Man vs. Machine :
Technological Promise and Political Limits of Automated Regulation Enforcement

Oliver R. Browne, Ludovica Gazze, Michael Greenstone & Olga Rostapshova

(This paper also appears as CAGE Discussion paper 646)

December 2022 No: 1440

Warwick Economics Research Papers

ISSN 2059-4283 (online)
ISSN 0083-7350 (print)

Man vs. Machine: Technological Promise and Political Limits of Automated Regulation Enforcement *

Oliver R. Browne, Ludovica Gazze, Michael Greenstone, Olga Rostapshova

September 28, 2022

Abstract

New technologies allow perfect detection of environmental violations at near-zero marginal cost, but take-up is low. We conducted a field experiment to evaluate enforcement of water conservation rules with smart meters in Fresno, CA. Households were randomly assigned combinations of enforcement method (automated or in-person inspections) and fines. Automated enforcement increased households' punishment rates from 0.1 to 14%, decreased water use by 3%, and reduced violations by 17%, while higher fine levels had little effect. However, automated enforcement also increased customer complaints by 1,102%, ultimately causing its cancellation and highlighting that political considerations limit technological solutions to enforcement challenges.

Keywords— Field Experiment; Automated Enforcement; Remote Sensing; Water Conservation
JEL CODES: Q25; K42.

*Contact: Oliver R. Browne: oliver.browne@brattle.com; Ludovica Gazze: ludovica.gazze@warwick.ac.uk; Michael Greenstone: mgreenst@uchicago.edu; Olga Rostapshova: oro@uchicago.edu. We thank Valentina Bosetti, Stefano Caria, Tatyana Deryugina, Clement Imbert, Koichiro Ito, Eliana La Ferrara, Ashley Langer, Leslie Martin, Erica Meyers, Louis Preonas, and participants at TWEEDS, Midwest Energy Fest, AERE, AFE, ASSA-AERE, BSE Insights Workshop, 13th Conference on the Economics of Energy and Climate, and seminars at Bocconi University, Bologna University and and University of Roma Tor Vergata. Iris Song, Jackson Reimer, Aditya Jain, Leon Beaufils, Monica Kim, and Tejomay Gadgil provided excellent research assistance. Our partners in Fresno provided endless support in implementing the experiment. We are grateful to the Laura and John Arnold Foundation and the Smith Richardson Foundation for their generous research support. This experiment was pre-registered on June 14th, 2018 on the Open Science Framework Preregistry with DOI: "10.17605/OSF.IO/XYFU3" and in the AEA RCT Registry with DOI: "10.1257/rct.8290-1.0". All remaining errors are our own.

1 Introduction

New technologies are often hailed as the solution to governments' failures to achieve compliance with laws and regulations in environmental (Alm and Shimshack, 2014; Duflo, Greenstone, Pande, and Ryan, 2013; Gibson, 2019; Reynaert and Sallee, 2021; Vollaard, 2017; Zou, 2021), tax collection (Sarin and Summers, 2020), and many other domains. These technologies drive the marginal cost of monitoring to virtually zero, thus increasing the probability of detection — a key parameter in Becker's (1968) canonical model of crime; they include speed cameras, biometric cards, automatic air pollution monitors, and satellites, and have inhibited speeding, corruption, pollution, deforestation, and tax evasion (Dušek and Traxler, 2022; Muralidharan and Sukhtankar, 2021; Greenstone, He, Jia, and Liu, 2020; Ferreira, 2021; Casaburi and Troiano, 2016). Indeed, a growing literature emphasizes that lowering reliable monitoring costs is key to significantly improving compliance with environmental and workplace regulations, as well as bill payment (Duflo, Greenstone, Pande, and Ryan, 2018; Banerjee, Duflo, and Glennerster, 2008; Meeks, Omuraliev, Isaev, and Wang, 2020). Yet, policymakers frequently choose not to adopt these technologies and when they do, they use them inconsistently. Some consider this a puzzle, while others point to political costs, such as an erosion of political capital (e.g., Brollo, Kaufmann, and La Ferrara, 2019).

This paper provides a rare opportunity to study both the benefits of one such technology *and* the political costs of its resulting perfect detection of violations. The context is the use of "smart meters" to enforce residential outdoor water use regulations in Fresno, CA. These regulations help cities cope with increasing drought conditions due to climate change (Diffenbaugh, Swain, and Touma, 2020). Utilities typically do not price water at marginal social cost, for political and ethical reasons and instead rely on non-price mechanisms to manage consumption (Brent, Cook, and Olsen, 2015; Brent and Ward, 2019). Residential outdoor watering restrictions are a primary mechanism because lawn irrigation is the single largest end-use of residential water (Hanak and Davis, 2006). Fresno like almost all other cities with outdoor watering restrictions, has relied on "water cops" to monitor compliance with these restrictions, despite having smart meters since 2013. Yet, violations were rampant and punishments were rare: 68% of households violated these restrictions at least once in the summer of 2016, yet only 0.4% of violations were sanctioned (Table 1).[1]

We implemented a randomized field experiment with the nearly 100,000 households in Fresno, introducing automated enforcement of outdoor water use restrictions via smart meters that enabled perfect detection of violations. Fresno was one of the first large municipal water utilities with universal smart meter adoption among single-family residential customers. We worked with the city to design and implement an evaluation that experimentally varied both the enforcement method — whether households were newly subject to automated enforcement via smart meters or continued the status quo of in-person inspections by water cops (i.e., detection rates of 100% versus 0.4%) — and the magnitude of the fines that violating households faced. For households in the automated group, the experiment also varied the 'excessive water use' threshold that triggers warnings and fines. The experiment took place between July and September 2018, during peak outdoor watering season.

[1] We define violations as water use above 300 gallons/hour during prohibited hours. This is the threshold the city eventually selected for automated enforcement.

We measured treatment effects both on water use and compliance using continuous smart meter data and on the political fallout using call data. To measure the latter, we collected information on customer phone calls to the city's Department of Public Utilities (DPU). More broadly, the experiment provided a perhaps unprecedented opportunity to vary the key parameters – probability of detection and penalty - that determine the cost of committing a crime in Becker's model, and empirically estimate the effects of these parameters on violations and backlash against the program.

There are three primary findings. First, automated enforcement greatly increased enforcement as well as compliance with the law. While the share of households fined for non-compliance grew from 0.1 to 14.3%, the increase in fines is the net effect of the effects of automated enforcement on compliance and on detection of violations. The improved enforcement reduced violations by 17% and violating households by 8% per month. Second, automated enforcement decreased summer water consumption by about 3%. If scaled citywide, the 3-month experiment would have achieved 20% of the annual reductions in residential water use that Governor Gavin Newsom requested of California residents on July 8, 2021, in response to another drought. The automated enforcement treatment also reduced water consumption after the pilot, suggesting even larger potential water conservation effects. Third, automated enforcement created political backlash that ultimately led to the program's termination, as the number of households calling the utility increased by 654% and calls identifiable as complaints and disputes of enforcement actions increased by 1,102%.

There are several other findings. Lower fines did not affect the frequency of violations, water consumption, or customer complaints. Moreover, the percentage effect of automated enforcement on water conservation was roughly constant across the distributions of baseline water consumption and income. However, heavy water users and the wealthy complained more frequently. Finally, warnings and fines caused immediate reductions in water consumption and increases in customer complaints. Together, these findings show that households understood the mechanics of automated enforcement and adjusted their behavior accordingly.

Yet, despite attempts to be customer-friendly (e.g., grace periods implicit in the fine schedule), the automated enforcement program ultimately did not survive. Public backlash, including customer calls, led the city to implement a fine moratorium, weaken the conservation rules, and finally institute new rules that essentially outlawed automated enforcement of violations. In practice, the city returned to relying on water cops and in-person inspections. This experience serves as a cautionary tale about the limits of new technologies to solve compliance problems and underscores the need for research to identify the settings where they can succeed.

This research contributes to three main strands of the literature. First, a growing body of work discusses the difficulties in successfully enforcing environmental regulations globally and the resulting flouting of laws and standards (Duflo, Greenstone, Pande, and Ryan, 2013; Duflo, Greenstone, Pande, and Ryan, 2018; Reynaert and Sallee, 2021; Gibson, 2019). A related literature explores the promises and failures of new technologies for monitoring compliance with environmental and workplace regulations (Zou, 2021; Fowlie, Rubin, and Walker, 2019; Mu, Rubin, and Zou, 2022; Banerjee and Duflo, 2006; Dhaliwal and Hanna, 2017). Our findings also highlight the importance of designing processes that maximize public support for adoption of new technology, in line with related literature (Blumenstock, Callen, Faikina, and Fiorin, 2022; Bossuroy, Delavallade, and Pons, 2019; Muralidharan and Sukhtankar, 2021; Atkin et al., 2017). Second, the paper adds to an extensive literature on crime that finds that people respond to expected future punishment (Bar-Ilan and Sacerdote, 2004; Drago, Galbiati, and Vertova, 2009), the perceived auditing probability (Kleven et al., 2011), and past punishment (Kuziemko, 2013; Maurin and Ouss, 2009; Haselhuhn, Pope, Schweitzer, and Fishman, 2011; Dušek and Traxler, 2022). Third, it contributes to a growing literature studying water conservation policies, including nudges, peer comparisons, monitoring and enforcement of outdoor water use restrictions, and combinations of policy instruments (Jessoe, Lade, Loge, and Spang, 2021; Pratt, 2019; Wichman, Taylor, and Haefen, 2016; Halich and Stephenson, 2009; Kenney, Klein, and Clark, 2004; Renwick and Green, 2000; Michelsen, McGuckin, and Stumpf, 1999; Hahn, Metcalfe, Novgorodsky, and Price, 2016).[2]

2 Experiment and Data

2.1 Experiment

We partnered with Fresno, California's fifth largest city, which by 2013 installed smart water meters at all 114,508 single-family households. These smart meters measure water consumption and transmit data every fifteen minutes, but do not allow for water use scheduling. Prior to our experiment, Fresno was using smart meters for billing and leak detection.

Water meters are controversial. Fresno residents fiercely resisted the implementation of smart meters in private homes. In 2006, the San Joaquin Taxpayers Association filed a lawsuit to prevent the installation of residential water meters, which was dismissed. Four years later, Fresno's city charter prohibited the use of smart meters for billing even while allowing for installation. The city finally agreed to metered residential water billing when state and federal authorities threatened to withhold the city's deliveries of Central Valley Project water.

Fresno has had summer outdoor watering restrictions since the mid-1990s, restricting outdoor watering to three nights per week. To detect violations of these restrictions, Fresno had five part-time water cops who issued 3,113 fines in 2016.[3] To limit fine burden and political discontent, Fresno only sanctioned the first violation in a month. For the first, second, and third month with violations, households were charged fines of \$0, \$50, \$100, respectively (baseline schedule). The fine levels for the second and third month in violation are comparable to a monthly water bill, which averaged \$79.29 in the summer of 2017. Households were notified of violations within days, and fines accrued on the following month's water bill. Fresno also offered "water audits" and "timer tutorials" to evaluate potential leaks and reset automated lawn sprinkler timers to comply with the watering schedule.

[2] The most comparable study is West, Fairlie, Pratt, and Rose (2021) which examines a similar excessive-use-threshold policy that was announced but never went into place. Without prior announcement, the city identified households that would have been in violation during one earlier week had the policy been in place, and notified them of this hypothetical violation and of the fine schedule. These households, about a third of accounts, reduced water use by 31% thereafter. In contrast, our study evaluates the citywide, average, treatment effects of an enacted policy, that is the policy-relevant parameter.

[3] Authors' calculations based on most recent violation data available (Browne, Gazze, and Greenstone, 2021).

We worked with city officials to implement a randomized field experiment, conducted between July and September 2018, to inform the citywide implementation of automated enforcement and the new fine schedule. We randomly assigned all single-family households in the city to one of twelve experimental groups, varying along two cross-randomized dimensions: 1) enforcement method: automated vs. in-person detection of violations, and 2) the schedule of fines. Because smart meters do not distinguish between outdoor and indoor water use, automated enforcement was based on a definition of 'excessive water use' above which the household is presumed to be using water outdoors. The city's regulation stipulated an 'excessive water use' threshold of 300 gallons/hour. We further randomized households in the automated group into one of three 'excessive water use' thresholds: 300, 500, or 700 gallons/hour. In terms of fines, households could either face the baseline schedule, or fines at 50% and 25% of the baseline levels.[4]

The city announced the program as a pilot in June 2018 through a media campaign. Each household received a mailer explaining the upcoming three-month pilot, during which each home was assigned an enforcement mechanism via a lottery system (Figure A1 presents an example). The mailer announced the recipient's assigned method and fine schedule. If a violation was detected, the city sent the household a letter within days specifying the time and magnitude (i.e., hourly water use in case of automated enforcement) of the violation, the detection method (including the excessive use threshold that the household faced), and the fine schedule (Figure A2 presents an example). Fines, if any, were charged on the following month's water bills. The first notifications were sent on July 18, 2018. The city did not sanction violations between August 1 and August 12, 2018, to enable its customer service to catch up with the backlog of customer calls received.[5] Figure A3 overlays the experimental timeline with trends in water use (Panel A) and calls (Panel B) in the automated and non-automated groups.

We underscore that this experiment is designed to infer the short-term effects of automated enforcement and fine levels because it only ran for three months. We conjecture that the impacts of a permanent program could be larger. For example, anticipation that the program would not be scaled up might have prevented households' adoption of technologies like sprinkler timers or lawn-to-turf conversions. Moreover, if the policy had been in place for longer, it might have become more effective over time, as more households learned about it and had more time to adjust behavior. Finally, we cannot exclude that the fine amounts were less salient than other features of the experiment as fines were only levied for two out of three experimental months and accrued on the subsequent month's water bill (although households were notified within days of the violation).

2.2 Data

Starting with the Fresno's population of 114,508 single-family residential households, we restricted the experimental sample in three ways. First, we include only households with positive water use under 216,000 gallons/month in April 2017, and exclude households that have had their water shut off or used more than 300 gallons/hour on average. Second, we excluded accounts that could not be matched to a single-family parcel in the assessor files, which we had

[4]Outreach materials only stated the assigned fine schedule without referencing the baseline schedule.

[5]The city introduced another fine moratorium on October 1, 2018 which was eventually lifted in May 2020 without fines being paid.

linked with data from the American Community Survey (years 2010-2014) using Census block group identifiers, because median household income at the Census block group is a variable of interest for heterogeneity analysis.[6] Third, we excluded households who changed their street address in May 2018. These restrictions dropped 24,927 households. We exclude an additional 677 households with missing water use data. Our final analysis sample consists of 88,904 households, observed on average 87 of the 92 pilot days, with the occasional smart meter malfunction preventing a balanced panel.

To log customer calls to the Department of Public Utilities, we hired representatives to staff a dedicated phone line and categorize incoming calls into forty mutually exclusive groups based on reason for the call. The calls fell into four broad categories: complaints and disputes, service requests, opt outs, and "other". Table A1 shows the share of calls in each group received between July and September 2018.

The smart meter data runs from January 2017 through February 2019. The call data covers the June 2018 through February 2019 period.

