Promoting Energy Conservation: A Field Experiment on Peer Comparisons and Rate Structures

Todd L. Cherry,[a] David M. McEvoy,[b] and Tanga M. Mohr[c]

ABSTRACT

We compare the effectiveness of social comparison nudges on energy consumption when residents pay for electricity and when electricity is included in monthly rent. Using a randomized control trial, our intervention uses digital messages (text and emails) to provide residents with home energy reports comparing their recent electricity usage with similar households. Our design allows us to investigate the pecuniary and non-pecuniary impacts of a widely-used behavioral nudge. The average treatment effects suggest that peer comparison nudges are less effective for non-ratepaying customers, implying that cost-saving motives play an important role.

Keywords: Behavioral nudge, Energy conservation, Social comparison, Social norms

https://doi.org/10.5547/2160-5890.14.1.tche

1. INTRODUCTION

Climate policies aimed at reducing greenhouse gas emissions have prompted electric and natural gas utilities to spend billions of dollars each year on programs to reduce energy use and increase investment in energy efficiency (US EIA 2019). Researchers have examined the wide-ranging efforts to identify the most cost-effective measures. Many programs provide subsidies for physical investments that improve energy efficiency, such as replacing lightbulbs, upgrading appliances and improving insulation, but other programs apply insights from the behavioral sciences to nudge consumers to reduce energy use. One popular behavioral intervention that has shown promise is the home energy report, which provides households personalized information on energy use, social comparisons and energy conservation information. There is no mystery to why these types of nudges are appealing. If effective, information interventions are an inexpensive, simple and minimally intrusive way to promote energy conservation. In this study, we provide new evidence on the efficacy of behavioral nudges by exploring the differential impacts of social comparisons under two pricing mechanisms: one in which residents pay for electricity at the margin and another in which electricity is included in the fixed monthly rent (non-ratepaying residents). The research design also provides some

[a] Department of Economics, University of Wyoming. E-mail: tcherry@uwyo.edu.
[b] Corresponding author. Department of Economics, Appalachian State University, 416 Howard St., Peacock Hall, Boone, NC 28608. Phone: 828-262-6126. E-mail: mcevoydm@appstate.edu.
[c] Department of Economics, Appalachian State University. E-mail: mohrtm1@appstate.edu.

perspective on the relative impact of behavioral nudges to simple differences in electricity pricing schemes.

The use of peer-comparisons in home energy reports attracts attention from policymakers and researchers because they tend to work. Schultz et al. (2007) and Nolan et al. (2008) were the first to explore the effectiveness of providing households with peer comparisons of electricity usage with samples of California residents. When the information was coupled with an injunctive message signaling it was good to be a low user (smiley face) and bad to be a high user (frowning face), on average people reduced their electricity consumption. Moreover, Nolan et al. (2008) showed that peer comparison information was more effective than messages promoting environmental friendliness, social responsibility or financial self-interest as motives for conservation. The next wave of research explored the effectiveness of peer comparisons at reducing electricity use using large-scale field experiments in the United States with the support of the company OPower, which is hired by utilities to mail customers Home Energy Reports (e.g., Allcott and Mullainathan 2010; Allcott 2011; Ayers et al. 2013; Costa and Kahn 2013; Allcott and Rogers 2014). The predominant result from this collection of studies is that, on average, these simple behavioral interventions cause a reduction in electricity usage of roughly two percent or less, at least in the short to medium term.[1,2]

What is not entirely clear is *why* peer comparisons motivate people to change their behaviors. In the set of studies just described, there are both pecuniary and non-pecuniary incentives to reduce consumption of electricity. One possibility is that the information interventions serve as consistent reminders for customers about their spending on electricity and potential for savings. Another possibility is that peer comparisons appeal to a consumer's desire not to be out of the social norm (see Benabou and Tirole 2003, 2006; Levitt and List 2007). Other possible explanations include competitiveness (trying to be part of the "good" group) and moral suasion (reducing consumption is the right thing to do) (see Ito et al. 2018). A study by Pellerano et al. (2017) explores the competing effects of normative and pecuniary motives in messaging. Like others, they report a strong influence of peer comparisons on reducing energy usage, but observe a crowding out (relative increase in usage) when coupled with messages also appealing to financial self-interest.

