Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage

Ian Ayres*
Yale Law School

Sophie Raseman** and Alice Shih†

By providing feedback to customers on home electricity and natural gas usage with a focus on peer comparisons, utilities can reduce energy consumption at a low cost. We analyze data from two large-scale, random-assignment field experiments conducted by utility companies providing electricity (the Sacramento Municipal Utility District [SMUD]) and electricity and natural gas (Puget Sound Energy [PSE]), in partnership with a private company, Opower, which provides monthly or quarterly mailed peer feedback reports to customers. We find reduction in energy consumption of 1.2% (PSE) to 2.1% percent (SMUD), with the decrease sustained over time (7 months [PSE] and 12 months [SMUD]). (*JEL* C44, D03, L94, Q41).

1. Introduction

In this article, we analyze two natural field experiments conducted by a third party on a total of approximately 170,000 household customers of two utilities, the Sacramento Municipal Utility District (SMUD) and Puget Sound Energy (PSE). These utilities, in partnership with a private company, Opower (formerly Positive Energy), randomly assigned a subset of these households to periodically receive mailed "home energy reports" comparing their energy usage to that of nearby neighbors in similarly sized houses. We find that households receiving Opower's reports make significant and lasting reduction in their energy consumption.

Since the circulation of our working paper analyzing the present experiments (which was the first academic paper to analyze the Opower data), several academic studies have analyzed Opower's results in different cities

*Townsend Professor, PO Box 208215, New Haven, CT 06511, USA.
Email: Ian.Ayres@yale.edu
**Email: Sophie.Raseman@aya.yale.edu
†Email: Alice.Shih@aya.yale.edu
Robert Cooter and seminar participants at Georgetown, Michigan Missouri, Pennsylvania, and Yale Universities provided helpful comments. Special thanks to Tyler Curtis and Alex Laskey for data and for responding to numerous questions.

The Journal of Law, Economics, and Organization, Vol. 29, No. 5
doi:10.1093/jleo/ews020
Advance Access published on August 20, 2012

(Costa and Kahn 2010; Allcott 2011; Allcott and Mullainathan 2011). Collectively, these Opower studies add to an extensive literature on the power of social norms to influence behavior in a variety of settings (see Cialdini and Goldstein 2004, for a review). This evidence extends to affecting "green" behaviors. For example, Goldstein et al. (2008) found that peer comparison information could increase towel reuse by hotel guests.

Changing one's behavior to conform to the behavior of others can be consistent with either "norm to conform" preferences or informational hypotheses. Learning that peers consume less (more) energy could increase (decrease) feelings of guilt about contributing to social problem and thereby impact private preferences and motivations to conserve. Alternatively, learning the behavior of peers might provide information about the possibility of alternative consumption choices and the relative benefits of those choices (Cooter et al. 2008). This study, like the other peer information studies, is not well structured to distinguish between these preference and information theories of behavioral change. For example, if households react to information that they are consuming less energy than their peers by increasing their usage, this change might be caused either by decreased guilt or by the households' making a Bayesian inference that they are missing out on valuable consumption opportunities.

In a literature review of the effect of feedback on home energy consumption, Fischer (2008) notes that of the dozen studies that she reviews that test the impact of peer comparison information, none had shown an effect on usage. She attributes this failure to the "boomerang" problem, where informing individuals of typical peer behavior causes those who have lower-than-average peer consumption to respond to peer information by increasing their home energy consumption. Cialdini et al. (1991) show that combining injunctive norms (norms that express social values rather than actual behavior) with descriptive norms can neutralize the boomerang effect. Schultz et al. (2007) conducted a randomized field study in San Marcos, CA, of the effectiveness of social norms messaging (alongside energy-saving tips) to reduce home energy consumption. See also Nolan et al. 2008. They found that combining the descriptive and injunctive messages (in this case, the emoticons ☺ and ☹) lowered energy consumption and reduced the undesirable boomerang effect.

The Opower experiments detailed in this article build on the findings of the San Marcos study. As in the San Marcos study, the Opower reports use descriptive norms as well as injunctive norms, such as ☺ emoticons, to reduce consumption and in order to counteract the boomerang effect. The Opower experiments reported here, however, go beyond the San Marcos experiment in a number of ways. First, the Opower experiments have a significantly larger sample size than that of the San Marcos experiment, which included 290 treated households, compared to 35,000 in the SMUD study and 40,000 in the PSE study. Second, the

Opower studies also allow us to test multiple new aspects of the dynamics of energy use feedback:

- Measuring longer term impacts. Whereas the San Marcos study's observation period was only 1 month, the SMUD and PSE experiments encompass 12 and 7 months of post-treatment data, respectively.
- Measuring daily impacts. Unlike the San Marcos study, the PSE experiment gives access to daily energy readings.
- Measuring impacts on both electricity and natural gas. The PSE experiment tested the effect of feedback on both electricity and natural gas usage, allowing for a fuller picture of household energy use.
- Measuring impacts of different message frequencies (quarterly versus monthly), different report content, and different envelop sizes.

Moreover, the Opower experiments were conducted using a more realistically scalable intervention. Instead of mailed reports, the San Marcos study used hanging doorknockers with hand-drawn emoticons. Together, the SMUD and PSE experiments provide compelling evidence that properly framed peer comparisons (combined with historical usage information and personalized tips) can predictably lower energy consumption, particularly of the highest energy using households.

The experiments analyzed here contribute to a growing literature on the impact of different forms of disclosure and information on energy use. Studies testing the effect of customer feedback on residential energy consumption have been promising, although these studies have often had small sample sizes and methodological limitations (see literature reviews by Abrahamse et al. 2005; Darby 2006; Fischer 2008; Ehrhardt-Martinez et al. 2010; and Faruqui et al. 2010). The Opower studies add to the evidence that social norms messages are one powerful form of feedback for energy consumers. There is also an emerging literature showing that social norms messages can promote other forms of green behavior, including reducing residential water usage (Ferraro and Price 2011) and installing energy-efficient light bulbs (Herberich et al. 2011).

The Opower results also provide evidence that disclosure can impact consumer behavior. Federal and state policymakers frequently mandate information disclosure as a tool to regulate consumer markets. The empirical evidence on disclosures, however, is inconsistent. For example, public hygiene grades for restaurants have influenced consumer choices about where to eat and led to fewer foodborne illness hospitalizations in Los Angeles (Jin and Leslie 2003). Yet a number of empirical investigations of disclosures have found that they have limited or no impact (e.g. see Ben-Shahar and Schneider 2011, for a review of the failures of mandatory disclosure policies in a number of important US consumer markets). The varied track record of disclosures suggests that their effectiveness will depend heavily on the framing, content, and delivery of the information being conveyed, and that it is critical that disclosure policies be informed by rigorous empirical analysis of which approaches are the most effective.

2. SMUD Experiment

We begin by analyzing the results of a field experiment to assess the impact of home energy report on energy consumption in the SMUD.

2.2 Experimental Design

The SMUD messaging experiment began in April 2008 and is still ongoing; the results presented in this paper cover the period from April 2008 through April 2009.[1] The sample includes approximately 84,000 households who are customers of SMUD. To select participants, Opower filtered by census tract within SMUD's customer base to maximize the number of single family homes with more than 12 months of billing history, that were on standard rate plans (nonmedical rate, nonphotovoltaic), and that had a matching parcel record with details about the home, such as house size and value.

Once participants were selected, the randomization process implemented a "batch" assignment: 959 batches of census blocks were randomly assigned to the treatment and control groups. These "batch blocks" consist of 50–100 homes. 35,000 households were in batch blocks assigned to the treatment group, and 50,000 were in batch blocks assigned to the control. Opower used this assignment methodology to increase the likelihood that neighbors would receive reports and have the opportunity to discuss the reports with each other, thereby increasing the motivation for taking actions to reduce home electricity consumption.

All members of the treatment group received home energy reports on a periodic basis. Each home energy report contains four key personalized components: (1) Current period neighbor comparison: A bar chart comparing the household's recent electricity use to a group of comparable neighbors and "efficient neighbors," with both normative and injunctive messages designed to motivate action;[2] (2) Historical peer comparison: A chart comparing the household's electricity usage to its comparable

1. All the data in this article, including data originally obtained from the utilities themselves as well as from third parties, was generously provided to the authors by Opower, who were responsible for designing and conducting the experiments. The authors of this article have received no compensation from and have no financial interest in Opower. SMUD has contracted with ADM & Associates to independently assess the success of the program. In addition, Opower engaged Summit Blue to independently evaluate the SMUD data. PSE also plans to select a third party to conduct independent program measurement and verification analysis.

2. The SMUD energy reports initially included between 1 and 3 smiling or frowning emoticons, but the frowning emoticons were discontinued shortly after the experiment after SMUD received a handful of complaints, and frowning emoticons were not used in the Puget Sound experiment.

neighbors and "efficient neighbors" over the last 12 months; (3) Personal historical comparison: A section comparing the household's usage in the current year by month with that household's usage in the same months from the previous year; (4) Targeted energy efficiency advice: tips selected based on the household's energy use pattern, housing characteristics, and household demographics. All reports were printed in color on a single 8½" × 11" sheet of paper. Examples of the elements of the front page of this report are included in Appendix Figure A1. Because of the multi-faceted nature of the intervention, the experiment cannot distinguish the impact of the "peer comparison" information from the impacts of the other elements of the report.

The 35,000 treatment households were then assigned to different sub-treatment groups to receive the reports either monthly or quarterly based on historical usage levels: the 25,000 households with higher consumption levels were assigned to the monthly frequency group, while the 10,000 households using less energy (<21.85 kWh/day) were assigned to receive the report quarterly.[3]

SMUD provided the basic data on energy consumption, including historical billing information dating back to January 1, 2006 (over two years before the beginning of the treatment in April 2008). Data on household parcel characteristics (such as square footage and home values) comes from the Sacramento County Assessor's Office.[4] Household-level demographic data (such as estimated income level and length of residence) comes from private direct marketing and data aggregation service databases.

