and after the implementation of ABR - 1st kinkpoint

	pre-ABR price	Post-ABR price			
		Marginal Price		Average Price	
		Control	Treatment	Control	Treatment
All HH	1.88	4.33	7.91	1.95	2.80
HH +/- 0.5				2.19	2.79
HH +/-0.25				2.48	2.77
HH +/-0.10				2.61	2.69

Table 8: Comparison of Prices before and after the implementation of ABR - 3rd Kinkpoint

We focus our analysis on the $\pm$10% neighborhoods of these two thresholds to avoid picking up our main treatment effect(Nataraj and Hanemann, 2011).[14] As in our original analysis, we assign households to treatment and control groups based on the average ratio of water use to each kink point in the year prior to the tariff change. Results from the first test are presented in the first column in Table 9.[15] The false treatment effect is not statistically significant. That is, households that are consuming above the first kink point before the tariff

[14]The distance between kink points is relatively small for some households. Therefore, if we consider larger samples for the falsification tests, we may include as treatment households some that are already treatment households in our original analysis, in which case the estimated treatment effect for the robustness check would be confounded.

[15]Difference in mean tests of the observable covariates for the false treatment and control groups are shown in Table A2 in the Appendix. The results indicate that there are no significant differences in observable characteristics with the exception of the *Irrigated area* for households in the $\pm$10% neighborhoods of the third kink point. However, this result does not invalidate the falsification test.

change do not update their target level and do not change their convergence rate towards the target over time.

	1^{st} kink point	3^{r} kink point
Intercept (α_1)	-0.016^{***}	-0.062^{***}
	(0.004)	(0.003)
Post (α_2)	-0.048^{***}	-0.067^{***}
	(0.006)	(0.005)
Treatment (α_3)	0.051^{***}	0.042^{***}
	(0.006)	(0.006)
Treatment x Post (α_4)	-0.010	0.004
	(0.009)	(0.010)
Diff$_{t-1}$ (β_1)	-0.385^{***}	-0.436^{***}
	(0.009)	(0.008)
Diff$_{t-1}$ x Post (β_2)	0.080^{***}	0.115^{***}
	(0.013)	(0.011)
Diff$_{t-1}$ x Treatment (β_3)	-0.013	0.083^{***}
	(0.013)	(0.014)
Diff$_{t-1}$ x Post x Treat (β_4)	0.022	-0.001
	(0.018)	(0.020)

Standard errors in parentheses
***p<0.01

Table 9: Robustness checks

We repeat the falsification test considering the average ratio of water use to the third kink point in the year prior to the tariff change as the threshold. Again, we find that the false treatment effect is not significant: this test also indicates that the tariff change did not induce any adjustment or convergence towards the third kink point. Although we observe a general decrease in water use with respect to these thresholds, we only find a significantly larger reduction for treatment households when we consider the water budget (second kink point) as the threshold. Taken together, these test results provide strong support for the validity of our approach. While we observe a general decrease of water use relative to the different thresholds and a decrease in the convergence rate towards the target level, as indicated by α_2 and β_2, respectively, we only identify a significant treatment effect in both α_4 and β_4

when we consider the water budget as threshold.

At first glance, our results may seem to contradict those of Nataraj and Hanemann (2011), our main reference study. Their results indicate that consumers respond to changes in prices, particularly to the introduction of a new block in the increasing block rate structure, whereas in our study we find that consumers react to the efficiency signal rather than to the price structure. However, caution is warranted when directly comparing these results because the price structures analyzed differ in the information provided to consumers. In particular, the price structure analyzed in Nataraj and Hanemann (2011) is a standard IBR with fixed block sizes across time and consumers. As discussed in the introductory section, our analysis focuses on the implementation of ABR, which is a type of IBR with household-specific block sizes, providing the consumer information about the efficient level of water use. In this sense, our results could be compared to the findings of Nemati et al. (2018). Their results indicate that social norms-based conservation programs such as digital and web-based water consumption analytics platforms cause significant reductions in water use, whereas households in their sample have no response to the IBR structure.