We add to this body of research by conducting a randomized control trial (RCT) that directly compares the effectiveness of peer-comparison information with and without the possibility of monetary incentives. Our experimental design explores the intervention in samples with two cost structures—one in which customers pay per kilowatt hour and one in which electricity is included in the fixed monthly rent. Non-ratepaying customers are a significant fraction of the US rental market. Based on the 2015 Residential Energy Consumption Survey (U.S. Energy Information Administration), approximately 35.6% of U.S. households are renters, and 15.4% of them have some or all of their electricity use included in their rent or condo fees. According to the 2013 Census, an estimated 17% of US rental housing have leases that do not require tenants to pay for heating (US Census 2013).

A few studies have explored how peer-comparison information impacts electricity consumption or thermostat settings for non-ratepaying residents (Delmas and Lessem 2014; Crago et al. 2020; Myers and Souza 2020), and the results suggest that the information nudge

1. See Buckley (2020) for a meta-analysis of studies examining behavioral nudges, and see Alcott (2015) for a summary of average treatment effects for 111 randomized control trials involving OPower.

2. Average residential electricity consumption in the US in 2019 was approximately 877 kWh per month (www.eia.gov) so a 2% reduction would on average translate to approximately 17.5 kWh per month if the effect was lasting.

has limited influence on energy consumption. These studies sample students living in a single dormitory (Delmas and Lessem 2014; Myers and Souza 2020) or a single graduate housing complex (Crago et al. 2020), and therefore the entire sample of subjects do not pay for their marginal usage. Our experimental design is novel in that we are able to compare the effectiveness of the same behavioral nudge across both marginal cost structures with the intent to directly isolate how much of the average treatment effect is attributed to pecuniary and non-pecuniary motives.

Apart from having both cost structures, our experimental design is unique relative to existing field experiments in other important ways. The samples are drawn from dozens of off campus student apartment complexes in Boone, NC. Most of the town is serviced by New River Light and Power (NRLP), an electric utility owned and operated by the local university—Appalachian State University. This research project is a collaborative effort with NRLP, which allows the research team to have unique access to detailed information about residential buildings/units as well as contact information for the residents who agreed to receive digital messages from NLRP.[3] Residents in our sample received the peer comparison information by both texts and emails. Other studies have demonstrated the effectiveness of text messages to inform residents about peak energy demand times and pricing (Jessoe and Rapson 2014; Royal and Rustamov 2018). Henry et al. (2019) found that home energy reports that include social comparisons can be effective when delivered via email. We believe ours is the first study to explore the use of text messaging to deliver social comparison information.

We tracked daily usage in kilowatt hours (kWh) for 357 units over 272 days. Roughly two-thirds of residents in our sample pay at the margin for their electricity and one-third have electric service included in their fixed monthly rent. When controlling for other unit-level characteristics, those who pay a fixed rate for electricity use about 33 percent more than customers in similar units who pay at the margin. This finding is significant and consistent with previous studies. For example, Levinson and Niemann (2004) show that households with electricity included in the rent have higher winter thermostat settings compared to ratepaying households. Munley et al. (1990) compare electricity use between renters that were randomly assigned as ratepayers or non-ratepayers and show average usage is 28 percent lower for those paying at the margin. However, we cannot rule out the possibility that high-use customers in our study self-selected into units that do not charge at the margin for electricity. If this type of selection bias exists, then the baseline difference between average usage cannot entirely be attributed to the difference in rate structures; that is, the high users would increase the average under either rate structure. A related and interesting behavioral finding that we add is that when the person paying the monthly electric bill is not the person residing in the ratepaying unit (e.g., a parent), residents use significantly more electricity compared to when one of the residents is the bill payer; that is, they closely mirror those customers who have electric included in the rent.

In terms of the behavioral nudge, we find that when residents pay at the margin the information interventions lead to an average treatment effect of approximately a 4.69 percent reduction in electricity use. Although the magnitude of the effect is high compared to previous field experiments (e.g., Allcott 2011; Ferraro and Price 2013; Allcott 2015; Allcott and Kessler 2019; Henry et al. 2019; Jessoe et al. 2021), the effect is only weakly significant in our study.