2.3 SMUD Results

Appendix Table A1 investigates whether the sample is well-balanced between the control and treatment groups. The random assignment at the census batch level, as opposed to the individual household level, allowed some statistically significant differences in some pretreatment variables. For example, the households in the treatment group on average were 16

3. All households (treatment and control) were also randomly assigned to one of two different report template groups and one of two different envelop-size groups. The two report template groups were "graphical" and "graphical + narrative." Both templates included the same core elements, including graphs with feedback information, but the narrative version (shown in Appendix Figure A1) included a blurb of text explaining the charts, reinforcing the normative messages, and highlighting tips on how to save energy (including both mentioning tips in the blurbs and pointing the reader toward the personalized tips section on the back of the report). The two envelop types tested included a standard business ″#10″ envelop (similar to the envelop used to deliver SMUD customer bills) and a larger 6″ × 9″ envelop. All envelops displayed the SMUD logo and return address. Envelop size did not affect the envelop content, which was always printed on 8½″ × 11″ paper; but folded differently to accommodate the different envelop sizes. These subgroup randomized tests of template and envelop size were not statistically significant and not reported here.

4. The heating fuel type was derived from the customers' rate codes, as SMUD offers lower rates to households with electric heat.

square feet smaller and used 0.3 kWh/day more in 2006 than the average control group households.[5]

Households in the treatment group that complained about receiving the Opower reports or who asked to stop receiving the report were allowed to opt out of the treatment. Only 2% of the treatment group opted out of the experiment. The following regressions, which retain these observations and which only control for pretreatment variables, should be interpreted as "intent to treat" effects. Unreported treatment on the treated estimates were of similar in size and significance. In addition, similar proportions of treatment and control households (8% and 7%, respectively [$p = 0.10$]) moved after the experiment began, and closed their SMUD accounts.

Figure 1 reports the results from monthly regressions on approximately 84,000 household observations where the log of monthly average kWh/day was regressed on a treatment group indicator and a constant. As shown in Figure 1, the treatment group's energy consumption (relative to the control group) moved erratically before the start of the experiment (indicated by a vertical line marking April 2008). For example, the treatment group used more electricity than the control group in February 2007 and less in June 2007, and these differences were statistically significant. Still, even before other factors are controlled for, there was a significant drop in energy usage for the treatment group relative to the control for all the months following the initial report mailing.

To account for factors besides the reports that may be driving the change in energy usage, we control for house characteristics (square feet, age of house, presence of pool or spa, house value, gas user, census tracts), household demographics (energy usage in 2006,[6] length of residence at particular house, number of residents, income, age, affluence), and the average number of cooling degree and heating degrees per day in each billing cycle.[7] Figure 2 shows that after controlling for these characteristics there was no systematic difference in energy usage between the treatment and control groups before the experiment began. With the exception of one month in the pretreatment stage, the difference between the energy usage of the control and treatment groups is statistically insignificant, straddling 0%. After the first reports arrived around April 15, 2008, we observe a significant drop in the electricity consumption of treatment

5. A parallel analysis (also reported in Appendix Table A1) of the subrandomization of envelop size and the graphical/narrative template within the treatment group shows that the data was well balanced between these groups.

6. We also constructed an indicator variable called "Tier 2 Pricing in 2006" that was set equal to one, if a household's average energy usage per day in 2006 placed the household into the utility's higher marginal price for energy.

7. Cooling and heating degree days are based on deviations from a base temperature of 65°. For example, a day with an average temperature of 68° will count as three cooling degree days. Similarly, a day with an average temperature of 62° will count as three heating degree days. Cooling and heating degree days are then divided by the duration of the billing cycle to generate a daily average.

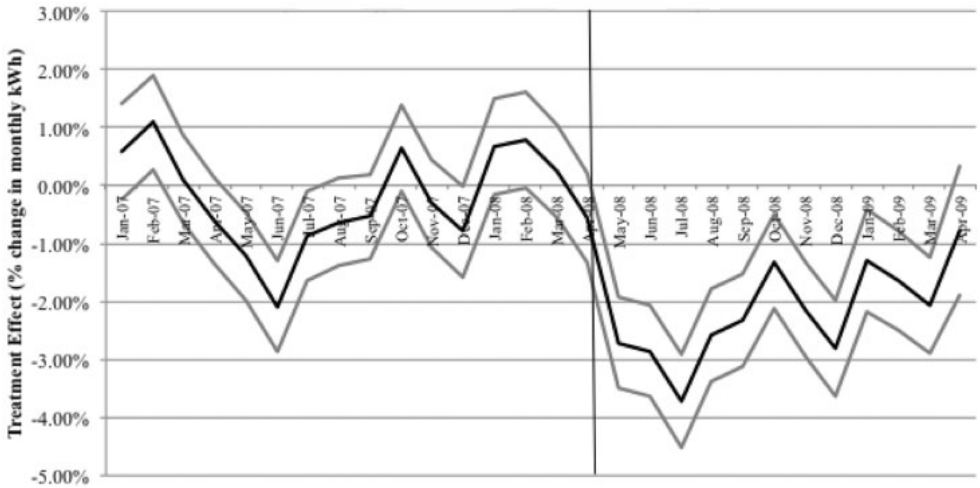

Figure 1. SMUD Treatment Effect (% change in kWh—without controls). Notes: 95% confidence intervals shown in gray. Vertical lines indicate first mailing. OLS regression on natural log of kWh/day clustered on household with same controls as in Table 1.

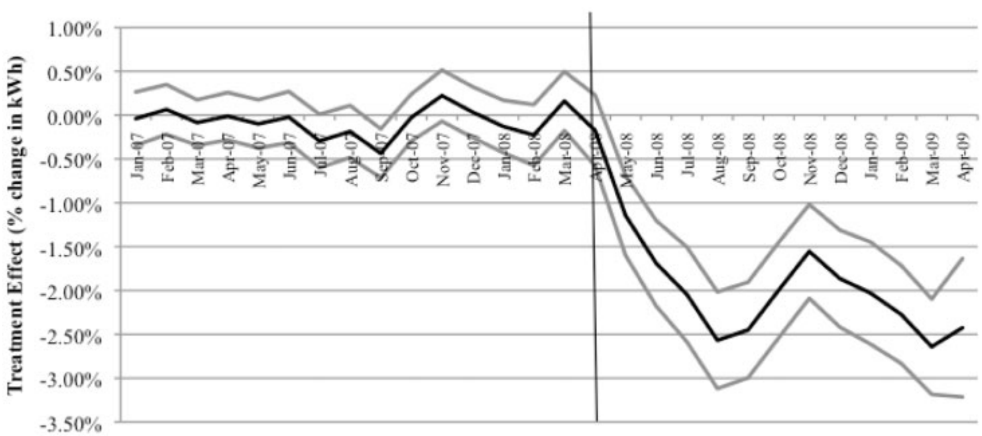

Figure 2. SMUD Treatment Effect (% change in kWh—with controls). Notes: 95% confidence intervals shown in gray. Vertical lines indicate first meeting. OLS regression on natural log of kWh/day clustered on household with same controls in as Table 1.

households relative to control households, on the order of 1% in May 2008. There is a steady decline until August 2008, when the treatment group saw a reduction in electricity usage by more than 2.5%. The gap between the usage levels of the control and treatment groups then narrows in the fall months (September–November 2008), though the reductions made by the treatment group are still significantly negative relative to the control group. After November 2008 the effect of the treatment grows in all months except April 2009, with the greatest reduction in electricity consumption since the beginning of the experiment (greater than 2.5%) occurring in March 2009, almost a year after the study first began.

To simultaneously investigate the impact of treatment across different months, we then "stacked" the household-month data and again regressed the log of average monthly kWh/day for individual households on the controls reported in Table 1 (calculating standard errors by clustering on household IDs). The interaction between "Treatment" and the "After First Mailing (April 2008)" variable captures the effect of being in the treatment group after the start of the experiment. The average effect of the treatment on energy reduction is significant and robustly estimated in Table 1 at about 2%, with or without ancillary controls and in a third specification in which we control for 83,695 household fixed effects.

Table 2 tests whether there are heterogeneous treatment effects by re-running the household fixed-effects regression from Table 1 but adding various interaction terms. We find no evidence in either specification that households with larger pre-experiment energy usage per square foot (kWh per square foot per day in 2006) exhibited larger energy reduction after receiving the peer message. Larger houses, however, did exhibit more pronounced treatment effects. The last column of Table 2 indicates that a one standard deviation increase in pre-experiment energy usage per

Table 1. SMUD OLS Regression of Log Household Monthly Average kWh/day

	No controls $n=2{,}262{,}815$	With controls $n=1{,}585{,}490$	With household fixed effects $n=2{,}262{,}815$
Treatment household	−0.0013	−0.0006	
Post April 2008 (first mailing)	−0.0183**	0.0784***	0.0510**
Treatment × post April 2008	−0.0199	−0.0215***	−0.0202***
Cooling degree days (per billing cycle)		0.0021***	0.0021***
Heating degree days (per billing cycle)		0.0006***	0.0006***
Narrative template		0.0013	
6 × 9 envelop		0.0009	
Quarterly report recipients		−0.1163***	
House sq. ft. (in 100s)		0.0034***	
House age		0.0002**	
Pool		0.0457***	
Spa		−0.0030	
House value (in $100,000s)		0.0042***	
Gas heat		0.0317**	
kWh/day usage in 2006		0.7595***	
Tier 2 price in 2006		0.0296***	
Length of residence		−0.0006***	
Number of residents		0.0079***	
Head of household age effects	No	Yes	No
Income quartile effects	No	Yes	No
Affluence effects[a]	No	Yes	No
Proprietary segment effects[b]	No	Yes	No
Census tracts fixed effects	No	Yes	No
Monthly fixed effects	No	Yes	Yes
Household fixed effects	No	No	Yes
R^2	0.0007	0.7061	0.7858

Regressions clustered on households with Huber-White robust standard errors. [a]Ten Affluence groups were created by Direct Group. [b]Proprietary segment groups created by OPOWER based on house characteristics.*, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

Table 2. Test for Heterogeneous Treatment Effects in SMUD OLS Regression of Log Household Monthly Average kWh/day