Table 1 reports on the randomization. We allocated 45% of our sample to the non-automated, baseline fine group, the control group (Panel A).[7] Each of the other eleven treatment groups includes 5% of sample households. Panel B of Table 1 reports differences between these groups and the control group. The treatment groups appear balanced in terms of baseline water use, violation, and clearance rate. Columns 1 and 2 of Table 1 report the randomization sample and analysis sample group sizes, respectively.

In 2017, the year before the experiment, control households used less than 600 gallons/day on average (Column 3). The typical household exceeded the 300 gallon/hour limit (which went into effect in 2018) by about 28 gallons/hour (Column 4) for about 0.141 hours per day or 1 hour per week (Column 5). Thus, there would have been about 12,000 violations per day if the excessive water use threshold had been in force in 2017.

Columns 6 and 7 explore the degree of compliance with water restrictions and the frequency of punishment using the most recent available enforcement data from 2016. Strikingly, 68.3% of households exceeded 300 gallons/hour when outdoor watering was illegal at least once during the summer of 2016. Yet, the water cops only issued violations to 0.4% of households.

To study potential heterogeneous effects of different treatments, we stratified the randomization based on being above or below the median of 1) baseline water use during April 2017 and 2) median income of the Census block group (Table 1, Columns 8 and 9). We additionally stratified the randomization by city council district. Finally, Column 10 reports opt-out rates by treatment group. Opt-outs would become subject to the 'harshest' automated enforcement group, at baseline fine and 300 gallon/hour threshold. 0.5% of households opted out, and these had higher baseline water use and violation rates (Table A2). Because treatment groups saw higher dropout rates, we present Intention-To-Treat (ITT) treatment effects that include opt-outs in their randomly assigned groups.

[6]The average number of sample households per Census block group is 287.

[7]We based the allocation across automated and non-automated groups on the city's capacity to handle calls.

3 Empirical Analysis

We assess the impact of the experimental treatments on compliance, water use, and customer calls. First, we estimate the average effects of automated enforcement across treatment groups. Second, we explore heterogeneity, including differences in treatment responses across households with different characteristics and responses to different thresholds and fine treatments. Third, we investigate how violating households respond to enforcement actions.

3.1 The Effects of Automated Enforcement

To study the effects of automated enforcement on compliance, water use, and customer calls, we pool all automated groups to estimate the following regression equation for the months July-September 2018:

$$y_{it} = \alpha + \beta \text{Automated}_i + \sum_{j \in \{25,50\}} \gamma_j \text{In-Person} \times \text{Fines}_i^j + \varepsilon_{it} \tag{1}$$

where y_{it} is an outcome for household i in month t, Automated$_i$ is an indicator for household i's assignment to an automated enforcement treatment, and In-Person×Fines$_i^{25}$ and In-Person×Fines$_i^{50}$ are indicators for household i's assignment to in-person enforcement treatments with fines at 25% and 50% level relative to the baseline schedule. The automated enforcement indicator restricts the effect of the nine automated treatments to be equal. Due to random assignment, the OLS estimator $\hat{\beta}$ captures the causal effect of automated enforcement. We cluster standard errors at the level of randomization: household. For our main results, we also report Westfall-Young stepdown adjusted p-values that correct for multiple hypotheses testing (Jones, Molitor, and Reif, 2019).

Table 2 presents estimates from equation (1), reporting the average differences in outcomes between the automated enforcement treatment groups and the control group during the pilot. Columns 1a-1e show effects on enforcement actions and compliance behavior, Column 2 details effects on water use, and Columns 3a-3b document the costs of automated enforcement in terms of customer calls, which the city interpreted as a measure of discontent with the policy. The in-person inspection treatment effects are reported in Table A3. Panel A of Table A4 probes the robustness of the results controlling for baseline water use data from summer 2017 and Panel B includes household and month fixed effects, as specified in our pre-analysis plan.[8]

The automated enforcement treatment increased compliance with the water conservation policy through its reliable detection of violations. The average number of violations per month decreased from 3.7 to 3.1 (Column 1a) and the share of households that exceeded 300 gallon/hour at least once in a month from 51% to 47% (Column 1b). Thus, violations decreased by 17% and 8% fewer households violated in any given month.[9]

Although violations remained high in the treatment group, the detection and punishment of violations increased dramatically. Households in the automated enforcement groups were more likely to receive warnings and fines, by

[8] Table A5 reports estimates and standard errors that account for the covariate-adaptive stratified randomization following Bugni, Canay, and Shaikh (2019).

[9] Table A6 estimates the difference in the number of months with violations across automated and control households. Households in the automated group were 3% less likely to ever violate and almost 20% less likely to violate all three pilot months.

1,715% and 14,200% respectively (Table 2, Columns 1c-1d) and paid $7.43 more per month (Column 1e), 9% of an average monthly bill. If the pilot were scaled citywide and the treatment effects remained constant, 16,374 households would be fined per month, and Fresno would collect $2.55 million (=$850,622*3) in fines over the summer.

Further, automated enforcement decreased water consumption by 2.9% and these estimates are precise and robust (Table 2, Column 2).[10] This decline was partly driven by decreases in heavy consumption hours: automated enforcement reduced the number of hours with use between 500-699 gallons and above 700 gallons by 19% and 25%, respectively (Table A8).[11]

Panel A of Figure 1 reports the results from estimating equation (1) but allows the treatment effect to vary for each month of the sample, including the months before (when we would expect no effect) and after the pilot. The treatment effect increased from a water consumption reduction of 1.5% in July, when fines had not been issued yet, to 4% in September, when 17.1% of households had received a warning or a fine. When the outcome variable is an indicator for whether a household had at least one violation in a month, the treatment effect increased over the course of the summer from 1.2% in July to 5.9% in September (Panel B). These findings are consistent with the possibility that households' knowledge of the new enforcement program increased over time and with the possibility that households decreased their water consumption after being sanctioned. Section 3.3 explores the latter possibility further.[12]

Notably, Panel A of Figure 1 shows statistically significant declines in water use in October-November 2018 (-292.3 gallon/month per treated household), even after the summer water restrictions were removed. Further, the treatment effect is still evident but smaller in December through February 2019 (when our data ends): -50.1 gallon/month per household on average with a standard error of 33.4. This finding of a conservation treatment effect that lasts beyond the treatment was also found in recent work about energy consumption in Japan (Ito, Ida, and Tanaka, 2018) and in Brazil (Costa and Gerard, 2021).

Keeping in mind *annual* water reductions mandated by the state of California, our results suggest the policy's effect on conservation extended beyond the three months when the experiment ran. The summer estimates indicate that scaling up the policy citywide would save an estimated 174 million gallons of water during the three months the pilot ran. Assuming the same effects would manifest throughout the seven summer months when outdoor water restrictions are in place, we estimate savings of 334 million gallons of water per summer. Moreover, including post-pilot effects on conservation in winter and assuming persistence beyond February at the same levels, the program would save an estimated 61 million gallons in the five winter months, leading to a total of 394 million gallons of water annually.

Columns 3a-3b of Table 2 report on the effect of automated enforcement on calls to DPU. Automated enforcement

[10] Table A7 investigates peer effects of automated enforcement. We find no evidence that households in Census blocks with a higher proportion of households in automated enforcement disproportionately reduced water use.

[11] Fine levels appear to have inconsistent effects in the case of non-automated enforcement. Table A3 reveals that the group assigned to the halved fine schedule decreased water consumption, while the group assigned to the 25% fine schedule increased water consumption by a statistically insignificant amount, potentially due to a slight imbalance in the randomization. These estimates are not statistically significant after correcting for multiple hypotheses testing.

[12] Treatment households do not appear to shift water consumption from banned to permitted hours (Table A9).

increased the monthly count of customers who called at least once by 654% (Table 2, Column 3a), and increased the monthly count of customers who called with a complaint or to dispute an enforcement action at least once by 1,102% (Table 2, Column 3b). Even conditional on receiving a violation warning, households in the automated enforcement group were 56% (6.3 percentage points) more likely to call DPU than control households (Figure A4). Automated enforcement generated 1,747 additional calls to DPU during the three-month experiment. If the city were to scale up automated enforcement citywide, we estimate that it would lead to 4,090 additional calls over the summer period. The political costs of these calls were substantial, leading, for example, to an enforcement moratorium in early August to allow DPU to catch up with the call backlog. Moreover, our partners reported that many customers also called City Council members complaining about the new system, which ultimately halted the scale-up of the automated enforcement program in 2019. Some city council members raised concerns about "Big Brother" types of policies when discussing the automated enforcement pilot.

Panel C of Figure 1 is constructed like Panels A and B, but using households' call probability as the outcome. The automated enforcement treatment caused a sharp increase in calls during the pilot. This treatment effect had disappeared by December 2018, with the lag presumably due to households' lag in receiving and responding to bills.

3.2 Heterogeneous Effects

In this section, we examine whether responses vary across households with differing characteristics and assess the impact of variation in fine levels and 'excessive water use' thresholds in explaining the impacts of automated enforcement.

First, we test whether automated enforcement had heterogeneous treatment effects across the two stratification variables, income and baseline use, as well as across baseline propensity to use excessive water. Specifically, we estimate a version of equation (1) that adds interactions of the automated indicator with indicators for each decile of the characteristic of interest (except the 5th) and includes deciles fixed effects. Thus, the coefficients on these interactions measure whether the treatment effect varies for these deciles relative to the 5th.

The causal effect of automated enforcement on water consumption did not vary proportionally to baseline use (Figure 2, Panel A). In contrast, the effect on the propensity to call sharply increased with baseline use. Panels B and C similarly show that the treatment effect on water use was relatively homogeneous across income and violation levels, but the wealthy and likely violators were responsible for a disproportionate share of the increase in calls and presumably of the political resistance to automated enforcement.

Second, we exploit experimental variation in fines and thresholds within the automated enforcement treatment to examine their influence on water use and propensity to call DPU. To do this, we estimate the following equation:

$$\begin{aligned} y_{it} = {} & \alpha + \beta_1 \text{Automated}_i + \beta_2 \text{Automated500}_i + \beta_3 \text{Automated700}_i + \\ & \beta_4 \text{Automated} \times \text{Fines50\%}_i + \beta_5 \text{Automated} \times \text{Fines25\%}_i + \\ & \sum_{j \in \{25,50\}} \gamma_j \text{In-Person} \times \text{Fines}_i^j + \varepsilon_{it} \end{aligned} \tag{2}$$

that characterizes the nine automated enforcement treatments with five indicators. Specifically, the Automated_i indicator captures the effect of the city's default automated policy which had a 300 gallon/hour threshold and the standard fine schedule. The four other indicators measure the effects of increasing the threshold to 500 or 700 gallon/hour or reducing the fine to 50% or 25% of the standard schedule.[13] We expect that higher excessive use thresholds and lower fines lead to less conservation and fewer calls.

Figure 2, Panel D plots estimates of β_2, β_3, β_4, and β_5, thereby displaying the effect of deviations from Fresno's default automated policy (which reduced water use by 6.3% and increased the probability of calling by 2 percentage points). Increases in the excessive use threshold monotonically decreased the treatment's effect on water conservation and the probability of calling. It seems reasonable to conclude that households understood this design feature of the policy.

In contrast, there is little evidence that reductions in fines affected water conservation or the call probability in an economically meaningful way.[14] Our previous work using 2013-2016 data from Fresno estimates that a 1% increase in marginal water rates leads to a decrease in water use of 19% (Browne, Gazze, and Greenstone, 2021). However, a direct comparison with this finding is challenging and probably inappropriate, because the fines are for discrete events (i.e., exceeding the hourly water consumption thresholds), rather than a tariff on all water consumption.

3.3 Households' Behavior after Warnings and Fines

In an unconstrained setting, we would have also experimentally assigned enforcement actions in response to violations of the excessive use threshold to learn their contribution to the treatment effects for our outcomes of interest. As an imperfect substitute, we conduct "event study"-style analyses of the impact of receiving warnings and fines. In the economics of crime literature, the response to the *application* of punishments is called the "specific deterrence" effect and conceptually contrasts with the *overall* effect of a change in enforcement actions that also includes changes in the probability of a violation, i.e., "general deterrence" (Glueck, 1928). A large literature finds important specific deterrence effects in other settings (Kuziemko, 2013; Maurin and Ouss, 2009; Haselhuhn, Pope, Schweitzer, and Fishman, 2011; Dušek and Traxler, 2022).

We exploit across-household variation in the timing of violations in an event-study design with household and

[13] Table A10 reports the coefficients on the fully specified model, that is:

$$y_{it}=\sum_{z\in\{300,500,700\}}\sum_{j\in\{25,50,100\}}\beta_{zj}\text{Auto}_z^j+\sum_{j\in\{25,50\}}\gamma_j\text{In-Person}\times\text{Fines}_i^j+\varepsilon_{it} \quad (3)$$

The table also reports the p-value of a test of the difference between the restricted model in equation (2) and the full model in equation (3). In other words, we test H_0 : $\hat{\beta}_{300,100}-\hat{\beta}_{300,50}=\hat{\beta}_{500,100}-\hat{\beta}_{500,50}=\hat{\beta}_{700,100}-\hat{\beta}_{700,50}$ and $\hat{\beta}_{300,100}-\hat{\beta}_{300,25}=\hat{\beta}_{500,100}-\hat{\beta}_{500,25}=\hat{\beta}_{700,100}-\hat{\beta}_{700,25}$ by computing $F=(SS_{red}/SS_{full})$ for household-month level regressions. For a level of significance α, we reject H_0 if F is larger than the upper $1-\alpha$ percentile in the $F(N_{clusters}-1, N_{clusters}-1)$ distribution. SS_{red} and SS_{full} are the residual sums of squares from the parsimonious and the full specification, respectively and $N_{clusters}$ is the number of clusters. For Columns 1c-1d, we use the formula: $F=(SS_{redfull}/s)/(SS_{full}/df_{full})$ where $s=df_{redfull}$, df_{red} and df_{full} are the degrees of freedom from the parsimonious and full model, and we use the $F(s, df_{full})$ distribution. Based on this test and α=0.1, we cannot reject the null hypotheses that the fine-threshold interactions do not matter but for Columns 1c-1d.

[14] Albeit small, the coefficient on the effect of a 25% fine on the call probability is statistically significant at conventional levels.

week fixed effects to estimate the effect of enforcement actions on water use and probability of calling DPU. This specification assumes that after controlling for time-invariant propensity to violate, the exact timing of the violations is as good as random. Specifically, we estimate the following equation:

$$y_{it}=\alpha+\sum_{j=-12}^{12}\sum_{a\in\{1,2,3\}}\beta_j^a I_{it}(j \text{ Weeks Post Violation } a)+\gamma_i+\gamma_t+\varepsilon_{it} \quad (4)$$

where y_{it} is an outcome for household i in week t, $\mathrm{I}_{it}(j$ Weeks Post Violation $a)$ is an indicator for week t being j weeks before or after household i received an enforcement action (with the indicators for weeks $j=-12$ and $j=12$ equaling 1 for weeks -12 and earlier and 12 and later, respectively), and γ_i and γ_t are household and week fixed effects respectively. $\mathrm{I}_{it}(j$ Weeks Post Violation $a)$ is constructed for Warning, Fine 1 and Fine 2 and equals zero for households that did not receive that enforcement action. We have at most ten weeks after the second fine. The estimated β_j^as associated with these indicators capture households' response to enforcement actions over time, where time is indexed relative to the violation. In contrast to the treatment effects in Sections 3.1 and 3.2, these estimates are not based on experimental variation and therefore not guaranteed to be causal.