When we consider the sample of residents that do not pay at the margin for their electricity, we find the intervention caused an average treatment effect of roughly an 0.88 percent in-

3. Customers that agreed to receive digital messages (text and email) from NRLP were not informed about this particular research study.

crease in electricity use, but the effect is not statistically different from zero. Our results provide suggestive evidence that financial motives play an important role in nudging customers to reduce electricity consumption. We note our study is underpowered and should be appropriately considered with the broader literature. The experimental design was informed by an ex-ante power analysis that indicated sufficient power to detect a roughly two percent average treatment effect if one existed. Ultimately, the usage in our sample was significantly more variable than predicted (i.e., high standard deviation), and our sample size was smaller than expected. Therefore, we cannot rule out the possibility of a false negative with respect to the estimated average treatment effect for non-ratepaying customers. Likewise, low powered studies can also lead to false positives, which means caution should be used when interpreting the 4.69 percent reduction in electricity usage by ratepaying customers (see Ioannidis et al. 2017).

The next section describes the experimental design and testable hypotheses. The results section follows and we conclude with a discussion about the overall findings, limitations and policy implications.

2. OVERVIEW OF THE EXPERIMENT AND HYPOTHESES

The primary research objective is to explore the relative effects of peer comparison nudges, delivered digitally, on electricity consumption with and without monetary incentives. To explore this, we draw from the population of residents who are serviced by NRLP. While NRLP services the entire Appalachian State University campus and some residential homes, our sampling was restricted to off-campus apartment complexes.

The population of apartment complexes serviced by NRLP was segmented into units that are charged electric bills per kWh (marginal cost > 0) and those that have their electric bill (and other utilities) covered in their monthly rents (marginal cost = 0). We refer to these as the MC > 0 and MC = 0 conditions.[4] Apartment complexes are binary in the sense that residents are responsible for paying electric bills or they do are not, and residents do not have a choice of the contract type with a property.

The experimental design consists of a randomly assigned control group and a treatment group for each MC condition. The control group is a set of apartments for which the researchers could monitor daily electricity usage but were not part of any intervention treatment. For the treatment group, the researchers sent periodic text and email messages to residents informing them of their recent usage and how it compared to other similar units in their proximity. In these messages, comparisons were made relative to the average electricity usage of their neighbors as well as the bottom 20th percentile (which we call efficient users). We refer to this treatment as the *peer comparison* (*PC*) treatment. An example of a text message to recipients in the PC treatment is found in Figure 1.[5]

The participant's usage is reported for the past month (previous four weeks), and a direct comparison is made relative to the average and to the efficient 20th percentile electricity users in a recipient's peer group. A link was provided that takes users to a dedicated website with tips

4. While residents in apartments in the MC=0 condition do not pay electric bills, the owner/operator of the complex pays NRLP for the electricity used in all units combined.

5. The information sent via text corresponds to the peer comparison information included in the home energy reports (HERs) mailed to customers in many other randomized control trial studies (e.g., Allcott 2011, 2015; Ferraro and Price 2013; Allcott and Kessler 2019). Since we could not include graphics in the text message, the reports we sent customers are simpler and more direct compared to other studies.

FIGURE 1
Example of text message in the peer comparison (PC) treatment

on how to reduce electricity usage. Participants were simultaneously sent email messages with the same information.

To indicate when control or treated groups are part of the MC = 0 condition, we add a "0" to the treatment abbreviation. The full experimental design is summarized in Table 1.

TABLE 1
Experimental design

	Control	Peer comparison (PC)
MC > 0	C	PC
MC = 0	C0	PC0

Our results will focus on two primary hypotheses. First, the peer comparison will lead to less energy consumption. This is based on the growing literature that examines the efficacy of peer comparisons. Many of these studies find evidence of reductions of roughly two percent in response to messages like the ones in our peer comparison (PC) treatments. Previous work largely examines peer comparisons when customers pay at the margin for electricity (i.e., MC > 0), which introduces a pecuniary incentive for consumers to reduce usage. In our MC = 0 comparison, we are able to observe the influence of non-pecuniary incentives, such as conforming to a social norm (see Benabou and Tirole 2006; Levitt and List 2007), competitiveness or reducing the moral burden or stigma of being a high user (Ito et al. 2018). Although recent studies by Delmas and Lessem (2014), Crago et al. (2020) and Myers and Souza (2020) suggest that social comparison message are not effective when residents do not pay directly for electricity, following the literature on conforming to social norms we hypothesize that the nudge motivates energy conservation relative to the control.