	I	II	III	Std. Error.	Effect of 1 Std. Dev. Increase (based on specification III)
T × post (April 2008)	−0.020***	0.025***	0.154**		
T × post × Quarterly		0.007*	0.004		
T × post × house value (in $100,000)		0.002*	0.002	1.458	0.003
T × post × kWh per sq. ft. per day in 2006		−0.069	0.010	0.009	0.000
T × post × tier 2 price in 2006			−0.191***		
T × post × length of residence			0.000	9.929	−0.001
T × post × spa			−0.014**		
T × post × house sq. ft. (in 100s)			0.033	5.968	0.196
T × post × house age			0.000	19.603	0.003
T × post × 6" × 9" envelop			0.005		
T × post × narrative			0.003		
T × post × number in residence			−0.001	0.437	0.000
T × post × pool			0.012**		
T × post × cooling degree days			−0.002***	3.017	−0.007
T × post × heating degree days			0.000	7.341	0.002
T × post × gas heat			0.006		
Monthly fixed effects	Yes	Yes	Yes		
Household fixed effects	Yes	Yes	Yes		
Observation	2,262,815	2,262,767	1,585,490		
R^2	0.79	0.79	0.79		

Variations on third specification in Table 1. Regressions clustered on households with Huber–White robust standard errors.*, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

square foot is associated with a 1.9 percentage points reduction in energy usage for the treated households. But relatedly, the third specification did find statistically larger energy-reduction treatment effects for households whose average energy consumption in 2006 placed them in the higher of two pricing tiers (where the marginal price per kWh was ~70% higher, averaging about 15 versus 9 cents per kWh). The more complete specification (in Column III) reports larger energy-reducing treatment effects for households with spas, and when there were more cooling degree days. In contrast, we find among households with a swimming pool that the marginal estimated effect of receiving an energy message was to increase energy usage by 1.2 percentage points.

To further explore the interaction between treatment and pre-experiment household energy usage, we categorized the households into deciles of pretreatment energy usage per house square foot (based on the available usage data for the 15 months prior to experiment), and reran the Table I "With Controls" regression adding these decile indicators fully-interacted with the Post and Treatment variables. Figure 3 reports reductions in energy usage of treated households relative to untreated households in the same decile. Figure 3 shows that treatment effects tended toward larger percentage reductions for households with larger pretreatment usage. The lowest 5 deciles were estimated to have smaller than average energy reductions, while 3 of the 5 highest deciles had higher than average energy reductions. The figure presents no evidence that treated households in the lower deciles of pretreatment energy usage exhibited a "boomerang" effect—increasing their energy usage when they learned that there consuming less electricity than their peers (Cialdini et al. 1991). The second to lowest decile was estimated to have a treatment effect that was statistically indistinguishable from zero (at a 5% confidence level), and no deciles exhibited a statistically significant increase in energy usage. The absence of a boomerang effect is consistent the findings of Schultz et al. (2007) in the San Marcos study, where lower-consuming households who received peer reports coupled with emoticons did not increase their energy usage. Even though the regression behind Figure 3 controlled for whether the household was a quarterly recipient, these decile results should be viewed with caution since the lower decile households were dominantly in batch blocks that were nonrandomly assigned to receiving quarterly reports. The lack of a boomerang effect might be confounded with the assignment of the lowest pretreatment energy user to a different kind of treatment.

We also investigated whether the impact of the treatment on energy usage decayed with the time since the last message (what we will call a "staleness" effect) or with the total number of peer reports a household has received (what we call a "routinization" effect). To test for these heterogeneous treatment effects, we reran the household fixed effects regression on just the quarterly data from Table 1 and interacted the core treatment regressor (Treatment $\times$ Post-April 2008) with three new

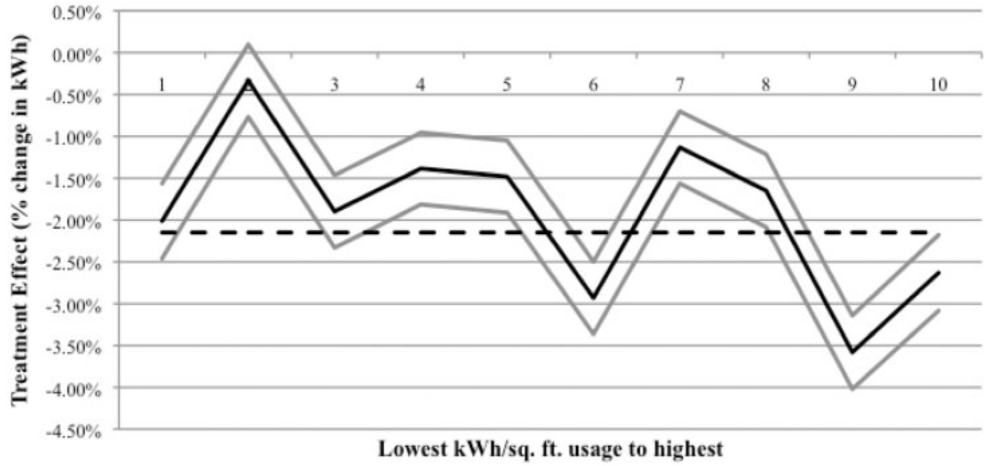

Figure 3. SMUD Treatment Effect (% change in kWh) by Pretreatment kWh/sq. ft. Usage by Decile. Notes: 95% confidence interval is shown. Horizontal dashed line indicates average change in kWh. OLS regression on natural log of kWh/day clustered on household id.

variables: "Numbers of Messages", "One Billing Cycle After the Last Message", and "Two Billing Cycles After the Last Message."[8] Table 3 reports the results of our staleness and routinization tests for both linear and quadratic "Number of Messages" specifications. In the third specification, the negative (and marginally significant) coefficient on the "Number of Mailings" variable suggests that instead of leading to routinization, repeated messages tended to reinforce the energy-reducing treatment effects. The repetition of up to 14 monthly messages in our sample tended to enhance energy reduction rather than causing households to grow habituated or numb to the message.

The insignificant coefficients on the energy usage coming one or two months after the last energy message indicate that the energy reducing treatment effects were not becoming more muted over time—suggesting that for households who only received the message once per quarter that the message's power was not becoming staler as the memory of the last message fades with time.

Table 4 estimates the potential yearly impact of the reports on both dollars saved and energy conservation if SMUD were to send the reports to all of the households in its customer base. At an average reduction of 2.35% for monthly recipients, the reports would reduce consumption 187 kWh per year per household for a total savings of about $26 a year per household. These average treatment effects correspond to a household reducing its use of a 75 watt bulb by about 6.8 h/day (or reducing air conditioner usage by about half an hour). While the modest percentage

8. For quarterly households (who received three bills per quarter), the billing cycle that included the date the energy message was received was the omitted category.

Table 3. SMUD OLS Regression of Log Household Monthly Average kWh/day, Clustering on Household ID with Robust Standard Errors (quarterly households only)

	With household fixed effects $n=656{,}403$	With household fixed effects $n=656{,}403$	With household fixed effects $n=656{,}403$
Treatment × post April 2008 (first mailing)	−0.01472***	−0.01372***	−0.00906***
Treat × post × no. of messages		−0.00001	−0.00673*
Treat × post × (no. of messages)^2			0.00151
Treat × post × one billing cycle after last message		−0.00107	−0.00022
Treat × post × two billing cycles after last message		−0.00271	−0.00173
Month fixed effects	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes
R^2	0.61	0.61	0.61

Regressions clustered on households with Huber-White robust standard errors. "One Billing Cycle After Last Message" refers to the bill received following the last energy report. "Two Billing Cycles After Last Message" refers to the secnd bill received following the last energy report.*, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

Table 4. SMUD Projected Cost Savings and Environmental Impact

	Monthly and quarterly weighted effects
Per household	
Reduction kWh/day	0.51
Reduction kWh in a year	187.20
Total savings in a year	$25.74
Savings per mailing	$2.78
For customer base of SMUD	
Annual kWh reduction	110,917,005
Annual reduction in metric tons CO2[a]	79,638
Annual savings	$15,250,601
Estimated short-term price increase necessary to produce comparable energy reduction[b]	7.02%
Estimated long-term price increase necessary to produce comparable energy reduction[b]	1.57%

[a]Annual reduction in metric tons CO2 was calculated by first calculating the average annual reduction for monthly recipient and quarterly recipients (using treatment effects weighted by prevalence of monthly and quarterly household recipients). The weighted average reduction was multiplied by the average kWh/year in 2008 to calculate the average yearly household reduction in kWh. This result was multiplied by the number of households in Sacramento to calculate the total annual kWh reduction in the city. The annual kWh reduction for the city of Sacramento was then converted to CO2 metric tons (assuming 7.18×10^{-4} metric tons CO2 / kWh calculated by the EPA). [b]Based on Bernstein and Griffin's (2005) regional price elasticity of electricity calculations

reductions (and associated behavioral changes) make the estimated impacts of the experiment more plausible, it should be remembered that some of the report recipients are likely to have discarded the unopened energy reports as junk mail (despite the fact that they arrived in envelops bearing the SMUD logo and return address). Using regional elasticity estimates (Bernstein and Griffin 2005), we find that to induce equivalent energy reductions, an energy excise tax would require a price increase of 7.02% in the short run (or 1.57% in the long run).

If the nearly 593,000 households in SMUD's customer base received reports using the same formula by which SMUD treatment households were assigned to monthly or quarterly reports and if the average percentage treatment effects could be validly extrapolated to the city as a whole, we could expect to see a reduction of over 110 million kWh in a year—the energy equivalent of conserving nearly 80,000 metric tons of carbon emissions. SMUD customers city-wide would save on their energy bills in aggregate over $15.2 million. But such extrapolations should be viewed with caution. The 13 months of this experiment cannot speak to long term treatment effects. The treatment effects on single family detached houses included in the experiment may differ from those that would pertain in the larger universe of SMUD accounts (as the SMUD system-wide average usage is about 2000 kWh per year lower than the average for households in the experiment).