Panel A of Figure 3 reports estimates of equation (4) where the outcome variable is the logarithm of daily water use, averaged over each week. Water use decreases sharply nearly immediately after each enforcement action: within two weeks of the action's mailing, water use falls below the levels that prevailed a month before the violation.[15] In each case, however, there are large relative increases in water consumption in the few weeks immediately preceding the enforcement action; these increases are consistent with several explanations that might naturally lead to reversion to prior demand levels, including a temporary increase in demand due to a special event or a leak, underscoring the limits of this non-experimental analysis. It seems reasonable to conclude that specific deterrence from enforcement actions is important in this setting and spurs households to make lasting changes in behavior or physical investments that influence consumption at least three months later. It is more difficult to confidently parse out the effect of these actions beyond the first few weeks, because whether and how quickly such reductions would have occurred in the absence of the enforcement actions is less clear.

Panel B of Figure 3 reports estimates of equation (4) where the outcome is an indicator that equals one if the household called DPU that week. The enforcement actions lead to large increases in the probability of a call, especially so for fines. The call rates remain elevated several weeks after fines. As with water consumption, there is evidence of an upward pre-trend in call rates (potentially for service), so here too we conclude that the notifications likely caused an immediate increase in calls but their influence on calls throughout the pilot cannot be conclusively isolated.[16][17]

[15]The sharp decreases in water use when notices of violation are sent to households for second and third violations but before fines accrue suggest at least some of the behavior change is driven by the notices rather than the actual fine.

[16]Figure A5 estimates equation (4) on the subsample of households in the automated group and finds virtually identical results as most warnings and fines were received by these households.

[17]Figures A6 and A7 explore trajectories of water use and calls for users who receive 1,2,3 enforcement actions, and users who first received an enforcement action in month 1, 2, 3 of the experiment. Figure A6 reveals that the bulk of the calls in the automated group are from one-time violators (Panel A) while most calls in the non-automated group are from households who never violate (Panel B). Figure A7 shows that calls in the automated group increase sharply in the month households first violate (Panel A), which is not the case for the non-automated group (Panel B). In contrast, water use shows less clear patterns, likely due to a combination of increasing pre-trends and behavior change highlighted in Figure 3.

We explore the contribution of enforcement actions to the overall treatment effects in Appendix B, which decomposes the overall effect of automated enforcement on water consumption by violation status, before and after violations. The associational evidence in table B1 suggests that most of the reduction in water consumption comes from the nearly 64% of households that never violated, but enforcement actions reduced water consumption of violating households. Together with the results in this section, these findings are consistent with the idea that both the threat of penalties and the penalties themselves produced the automated enforcement group's reduction in water consumption.

4 Conclusion

This paper presents results from what we believe is the first field experiment to study the impact of automating the enforcement of local environmental regulations. To our knowledge, it is also the first experiment to randomize both detection methods and sanctions, and the first to do so in a context where compliance is perfectly observed, and on a representative population at the city level.

We study the use of smart meters to enforce outdoor water-use restrictions. We find that automated enforcement increased the punishment of violations, decreased violations by 17%, and reduced water consumption by 3% during the pilot (with evidence of longer-lasting conservation impacts). However, these benefits came with substantial costs to Fresno city government as calls (mostly complaints and disputes) to the city increased by 654%.

Despite the gains in water conservation, the political backlash caused Fresno to reverse its plan to scale automated enforcement of water use regulations citywide. First, the city enacted a fine moratorium the day after the pilot concluded which remained in effect until May 2020. Second, in April 2019 Fresno's Council voted unanimously to 1) lower the maximum financial penalty from \$200 to \$100, 2) increase the permitted hours of outdoor water use, 3) relax the excessive water use threshold from 300 to 400 gallons/hour, and 4) stipulate that fines cannot be imposed based on meter readings – effectively disallowing automatic enforcement. These actions make apparent the very real political constraints that prevented implementation of this technological solution to water conservation in Fresno.

Importantly, our experiment does not speak to the optimality of outdoor watering regulations *per se*. As environmental agencies and private actors increase adoption of remote sensing and continuous monitoring technologies to achieve water conservation and other goals, policymakers must adapt aging regulations to reflect the availability of high-frequency, real-time data. In doing so, it is clear that the environmental and conservation benefits must be weighed against the political costs associated with perfect detection of violations.

References

Allcott, Hunt and Todd Rogers (2014). "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation". In: *American Economic Review* 104 (10), pp. 3003–3037.

Alm, James and Jay Shimshack (2014). "Environmental Enforcement and Compliance: Lessons from Pollution, Safety, and Tax Settings. Foundations and Trends". In: *Foundations and Trends(R) in Microeconomics* 10 (4), pp. 209–274.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen (2017). "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan". In: *The Quarterly Journal of Economics* 132.3, pp. 1101–1164.

Banerjee, Abhijit V. and Esther Duflo (2006). "Addressing Absence". In: *Journal of Economic Perspectives* 20 (1), pp. 117–132.

Banerjee, Abhijit V., Esther Duflo, and Rachel Glennerster (2008). "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System". In: *Journal of the European Economic Association* 6 (2-3), pp. 487–500.

Bar-Ilan, Avner and Bruce Sacerdote (2004). "The Response of Criminals and Noncriminals to Fines". In: *Journal of Law & Economics* 47 (1), pp. 1–17.

Blumenstock, Joshua, Michael Callen, Anastasiia Faikina, and Stefano Fiorin (2022). "Modernizing the State During War: Experimental Evidence from Afghanistan". Working Paper.

Bossuroy, Thomas, Clara Delavallade, and Vincent Pons (2019). "Biometric Tracking, Healthcare Provision, and Data Quality: Experimental Evidence from Tuberculosis Control". In: *National Bureau of Economic Research Working Paper Series No. 26388.*

Brent, Daniel, Joseph Cook, and Skylar Olsen (2015). "Social Comparisons, Household Water Use, and Participation in Utility Conservation Programs: Evidence from Three Randomized Trials". In: *Journal of the Association of Environmental and Resource Economists* 2 (4), pp. 597–627.

Brent, Daniel and Michael Ward (2019). "Price Perceptions in Water Demand". In: *Journal of Environmental Economics and Management* 98, p. 102266.

Brollo, Fernanda, Katja M. Kaufmann, and Eliana La Ferrara (2019). "The Political Economy of Program Enforcement: Evidence from Brazil". In: *Journal of the European Economic Association* 18 (2), pp. 750–791.

Browne, Oliver R., Ludovica Gazze, and Michael Greenstone (2021). "Do Conservation Policies Work? Evidence from Residential Water Use". In: *Environmental and Energy Policy and the Economy.* Ed. by M. Kotchen, J.H. Stock, and C. Wolfram. Vol. 2. Chicago: University of Chicago Press, pp. 190–225.

Bugni, Federico A., Ivan A. Canay, and Azeem M. Shaikh (2019). "Inference under covariate-adaptive randomization with multiple treatments". In: *Quantitative Economics* 10.4, pp. 1747–1785.

Casaburi, Lorenzo and Ugo Troiano (2016). "Ghost-House Busters: The Electoral Response to a Large Anti–Tax Evasion Program". In: *The Quarterly Journal of Economics* 131.1, pp. 273–314.

Costa, Francisco and François Gerard (2021). "Hysteresis and the Welfare Effect of Corrective Policies: Theory and Evidence from an Energy-Saving Program". In: *Journal of Political Economy* 129 (6), pp. 1705–1743.

Dhaliwal, Iqbal and Rema Hanna (2017). "The devil is in the details: The successes and limitations of bureaucratic reform in India". In: *Journal of Development Economics* 124.C, pp. 1–21.

Diffenbaugh, Noah S., Daniel L. Swain, and Danielle Touma (2020). "Anthropogenic warming has increased drought risk in California". In: *Proceedings of the National Academy of Sciences* 112 (13), pp. 3931–3936.

Drago, Francesco, Roberto Galbiati, and Pietro Vertova (2009). "The Deterrent Effects of Prison: Evidence from a Natural Experiment". In: *Journal of Political Economy* 117 (2), pp. 257–280.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan (2013). "Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India". In: *The Quarterly Journal of Economics* 128 (4), pp. 1499–1545.

— (2018). "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India". In: *Econometrica* 86 (6), pp. 2123–2160.

Dušek, Libor and Christian Traxler (2022). "Learning from Law Enforcement". In: *Journal of the European Economic Association* 20.2, pp. 739–777.

Ferreira, Alipio (2021). "Satellites and Fines: Using Monitoring to Target Inspections of Deforestation". In.

Fowlie, Meredith, Edward Rubin, and Reed Walker (2019). "Bringing Satellite-Based Air Quality Estimates Down to Earth". In: *AEA Papers & Proceedings* 109, pp. 283–288.

Gibson, Matthew (2019). "Regulation-Induced Pollution Substitution". In: *The Review of Economics and Statistics* 101 (5), pp. 827–840.

Glueck, Sheldon (1928). "Principles of a Rational Penal Code". In: *Harvard Law Review* 41 (4), pp. 453–482.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu (2020). "Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution". In: *National Bureau of Economic Research Working Paper Series No. 27502.*

Hahn, Robert, Robert D Metcalfe, David Novgorodsky, and Michael K Price (2016). "The Behavioralist as Policy Designer: The Need to Test Multiple Treatments to Meet Multiple Targets". In: *National Bureau of Economic Research Working Paper Series No. 22886.*

Halich, Greg and Kurt Stephenson (2009). "Effectiveness of Residential Water-Use Restrictions Under Varying Levels of Municipal Effort". In: *Land Economics* 85 (4), pp. 614–626.

Hanak, Ellen and Matthew Davis (2006). "Lawns and Water Demand in California". In: *California Economic Policy* 2 (2), pp. 1–22.

Haselhuhn, Michael P., Devin G. Pope, Maurice E. Schweitzer, and Peter Fishman (2011). "The Impact of Personal Experience on Behavior: Evidence from Video-Rental Fines". In: *Management Science* 58 (1), pp. 52–61.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka (Feb. 2018). "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand". In: *American Economic Journal: Economic Policy* 10 (1), pp. 240–67.

Jessoe, Katrina, Gabriel E. Lade, Frank Loge, and Edward Spang (2021). "Residential Water Conservation During Drought: Experimental Evidence from Three Behavioral Interventions". In: *Journal of Environmental Economics and Management* 110 (102519).

Jones, Damon, David Molitor, and Julian Reif (2019). "What do Workplace Wellness Programs do? Evidence from the Illinois Workplace Wellness Study". In: *The Quarterly Journal of Economics* 134 (4), pp. 1747–1791.

Kenney, Douglas S., Roberta A. Klein, and Martyn P. Clark (2004). "Use and Effectiveness of Municipal Water Restrictions During Drought in Colorado". In: *Journal of the American Water Resources Association* 40 (1), pp. 77–87.

Kleven, Henrik Jacobsen, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez (2011). "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark". In: *Econometrica* 79 (3), pp. 651–692.

Kuziemko, Ilyana (2013). "How Should Inmates Be Released From Prison? An Assessment of Parole Versus Fixed Sentence Regimes". In: *The Quarterly Journal of Economics* 128 (1), pp. 371–424.

Maurin, Eric and Aurelie Ouss (2009). "Sentence Reductions and Recidivism: Lessons from the Bastille Day Quasi Experiment". In: *IZA Discussion Paper Series No. 3990.*

Meeks, Robyn, Arstan Omuraliev, Ruslan Isaev, and Zhenxuan Wang (2020). "Smart Meters and the Benefits from Electricity Quality Improvements". Working Paper.

Michelsen, An M., J. Thomas. McGuckin, and Donna Stumpf (1999). “Nonprice Water Conservation Programs as a Demand Management Tool”. In: *Journal of the American Water Resources Association* 35 (3), pp. 593–602.

Mu, Yingfei, Edward Rubin, and Eric Zou (2022). “What’s Missing in Environmental (Self-)Monitoring: Evidence from Strategic Shutdowns of Pollution Monitors”. In: *National Bureau of Economic Research Working Paper Series No. 28735.*

Muralidharan Karthik, Paul Niehaus and Sandip Sukhtankar (2021). “Identity Verification Standards in Welfare Programs: Experimental Evidence from India”. In: *National Bureau of Economic Research Working Paper Series No. 26744.*

Pratt, Bryan (2019). “A Fine Is More Than a Price: Evidence from Drought Restrictions”. Working Paper.

Renwick, Mary E. and Richard D. Green (2000). “Do Residential Water Demand Side Management Policies Measure Up? An Analysis of Eight California Water Agencies”. In: *Journal of Environmental Economics and Management* 40 (1), pp. 37–55.

Reynaert, Mathias and James M. Sallee (2021). “Who Benefits When Firms Game Corrective Policies?” In: *American Economic Journal: Economic Policy* 13 (1), pp. 372–412.

Sarin, Natasha and Lawrence H. Summers (2020). “Understanding the Revenue Potential of Tax Compliance Investment”. In: *National Bureau of Economic Research Working Paper Series No. 27571.*

Vollaard, Ben (2017). “Temporal displacement of environmental crime: Evidence from marine oil pollution”. In: *Journal of Environmental Economics and Management* 82, pp. 168–180.

West, Jeremy, Robert W. Fairlie, Bryan Pratt, and Liam Rose (2021). “Automated Enforcement of Irrigation Regulations and Social Pressure for Water Conservation”. In: *Journal of the Association of Environmental and Resource Economists* 8 (6), pp. 1179–1207.

Wichman, Casey J., Laura O. Taylor, and Roger H. von Haefen (2016). “Conservation Policies: Who Responds to Price and Who Responds to Prescription?” In: *Journal of Environmental Economics and Management* 79, pp. 114–134.

Zou, Eric (2021). “Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality”. In: *American Economic Review* 111 (7), pp. 2101–2026.