The existing studies that explore the nudge when MC = 0 do not have a corresponding MC > 0 treatment, so they cannot compare the intervention across cost structures. Because our study has both cost structures we can make this comparison. Our second hypothesis is that the treatment effect for MC > 0 is greater than for MC = 0 because of the financial incentive to reduce electricity usage when consumers pay directly.

3. METHODOLOGY AND PROTOCOLS

3.1 Treatment assignment

Our assignment of individual apartments to the control group and treatment groups followed a stratified sampling process. Working with NRLP, we started by segmenting complexes by the cost structure—electricity is paid at the margin or electricity service is included in the rent.[6]

The next step was to stratify the units by size (i.e., studio, 1, 2, 3+ bedrooms). We randomly assigned units from each stratum to the control and treatment conditions. Ideally, we would include all units serviced by NRLP. However, the ability to send messages was restricted to a subset of the population that had voluntarily provided their phone number and email to receive future text and email messages (about energy-related information) from NRLP.[7] From this group we had access to a sample of 505 residents in 357 unique apartments. The total number of units and participants by unit type and cost structure is provided in Table 2. Note that the distribution of available units and participants differs by cost structure. In particular, the largest units were mostly from complexes that had electricity included in the monthly rent. There were fewer units and participants in MC = 0; therefore, the samples are smaller and unbalanced. For each unit in the study, we gathered the total square footage, heating fuel type, heating system type (when available) and year the unit/complex was built. For our sample, 98% of all units use electric heating.

TABLE 2
Units and number of participants (in parentheses) by experimental group

	Control	Peer comparison (PC)	Totals
MC > 0			
1 Bedroom	25 (28)	23 (23)	48 (51)
2 Bedroom Small	31 (48)	30 (41)	61 (89)
2 Bedroom Large	56 (74)	60 (82)	116 (156)
3+ Bedrooms	6 (6)	6 (8)	12 (14)
Total	**118 (156)**	**119 (154)**	**237 (310)**
MC = 0			
2 Bedroom Large	12 (15)	33 (41)	25 (56)
3+ Bedrooms	31 (55)	44 (84)	75 (139)
Total	**43 (70)**	**77 (125)**	**120 (195)**

3.2 Messages and data

We started gathering daily electricity usage by unit on September 2nd, 2018. For the first 48 days, participants received no messages from NRLP. That means that while units had already been allocated into treatments, for the first 48 days there was no behavioral intervention. Table 3 provides the average electricity usage by treatment group and cost structure during this

6. NRLP has access to meter readings for all units they service, even those that are not direct customers because the resident's electricity is included as part of their rent (i.e., MC = 0).

7. NRLP had not utilized the text messaging feature prior to this study, but would send bulk (i.e., non-targeted) email with energy tips. This legal requirement to opt-in potentially effects the external consistency of our results if those who opt in are different from those who do not. Because units are randomized into control and treatment groups, this is not a concern for the internal consistency of our results. Also, the choice to opt-in was not for this particular study, but to receive messages and allow use of data generally.

pre-intervention period. The numbers reveal a stark difference in usage across cost structures, with the average usage being markedly higher in the units in which the marginal cost of electricity usage is zero. Part of the difference in average usage could be attributed to the skewed distribution of larger units in the MC = 0 treatment (see Table 2). However, the difference in average usage between cost structures remains statistically significant when controlling for bedrooms, number of participants and building age. As would be expected from randomized allocations, the average usage within each cost structure is similar by treatment group. However, the average daily usage in the control group is 3.48% higher for MC > 0 and 2.67% higher for MC = 0 than the corresponding peer comparison treatments. The initial difference in average usage suggests that the random assignment did not fully mitigate differences between control and treatment groups. We attempt to control for this in our analysis of average treatment effects in the results section.