Table 1's treatment estimates imply average household energy savings per mailing of $2.78. By combining this evidence of household energy savings with information about the costs of producing and mailing the energy reports, we estimate that on average the mailings cost 4.94 cents per kilowatt-hour saved. This estimate of cost effectiveness is on par or lower than those of other energy efficiency programs (Auffhammer et al. 2007; Friedrich et al. 2009; Allcott 2011). Since, as shown in Figure 3, higher energy users made significantly larger reductions in energy, it appears likely that targeting reports at only higher energy consumers would be even more cost effective. In an analysis conducted across a number of Opower experiments, Allcott and Mullainathan (2010) estimate that the Opower-type treatment costs an electric utility 2.5 cents/kWh saved. An important question for further study is how effective messages similar to Opower's can be through electronic communication channels, such as delivery through email, a web browser, or to a mobile phone or other mobile device. Such an approach is promising as it has the potential to significantly reduce the cost and increase the efficiency of the intervention.

3. PSE Experiment

3.1 Experimental Design

In October 2008, Opower launched another home energy report experiment, this time working collaboratively with PSE in King County, WA. There were three major differences between the experimental designs of the SMUD and PSE studies: first, the PSE reports encompassed both electricity and natural gas, allowing for a more complete picture of each household's energy use; second, the PSE study randomly assigned households to either monthly or quarterly report frequency;[9] and third, the study used household-level randomization, instead of less balanced batch-level randomization used in the SMUD study.

The PSE experiment consisted of approximately 84,000 homes randomly assigned to control and treatment groups. These homes were chosen from PSE's 1.3 million residential customers who met the following criteria:

- Single family homes located in King County, WA
- Exactly one active electric account and one active gas account with PSE
- History for both gas and electric accounts dating to January 2007
- Matched parcel record available from the King County Assessor's data
- Not identified by the King County Department of Assessments as having solar heat

"All electricity" households that rely on electricity for heating were excluded from the experiment to focus more cleanly on households that rely primarily on gas for heating. These filters created a pool of approximately 100,000 households that were eligible to participate in the program. Additional exclusions were made to eliminate homes with distant neighbors or with unusual home sizes (so that neighbor comparisons would be more meaningful) and homes that used relatively little energy (less than approximately 23,000 kWh per year). In order to test the effect of the frequency of the reports on home energy consumption, households were also randomly assigned to receive the report on a monthly or quarterly basis in the ratio of 3:1. Sample elements from the front page of this report are included in Appendix Figure A2. Unlike in the case of SMUD, the PSE report included energy information regarding both electricity and natural gas consumption. The report began with a Combined Energy Cost (CEC) comparison, in addition to two charts tracking the last 12 months of household kWh and therm consumption relative to nearby neighbors in similar size homes.[10]

9. Unlike in the SMUD experiment, all of the PSE reports used the same "graphical" template and standard-business envelop size.

10. The combined energy cost (CEC) is an estimate of the cost of electricity and gas used by the household. On the reports the combined energy cost was reported in terms of a price-weighted index (PWI), where PWI = 12.51*therms + kWh. The factor 12.51 represents the kilowatt-equivalent price of one additional therm for a PSE customer. An estimate of the CEC can then be found by multiplying the PWI by the approximate price of 1 kWh, 8 cents. The combined energy cost does not exactly reflect the relative costs to the households because the actual pricing formula took into account other factors (e.g., fixed costs).

3.2 PSE Results

Appendix Table A2 shows that the randomization was successful in producing treatment and control households with similar pretreatment attributes.[11] Only 1% of the treatment group opted out of receiving the reports, which, as in the SMUD experiment, suggests that the following intent to treat estimates will be nearly identical to treatment on the treated effects. About 2.3% of both the control and treatment households closed their accounts with PSE during the experiment because they moved away.

Figure 4a and b reports the results of regressions of the log of monthly average kWh per day and therms per day usage on a treatment indicator and a constant. Unlike SMUD (Figure 1), where census-tract level randomization created some substantial pre-experiment differences between treatment and control households, the PSE data show no substantial differences in pre-experiment usage. All differences in energy usage between the control and treatment groups for pre-experiment usage, as expected, were statistically insignificant and close to 0%. Figure 4a and b shows that the treatment households reduced their use of both electrical and natural gas energy relative to the control households in November 2008 (the first full month after the reports were sent out on October 20).

Table 5 displays the results of stacked regressions (analogous to the SMUD regressions displayed in Table 1) on approximately 1.5 million household-month observations. The regressions are run on the log of three measures of energy use: average monthly kWh per day, average monthly therms per day, and the average monthly CEC. As in the SMUD Table 1 analysis, we report the results of parallel regressions with and without controls for house demographics (such as square footage, age of house, house value), household demographics (such as past energy usage), month, and cooling degree days and heating degree days. As with the SMUD data, the estimated treatment effects are quite robust to the inclusion of ancillary controls. On average, households in the treatment group reduced kWh usage by 1.2%, therm usage by 1.2% or 1.3%, and a combined price-weighted usage by 1.2% compared with the control group.

One potential explanation for why the estimated PSE treatment effect is smaller than the SMUD average is that the PSE experiment has been running for a shorter time. But restricting the SMUD data to include just the first 8 months of postexperiment energy usage had no impact on the size or significance of the estimated treatment effects when we reran the regressions in Table 1. Alternatively, the treatment effect might be larger in SMUD if PSE households tended before the experiment to be more energy conscious. We find some support for this hypothesis as we

11. The table does reveal some statistically significant differences between the randomly assigned monthly or quarterly groups, but the raw differences in levels were not substantial (for example, in 2007 the average kWh per day was 30.2 and 30.5 for the monthly and quarterly households, respectively).

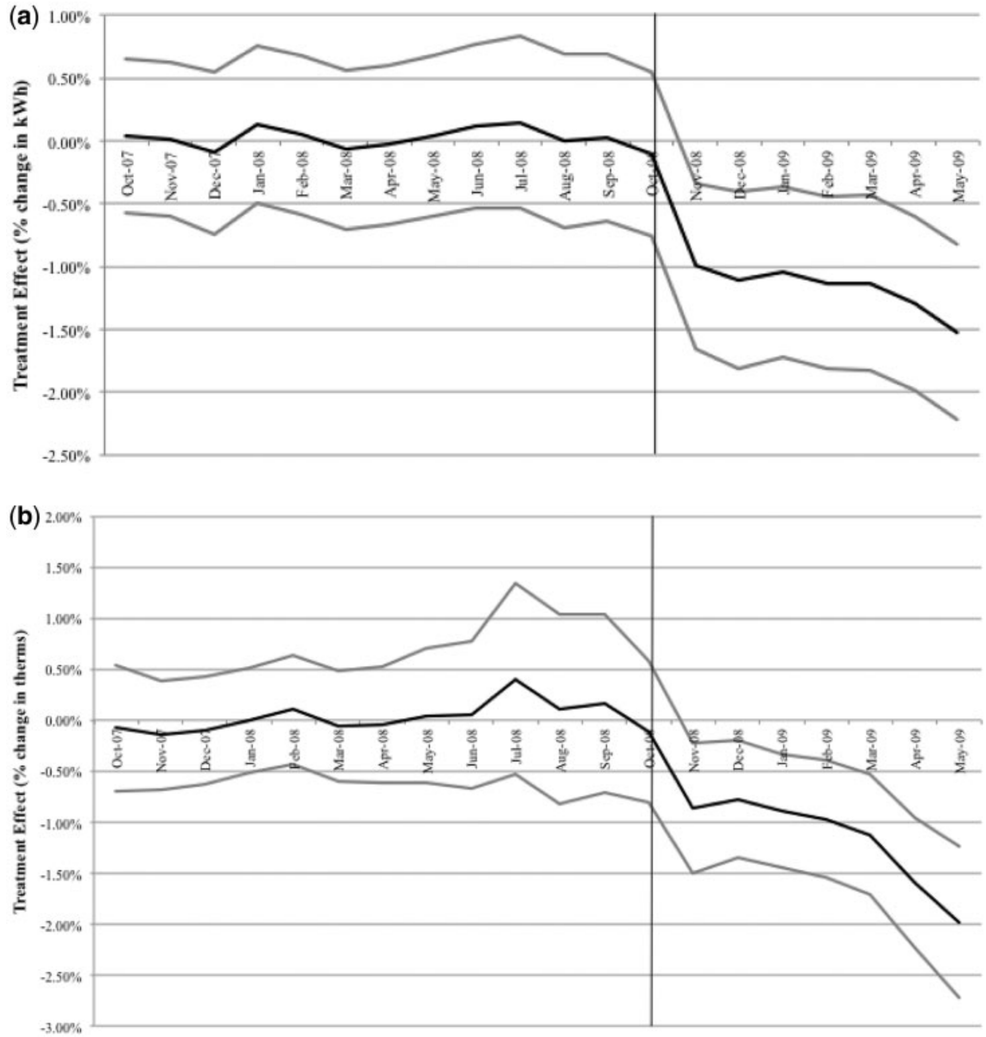

Figure 4. PSE Treatment Effect (% change in kWh and therms—without controls). Notes: 95% confidence intervals shown in gray. Vertical lines indicate first mailing. OLS regression on natural log of kWh/day clustered on household id with same controls as in Table 3.

estimated lower pretreatment electricity usage (as measured by kWh per square feet per cooling degree day) in Puget Sound than in Sacramento (restricting our attention to households in both cities that used gas heat) (see Ayres and Nalebuff 2005). This, at least weakly, suggests that the size of the treatment effect will be smaller in cities where there is greater pre-experiment conservation. But there are substantial differences in two cities' climates, with Sacramento averaging more than six times the number of cooling degree days (National Climate Data Center, http://www.ncdc.noaa.gov/oa/climate/online/ccd/nrmcdd.html)

Table 6 tests for heterogeneous treatment effects (analogous to the decomposition analysis performed for SMUD in Table 2) in the PSE data by rerunning the household fixed effects specification from Table 5, this time adding various interaction terms with the treatment variable of interest (Treatment $\times$ Post). We find that households with larger pretreatment