Tables

Table 1: Summary Statistics and Balance Check

Dependent Variable	Original group size (N)	Group size (N)	Avg. daily water use (gal/day)	Excess water use during banned hrs. (gal/hr)	Number of banned hrs. >300 gal/hr per day	Share HH with any violations, Jul - Sep 2016	Violation clearance rate	Share high users	Share in high income block group	Dropout rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Control Group										
Non-automated Baseline Fine	40,311	40,008	575	28	0.141	0.683	0.004	0.501	0.500	0.003
Panel B: Treatment Groups			***Difference Relative To Non-Automated Baseline Fine***							
Non-automated										
50% Fine Level	4,479	4,445	-7.832	0.010	0.001	0.021***	-0.001	-0.003	-0.000	-0.001**
			(6.301)	(1.085)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
25% Fine Level	4,479	4,440	7.222	2.130*	0.005	0.006	-0.001	-0.002	-0.003	0.000
			(6.417)	(1.270)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
Automated: 300 gal/hr										
Baseline Fine	4,479	4,449	-0.218	1.935	0.002	0.004	0.000	0.001	-0.000	0.009***
			(6.416)	(1.196)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.002)
50% Fine Level	4,479	4,440	-5.428	-0.341	-0.001	-0.004	0.002	0.002	-0.000	0.006***
			(5.965)	(1.034)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
25% Fine Level	4,479	4,445	-3.341	-0.530	-0.000	-0.010	0.000	-0.003	-0.004	0.004***
			(6.254)	(1.049)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
Automated: 500 gal/hr										
Baseline Fine	4,479	4,455	-3.196	-0.535	-0.004	-0.014*	-0.001	0.001	0.000	0.005***
			(6.144)	(1.132)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
50% Fine Level	4,479	4,443	2.467	-0.313	-0.003	-0.004	0.000	-0.000	0.002	0.004***
			(6.166)	(1.133)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
25% Fine Level	4,479	4,438	11.553*	1.010	0.009*	0.007	-0.000	0.001	-0.003	0.002**
			(6.792)	(1.242)	(0.005)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
Automated: 700 gal/hr										
Baseline Fine	4,479	4,449	-1.160	0.085	0.001	0.001	-0.000	-0.002	0.001	0.005***
			(6.284)	(1.084)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
50% Fine Level	4,480	4,443	8.753	-0.896	-0.001	-0.001	-0.000	-0.002	-0.004	0.005***
			(6.504)	(1.075)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
25% Fine Level	4,479	4,449	7.200	0.492	0.005	-0.006	0.000	-0.002	0.002	0.003**
			(6.647)	(1.185)	(0.004)	(0.008)	(0.001)	(0.008)	(0.008)	(0.001)
P-value of joint F test	N/A	N/A	0.429	0.687	0.602	0.140	0.751	1	1	<0.001

Notes: This table reports baseline characteristics in the control group (Panel A) and differences between each treatment group and the control group for those characteristics (Panel B). Columns 1 and 2 report the size of treatment groups at the point of randomization and analysis respectively. Columns 3 - 5 include data from July - September 2017; Columns 6 - 7 include 2016 data for which we only matched 90% households, evenly distributed across treatment groups; Columns 8 includes data from April 2017; Column 9 includes data from the Census Bureau's 2010-2014 5-year American Community Survey; Column 10 includes data from July - September 2018. Baseline fine amounts are \$0, \$50, \$100, \$200 for the first, second, third and fourth (and thereafter) violations in a year. Violation clearance rate is the ratio of number of warnings and fines received to number of hours in violation during July - September 2016. High users are defined as having water use above the median in April 2017. High-income block groups have median income above the median block group in the city. The dropout rate in Column 10 is computed using the group sizes reported in Column 2 as the denominator. The model utilized in the joint F-test is the column's respective covariate on binary indicators for each treatment group, omitting the non-automated baseline fine group as the intercept. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

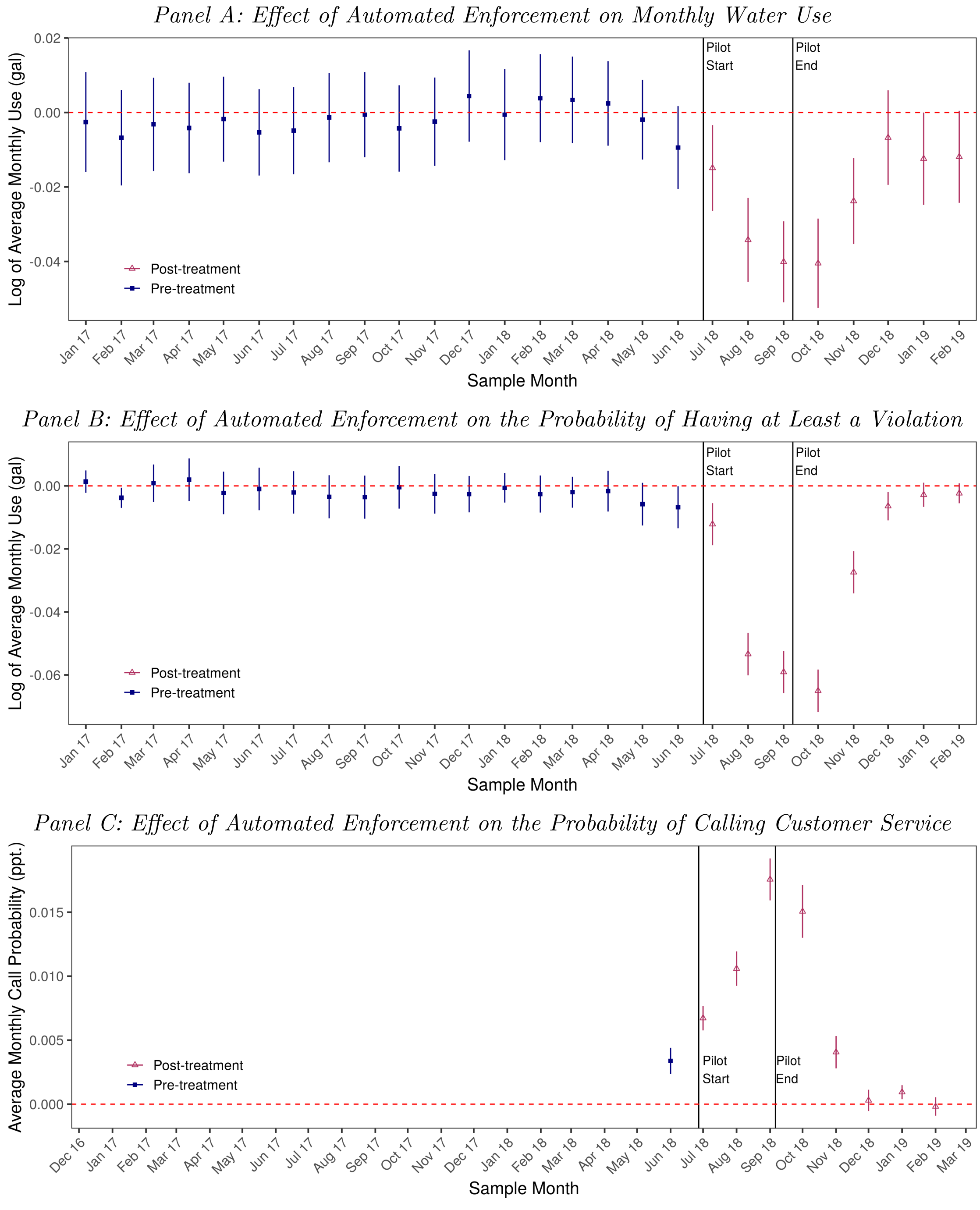
Table 2: Effect of Automated Enforcement on Compliance, Water Use, and Customer Contact

	Enforcement & Compliance					*Water Use*	*Customer Contact*	
Dependent Variable	Violations in a month (300+ gal/hr)	Violated at Least Once in a month	Received a Warning	Received a Fine	Fine Levied ($/month)	Log of Monthly Water Use (gal)	Called at Least Once in a month x 100	Called at Least Once in a month Disputes/Complaints x 100
	(1a)	(1b)	(1c)	(1d)	(1e)	(2)	(3a)	(3b)
Automated Enforcement	-0.636***	-0.041***	0.343***	0.141***	2.35***	-0.029***	1.151***	0.792***
	(0.046)	(0.003)	(0.003)	(0.002)	(0.04)	(0.005)	(0.041)	(0.032)
Summer Effect if Pilot Applied Citywide (3 Months)	-218,547***	-14,219***	117,710***	48,406***	807,022***	-174,042,096***	395,290***	271,976***
	(15,755)	(1,004)	(860)	(604)	(13,423)	(31,485,487)	(14,204)	(11,048)
Control Mean	3.736	0.509	0.020	0.001	0.03	9.492	0.207	0.123
Adjusted p-value	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001	0.001	$<$0.001	$<$0.001
N	261,961	261,961	88,904	88,904	261,961	261,311	261,961	261,961

Notes: This table reports household level (Columns 1c-1d) and household-month level (Columns 1a,1b,1e-3b) regression coefficients on the effects of automated enforcement relative to the non-automated baseline fine group on the outcomes described in each column. Included in each regression are indicators for non-automated, alternative fine levels the coefficients for which are reported in Appendix Table A3. Each regression includes data for the months of July-September 2018 only. Column 2 drops observations with zero monthly use. Column 3b distinguishes the impact on customer disputes and complaints of utility enforcement actions. Watering schedule violations are defined as households exceeding 300 gal/hr during banned hours. The outcome variables in Columns 3a and 3b are multiplied by 100 for ease of visualization. Summer effect is calculated by multiplying each coefficient by the number of households in Fresno (114,508) and the number of pilot months in the summer (3). Adjusted p-value reports the multiple inference adjusted p-value of the estimates, controlling for the family-wise error rate. This correction is done by defining a family of hypotheses that includes each of the eight outcome variables reported in the table. In this case, our null hypothesis is $\beta = \gamma_{25} = \gamma_{50} = 0$. The adjusted p-values are calculated over 10,000 iterations of the Westfall-Young free step-down resampling algorithm (Jones, Molitor, and Reif, 2019). Standard errors are clustered at the household level except for columns 1c-1d which report heteroskedasticity robust standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

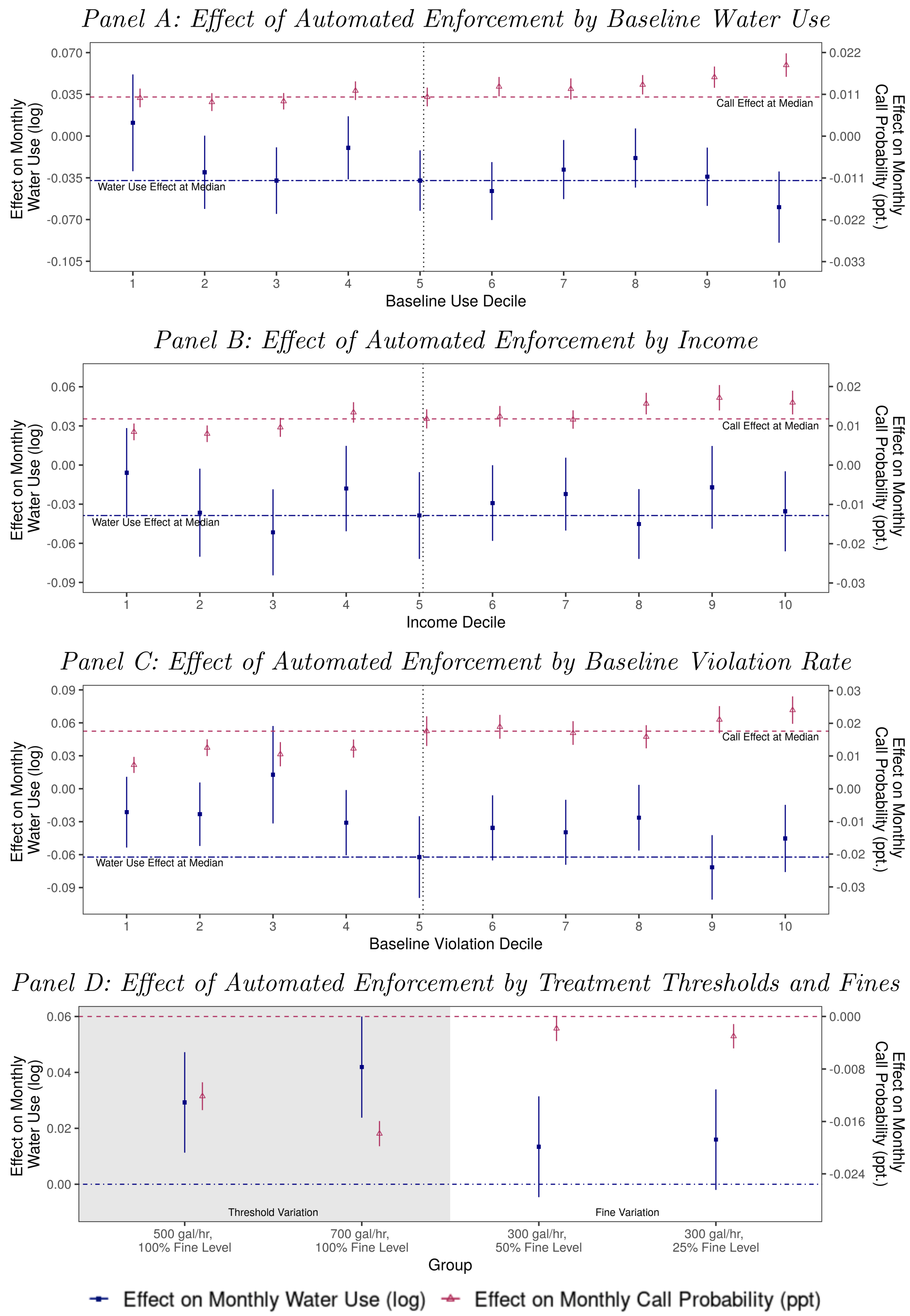
Figures

Figure 1: Effect of Automated Enforcement Over Time



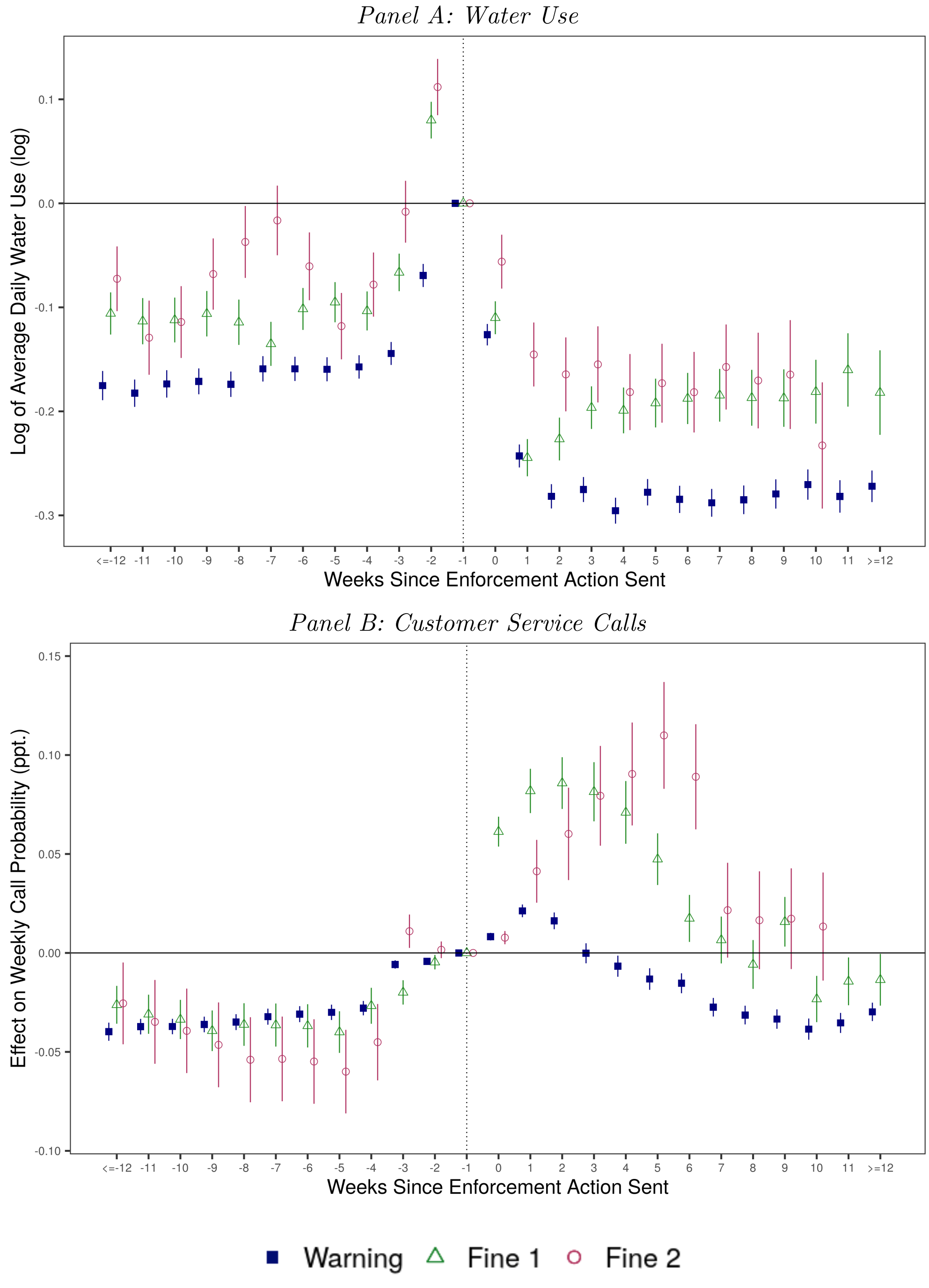
Notes: This figure plots month-by-month coefficient estimates of the effect of automated enforcement on the logarithm of a household's monthly water use (Panel A), the probability of a household violating at least once in a month (Panel B), and the probability of a household calling customer service (Panel C). Panels A and B include data from January 2017 to February 2019. Panel C includes data from June 2018 to February 2019 due to limited data availability. All panels include sample month fixed effects and indicators for households subject to visual inspection and alternative fine schedules. The vertical bars represent 95% confidence intervals based on standard errors clustered at the household level.