Across the control and treatment groups, the average number of bedrooms, square footage or age of the unit are not statistically different (Table 3).

TABLE 3
Average daily electricity usage (kWh) before interventions (first 48 days) and apartment characteristics by payment method

	Control	Peer comparison (PC)	p-value (2-tail)
MC > 0			
Usage (kWh)	15.15	14.64	.0029[a]
	(10.02)	(7.90)	
Number of bedrooms	1.77	1.84	.4323
	(.54)	(.70)	
Square Feet	728.87	734.86	.8351
	(180.09)	(250.47)	
Age	37.12	36.24	.6226
	(14.15)	(13.06)	
MC = 0			
Usage (kWh)	29.26	28.50	.0183[a]
	(10.61)	(12.28)	
Number of bedrooms	3.40	3.09	.0952
	(.89)	(.97)	
Square Feet	1095.91	1039.35	.1075
	(165.89)	(189.90)	
Age	17.30	17.16	.6765
	(1.73)	(1.87)	

Notes: Standard deviations in parentheses.
[a] Each observation for usage is daily usage for each unit. If we instead consider the average for each unit over the 48 days as the level of observation, this difference is no longer significant.

On day 49, the participants in the peer comparison (PC) treatment received the first information message. Each participant was simultaneously sent a text and an email with the relevant peer comparison information (see Figure 1). From September 2 to May 30 (length of the study) eight messages in total were sent, staggered about 30 days apart. The messages provided the electricity usage (kWh) in the unit over the previous four weeks, how that compared to the most efficient units (20th percentile) and to the average units in their peer comparison group. Peer comparison groups were defined by price structure and apartment size (see first column in Table 2). For clarity, there were four peer comparison groups for MC>0 (one bedroom, two bedroom small, two bedroom large, and 3+ bedrooms) and two peer comparison groups for MC=0 (two bedroom large and 3+ bedrooms).

A few studies use email to deliver home energy reports or appliance-specific energy feedback (Brulisauer et al. 2020; Henry et al. 2019; Myers and Souza 2020), but to our knowledge we are the first to use text messaging. Although we expect that the student demographic is more likely to read and respond to text communication, we simultaneously delivered the information via text and email. This strategy was used to increase the likelihood participants were exposed to the intervention messages.[8] We collected data for 272 days in total, resulting in 96,747 unit-level observations. As described in Table 2, some units have multiple people receiving messages from NRLP which leads to 136,855 participant-level observations. We control for the multiple unit participants in our analysis.

4. RESULTS

Figure 2 illustrates the higher pre-treatment daily usage in units that do not pay a marginal cost for electricity. For participants with MC > 0, we also have data on whether the electricity bills were paid by someone other than the tenant (e.g., parents), which we consider a pseudo MC = 0 condition. In these situations, the tenant is charged at the margin for electricity but is not the person directly responsible for paying. While the number of units in our sample who have a non-tenant payer is small (37 units), the difference in usage between those who pay their own bill or have non-resident bill payers is significant ($p < 0.001$).[9]

It is possible that the differences in pre-treatment usage we observe are not entirely driven by the difference in electricity pricing. For example, differences in building characteristics between MC = 0 and MC > 0, or the self-selection of high-use individuals to MC = 0 properties could explain some of what we observe. Although we cannot measure the extent of self-selection, there are reasons this potential confound may be limited for our particular sample. First, there is consistency in heating sources in our sample with 98% of units using electricity for heating; second, the MC = 0 units are newer (and likely more efficient). Third, our participants are primarily university students looking for housing in a supply constrained market.