Table 5. PSE OLS Regression of Natural log of kWh/day, therms/day, CEC/day Clustering on Household ID

	kWh/day (without controls) $n = 1{,}567{,}395$	kWh/day (with controls) $n = 1{,}567{,}143$	kWh/day (with household fixed effects) $n = 1{,}567{,}395$	therms/day (without controls) $n = 1{,}567{,}395$	therms/day (with controls) $n = 1{,}567{,}143$	therms/day (with household fixed effects) $n = 1{,}567{,}395$	CEC/day (without controls) $n = 1{,}567{,}395$	CEC/day (with controls) $n = 1{,}567{,}143$	CEC/day (with household fixed effects) $n = 1{,}567{,}395$
Treatment household	0.0003	−0.0005		0.0004	0.0007		0.0007	0.0003***	
After first mailing (October 2008)	0.0433***	0.0847**	0.0603	0.4366***	−1.0635***	−1.0487***	0.2175***	−0.2597***	−0.2635***
Treatment × after first mailing	−0.0123***	−0.0123***	−0.0122***	−0.0125***	−0.0121***	−0.0120***	−0.0124***	−0.0122***	−0.0121***
Cooling degree days		−0.0019***	−0.0017***		−0.0058***	−0.0059***		−0.0042***	−0.0042***
Heating degree days		−0.0002	−0.0001		0.0052***	0.0051***		0.0015***	0.0015***
House sq. Ft. (in 1000s)		1.17E-05***			−1.03E-06			1.22E-05***	
House age		0.0001**			−0.0009***			0.0002***	
House value (in $100,000s)		−3.90E-09			1.22E-08**			1.66E-08***	
Quarterly recipient		−0.0004			0.0003			0.0001	
Therms usage per day in 2007		−0.0013			1.0014***			0.4115***	
kWh usage per day in 2007		0.9338***			−0.0039*			0.4962***	
Tier 2 price in 2007		−0.0032*			0.0044***			−0.0477***	
Proprietary segment effects[a]	No	Yes	No	No	Yes	No	No	Yes	No
Month fixed effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
R^2	0.00	0.72	0.83	0.06	0.85	0.88	0.04	0.83	0.88

Regressions clustered on households with Huber-White robust standard errors. [a]Proprietary segment groups created by Positive Energy based on house characteristics. *, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

energy usage per square foot exhibited larger energy-reducing treatment effects. Larger, more valuable, and older households had smaller energy-reduction treatment effects.

In the SMUD experiment (as shown in second specification of Table 2), those who received the report monthly saved more electricity than those who received it quarterly (although the result was only significant at 10% level). However, in SMUD only the lower (pretreatment) energy-using households were assigned to the quarterly treatment group, leading to the possibility that the smaller estimated quarterly treatment effect was driven by lower pretreatment energy usage. In the PSE experiment, however, with randomized monthly and quarterly recipients, we are better able to gauge the causal impact of report frequency. Table 6 shows that in the PSE data, quarterly treatment effects are statistically indistinguishable from monthly effects.

Figure 5 further explores the heterogeneous treatment effects for different levels of pretreatment usage by rerunning the CEC "with controls" regressions of Table 5 but adding treatment interactions for each of the pretreatment energy usage deciles. One finds (as in the Figure 3 analysis of SMUD data) that households with more pretreatment energy use per square foot exhibit larger energy reduction—with four of the five highest deciles exhibiting larger than average energy reductions and three of the five lowest deciles exhibited lower than average energy reductions. Figure 5 indicates that Sacramento households also did not exhibit boomerang effects. The treatment effects of two lowest Puget Sound deciles were not statistically different from zero, and all the other deciles were estimated to have statistically significant energy-reducing treatment effects. Rerunning this regression with just monthly data produced qualitatively similar results (with the two lowest deciles not statistically different from zero).

One advantage of the Puget Sound experiment is that PSE collects *daily* data on energy usage, with the aid of an automated meter read system called CellNet. Figure 6 reports the results of a regression (using "with controls" specification of Table 5) of household-day energy usage where the treatment variable (Treatment × After First Mailing) was interacted with day of week indicators. The figure shows that Sunday (the day of the week with the highest energy consumption) has the largest treatment effect, and that 38% of the estimated treatment effect occurs between 12:00 AM Sunday morning and 11:59 PM Monday night. It may be that the energy savings is even more tightly concentrated during the weekend, with the bulk of the "Monday" savings occurring during the night between Sunday and Monday. The evidence that most of the savings is happening on two contiguous days roughly overlapping with the weekend suggests that the primary impact of the energy reports may not be driven by certain types of durable conservation efforts (such as a more energy efficient refrigerator). But it is impossible with our limited data to empirically distinguish among a number of alternative hypotheses for

Table 6. PSE OLS Heterogeneous Treatment Effects on kWh/day, Therms/day, CEU/day Clustering on Household ID, Household Fixed Effects

	kWh/day I $n=1,567,395$	kWh/day II $n=1,567,143$	kWh/day III $n=1,567,143$	therms/day I $n=1,567,395$	therms/day II $n=1,567,143$	therms/day III $n=1,567,143$	CEC/day I $n=1,567,395$	CEC/day II $n=1,567,143$	CEC/day III $n=1,567,143$
T × post	−0.012***	0.003	−0.049***	−0.012***	−0.006	−0.066***	−0.012***	−0.004	−0.042***
T × post × quarterly		0.003	0.003		−0.002	−0.002		0.002	0.002
T × post × house value (in $100ks)		−1.41E-05*	−2.25E-05*		−5.79E-06	−2.30E-05		−1.06E-05	−1.74E-05*
T × post × kWh in 2007 per sq. ft.		−0.730***	−0.997***		−0.191	−0.815***		−0.345*	−0.690***
T × post × tier 2 price in 2007			−0.006			0.003			−0.001
T × post × therms in 2007 per sq. ft.			5.181			−4.645			−1.626
T × post × HDD			6.93E-05			1.07E-04			9.65E-05
T × Post × CDD			0.061			0.192			0.083
T × post × house sq. ft. (in 100s)			0.001***			0.001***			0.001***
T × post × house age			0.001***			0.002***			0.001***
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.83	0.83	0.83	0.88	0.88	0.88	0.88	0.88	0.88

Regressions clustered on households with Huber–White robust standard errors.*, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

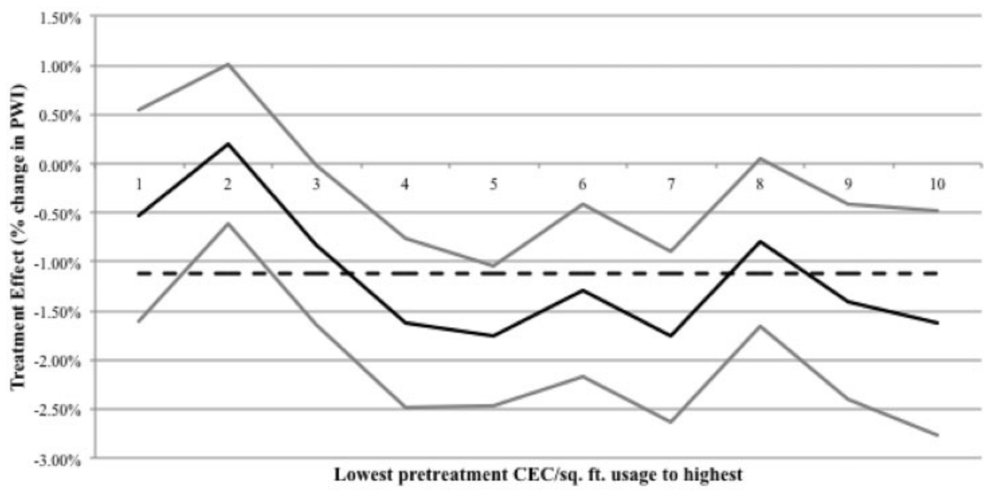

Figure 5. PSE Treatment Effect (% change in CEC) by Pretreatment CEC/sq. ft. Deciles. Notes: 95% confidence intervals shown in gray. Horizontal dashed line indicate first mailing. OLS regression on natural log of CEC/day clustered on household id with same controls as in Table 5.

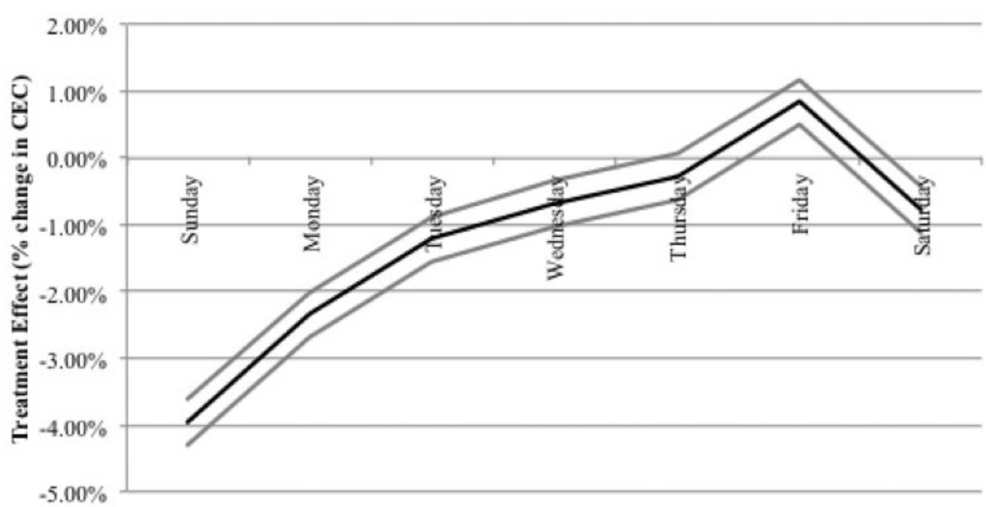

Figure 6. PSE Treatment Effect (% change in CEC) Day of the Week. Notes: 95% confidence interval is shown. OLS regression on natural log of CEC/day clustered on household id with same controls as in Table 3.

heightened weekend conservation (including, for example, the possibility of increased mindfulness of energy consumption on the weekends).[12] The statistically significant positive treatment effect on Friday also suggests the possibility of temporal substitution—for example, shifting chores (e.g. laundry) from the weekend to Friday. The weekend skew of the treatment

12. On the other hand, increased savings on the weekends could be the result of disproportionate use on weekends of durables (such as new energy-efficient washing machines).

effects also highlights the possibility that the reductions in utility usage do not necessarily mean that the individuals in the household reduced their overall personal energy consumption. For example, it is possible that treatment households substitute toward energy consumption outside the household simply choosing to engage in more weekend activities outside the house. These daily treatment effects are puzzling and resist any simple theoretical explanation.