Figure 2: Effect of Automated Enforcement on Cumulative Compliance and Water Use, by Baseline Use, Income, Baseline Violation Rate, and Treatment



Notes: Panels include data from July 2018 to September 2018. Panel A plots the estimated effect of automated enforcement on the log of monthly water use (blue) and call probability (red) by households' baseline water use deciles. Panel B and C plot the same estimates by households' block group median income decile and baseline violation rate deciles, respectively. The regressions in Panels A-C include an indicator for households assigned to automated enforcement, decile indicators, and interactions between each decile and the automated group indicator, omitting the fifth decile as the reference category. Panels A-C are the linear combination of the decile's regression output and the coefficient from the automated group indicator. The horizontal line with an alternating pattern is the effect on water use at the median decile. The dashed horizontal line is the effect on call probability at the median decile. Baseline water utilization ranges from 6-146,940 gallons with a median of 7,640 gallons. Household block group median annual income ranges from \$11,334-\$153,194 with a median of \$56,348. Baseline violation rate ranges from 1-867 hours in violation during the baseline period with a median of 9 hours. Panel D plots the estimated effects of assignment to each automated enforcement threshold and fine level relative to the default automated enforcement policy of 300 gal/hr and baseline fine level (i.e. the β_2 and β_3 coefficients from equation (2) in the gray graph area, and β_4 and β_5 coefficients from equation (2) in the white graph area). Each regression includes indicators for visual inspection and alternative fine levels. Baseline water use data is from January 2017 to February 2018, the period used for the stratified randomization. Baseline violation rate deciles are defined using data from July 2017 to September 2017. The vertical bars represent 95% confidence intervals based on household-level clustered standard errors.

Figure 3: Effect of Fines and Warnings on Water Use and Customer Service Calls



Notes: This figure plots coefficient estimates from a household-week level regressions of an outcome variable on indicators for each week's timing relative to receipt of first, second, and third notices, that is first warning, and first and second fines. The outcome variable in the top panel is the log of average daily water use. The outcome variable in the bottom panel is the probability of calling customer service. Week 0 denotes the calendar week in which the warning/fine was sent, and week -1 is omitted as the reference category. The sample includes data from January 2017 through February 2019 for water use and from June 2018 to February 2019 for probability of calling DPU. The sample drops notices waived ex-post. The regression includes household and week fixed effects. The vertical bars represent 95% confidence intervals based on standard errors clustered at the household level.

A Appendix Tables and Figures

A.1 Appendix Tables

Table A1: Calls to Customer Service by Topic of Conversation

Topic of Conversation (1)	Count (2)	Percent of Category (3)
Complaints/Disputes	**1,370**	
Misc. complaint regarding a notice but no formal dispute	343	25%
Request to review notice or meter reading	303	22.1%
General dispute or appeal of notice	231	16.9%
Dispute - Filling, draining or using a pool, pond or home spa	176	12.8%
Request to review date and time of violation	139	10.1%
Request to review notice, offered materials on customer service portal	72	5.3%
Dispute - Water is being stolen	22	1.6%
Dispute - Hired gardeners	16	1.2%
Dispute - Not at home at time of violation	15	1.1%
Request to review notice, resident noted pipe leak	14	1%
Dispute - Followed prescribed watering schedule	14	1%
Dispute - No longer living at the address	9	0.7%
Dispute - Rental property without ability to change water use	8	0.6%
Dispute - Landscaping Property	3	0.2%
Dispute - Watering new garden after planting	2	0.1%
Dispute - Meter system has been set incorrectly	2	0.1%
Dispute - Preparing house for sale	1	0.1%
Misc.	**403**	
Misc. clarifying question	295	73.2%
Misc. question about the pilot	94	23.3%
Misunderstanding regarding pilot design	6	1.5%
Questions about billing and past usage	5	1.2%
Clarifying email or mailing address for pilot communications	2	0.5%
Callback request	1	0.2%
Service Request	**252**	
Request for help managing a sprinkler timer	98	38.9%
Leak survey request	49	19.4%
Notification of known infrastructure repair request	35	13.9%
Sprinkler timer has been set incorrectly	32	12.7%
Notifying utility of broken pipe	16	6.3%
Notifying utility of sprinkler malfunction	11	4.4%
Sprinkler/timer inspection request	7	2.8%
Notifying utility of broken sprinkler timer	3	1.2%
Notifying utility of broken water meter	1	0.4%
Opt out	**56**	
Request to opt out of the pilot	25	44.6%
Opt out confirmation request	13	23.2%
Initially requested to opt out but decided to remain in program	11	19.6%
Confused about use thresholds after opting out	7	12.5%

Notes: This table tabulates the topics of conversation for all calls to the utility's customer service department between July and September 2018. The calls are distinguished into four broad categories listed in bold.

Table A2: Characteristics of Opt-out Households

Dependent Variable	Total water use (gal/day)	Excess water use during banned hrs. (gal/hr)	Number of banned hrs. >300 gal/hr per day	Share HH with any violations, Jul - Sep 2016	Violation clearance rate	Share high users	Share in high income block group
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Opt-out Households	75***	2	0.0132	0.073***	0.001	-0.003	0.050**
	(20)	(3)	(0.011)	(0.021)	(0.003)	(0.023)	(0.023)
All Other Households	570	28	0.140	0.683	0.004	0.501	0.499
N	84,793	84,793	84,793	76,531	52,298	84,793	84,793

Notes: This table reports household level regression coefficients on the characteristics of households that opted out of the experiment. Columns 1 - 3 include data from July - September 2017; Columns 4 - 5 include 2016 data for which we only matched 90% households; Column 6 includes data from April 2017. Column 7 includes data from the Census Bureau 2010-2014 5-year American Community Survey. Violation clearance rate is the ratio of number of warnings and fines received to the number of hours in violation during July-September 2016. High users are defined as having baseline daily water use above the median. High-income block groups have median income above the median block group in the city. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Effect of Non-automated Enforcement, Alternative Fine Levels on Compliance, Water Use, and Customer Contact

	Enforcement & Compliance					*Water Use*	*Customer Contact*	
Dependent Variable	Violations in a month (300+ gal/hr)	Violated at Least Once in a month	Received a Warning	Received a Fine	Fine Levied ($/month)	Log of Monthly Water Use (gal)	Called at Least Once in a month x 100	Called at Least Once in a Month Disputes/Complaints x 100
	(1a)	(1b)	(1c)	(1d)	(1e)	(2)	(3a)	(3b)
Panel A: Base Specification								
50% Fine Level	-0.120	-0.005	-0.003*	-0.001***	-0.02**	-0.026**	-0.077**	-0.024
	(0.103)	(0.007)	(0.002)	(0.000)	(0.01)	(0.012)	(0.034)	(0.029)
Adjusted p-value	0.953	0.862	0.268	0.081	0.067	0.264	0.221	0.949
25% Fine Level	0.109	-0.000	0.001	-0.000	-0.02***	0.007	-0.023	-0.016
	(0.108)	(0.007)	(0.002)	(0.001)	(0.01)	(0.012)	(0.043)	(0.030)
Adjusted p-value	0.984	0.905	0.880	0.880	0.008	0.953	0.953	0.953
Control Mean	3.736	0.509	0.020	0.001	0.03	9.492	0.207	0.123
N	261,961	261,961	88,904	88,904	261,961	261,311	261,961	261,961
Panel B: 2017 Water Use Controls								
50% Fine Level	-0.026	0.002	-0.002	-0.000	-0.02	-0.005	-0.101*	-0.040
	(0.089)	(0.007)	(0.002)	(0.001)	(0.04)	(0.008)	(0.052)	(0.040)
25% Fine Level	0.049	-0.001	0.000	-0.001	-0.02	0.003	0.033	0.017
	(0.090)	(0.007)	(0.002)	(0.001)	(0.04)	(0.009)	(0.060)	(0.045)
N	251,766	251,766	84,438	84,438	251,766	251,388	251,766	251,766
Panel C: Household and Month FEs								
50% Fine Level	-0.122	-0.002				-0.003		
	(0.104)	(0.006)				(0.009)		
25% Fine Level	-0.013	0.004				-0.003		
	(0.109)	(0.006)				(0.009)		
N	513,116	513,116				512,122		

Notes: This table reports household level (Columns 1c-1d) and household-month level (Columns 1a,1b,1e-3b) regression coefficients on the effects of non-automated enforcement with alternative fine levels relative to the non-automated baseline fine group on the outcomes described in each column. Included in each regression is an indicator for automated enforcement the coefficients for which are reported in Table 2. Panel A includes data for the months of July-September 2018 only. Panel B includes the same data but additionally controls for total water use between July-September 2017. Panel C includes data for the months of July-September 2017 and 2018 and includes fixed effects for household and sample month. The number of observations is not consistent between panels due to imbalanced data coverage during 2017. Column 2 drops observations with zero monthly use. Watering schedule violations are defined as households exceeding 300 gal/hr during banned hours. The outcome variables in Columns 3a-3b are multiplied by 100 for ease of visualization. Adjusted p-value reports the multiple inference adjusted p-value of the estimate, controlling for the family-wise error rate. This correction is done by defining a family of hypotheses that includes each of the eight outcome variables reported in the table. In this case, our null hypothesis is $\beta = \gamma_{25} = \gamma_{50} = 0$. The adjusted p-values are calculated over 10,000 iterations of the Westfall-Young free step-down resampling algorithm (Jones, Molitor, and Reif, 2019). Standard errors are clustered at the household level except for columns 1c-1d which report heteroskedasticity robust standard errors.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Effect of Automated Enforcement on Compliance, Water Use, and Customer Contact - Additional Controls

	Enforcement & Compliance					*Water Use*	*Customer Contact*	
Dependent Variable	Violations in a month (300+ gal/hr)	Violated at Least Once in a month	Received a Warning	Received a Fine	Fine Levied ($/month)	Log of Monthly Water Use (gal)	Called at Least Once in a month x 100	Called at Least Once in a Month Disputes/Complaints x 100
	(1a)	(1b)	(1c)	(1d)	(1e)	(2)	(3a)	(3b)
Panel A: 2017 Water Use Controls								
Automated Enforcement	-0.605***	-0.045***	0.347***	0.144***	2.17***	-0.030***	1.093***	0.745***
	(0.039)	(0.003)	(0.003)	(0.002)	(0.04)	(0.004)	(0.042)	(0.033)
N	251,766	251,766	84,438	84,438	251,766	251,388	251,766	251,766
Panel B: Household and Month FEs								
Automated Enforcement	-0.687***	-0.040***				-0.029***		
	(0.050)	(0.003)				(0.004)		
N	513,116	513,116				512,122		

Notes: This table reports household level (Columns 1c-1d) and household-month level (Columns 1a,1b,1e-3b) regression coefficients on the effects of automated enforcement relative to the non-automated baseline fine group on the outcomes described in each column. Included in each regression are indicators for non-automated, alternative fine levels the coefficients for which are reported in Appendix Table A3. Each regression in Panel A includes data for the months of July-September 2018 and controls for total water use between July-September 2017. Panel B includes data for the months of July-September 2017 and 2018, and includes fixed effects for household and sample month. The number of observations is not consistent between panels due to imbalanced data coverage during 2017. Column 2 drops observations with zero monthly use. Watering schedule violations are defined as households exceeding 300 gal/hr during banned hours. The outcome variables in Columns 3a-3b are multiplied by 100 for ease of visualization. Standard errors are clustered at the household level except for columns 1c-1d which report heteroskedasticity robust standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Effect of Automated Enforcement on Compliance, Water Use, and Customer Contact - Adjustment for Stratification

	Enforcement & Compliance					*Water Use*	*Customer Contact*	
Dependent Variable	Violations in a month (300+ gal/hr)	Violated at Least Once in a month	Received a Warning	Received a Fine	Fine Levied ($/month)	Log of Monthly Water Use (gal)	Called at Least Once in a month x 100	Called at Least Once in a month Disputes/Complaints x 100
	(1a)	(1b)	(1c)	(1d)	(1e)	(2)	(3a)	(3b)
Automated Enforcement	-0.635***	-0.041***	0.343***	0.141***	2.35***	-0.029***	1.151***	0.792***
	(0.030)	(0.002)	(0.002)	(0.002)	(0.03)	(0.003)	(0.040)	(0.031)
Non-Automated Enforcement								
50% Fine Level	-0.119*	-0.005	-0.003*	-0.001***	-0.02***	-0.025***	-0.077**	-0.024
	(0.068)	(0.005)	(0.002)	(0.000)	(0.01)	(0.007)	(0.034)	(0.029)
75% Fine Level	0.109	0.000	0.001	-0.000	-0.02***	0.007	-0.023	-0.016
	(0.071)	(0.005)	(0.002)	(0.001)	(0.01)	(0.007)	(0.043)	(0.030)
Control Mean	2.756	0.465	0.013	0.001	0.01	9.117	0.242	0.152
N	261,961	261,961	88,904	88,904	261,961	261,311	261,961	261,961

Notes: This table reports household level (Columns 1c-1d) and household-month level (Columns 1a,1b,1e-3b) regression coefficients on the effects of automated enforcement relative to the non-automated baseline fine group on the outcomes described in each column. Included in each regression are interaction terms between the automated monitoring indicator and each strata utilized in the randomization process. The reported estimates in the automated enforcement category are the weighted average of the coefficients from the full saturated model. Standard errors have been adjusted for the covariate-adaptive randomization, following Bugni, Canay, and Shaikh (2019). Indicators for non-automated, alternative fine levels are included as covariates. Each regression includes data for the months of July-September 2018 only. Column 2 drops observations with zero monthly use. Column 3b distinguishes the impact on customer disputes and complaints of utility enforcement actions. Watering schedule violations are defined as households exceeding 300 gal/hr during banned hours. The outcome variables in Columns 3a and 3b are multiplied by 100 for ease of visualization. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Effect of Automated Enforcement on Number of Months in Violation