5 Heterogeneous model

5.1 Incorporating heterogeneous consumer responses

The empirical model proposed thus far assumes that the effect of the tariff change is homogeneous across households. However, this is unlikely, and our estimation results suggest there is underlying heterogeneity in our sample. Such heterogeneity may arise from a number of sources. For example, households may differ in their abilities to react to the new tariff due to unobserved differences in water-using technologies or in their levels of understanding of the new tariff (as mentioned above). We can also expect that some households may not pay close attention to their water bill. Whatever the source, neglecting this heterogeneity can lead to biased parameter estimates and incorrect policy conclusions.

An increasing number of papers address the issue of heterogeneity in the residential water demand literature. Among the different methodologies, Latent Class Models (LCM) provide an intuitive approach not only to control for heterogeneity but also to reveal underlying "types" of consumers which can be helpful for deriving policy conclusions (Pérez-Urdiales et al., 2016). This is mainly because LCM is consistent with the way producers including public agencies (water or otherwise) tend to think about their customers when deciding which programs and/or services to offer. Specifically, producers tend to offer menus with discrete options that are targeted to various types of consumers. Examples from water policy include information campaigns targeted to specific customers, rebate programs with specific eligibility criteria, and pricing plans that are tailored to different user groups (e.g. residential, low income, commercial, agricultural). LCMs have been implemented previously in research areas closely related to this study. For example, they have been used in the economic growth literature to identify the so-called "convergence clubs" or groups of countries/regions with similar convergence patterns (Paap et al., 2005; Centorrino and Pérez-Urdiales, 2014). They also have been combined with a DID strategy by Deb and Vargas (2016) to measure a heterogeneous treatment effect on the implementation of calorie-labeling laws in the US.

Here we use LCM to identify potentially different responses to the implementation of ABR. LCM assumes that the sample of households is a finite mixture of J subpopulations (Cameron and Trivedi, 2005). As noted by Fernandez-Blanco et al. (2009), this method involves simultaneously estimating the convergence behavior of each group, and classifying households into different groups. Thus, our goal is to re-estimate Equation (8) while identifying different groups (or classes) of households. Convergence behavior is assumed to be the same within each class, but differences are allowed across classes. Class membership is probabilistic in a LCM, so we assume that each household belongs to each class j with a probability π_j. With this in mind, we can write the likelihood of observing our dependent variable for a single household Δz_{it} as:

$$f(\Delta z_{it} \mid \Delta z_{it-1}; \beta; \pi) = \sum_{j=1}^{J} \pi_j f_j(\Delta z_{it} \mid \Delta z_{it-1}; \beta_j) \tag{9}$$

If we assume a normal mixture, the sample likelihood function can be written as the sum of J likelihood normal distributions weighted by the probabilities of class membership:

$$l(\beta, \pi \mid \Delta z_{it}) = \prod_{i=1}^{N} \left(\sum_{j=1}^{J} (\pi_j \prod_{t=1}^{T} \frac{1}{\sigma_i} \phi(\epsilon_{it}/\sigma_i))^{d_{ij}} \right) \tag{10}$$

where d_{ij} denotes an indicator variable which takes value 1 if household i belongs to class j, and 0 otherwise. This likelihood function may be maximized using the Expectation-Maximization (EM) algorithm.[16] Maximization produces an estimated set of α and β coefficients as in table 3 for each class, where the number of classes is specified exogenously.

As discussed by Pérez-Urdiales et al. (2016), the usual way to proceed is to estimate models with a sequentially increasing numbers of classes. Then, in order to select the appropriate number of classes, we compute likelihood-based information criteria, such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC). These statistics reward the goodness of fit and penalize an increasing number of estimated parameters. Therefore, they are used to compare models with different numbers of classes, preferring the model that minimizes the criteria.