The more granular data also allows us to estimate treatment effects across the billing cycle. Figure 7 reports the results of a series of regressions (using the "with controls" specification of Table 5) calculating the treatment effect in terms of kWh and therms for particular weeks before and after the experiment began. Figure 7a reports the week by week treatment effects on kWh and therms for recipients of monthly reports. The vertical lines denote the approximate delivery dates of the reports.[13] The treatment effects for all 31 post-experiment weeks are statistically significant and negative. The treatment effects for therm usages for the weekly data becomes more modest in the final three weeks of the data, just as cooling degree days in the Spring started to increase—suggesting (albeit with very limited data) that the percentage treatment effect on gas may be more modest when home heating becomes less important. Figure 7b analogously reports the weekly treatment effects for quarterly report recipients on kWh and therms. After the first mailing, 52% of the treatment effects observed on therms were statistically lower than zero for therms, and 77% of the treatment effects for kWh were statistically significant ($p < .05$) reductions.

As in SMUD (Table 3), we again tested for staleness and routinization effects on quarterly households—using an interacted version of the fixed household effects specification in Table 5. However, Table 7 reports no statistically significant interactions. The size of the treatment effect was unaffected by either the number of the messages or the relative staleness of the message for quarterly report recipients.

Finally, Table 8 assesses the potential economic and environmental impact if reports were sent to all households in PSE's customer base. Per household, energy report recipients save on average about \$13 a year from kWh reduction, and \$11 a year from therms reduction, for a total of about \$24 of savings in a year. With over 930,000 households receiving electric service and over 681,000 households receiving gas service from PSE, PSE customers might stand to save \$22 million annually from these peer-comparison energy reports (about \$23 million from monthly reports or about \$21 million from quarterly reports). As with SMUD estimates, extrapolation of treatment effects from the experiment sample (single-family houses using both gas and electricity) to the broader

13. The timing of monthly reports was not evenly spaced across time because of holiday interruptions and other logistical issues. Printing and sending a particular mailing sometimes took multiple days to complete.

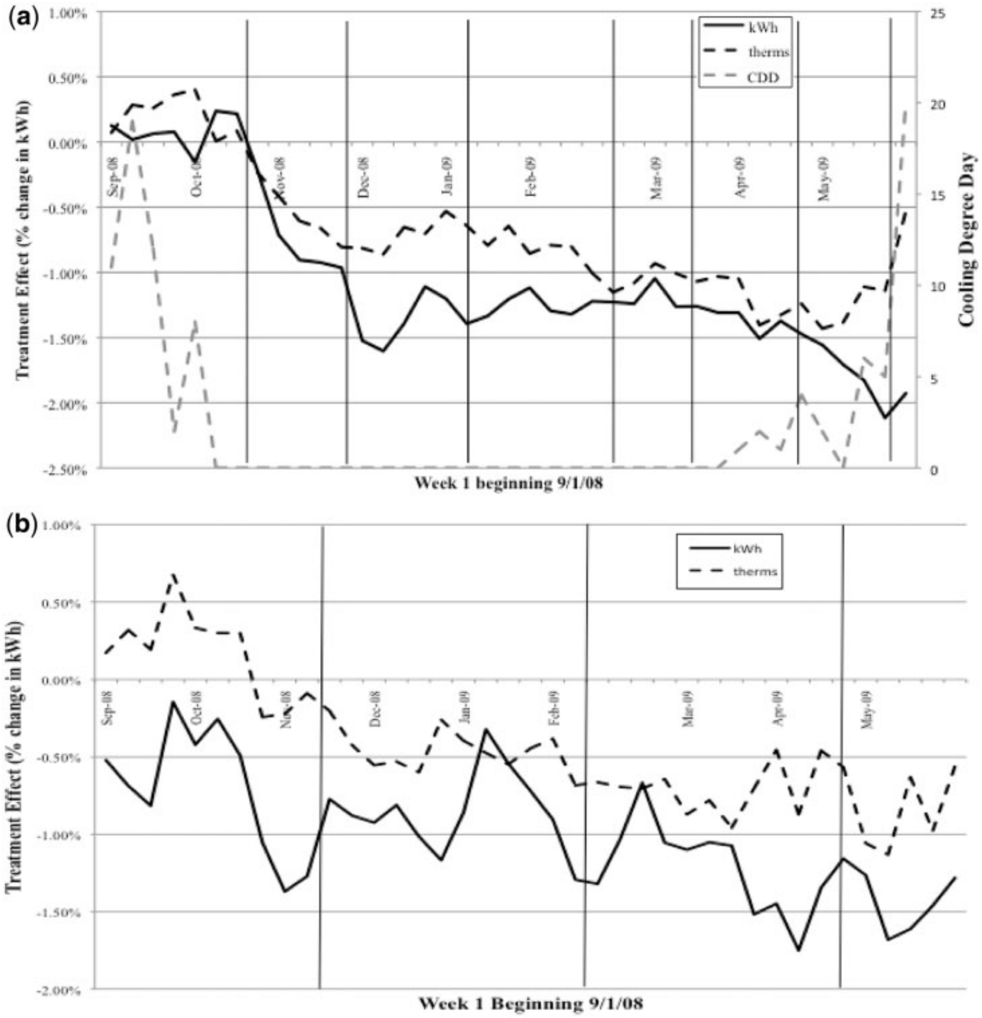

Figure 7. PSE Treatment Effect (% change in kWh and therms), (a) Monthly Recipients (with degree days) and (b) Quarterly Recipients. Notes: 95% confidence interval is shown. Vertical lines indicate mailings. OLS regression on natural log of CEC/day clustered on household id with same controls as in Table 3.

population of PSE households (93.6% of which relied on gas and electricity to heat their houses) should be viewed with caution. In environmental terms, this projected customer-base-wide savings from quarterly reports is the equivalent of over 100 metric tons of CO_2. PSE households saved on average \$3.06 per mailing (\$2.06 for the monthly reports and \$5.57 per mailing for the quarterly reports). While energy reduction as a percentage of consumption was lower in Puget Sound than in Sacramento, it was applied to a larger base (which included natural gas). The added savings in therms overall made the PSE intervention more cost effective with an overall cost per kilowatt-hour saved of just 1.78 cents. These average treatment effects in terms of both electricity and therm reduction are equivalent to a household reducing its use of a 75 watt bulb by about 16 h per day (or reducing air conditioner usage by about 72 min).

Table 7. PSE OLS Regression of Natural Log of kWh/day, Therms/Day, CEU/Day Clustering on Household ID (quarterly households only)

	kWh/day (with household fixed effects) $n=418{,}306$	kWh/day (with household fixed effects) $n=418{,}306$	kWh/day (with household fixed effects) $n=418{,}306$	therms/day (with household fixed effects) $n=415{,}577$	therms/day (with household fixed effects) $n=415{,}577$	therms/day (with household fixed effects) $n=415{,}577$	CEC/day (with household fixed effects) $n=418{,}379$	CEC/day (with household fixed effects) $n=418{,}379$	CEC/day (with household fixed effects) $n=418{,}379$
Treatment × after first mailing	−0.009***	−0.006	−0.006	−0.013***	−0.010**	−0.009	−0.008****	−0.004	−0.009
Treat × post × no. of messages		−0.002	−0.002		−0.002	−0.004		−0.003	0.004
Treat × post × (no. of messages)^2			−4.11E-05			0.0005			−0.0017
Treat × post × one billing cycle after last message		−1.36E-04	−1.50E-04		0.002	0.0021		−4.03E-05	−0.0006
Treat × post × two billing cycles after last message		0.001	0.001		0.0018	0.0020		0.00112	0.0006
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.82	0.82	0.82	0.88	0.88	0.88	0.87	0.87	0.87

Regressions clustered on households with Huber–White robust standard errors. *, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

Table 8. PSE Projected Cost Savings and Environmental Impact

	Monthly and quarterly weighted effects
Per household (kWh)	
Reduction (kWh/day)	0.39
Total kWh reduction in a year	140.85
Total kWh savings in a year	$12.97
Per Household (therms)	
Reduction therms/day	0.03
Total therms reduction in a year	10.14
Total therms savings in a year	$11.02
For customer base of PSE	
Annual savings per household	$24.00
Savings per mailing	$3.06
Annual savings for Puget Sound	$22,326,161
Annual savings in metric tons of CO2[a]	128,647
Estimated short-term price increase necessary to produce comparable energy reduction[b]	3.90%
Estimated long-term price increase necessary to produce comparable energy reduction[b]	0.87%

[a]Based on 7.18×10^{-4} metric tons CO2 / kWh and 0.005 metric tons CO2/therm calculated by the EPA.[b]Based on Bernstein and Griffin's (2005) regional price elasticity of electricity calculations.

Again using regional elasticity estimations (Bernstein and Griffin 2005), we find that to induce equivalent energy reductions with an energy excise tax, a price increase of 3.9% in the short run (or .87% in the long run) would be required.

4. Conclusion

Both the PSE and SMUD experiments reveal that Opower peer comparison reports cause significant reductions in home energy use. The PSE and SMUD experiments show that the effects of the report continue to be strong up to 7 and 12 months after the households begin to receive reports, respectively. The experiments analyzed here add evidence of external validity to the findings of the San Marcos study. They show that a peer comparison intervention (combined with personalized energy-saving tips, energy usage history and injunctive emoticons) can be feasibly scaled. In both experiments, households with higher pretreatment energy use per square foot saved more than households with lower pretreatment energy use.

The experiments also provide at best suggestive evidence about the types of behavior that may be driving energy reductions. The day-of-the-week treatment effects uncovered in the PSE data, for example, are inconsistent with certain durable conservation measures (such as a installing a more energy-efficient refrigerator) that would be expected to deliver more consistent levels of energy use reductions throughout the week.