	Never Violated (1)	Violated in One Month (2)	Violated in Two Months (3)	Violated in Three Months (4)
Automated Enforcement	0.010***	0.038***	0.018***	-0.066***
	(0.003)	(0.003)	(0.003)	(0.003)
Control Mean	0.329	0.177	0.158	0.335
N	88,904	88,904	88,904	88,904

Notes: This table reports household level regression coefficients on the effects of automated enforcement relative to the non-automated baseline fine group on the outcomes described in each column. Included in each regression are indicators for non-automated, alternative fine levels the coefficients for which are omitted. It includes data for the months of July-September 2018 only. Watering schedule violations are defined as households exceeding 300 gal/hr during banned hours. Standard errors clustered at the household level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A7: Peer Effects of Automated Enforcement

Dependent Variable	Log of Daily Water Use (gal)				
	(1)	(2)	(3)	(4)	(5)
Automated Enforcement	-0.030***	-0.031***	-0.034***	-0.031***	-0.034***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Share Automated		0.209	-0.038		
		(0.169)	(0.110)		
Share 300gal/hr Threshold				0.284	-0.044
				(0.208)	(0.146)
Share 500gal/hr Threshold				0.287	0.248
				(0.222)	(0.151)
Share 700gal/hr Threshold				0.050	-0.098
				(0.242)	(0.154)
N	7,466,297	7,466,323	7,466,323	7,466,323	7,466,323
Additional Controls			X		X

Notes: This table reports household-day level regression coefficients on the effects of automated enforcement relative to non-automated groups baseline fine group on the log of daily water use. Included in each regression are day fixed effects and indicators for non-automated, alternative fine levels the coefficients for which are omitted. Each regression includes data for the months of July-September 2018 only. The table also reports coefficients incorporating the share of the household's neighbors subject to automated enforcement and to one of the three water use thresholds. Neighbors are defined by Census block. Columns 3 and 5 contain household April 2017 water use and Census block group median income fixed effects. Standard errors in parentheses are clustered at the block level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8: Automation Impact on Number of Hours in Violation

	Violations (300+ gal/hr) (1)	Violations (300-499 gal/hr) (2)	Violations (500-699 gal/hr) (3)	Violations (700+ gal/hr) (4)
Automated Enforcement	-0.636***	-0.348***	-0.157***	-0.131***
	(0.046)	(0.031)	(0.016)	(0.015)
Non-Automated Enforcement				
50% Fine Level	-0.120	-0.115*	-0.015	0.011
	(0.103)	(0.065)	(0.036)	(0.037)
25% Fine Level	0.109	0.077	0.016	0.015
	(0.108)	(0.072)	(0.038)	(0.036)
Control Mean	3.736	2.386	0.822	0.528
N	261,961	261,961	261,961	261,961

This table reports household-month level regression coefficients on the effects of automated enforcement on the outcomes described in each column. Column 1 replicates Column 1a from Table 2. The sample includes data for the months of July-September 2018 only. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A9: Automation Impact on Average Hourly Use: Permitted Versus Banned Hours

Dependent Variable	Log of Average Water Use over a Month (gal/hr)		
	Overall	Permitted Hours	Banned Hours
	(1)	(2)	(3)
Automated Enforcement			
July	-0.012**	0.002	-0.028***
	(0.005)	(0.008)	(0.006)
August	-0.033***	0.006	-0.081***
	(0.005)	(0.008)	(0.006)
September	-0.039***	0.007	-0.083***
	(0.005)	(0.008)	(0.006)
Non-Automated			
50% Fine Level	-0.022*	-0.030*	-0.012
	(0.011)	(0.016)	(0.013)
25% Fine Level	0.007	0.014	0.003
	(0.012)	(0.016)	(0.013)
Control Mean	6.129	5.301	5.258
N	261,311	260,405	261,153
Average Number of Hours	667.8	178.4	489.4

Notes: This table reports household-month level regression coefficients on the effects of automated enforcement by month on the outcomes described in each column. Each regression includes month fixed effects. The sample includes data for the months of July-September 2018 only. Standard errors clustered at the household level in parentheses. This table also reports the average number of permitted or banned watering hours per month. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A10: Effect of Automated Enforcement on Compliance, Water Use, and Customer Contact - By Treatment Group

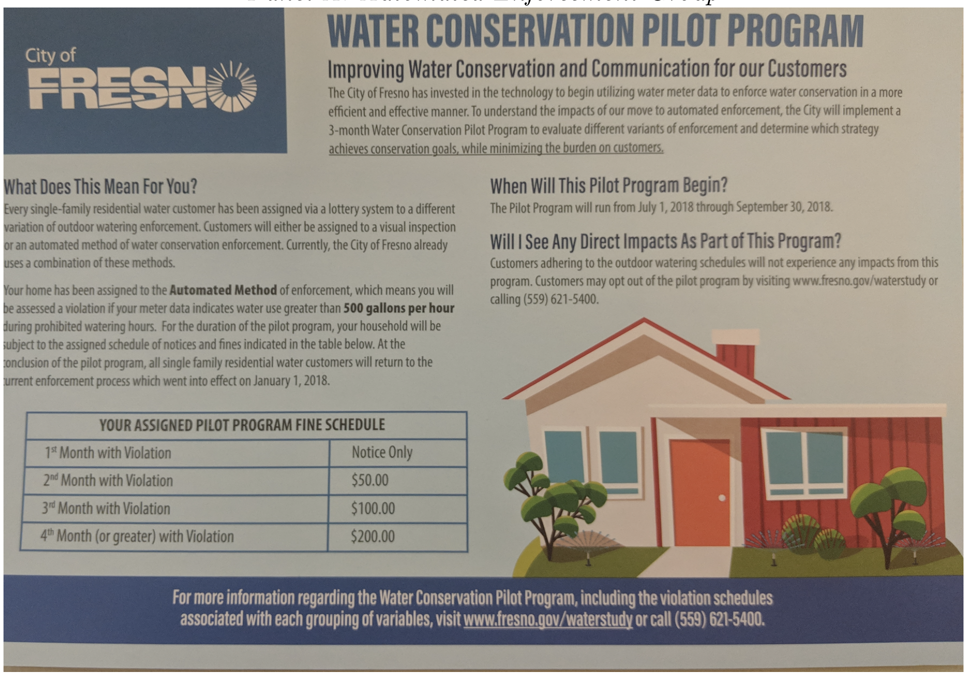
	Enforcement & Compliance					*Water Use*	*Customer Contact*	
Dependent Variable	Violations in a month (300+ gal/hr)	Violated at Least Once in a month	Received a Warning	Received a Fine	Fine Levied ($/month)	Log of Monthly Water Use (gal)	Called at Least Once in a month x 100	Called at Least Once in a Month Disputes/Complaints x 100
	(1a)	(1b)	(1c)	(1d)	(1e)	(2)	(3a)	(3b)
Automated: 300 gal/hr								
Baseline Fine	-0.937***	-0.081***	0.553***	0.263***	7.91***	-0.065***	2.608***	1.868***
	(0.093)	(0.006)	(0.007)	(0.007)	(0.24)	(0.012)	(0.171)	(0.134)
50% Fine Level	-0.885***	-0.074***	0.555***	0.266***	4.04***	-0.049***	1.875***	1.313***
	(0.089)	(0.006)	(0.007)	(0.007)	(0.12)	(0.012)	(0.144)	(0.114)
25% Fine Level	-0.957***	-0.057***	0.577***	0.277***	2.11***	-0.046***	1.970***	1.328***
	(0.085)	(0.006)	(0.007)	(0.007)	(0.06)	(0.012)	(0.146)	(0.114)
Automated: 500 gal/hr								
Baseline Fine	-0.857***	-0.044***	0.312***	0.099***	2.66***	-0.037***	0.966***	0.666***
	(0.089)	(0.006)	(0.007)	(0.005)	(0.14)	(0.012)	(0.107)	(0.085)
50% Fine Level	-0.597***	-0.035***	0.315***	0.120***	1.64***	-0.017	1.045***	0.739***
	(0.090)	(0.006)	(0.007)	(0.005)	(0.08)	(0.012)	(0.119)	(0.091)
25% Fine Level	-0.438***	-0.024***	0.319***	0.112***	0.78***	-0.017	0.794***	0.534***
	(0.095)	(0.006)	(0.007)	(0.005)	(0.04)	(0.012)	(0.099)	(0.078)
Automated: 700 gal/hr								
Baseline Fine	-0.418***	-0.022***	0.151***	0.041***	1.07***	-0.015	0.359***	0.244***
	(0.095)	(0.007)	(0.006)	(0.003)	(0.09)	(0.012)	(0.074)	(0.060)
50% Fine Level	-0.278***	-0.014**	0.151***	0.046***	0.65***	-0.011	0.466***	0.229***
	(0.096)	(0.007)	(0.006)	(0.003)	(0.02)	(0.012)	(0.079)	(0.056)
25% Fine Level	-0.358***	-0.021***	0.149***	0.044***	0.27***	-0.007	0.267***	0.198***
	(0.099)	(0.007)	(0.006)	(0.003)	(0.02)	(0.012)	(0.066)	(0.053)
Non-Automated								
50% Fine Level	-0.120	-0.005	-0.003*	-0.001***	-0.02**	-0.026**	-0.077**	-0.024
	(0.103)	(0.007)	(0.002)	(0.000)	(0.01)	(0.012)	(0.034)	(0.029)
25% Fine Level	0.109	-0.000	0.001	-0.000	-0.02***	0.007	-0.023	-0.016
	(0.108)	(0.007)	(0.002)	(0.001)	(0.01)	(0.012)	(0.043)	(0.030)
N	261,961	261,961	88,904	88,904	261,961	261,311	261,961	261,961
P-val*	0.497	0.497	0.056	0.008	0.152	0.500	0.492	0.493

Notes: This table reports household level (Columns 1c-1d) and household-month level (Columns 1a,1b,1e-3b) regression coefficients on the effects of each treatment group on the outcomes described in each column relative to the non-automated baseline fine group. The sample includes data for the months of July-September 2018. Column 2 drops observations with zero monthly use. Watering schedule violations are defined as households exceeding 300 gal/hr during banned hours. The outcome variables in Columns 3a-3b are multiplied by 100 for ease of visualization. Standard errors are clustered at the household level except for columns 1c-1d which report heteroskedasticity robust standard errors. This table also reports the P-value from the F-test that this fully specified model is equivalent to a more parsimonious specification that includes indicators of a household's assigned automated enforcement threshold and fine level but not their interaction (our null hypothesis H_0), according to the formula: $F = \frac{SS_{red}}{SS_{full}}$ for household-month level regressions. We reject H_0 if F is larger than the upper $1-\alpha$ percentile in the $F(N_{clusters}-1, N_{clusters}-1)$ distribution, where α is the level of significance, SS_{red} and SS_{full} are the residual sums of squares from the parsimonious and the full specifications respectively, and $N_{clusters}$ is the number of clusters. For household level regressions, we instead use the formula: $F = \frac{(SS_{red}-SS_{full})/s}{SS_{full}/df_{full}}$ where df_{red} is the degrees of freedom from the parsimonious model and df_{full} is the degrees of freedom from the full model, both equal to the number of clusters minus the number of parameters in this case. We define $s = df_{full} - df_{red}$ and reject H_0 if F is larger than the upper $1-\alpha$ percentile in the $F(s, df_{full})$ distribution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

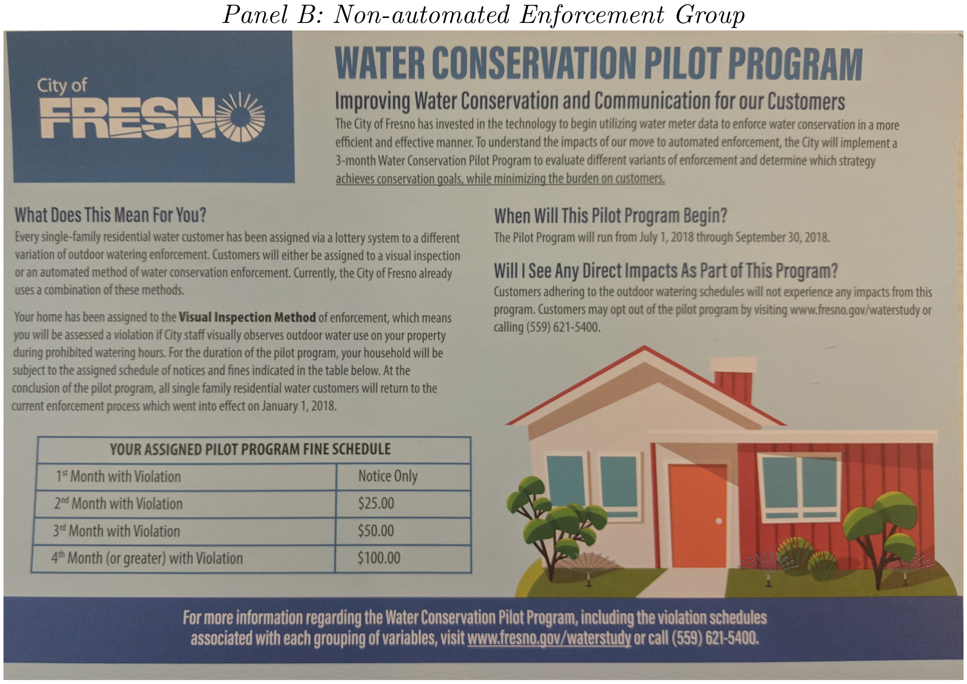
A.2 *Figures*

Figure A1: Mailer Announcing the Pilot in June

Panel A: Automated Enforcement Group



Panel B: Non-automated Enforcement Group



Notes: This figure displays the front page of the mailer sent at the beginning of June 2018 to all households in our experimental sample to announce the beginning of the pilot program in July. Panel A is a mailer to be delivered to a household subject to automated enforcement. Panel B is a mailer to be delivered to a household subject to visual enforcement.

Figure A2: Excessive Water Use Warning

NOTICE OF EXCESSIVE WATER USE:

Excessive Water Use During Restricted Day/Time

[Street Address], METER # [XXX]

Dear [Ratepayer Name]:

Our records indicate the property at this address has violated the City of Fresno's excessive water use provisions. Based on our records, the property associated with your account used more water than is allowed under the excessive use limit during restricted days or hours as measured by your water meter.

Recorded Violation(s):

Wednesday, 5/16/2018 6:59:00 AM - 490 Gallons used

This is your first month with a violation. This **NOTICE** of a violation has been recorded on your customer account. Subsequent violations are subject to fines in accordance with the fine schedule in place at the time the violation occurred.

The violation identified above is your first violation this month. If you incur additional violations this month, you will not receive another notice. You can look up your violation history at https://appdev.fresno.gov/waterpilot/lookup/. Login: [XXX]. Password: [XXX].