5.2 Convergence to efficient water use assuming heterogeneity

Using LCM to re-estimate our DID parameters, we focus on the most restrictive sub-sample (households in the ±10% neighborhood of the threshold) because we know the treatment can be considered randomly assigned in this neighborhood. We use Equation (10) estimate LCMs with 2, 3, 4 and 5 classes. We then compare AIC and BIC across these models to

[16]The reader is referred to Paap et al. (2005) for further details about the EM algorithm.

select the optimal number of classes. In our case, the 5-class model is preferred, and so we focus our discussion on these results. [17]

Table 10 presents the parameter estimates for the 5-class model. The mean posterior probability estimates, reported at the bottom of the table, characterize the probability that each household belongs to each class; this is also the proportion of the sample that we would expect to be in each class. Class 2 is the largest with 53.15% of the sample, followed by Class 1 with 18.30%, Class 3 with 12.46% and Class 5 with 11.40%. The residual 4.7% is in Class 4.

Turning to the DID estimates of interest, α_4 is significant and negative for all the classes except *Class 5*, implying that the implementation of ABR caused treatment households belonging to these classes to shift their target levels down. The other DID estimate, β_4, is significant for all the classes, taking negative values for *Class 4* and *Class 5*. This means the tariff change caused treatment households in these classes to converge to the target level faster, whereas it generated a decrease in the convergence rate for the remaining classes. We also find heterogeneity in the signs and magnitudes of the effects across classes for the remaining coefficients, making it difficult to provide clear intuition for the entire table of results. In order to gain some insights, we use these results to calculate the target levels and convergence rates for the different classes before and after the tariff change. Doing so also reveals whether treatment households consider the efficient level of water use as their target after the tariff change.

[17] The 6-class model did not converge.

	Class 1	Class 2	Class 3	Class 4	Class 5
Intercept (α_1)	-0.017***	-0.042***	-0.008***	-0.072***	-0.036***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Post (α_2)	-0.021***	-0.068***	-0.053***	0.027***	-0.159***
	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Treatment (α_3)	0.073***	0.053***	0.013***	0.069***	0.074***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Treatment x Post (α_4)	-0.022***	-0.007**	-0.047***	-0.052***	0.038***
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
Diff$_{t-1}$ (β_1)	-0.595***	-0.497***	-0.240***	-0.908***	-0.564***
	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Diff$_{t-1}$ x Post (β_2)	0.077***	0.147***	0.033***	0.633***	-0.203***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Diff$_{t-1}$ x Treatment (β_3)	-0.005	-0.059***	-0.114***	0.780***	0.116***
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Diff$_{t-1}$ x Post x Treat (β_4)	0.073***	0.058***	0.047***	-0.632***	-0.141***
	(0.011)	(0.009)	(0.011)	(0.011)	(0.011)
Mean posterior probabilities	0.1830	0.5315	0.1246	0.0470	0.1140

Standard errors in parentheses
***p<0.01

Table 10: Estimation results - LCM HH ±10%

Table 11 shows these calculated values. As in the homogeneous model, all of these values are statistically different from zero.[18] Households in the control group had target levels below their water budgets before the tariff change (indicated by negative values in row 1), across all classes. Moreover, control households experienced a downward shift in their target levels after the tariff change (larger negative values in row 3), except households in *Class 4* who were the most efficient pre-ABR and became relatively less efficient after the tariff change. Regarding the convergence rates for the control group, we observe a general decrease due to the tariff change (rows 2 and 4), with the exception of *Class 5*.

Households in the treatment group had target levels above their water budgets before the tariff change (indicated by positive values in row 5), with the exception of *Class 4*. After the tariff change, the targets for the treatment group decreased to efficient levels (negative values in row 7), with the exception of *Class 1*. This class was the most inefficient before the tariff change, and even though the distance between actual and efficient water use decreased, treatment households in this class remained inefficient after the tariff change. The adjustments by the other four classes did not lead to convergence to the budget, but rather to a level just below the budget. For instance, the water budget for treatment households belonging to the largest class, *Class 2*, after the tariff change is 19.74 ccf. Based on the estimated post-ABR target, the average level of water use for treatment households in this class is 18.49 ccf. This small difference is perhaps evidence of a desire to create a buffer against stochastic (accidental) inefficient consumption. Regarding the convergence rates for the treatment group, these decrease after the tariff change (with the exception of *Class 5*), similar to households in the control group.