The Opower experiments suggest that governmental entities looking for low-cost interventions to promote conservation should consider mandating or incentivizing utilities to provide peer comparison information—through separate mailings, in the regular utility bill, or through electronic communication channels. Although some utilities, such as those that are publically owned (like SMUD), or are private but regulated (like PSE), are beginning to provide such feedback, some utilities may not have adequate incentives to promote customer energy conservation. Similar levels of short-term energy reduction could alternatively be achieved by tax increases of 3–7%, but mandating peer information is more likely to be politically feasible.

Even without express regulatory mandates, peer-comparison information is quickly gaining substantial market penetration. At the moment, Opower is providing energy reports via paper mailings or web access to more than 10 million households at 46 different US utilities in 22 states (including 8 of the largest 10 utilities) (Allcott 2011). The Opower results have also attracted the attention of policymakers. President Barack Obama has highlighted Opower as a clean energy success story. The White House has cited to Opower's example in their support for investment in "smart grid" technologies that, among other things, enable sophisticated forms of energy customer feedback.[14] US legislators have introduced bills in Congress aiming to give consumers ready access to richer data on their energy usage (E-KNOW Act 2011). The UK government, also citing the Opower results, has announced that it is partnering with Opower and another utility to test the use of peer feedback to energy consumers. The government is also seeking a voluntary agreement with the country's utilities to provide comparative consumption information to their customers and has stated that if an agreement cannot be reached, it will seek legislation (Behavioral Insights Unit 2011).

Finally, the Opower experiments suggest that peer-comparison feedback, such as direct mailings, might prove an effective tool in a broad range of other situations. Schools might mail parents reports of how the absences or tardiness of their children compares to that of other students. A gym might inform its lazier patrons of how often typical members work out. Employers might inform low-saving employees how much more their peers are saving in the company 401(k) plan. Mutual fund complexes might inform investors about how much other customers pay in fees are

14. Remarks of President Barack Obama on Clean Energy Jobs, March 5, 2010, Opower offices, Arlington, Virginia, *available at* http://www.whitehouse.gov/the-press-office/remarks-president-clean-energy-jobs. The Administration released a policy platform on smart grid in June 2011 that cited to an earlier version of this working paper as evidence of the potential of enabling more advanced feedback for energy consumers (White House 2011).

on similar style mutual funds. As these preliminary examples show, the area of peer comparison feedback is ripe for innovation and experimentation.

Appendix A

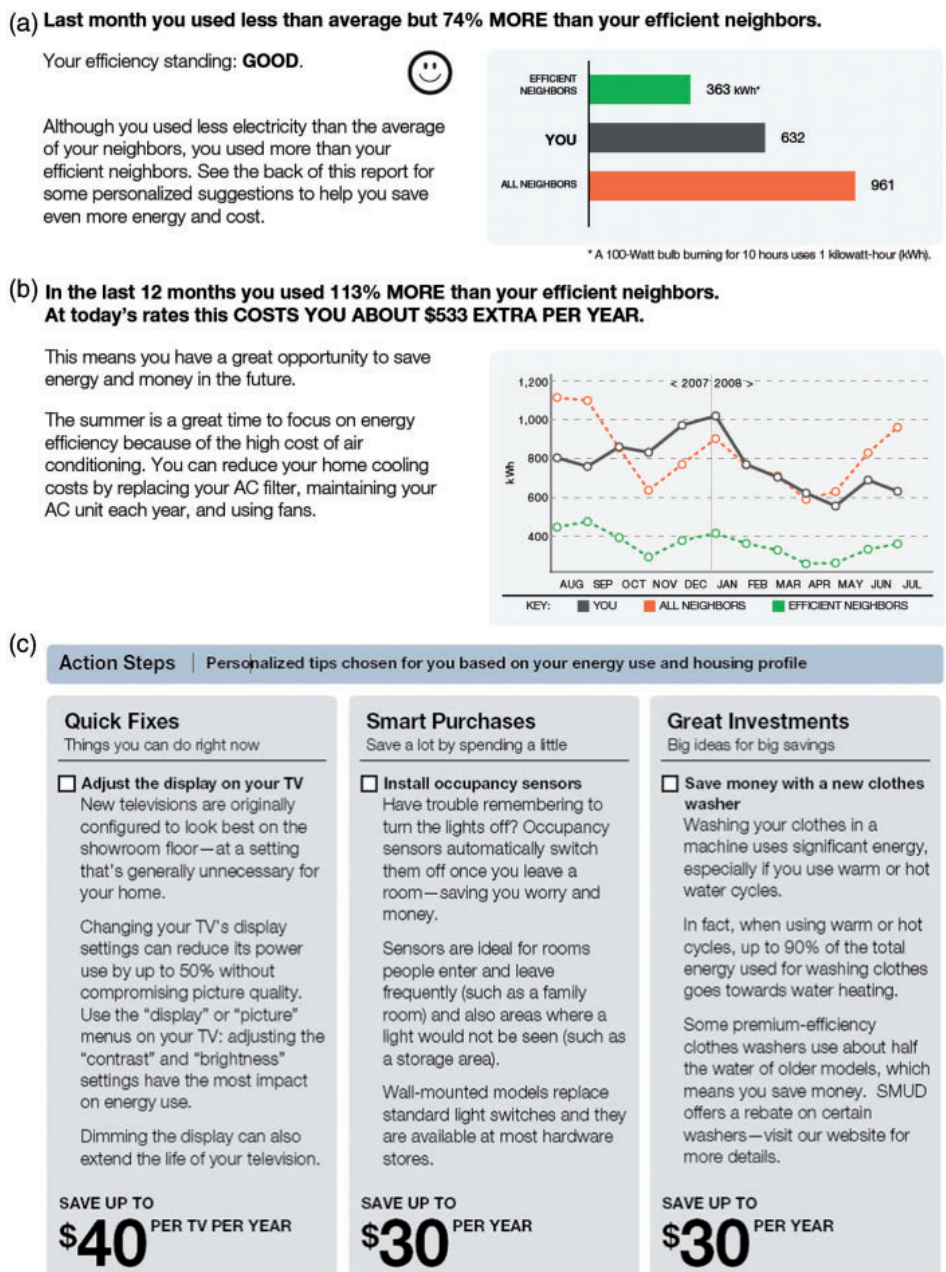

Figure A1. (a) SMUD Sample Report, Narrative Template, (b) SMUD Sample Report, Narrative Template, and (c) SMUD Action Steps Template.

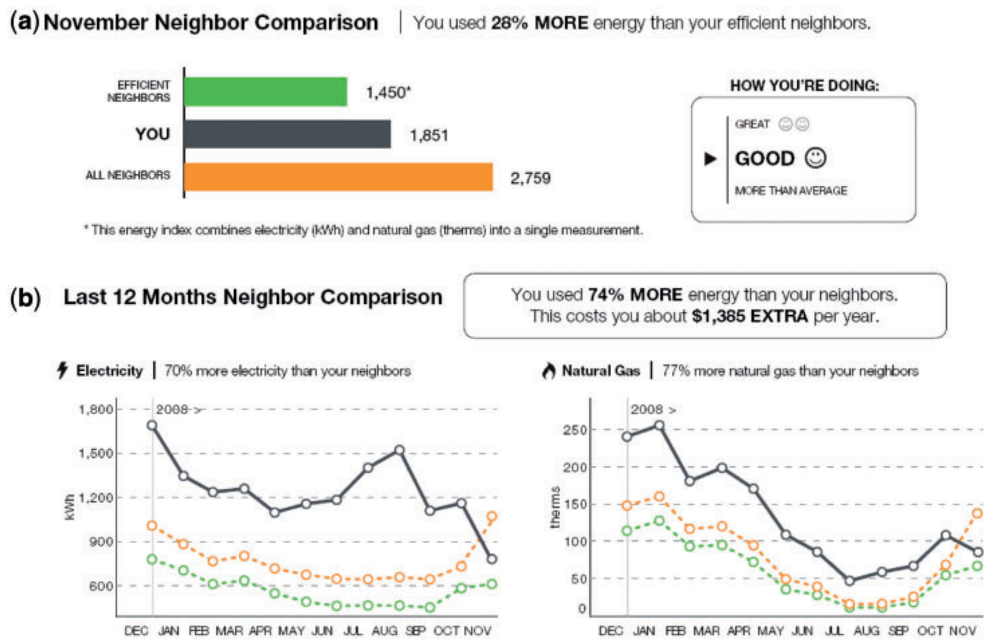

Figure A2. (a) and (b) PSE Sample Report, Narrative Template.

Table A1. Mean Comparison of all SMUD Pretreatment Variables

Variable name	Experiment $n=34{,}557$	Control $n=49{,}570$	Graphical $n=41{,}841$	Narrative $n=41{,}856$	#10 Envelop $n=42{,}276$	6×9 envelop $n=41{,}851$
House square foot	1737	1753***	1742	1732	1731	1743*
House age	35.73	36.92	35.79	35.66	35.62	35.83
Pool	0.21	0.22***	0.21	0.21	0.20	0.21*
Spa	0.04	0.04	0.04	0.04	0.04	0.04
House value	$213,584	$215,189	$214,336	$212,833	$212,478	$214,690
Gas heat	0.73	0.75	0.73	0.73	0.73	0.73
Account closed	0.08	0.07	0.07	0.08	0.08	0.07
Opt out	0.02	.	0.02	0.02	0.021	0.018*
Quarterly recipient	0.29	0.29*	0.29	0.29	0.29	0.29
kWh usage per day in 2006	31.95	31.65***	31.62	31.68	31.71	31.58
Length of residence	14.03	14.21**	14.11	13.94	13.99	14.06
Number at residence	1.93	1.93	1.93	1.94	1.94	1.93
Quartile 1 income group	0.11	0.11**	0.11	0.11	0.11	0.11
Quartile 2 income group	0.19	0.19	0.20	0.19**	0.20	0.19
Quartile 3 income group	0.16	0.16	0.16	0.16	0.16	0.17
Quartile 4 income group	0.23	0.23	0.23	0.23	0.23	0.23

(continued)