We are currently in a Pilot Program with the fine schedule below:

- 1st Month with Violation: Notice Only
- 2nd Month with Violation: [$12.50]
- 3rd Month with Violation: [$25.00]
- 4th Month with Violation: [$50.00]

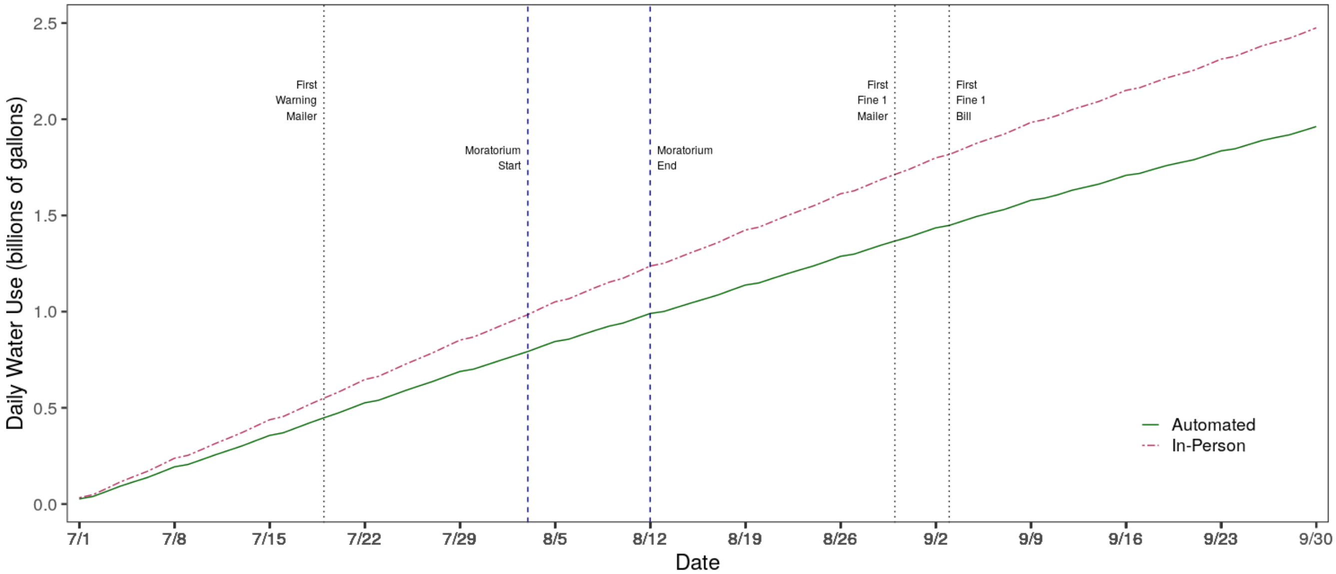
At the end of the pilot period, the Fine Schedule for this property will return to the standard Fine Schedule that went into effect on January 1, 2018. As a reminder, during the pilot your excessive use threshold during restricted days and hours is [300/500/700] gal/hr.

The City of Fresno offers a variety of free services for water utility customers, including water leak surveys and assistance with setting outdoor irrigation timers. For more information, and/or to arrange an appointment for one of these services, please contact Water Conservation at (559) 621-5400. To further support customers, the City recently launched EyeOnWater, a website and application, which allows customers to monitor their water consumption. Information about EyeOnWater can be found at https://fresno.eyeonwater.com.

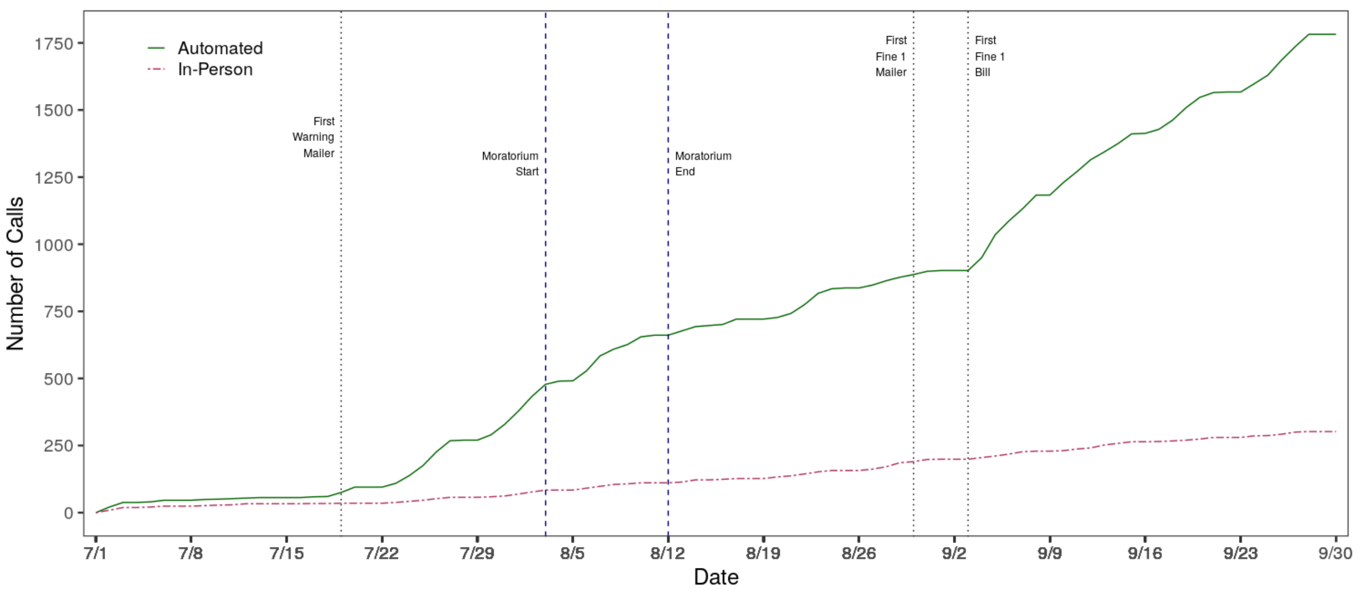
Notes: This figure displays a sample warning of violation sent by the City after a household violated the watering schedule.

Figure A3: Cumulative Water Use and Number of Calls Received by Treatment

Panel A: Water Use



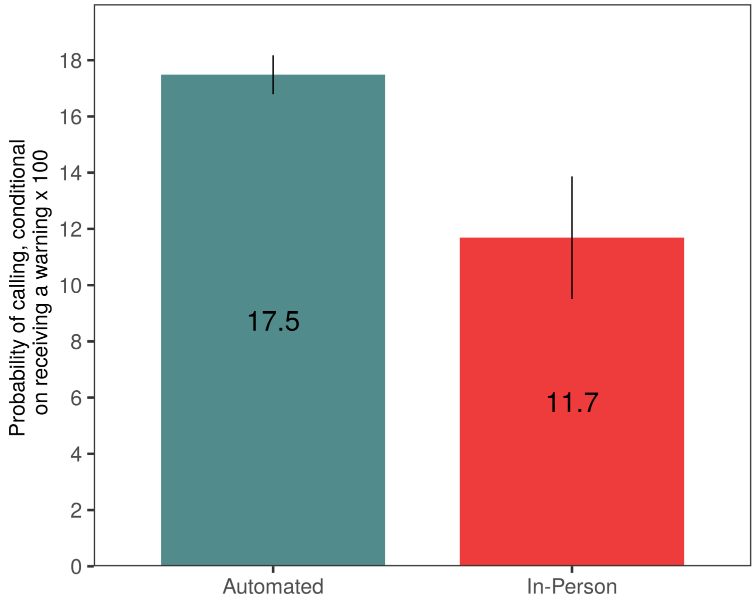
Panel B: Customer Service Calls



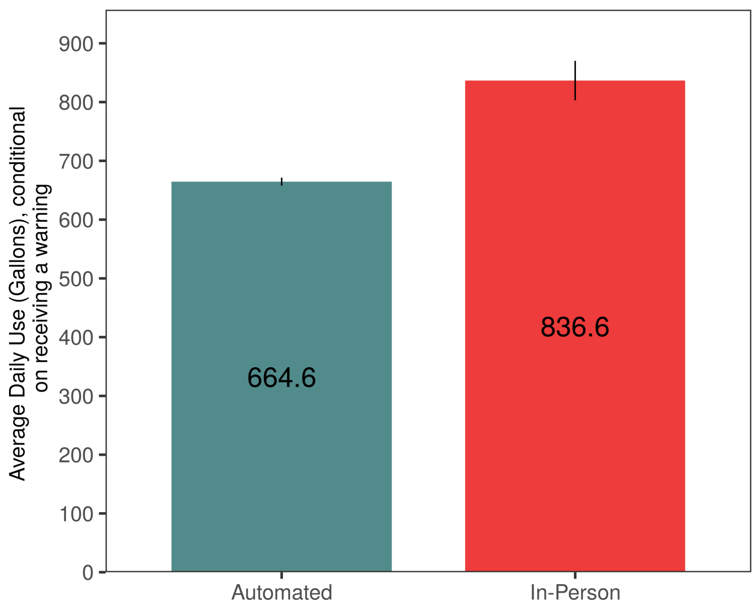
Notes: This figure plots the cumulative water use (Panel A) and number of customer service calls (Panel B) received by the utility, distinguished by monitoring technique. The black dotted lines refer to the dates on which the first Warning mailer was sent to residents, the first Fine 1 mailer was sent to residents, and the first water bill with a Fine 1 penalty was sent to residents. The blue dashed lines demarcate the Fine 1 moratorium between August 1 and August 12.

Figure A4: Probability of Calling and Water Use Conditional on Receipt of Warning, by Treatment

Panel A: Probability of Calling

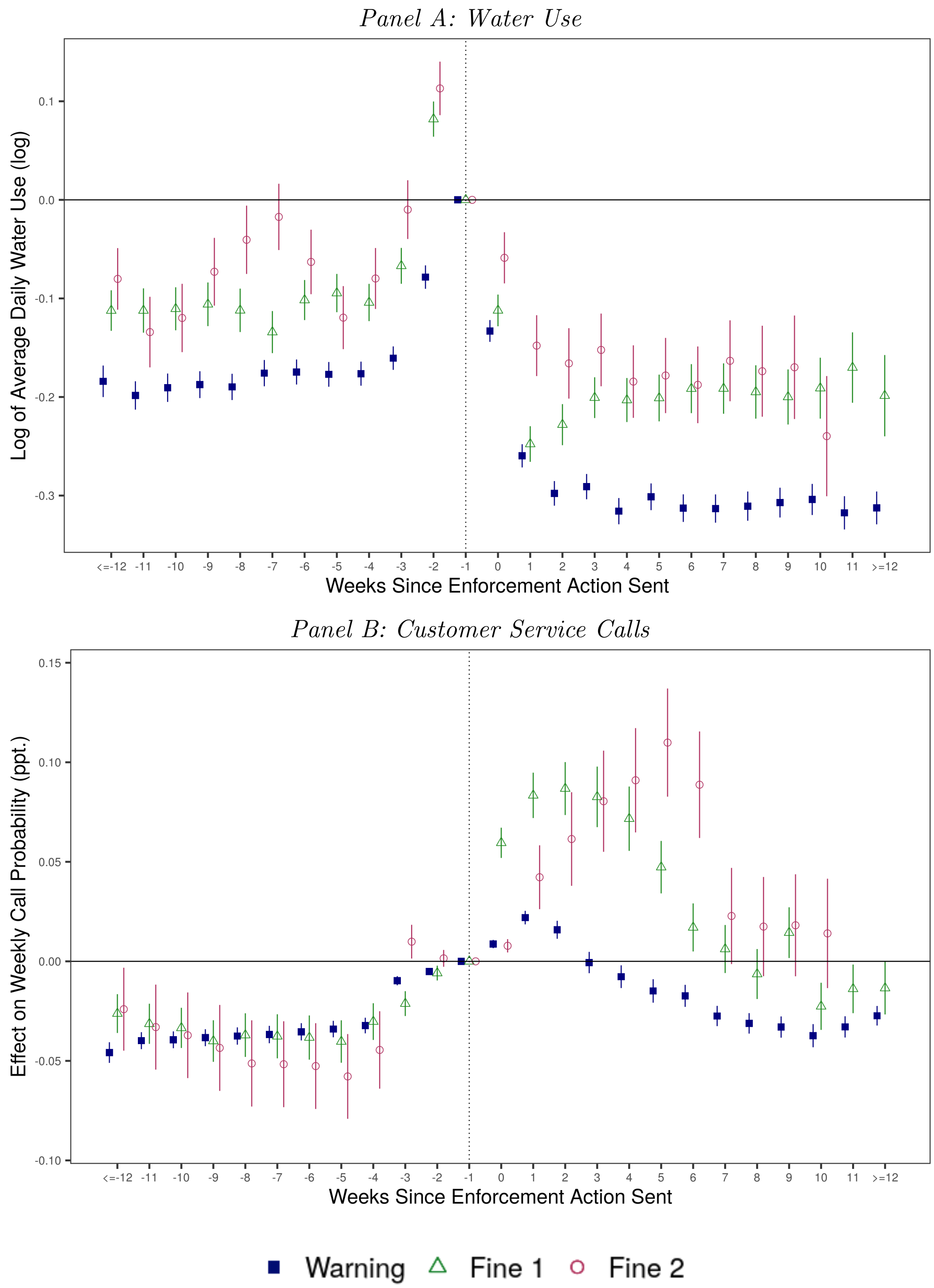


Panel B: Average Daily Water Use



Notes: Panel A plots the probability that a household in the automated (green) and non-automated (red) group calls utility customer service during July-September 2018, conditional on receiving a warning of violation. Values are multiplied by 100 for ease of visualization. Similarly, panel B plots average daily water use among households who received warnings in the automated (green) and non-automated (red) groups, conditional on receiving a warning of violation. The vertical lines in the plots represent 95% confidence intervals.

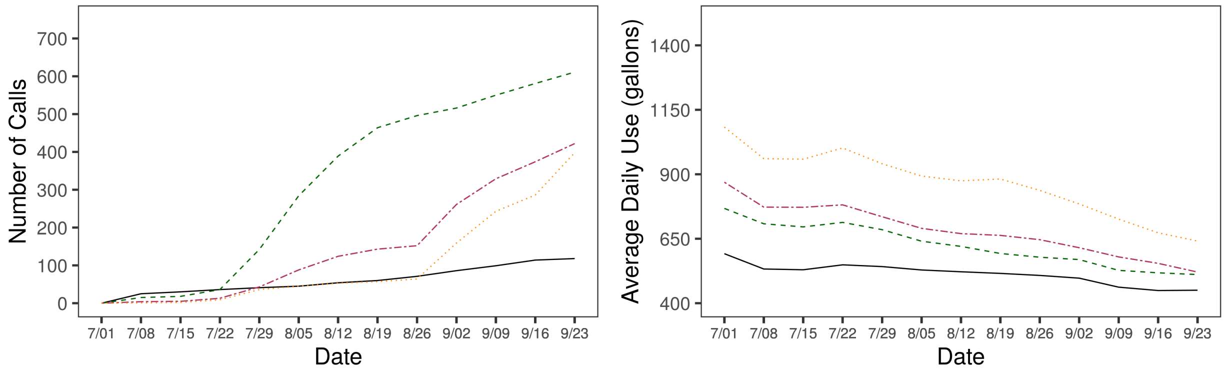
Figure A5: Effect of Fines and Warnings on Water Use and Customer Service Calls in Automated Enforcement Group



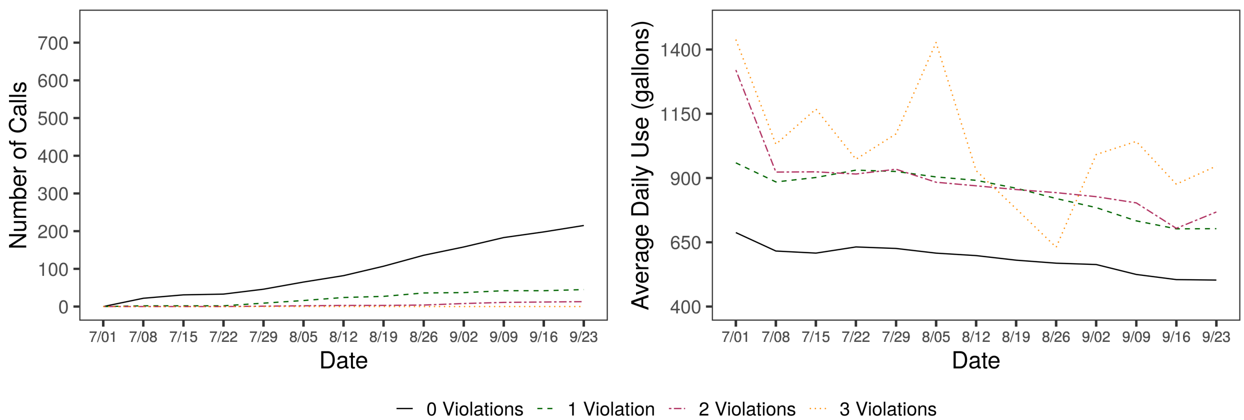
Notes: This figure plots coefficient estimates from a household-week level regression of an outcome variable on indicators for each week's timing relative to receipt of first, second, and third notices, that is first warning, and first and second fines. The outcome variable in the top panel is the log of average daily water use. The outcome variable in the bottom panel is the probability of calling customer service. Week 0 denotes the calendar week in which the warning/fine was sent, and week -1 is omitted as the reference category. The sample includes data from January 2017 through February 2019 for water use and from June 2018 to February 2019 for probability of calling DPU. The sample includes households in the automated enforcement group only. The sample drops notices waived ex-post. The regression includes household and week fixed effects. The vertical bars represent 95% confidence intervals based on standard errors clustered at the household level.