In general, results from the heterogeneous model indicate an increase in water use efficiency across classes, for both treatment and control groups. However, the nature of the change varies across classes. In order to facilitate a comparison, we show the policy-induced

[18]For this analysis, households are assigned to a particular class based on the most likely class membership, that is, the highest posterior probability.

			Class 1	Class 2	Class 3	Class 4	Class 5
Control	Pre-ABR	Intercept	-0.017	-0.042	-0.008	-0.072	-0.036
		Convergence Rate	-0.595	-0.497	-0.240	-0.908	-0.564
	Post-ABR	Intercept	-0.038	-0.110	-0.061	-0.044	-0.196
		Convergence Rate	-0.518	-0.350	-0.207	-0.274	-0.767
Treatment	Pre-ABR	Intercept	0.056	0.010	0.006	-0.003	0.038
		Convergence Rate	-0.600	-0.556	-0.354	-0.128	-0.448
	Post-ABR	Intercept	0.013	-0.065	-0.094	-0.028	-0.084
		Convergence Rate	-0.450	-0.352	-0.274	-0.127	-0.792

Table 11: Convergence rates and target levels- HH $\pm 10\%$

changes in target levels and convergence rates for the five classes within both treatment and control groups in Table 12. Lines 1 and 3 show that Class 5 experienced the strongest increase in water use efficiency, for both treatment and control groups. Moreover, the convergence rate for this class increased after the tariff change (becoming more negative). While we do not have information to characterize class membership, we speculate that Class 5 may be representative of households adopting water-efficient technologies after the implementation of ABR, as their behavior is consistent with this type of structural change, i.e., a large, fast reduction in water use. Classes 1, 2 and 3 appear qualitatively different from Class 5 but similar to each other, and exhibit more subtle adjustments that happen more slowly. These classes may include households changing their water use habits instead of their installed technologies, as habit changes tend to be less effective than new technologies (Inskeep and Attari, 2014) and require more time to establish due to the learning process. Differences across these three classes may reflect differences in the type of habit being changed, such as indoor versus outdoor. Class 4 is the smallest group and, as is typical of LCM, appears to include unusual responses not well characterized by the other more intuitive groups. Unfortunately, our data set does not contain information on habits and technologies to test these speculations, but this would be a logical extension for future work.

		Class 1	Class 2	Class 3	Class 4	Class 5
Control	Pre-PostABR Difference intercept	-0.021	-0.068	-0.053	0.028	-0.160
	Pre-PostABR Difference convergence rate	0.077	0.147	0.033	0.634	-0.203
Treatment	Pre-PostABR Difference intercept	-0.043	-0.075	-0.100	-0.025	-0.122
	Pre-PostABR Difference convergence rate	0.150	0.204	0.080	0.001	-0.344

Table 12: Differences in target levels and converge rates - HH $\pm 10\%$

6 Conclusions and policy implications

This paper investigates the effects of allocation-based rates on inefficient consumers in the residential water sector. By exploiting a change in tariff from a uniform rate to ABR for a Southern California case study, we are able to use a RDD approach to assign households to treatment and control status based on their relative water use efficiency just prior to the rate change. Then, we combine a DID methodology with an empirical strategy similar to that used to estimate convergence in a time series framework. By doing so, we are able to characterize the gradual adaptation by inefficient customers to this new tariff.

Focusing on the households in the $\pm 10\%$ neighborhood of the treatment threshold, the results of our homogeneous model indicate that there was a general increase in water use efficiency after the change in tariff. Moreover, this increase was significantly larger for households that consumed water inefficiently prior to the tariff change (treatment households). In terms of gradual adaptation, we find evidence that inefficient households converge faster to their new target efficiency levels after the tariff change. Moreover, our falsification tests reject the notion that these effects are due to discrete increases in marginal price ("kink points") in the ABR structure, and instead suggest that the information signal provided by the water budget is driving the observed behavior.