Table A1. Continued

Variable name	Experiment $n=34{,}557$	Control $n=49{,}570$	Graphical $n=41{,}841$	Narrative $n=41{,}856$	#10 Envelop $n=42{,}276$	6 × 9 envelop $n=41{,}851$
Age ≤24 (years)	0.00	0.00	0.00	0.00	0.00	0.00
Age 25–29 (years)	0.01	0.01	0.01	0.01	0.01	0.01
Age 30–34 (years)	0.03	0.03	0.03	0.03	0.03	0.03
Age 35–39 (years)	0.06	0.05	0.05	0.06	0.06	0.05
Age 40–44 (years)	0.07	0.07	0.07	0.07	0.07	0.07
Age 45–59 (years)	0.09	0.09	0.09	0.09	0.09	0.09
Age 50–54 (years)	0.100	0.096*	0.10	0.10	0.10	0.10
Age 55–59 (years)	0.093	0.089**	0.09	0.09	0.10	0.09*
Age 60–64 (years)	0.07	0.07	0.07	0.07	0.07	0.07
Age <65(years)	0.01	0.02	0.01	0.01	0.01	0.01
Age 65–69 (years)	0.05	0.05	0.05	0.05	0.05	0.05
Age 70–74 (years)	0.04	0.04	0.04	0.04	0.04	0.04
Age >75 (years)	0.07	0.07	0.07	0.07	0.07	0.08
January 2007	36.71	36.72	36.69	36.74	36.78	36.65
February 2007	33.10	32.95	33.04	33.17	33.13	33.08
March 2007	28.00	28.14	27.94	28.07	28.01	27.99
April 2007	24.67	24.95***	24.63	24.72	24.72	24.63
May 2007	25.44	25.89	25.39	25.49	25.50	25.38
June 2007	28.53	29.28	28.48	28.58	28.58	28.48
July 2007	36.92	37.32***	36.88	36.95	36.96	36.87
August 2007	36.80	37.13***	36.73	36.87	36.87	36.73
September 2007	37.78	38.01*	37.81	37.76	37.86	37.71
October 2007	25.70	25.63	25.68	25.72	25.78	25.62
November 2007	25.21	25.44**	25.15	25.27	25.28	25.14
December 2007	30.77	31.18***	30.69	30.86	30.79	30.76
January 2008	36.07	36.00	36.02	36.12	36.08	36.06
February 2008	32.81	32.75	32.75	32.87	32.87	32.76
March 2008	27.48	27.57	27.47	27.49	27.53	27.43
Affluence1	0.007	0.008**	0.01	0.01	0.01	0.01
Affluence2	0.031	0.029*	0.03	0.03	0.028	0.034***
Affluence3	0.16	0.15	0.16	0.16	0.16	0.16
Affluence4	0.10	0.10	0.10	0.10	0.10	0.10
Affluence5	0.17	0.17	0.17	0.17	0.17	0.17
Affluence6	0.10	0.08	0.10	0.10	0.10	0.10
Affluence7	0.08	0.08	0.07	0.08	0.08	0.07
Affluence8	0.036	0.038**	0.04	0.04	0.04	0.04
Affluence9	0.03	0.03	0.03	0.03	0.03	0.03
Affluence10	0.0014	0.0007***	0.0018	0.0009**	0.00	0.00
Greenergy	0.09	0.09	0.09	0.09	0.09	0.09
Electric heat	0.27	0.25	0.27	0.26	0.27	0.27

*, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

Table A2. Mean Comparison of all PSE Pretreatment Variables

Variable name	Experiment $n=34{,}891$	Control $n=44{,}121$	Monthly $n=24{,}949$	Quarterly $n=9949$
House square foot	2138.56	2139.99	2139.32	2136.68
House age	29.98	29.98	30.06	29.78
House value	$345,046	$346,041	$345,874	$342,971
Account closed	0.02	0.02	0.02	0.03
Opt out	0.01	.	0.01	0.00***
Therms usage per day in 2007	2.50	2.50	2.50	2.50
kWh usage per day in 2007	30.31	30.26	30.23	30.50*
Quarterly recipient	0.29	0.25***		
kwh/day use in ...				
October 2007	29.71	29.68	29.64	29.89
November 2007	33.29	33.24	33.23	33.46
December 2007	39.21	39.16	39.13	39.43
January 2008	35.68	35.58	35.58	35.93*
February 2008	32.68	32.61	32.60	32.89
March 2008	31.62	31.60	31.55	31.81
April 2008	29.26	29.25	29.19	29.41
May 2008	27.01	27.00	26.94	27.20
June 2008	26.98	26.98	26.91	27.17*
July 2008	26.16	26.16	26.09	26.33
August 2008	27.14	27.20	27.06	27.34
September 2008	26.60	26.62	26.55	26.72
therms/day use in ...				
October 2007	2.45	2.45	2.45	2.45
November 2007	3.69	3.69	3.69	3.68
December 2007	4.63	4.63	4.64	4.61
January 2008	5.07	5.07	5.08	5.05
February 2008	3.95	3.94	3.96	3.94
March 2008	3.84	3.84	3.85	3.83
April 2008	3.07	3.07	3.07	3.06
May 2008	1.61	1.61	1.61	1.61
June 2008	1.37	1.37	1.37	1.37
July 2008	0.66	0.66	0.65	0.67***
August 2008	0.66	0.66	0.65	0.67***
September 2008	0.96	0.96	0.96	0.97

*, **, and ***Significance at the 90%, 95%, and 99% levels, respectively.

References

Abrahamse, Wokje, Linda Steg, Charles Vlek, and Talib Rothengatter. 2005. "A Review of Intervention Studies Aimed at Household Energy Conservation," 25 *Journal of Environmental Psychology* 273–91.

Allcott, Hunt. 2011. "Social Norms and Energy Conservation," 95 *Journal of Public Economics* 1082–95.

Allcott, Hunt, and Sendhil Mullainathan. 2010. "Behavior and Energy Policy," 327 *Science* 1204–5.

———. 2011. "External Validity and Partner Selection Bias," Working Paper, NYU.

Auffhammer, Maximilian, Carl Blumstein, and Meredith Fowlie. 2007. "Demand-Side Management and Energy Efficiency Revisited," CSEM working paper (April).

Ayres, Ian, and Barry Nalebuff. 2005. "Peer Pressure," *Forbes* 135 (April 11, 2005).

Ben-Shahar, Omri, and Carl E. Schneider. 2011. "The Failure of Mandated Discourse," 159 *University of Pennsylvania Law Review* 647.

Bernstein, Mark A., and James Griffin. 2005. Regional Differences in the Price-Elasticity of Demand for Energy. The RAND Corporation Technical Report. Santa Monica, CA.

Behavioral Insights Unit. 2011. Behavior Change and Energy Use. <http://www.cabinetoffice.gov.uk/sites/default/files/resources/behaviour-change-and-energy-use.pdf>.

Cialdini, Robert. B., and Noah J. Goldstein. 2004. "Social influence: Compliance and conformity," 55 *Annual Review of Psychology,* 591–621.

Cialdini, Robert B., Carl A. Kallgren, and Raymond R. Reno. 1991. "A Focus Theory of Normative Conduct: a Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior," in M. P. Zanna, ed., *Advances in Experimental Social Psychology*. Vol. 24. 201–234. New York, NY: Academic.

Costa, Dora, and Matthew Kahn. 2010. "Energy Conservation Nudges and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment," NBER Working Paper No. 15939 (April).

Cooter, Robert, Michael Feldman, and Yuval Feldman. 2008. "The Misperception of Norms: The Psychology of Bias and the Economics of Equilibrium," 4(3) *Review of Law and Economics* 889–911.

Darby, Sarah. 2006. The Effectiveness of Feedback on Energy Consumption: a Review For DEFRA of the Literature on Metering, Billing and Direct Displays. Report for the UK Department for Environment, Food and Rural Affairs. http://www.defra.gov.uk/environment/climatechange/uk/energy/research/pdf/energyconsump-feedback.pdf.

Ehrhardt-Martinez, Karen, Kat A. Donnelly, and John A. "Skip" Laitner. 2010. Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities. American Council for Energy-Efficient Economy. https://www.aceee.org/research-report/e105.

E-KNOW Act (2011) S. 1029, 112th Congress. Introduced by Senator Mark Udall of Colorado and co-sponsored by Senator Scott Brown of Massachusetts.

Faruqui, Ahmad, Sanem Sergici, and Ahmed Sharif. 2010. The impact of informational feedback on energy consumption—a survey of the experimental evidence. The Brattle Group, Inc. http://www.brattle.com/_documents/uploadlibrary/upload772.pdf.

Ferraro, Paul J., and Michael K. Price. 2011. "Using Non-Pecuniary Strategies to Influence Behavior: Evidence from a Large Scale Field Experiment," NBER Working Paper No. 17189.

Fischer, Corinna. 2008. "Feedback on Household Electricity Consumption: a Tool for Saving Energy?," 1 *Energy Efficiency* 79–104.

Friedrich, Katherine, Maggie Eldridge, Dan York, Patti Witte, and Marty Kushler. 2009. "Saving Energy Cost-Effectively: A National Review of the Cost of Energy Saved through Utility-Sector Energy Efficiency Programs." ACEEE Report No. U092 (September).

Goldstein, Noah J., Robert B. Cialdini, and Vladas Griskevicius. 2008. "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels," 35(3) *Journal of Consumer Research* 472–82.

Herberich, David H., John A. List, and Michael K. Price. 2011. "How Many Economists does it take to Change a Light Bulb?: A Natural Field Experiment on Technology Adoption," working paper. http://econ-new.stanford.edu/files/Herberich4_13.pdf.

Jin, Ginger Zhe, and Phillip Leslie. 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," 118(2) *Quarterly Journal of Economics* 409–51.

Latané, Bibb, and J. M. Darley. 1970. *The Unresponsive Bystander: Why Doesn't He Help?* New York, NY: Appleton-Century-Crofts.

Nolan, Jessica M., Wesley P. Schultz, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius. 2008. "Normative Influence is Underdetected," 34 *Personality and Social Psychology Bulletin* 913–23.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius. 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms," 18 *Psychological Science* 429–34.

White House. 2011. "A Policy Framework for the 21st Century: Securing Our Clean Energy Future." http://www.whitehouse.gov/sites/default/files/microsites/ostp/nstc-smart-grid-june2011.pdf.