Figure A6: Water Use and Customer Service Calls by Violation History over Time

Panel A: Total Calls (Left) and Average Daily Water Use (Right) in Automated Enforcement Group



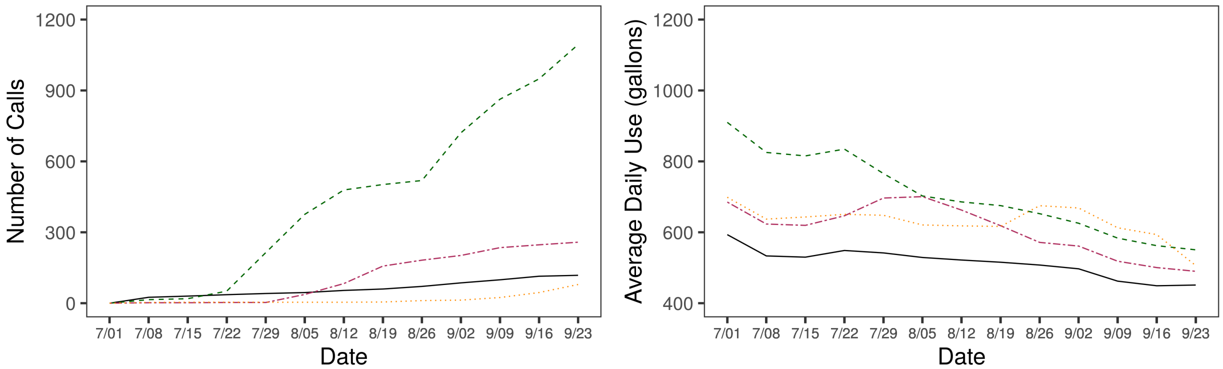
Panel B: Total Calls (Left) and Average Daily Water Use (Right) in Non-automated Enforcement Group



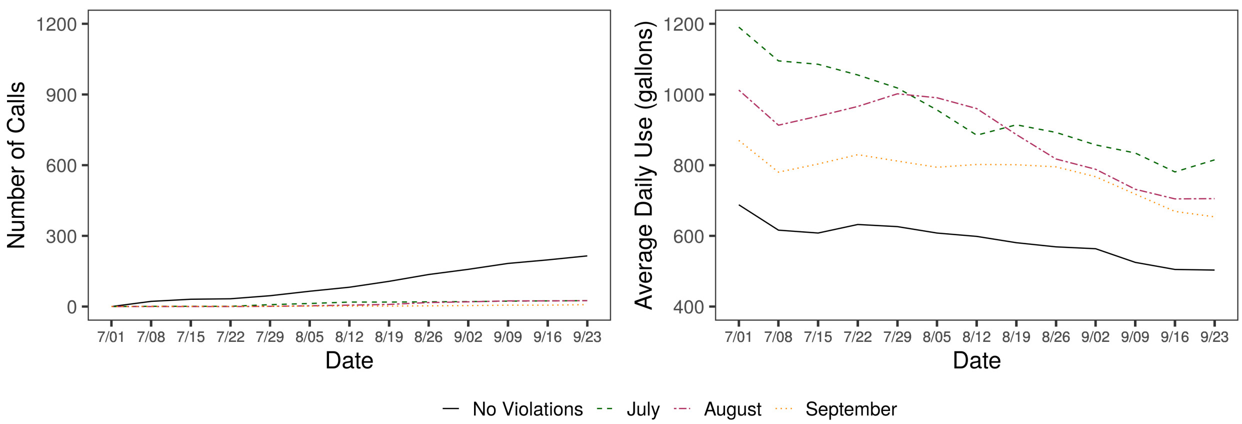
Notes: The figures plot cumulative customer service calls (left) and average daily water use (right) by week according to households' ex-post violation history (total number of enforcement actions received). Panel A includes households subject to automated enforcement, Panel B includes households subject to non-automated enforcement. All data are from July-September 2018.

Figure A7: Water Use and Customer Service Calls by Month of First Violation

Panel A: Total Calls (Left) and Average Daily Water Use (Right) in Automated Enforcement Group



Panel B: Total Calls (Left) and Average Daily Water Use (Right) in Non-automated Enforcement Group



Notes: The figures plot cumulative customer service calls (left) and average daily water use (right) by week according to the calendar month in which households' received their first enforcement action. Panel A includes households subject to automated enforcement, Panel B includes households subject to non-automated enforcement. All data are from July-September 2018 only.

B *Appendix: Decomposition*

This appendix decomposes the experimentally estimated reduction in water consumption due to automated enforcement by reporting treatment effects on water consumption among subgroups of the automated enforcement treatment group, where the subgroups are based on endogenous responses to the policy. Since the subgroups reflect households' responses, the results are associational and this must be considered an accounting exercise, in contrast to the causal estimates in Sections 3.1 and 3.2.

First, we estimate the following equation:

$$y_{it}{=}\alpha{+}\sum_{j=0}^{3}\beta_j \text{Automated}_i{*}I_i(j \text{ Violations}){+}\sum_{j\in\{25,50\}}\gamma_j \text{ In-Person}{\times}\text{Fines}_i^j{+}\varepsilon_{it} \tag{B1}$$

where y_{it} is an outcome for household i in month t, $Automated_i$ is an indicator for household i being randomly assigned to automated enforcement, $I_i(j \text{ Violations})$ are indicators for households who received j enforcement actions. $I_i(j \text{ Violations})$ is a vector of four indicators – zero violations, one violation (i.e., receipt of a warning only), two violations (i.e., warning and one fine), and three violations (i.e., warning and two fines). These independent categories, together, span the automated enforcement treatment group as households are only sanctioned for one violation per month in the three-month pilot. The estimated β_j's associated with these indicators provide estimates of each of these groups' water consumption relative to the control group.

We also estimate equation (B2) that further divides the three subgroups of violators into the periods before and after each violation, i.e.:

$$\begin{aligned}
y_{it}{=}&\alpha{+}\beta_0\text{Automated}_i{*}I_i(\text{Violated0Times}){+}\\
&\beta_1^{\text{BeforeWarning}}\text{Automated}_i{*}I_i(1\text{ Violation}){*}I_t(\text{BeforeWarning}){+}\\
&\beta_1^{\text{AfterWarning}}\text{Automated}_i{*}I_i(1\text{ Violation}){*}I_t(\text{AfterWarning}){+}\\
&\beta_2^{\text{BeforeWarning}}\text{Automated}_i{*}I_i(2\text{ Violations}){*}I_t(\text{BeforeWarning}){+}\\
&\beta_2^{\text{BetweenWarning\&Fine}}\text{Automated}_i{*}I_i(2\text{ Violations}){*}I_t(\text{BetweenWarning\&Fine}){+}\\
&\beta_2^{\text{AfterFine}}\text{Automated}_i{*}I_i(2\text{ Violations}){*}I_t(\text{AfterFine}){+}\\
&\beta_3^{\text{BeforeWarning}}\text{Automated}_i{*}I_i(3\text{ Violations}){*}I_t(\text{BeforeWarning}){+}\\
&\beta_3^{\text{BetweenWarning\&Fine}}\text{Automated}_i{*}I_i(3\text{ Violations}){*}I_t(\text{BetweenWarning\&Fine}){+}\\
&\beta_3^{\text{AfterFine}}\text{Automated}_i{*}I_i(3\text{ Violations}){*}I_t(\text{AfterFine}){+}\\
&\sum_{j\in\{25,50\}}\gamma_j\text{In-Person}{\times}\text{Fines}_i^j{+}\varepsilon_{it}
\end{aligned} \tag{B2}$$

where $I_t(\text{BeforeWarning})$, $I_t(\text{AfterWarning})$, $I_t(\text{BetweenWarning\&Fine})$, $I_t(\text{AfterFine})$ are indicators for months before and after a warning is sent to a household, after the warning but before the first fine (for repeat violators), and after the first fine, respectively.[18] Finally, both equations (B1) and (B2) include indicators for household i's assign-

[18]For three-time violators, the second fine is received outside the pilot period.

ment to an in-person enforcement treatment with fines at 25% or 50% of the baseline schedule (In-Person$\times$Fines$_i^{25}$, In-Person$\times$Fines$_i^{50}$).

Panel A of Table B1 reports estimates of equation (B1) in Column 1. Among households in the automated group, only those who never violate (63.5% of them) consume less water than households in the control group, specifically 15.5% less water. By contrast, households in the automated group who violated 1, 2, and 3 times consumed 9.4%, 22.5%, and 48.1% more water, respectively, than the average household in the control group, and account for 22.1%, 8.8%, and 5.6% of the automated group. Of course, there is a mechanical element to the finding about the three violating groups, because they are violators precisely because of their heavy consumption.

Column 2 of Table B1 reports estimates of equation (B2) and Column 4 details the number of months that the average household in each of the violating groups was in each of the subcategories.[19] None of the violating groups ever consumed less than the control group in a statistically meaningful way. At the same time, each notification of a violation is associated with a within group reduction in water consumption in all three of the violation groups, suggesting that households respond to the warning and to the fines. Finally, we note that because enforcement actions took time, these relative reductions in water use only accrued for relatively short periods (Column 4).

Next, we investigate whether the persistent water use decrease in the automated group can be attributed to technology investments (Allcott and Rogers, 2014). To do so, we divide the automated enforcement group into those that did and did not request major services such as timer tutorials and leak audits. Then we estimate the following two equations:

$$y_{it}=\alpha+\beta_0\text{Automated}_i*I_i(\text{NoServiceRequested})+\beta_1\text{Automated}_i*I_i(\text{ServiceRequested})+\sum_{j\in\{25,50\}}\gamma_j\text{In-Person}\times\text{Fines}_i^j+\varepsilon_{it} \tag{B3}$$

$$\begin{aligned}y_{it}=\;&\alpha+\beta_0\text{Automated}_i*I_i(\text{NoServiceRequested})+\\&\beta_1^{\text{BeforeService}}\text{Automated}_i*I_i(\text{ServiceRequested})*I_t(\text{BeforeService})+\\&\beta_1^{\text{AfterService}}\text{Automated}_i*I_i(\text{ServiceRequested})*I_t(\text{AfterService})+\\&\sum_{j\in\{25,50\}}\gamma_j\text{In-Person}\times\text{Fines}_i^j+\varepsilon_{it}\end{aligned} \tag{B4}$$

Panel B of Table B1 reports estimates of equations (B3) and (B4) in Columns 1 and 2. Among households in the automated group, only those who do not request major services use less water than the control group. For the 1% of households in the automated group requesting a service, water use is higher than the control group mean both before and after the request was made. These findings suggest that the automated enforcement group's persistent decrease in water use is not explained by permanent technology investments, at least those that we can measure.

Overall, Table B1's associational evidence suggests that the bulk of the reduction in water consumption comes from the nearly 64% of households that never violated the rules, but the detection of violations and resulting fines

[19] Rows within a compliance group in Column 4 might not sum to 3 because of missing observations for some household-months.

reduced water consumption of violating households over time. These findings are consistent with the idea that the threat of penalties and the penalties themselves produced the automated enforcement group's reduction in water consumption.

Table B1: Decomposition of Automated Enforcement Effect

	Log of Monthly Water Use (gal) (1)	Log of Monthly Water Use (gal) (2)	Share of Households in Automated Group (3)	Number of Months in Category (4)
Panel A: Enforcement & Compliance				
Automated Enforcement, Zero Violations	-0.155***	-0.155***	0.635	
	(0.006)	(0.006)		
Automated Enforcement, One Violation	0.094***		0.221	
	(0.008)			
Before Warning		0.188***		1.584
		(0.009)		
After Warning		-0.015		1.369
		(0.009)		
Automated Enforcement, Two Violations	0.225***		0.088	
	(0.010)			
Before Warning		0.355***		1.258
		(0.011)		
Between Warning & Fine 1		0.174***		1.368
		(0.011)		
After Fine 1		-0.031		0.368
		(0.019)		
Automated Enforcement, Three Violations	0.481***		0.056	
	(0.012)			
Before Warning		0.627***		1.025
		(0.012)		
Between Warning & Fine 1		0.488***		1.084
		(0.013)		
After Fine 1		0.303***		0.892
		(0.014)		
Panel B: Customer Contact				
Automated Enforcement, No Major Services Requested	-0.030***	-0.030***	0.992	
	(0.005)	(0.005)		
Automated Enforcement, Major Services Requested	0.203***		0.008	
	(0.038)			
Before Service Request		0.200***		2.032
		(0.039)		
After Service Request		0.282**		0.954
		(0.124)		
Control Mean	9.491	9.491		
N	261,311	261,311	40,011	

Notes: Panel A reports the log of average monthly water use among households with different ex-post compliance behavior in the automated enforcement group size relative to the non-automated, baseline fine group. Each regression includes data from July-September 2018 only. Column 2 reports coefficients from a regression that interacts ex-post compliance behavior with indicators for the months before and after the enforcement action (warning or fine) was sent to the household. Column 3 reports the share of households in the automated group in each compliance category. That is, those that received no enforcement actions, a warning, a warning and one fine, and a warning and two fines respectively during the experiment period. Column 4 reports the fraction of pilot months households in the automated groups in each compliance category spend before and after receiving each enforcement actions. Rows within a compliance group in Column 4 do not necessarily sum to 3 because of missing observations for some household-months. Panel B produces analogous estimates regarding household major service requests including interior and exterior audits, supply of hardware, and time tutorials. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$