When we allow for heterogeneity, we identify five different household classes. We find that the implementation of ABR generated an improvement in water use efficiency among treatment households for four of the estimated classes. The target levels for treatment households in these classes were positive (inefficient) prior to the tariff change and became negative (efficient) after the change. We also find evidence of heterogeneity in the target levels after the tariff change. While none of the new targets is statistically equal to zero, which would imply that consumers are targeting consumption at their water budgets, most of the classes have negative targets that are close to zero, implying these households are targeting water use levels just below their budgets. We also find that the tariff change had a heterogeneous effect on the convergence rates for treatment households in terms of both sign and magnitude. The convergence rates generally decrease, with the exception of the class experiencing the strongest downward shift in its target.

Our analysis suggests that ABR can be useful for policymakers wishing to improve water use efficiency. During the year prior to the tariff change, 26% of our sample households were consuming on average above their efficient level of water use. This percentage was reduced by more than half to 12% during the last year of our data. Our results also rebut concerns that ABR is too complex for customers to understand, suggesting instead that the efficiency signal provided by the water budget has a measurable effect on consumer behavior.

Policymakers would benefit from future work that investigates the defining characteristics of the latent consumer classes revealed by our study, so that future policies could be more effectively tailored to these different groups. Work that aims to understand the specific changes in habits and/or technologies implemented by different consumer classes to increase water use efficiency also would be helpful. If our speculative analysis is on track, then a large majority of households may be achieving improvements through habit change, while a much smaller percentage may be responding with new technologies. It would be particularly interesting to replicate our analysis with additional stated preference data on habits and technologies to better understand consumer class profiles and their adaptations to ABR, so that water providers can more effectively help their customers achieve lasting improvements in water use efficiency. Additional research analyzing how billing frequency may affect consumers' ability to respond to the signal provided by the water budget also would be valuable.

Appendices

A Difference in means tests

	Control	Treatment	t-test	p.value	Control	Treatment	t-test	p.value
		All HH				HH +/- 50%		
Income	307.80	323.45	-20.90	0.00	314.80	323.17	-10.38	0.00
Education	0.50	0.52	-6.49	0.00	0.50	0.52	-5.32	0.00
Irrigated area	3.11	2.66	19.56	0.00	3.05	2.73	12.84	0.00
Members	3.26	3.26	-0.19	0.85	3.32	3.28	2.28	0.02
No control households	7187				5143			
No treatment households	2454				2192			
		HH +/- 25%				HH +/- 10%		
Income	318.96	322.96	-3.96	0.00	320.46	321.99	-0.96	0.34
Education	0.51	0.51	-1.99	0.05	0.51	0.51	-0.36	0.72
Irrigated area	3.00	2.81	5.80	0.00	2.94	2.86	1.52	0.13
Members	3.35	3.30	1.97	0.05	3.36	3.30	1.47	0.14
No control households	2652				860			
No treatment households	1512				669			

Table A1: Difference in mean tests for observable covariates

	Control	Treatment	t-test	p.value	Control	Treatment	t-test	p.value
		1^{st} Block				3^{rd} Block		
Income	295.31	295.85	-0.16	0.87	324.91	329.46	-1.37	0.17
Education	0.50	0.49	1.12	0.26	0.53	0.52	0.59	0.55
Irrigated area	2.71	2.83	-1.00	0.32	2.51	2.17	3.23	0.00
Members	3.23	3.36	-1.66	0.10	3.23	3.15	1.05	0.30
No control households	207				274			
No treatment households	205				97			

Table A2: Difference in mean tests for observable covariates- Falsification test

B Panel unit root tests

		Control		Treatment	
		Pre-ABR	Post-ABR	Pre-ABR	Post-ABR
All HH	HT test	-449.3	-428.7	-247.3	-230.4
	p-value	0.000	0.000	0.000	0.000
	IPS test	-167.5	-175.0	-98.54	-92.08
	p-value	0.000	0.000	0.000	0.000
HH ± 50%	HT test	-374.7	-354.1	-236.1	-220.6
	p-value	0.000	0.000	0.000	0.000
	IPS test	-139.7	-144.5	-92.89	-87.62
	p-value	0.000	0.000	0.000	0.000
HH ± 25%	HT test	-270.6	-252.9	-194.7	-185.0
	p-value	0.000	0.000	0.000	0.000
	IPS test	-101.2	-102.5	-76.98	-73.98
	p-value	0.000	0.000	0.000	0.000
HH ± 10%	HT test	-155.7	-144.6	-130.9	-125.5
	p-value	0.000	0.000	0.000	0.000
	IPS test	-57.94	-57.68	-51.48	-49.82
	p-value	0.000	0.000	0.000	0.000

Table B1: Unit root tests

References

Allcott, H. (2011), 'Social norms and energy conservation', *Journal of Public Economics* **95**(9), 1082 – 1095.

Arbués, F. and Barberán, R. (2012), 'Tariffs for Urban Water Services in Spain: Household Size and Equity', *International Journal of Water Resources Development* **28**(1), 123–140.

Baerenklau, K. A., Schwabe, K. A. and Dinar, A. (2014), 'The residential water demand effect of increasing block rate water budgets', *Land Economics* **90**(4), 683–699.

Barro, R. J., Sala-I-Martin, X., Blanchard, O. J. and Hall, R. E. (1991), 'Convergence across states and regions', *Brookings Papers on Economic Activity* **1991**(1), 107–182.

Barro, R. and Sala-i Martin, X. (1992), 'Convergence', *Journal of Political Economy* **100**(2), 223–251.

Ben-David, D. (1993), 'Equalizing exchange: Trade liberalization and income convergence*', *The Quarterly Journal of Economics* **108**(3), 653–679.

Ben-David, D. (1996), 'Trade and convergence among countries', *Journal of International Economics* **40**(3), 279 – 298. Symposium on Growth and International Trade: Empirical Studies.

Bernard, A. B. and Durlauf, S. N. (1996), 'Interpreting tests of the convergence hypothesis', *Journal of Econometrics* **71**(1), 161 – 173.

Breitung, J. and Pesaran, M. H. (2008), *Unit Roots and Cointegration in Panels*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 279–322.

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

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R. (2018), *rdrobust: Robust Data-Driven Statistical Inference in Regression-Discontinuity Designs*. R package version 0.99.3.
URL: *https://CRAN.R-project.org/package=rdrobust*

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2015), 'Optimal data-driven regression discontinuity plots', *Journal of the American Statistical Association* **110**(512), 1753–1769.

Cameron, A. and Trivedi, P. (2005), *Microeconometrics: Methods and applications*, Cambridge University Press, New York.

CDWR (2010), 'California's Drought of 2007-2009: An Overview', *California Department of Water Resources (CDWR). Available at http://www.water.ca.gov/waterconditions/docs/DroughtReport2010.pdf* .

Centorrino, S. and Pérez-Urdiales, M. (2014), Convergence in the Spanish Regions: too little or too much European Structural Funds?, Technical Report 77, Fundación Alternativas, Estudios de Progreso.

Deb, P. and Vargas, C. (2016), Who benefits from calorie labeling? an analysis of its effects on body mass, Working Paper 21992, National Bureau of Economic Research.

Fernandez-Blanco, V., Orea, L. and Prieto-Rodriguez, J. (2009), 'Analyzing consumers heterogeneity and self-reported tastes: An approach consistent with the consumer's decision making process', *Journal of Economic Psychology* **30**(4), 622 – 633.

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

García-Rubio, M. A., Ruiz-Villaverde, A. and González-Gómez, F. (2015), 'Urban water tariffs in spain: What needs to be done?', *Water* **7**(4), 1456–1479.

Gleick, P. H. and Cain, N. L. (2004), *The World's Water 2004-2005: The Biennial Report on Freshwater Resources*, Vol. 2004-2005, Island Press.

Harris, R. D. and Tzavalis, E. (1999), 'Inference for unit roots in dynamic panels where the time dimension is fixed', *Journal of Econometrics* **91**(2), 201 – 226.

Im, K. S., Pesaran, M. and Shin, Y. (2003), 'Testing for unit roots in heterogeneous panels', *Journal of Econometrics* **115**(1), 53 – 74.

Inskeep, B. D. and Attari, S. Z. (2014), 'The water short list: The most effective actions u.s. households can take to curb water use', *Environment: Science and Policy for Sustainable Development* **56**(4), 4–15.

Lee, D. S. and Lemieux, T. (2010), 'Regression discontinuity designs in economics', *Journal of Economic Literature* **48**(2), 281–355.
URL: *http://www.jstor.org/stable/20778728*

Levin, A., Lin, C.-F. and Chu, C.-S. J. (2002), 'Unit root tests in panel data: asymptotic and finite-sample properties', *Journal of Econometrics* **108**(1), 1 – 24.

Mayer, P., De Oreo, W., Chesnutt, T. and Summers, L. (2008), 'Water budgets and rate structures: Innovative management tools', *Journal-American Water Works Association* **100**(5), 117–131.

Monteiro, H. and Roseta-Palma, C. (2011), 'Pricing for scarcity? an efficiency analysis of increasing block tariffs', *Water Resources Research* **47**(6).
URL: *http://dx.doi.org/10.1029/2010WR009200*

Nataraj, S. and Hanemann, W. M. (2011), 'Does marginal price matter? a regression discontinuity approach to estimating water demand', *Journal of Environmental Economics and Management* **61**(2), 198 – 212.

Nemati, M., Buck, S. and Soldati, H. (2018), 'Heterogeneous effects of real-time consumption analytics on residential water consumption', *Mimeo* .

Oldenburg, C. E., Moscoe, E. and Bärnighausen, T. (2016), 'Regression Discontinuity for Causal Effect Estimation in Epidemiology', *Current Epidemiology Reports* **3**(3), 233–241.

Paap, R., Franses, P. H. and van Dijk, D. (2005), 'Does africa grow slower than asia, latin america and the middle east? evidence from a new data-based classification method', *Journal of Development Economics* **77**(2), 553 – 570.

Pérez-Urdiales, M., García-Valiñas, M. A. and Martínez-Espiñeira, R. (2016), 'Responses to changes in domestic water tariff structures: A latent class analysis on household-level data from granada, spain', *Environmental and Resource Economics* **63**(1), 167–191.

Pinto, F. S. and Marques, R. C. (2015), 'Tariff recommendations: A Panacea for the Portuguese water sector?', *Utilities Policy* **34**, 36 – 44.

R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Sebri, M. (2014), 'A meta-analysis of residential water demand studies', *Environment, Development and Sustainability: A Multidisciplinary Approach to the Theory and Practice of Sustainable Development* **16**(3), 499–520.

Slaughter, M. J. (2001), 'Trade liberalization and per capita income convergence: a difference-in-differences analysis', *Journal of International Economics* **55**(1), 203 – 228.

U.S. Bureau of the Census (2014), 'County-to-County Migration Flows: 2007-2011', *Available at http://www.census.gov/topics/population/migration/guidance/county-to-county-migration-flows.html* .

Wooldridge, J. (2007), 'What's new in econometrics? lecture 10 difference-in-differences estimation', *NBER Summer Institute, available at: www. nber. org/WNE/Slides7–31–07/slides_10_diffindiffs. pdf, accessed December 2016* **9**, 2011.

Zetland, D. and Gasson, C. (2013), 'A global survey of urban water tariffs: are they sustainable, efficient and fair?', *International Journal of Water Resources Development* **29**(3), 327–342.