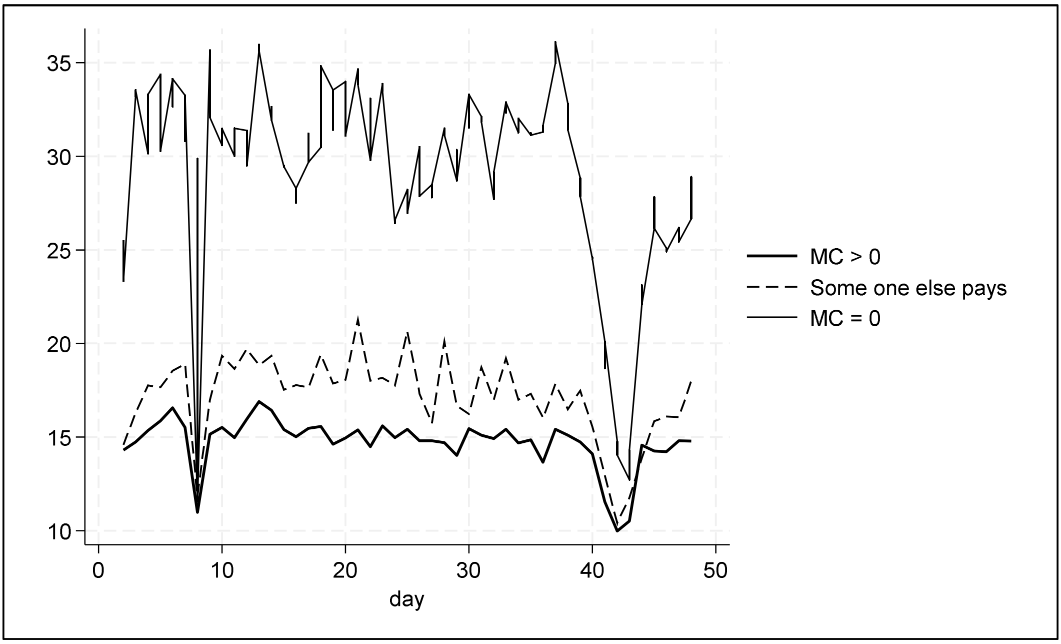
4.1 Average treatment effects

We start by establishing a baseline measure of usage by creating a variable for the average daily usage for the control group, in each cost structure (group *C* and *C0* in Table 1) throughout the duration of the study. Then, following Alcott (2011), we normalize daily usage by dividing it by the daily baseline of the control group, and multiplying it by 100 to facilitate the interpretation of results in percentage changes. There are a couple of reasons for doing this normalization. One is to view the results as a percentage change relative to the control group. Secondly, as shown in Table 3, the daily pre-treatment data in the control and treatment groups are not statistically equivalent. Benchmarking the treatment data by the control group

8. We could not observe whether participants viewed their text messages, we only received feedback in the limited cases in which the text messages bounced (incorrect number or provider). We documented these cases in the study and in many situations were able to fix the problem before the next message. We could, however, observe how many times the participant opened their email messages, and we documented this in our data as well.

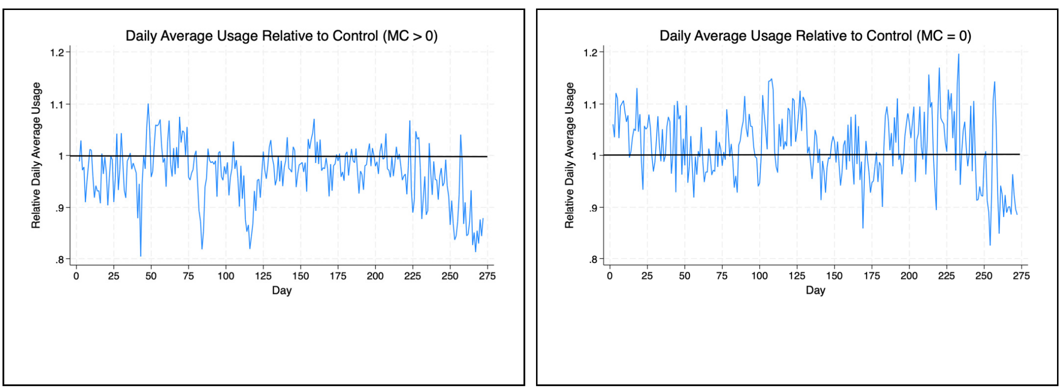
9. We thank an anonymous reviewer for pointing out that the observed difference in usage when "someone else pays" could also reflect particular types of residents that may be less concerned about energy conservation, even if they were responsible for paying their own bills. For example, students from affluent families that may not be responsive to electricity prices.

FIGURE 2
Pre-treatment electricity usage (kWhs) by cost structure



mitigates this imbalance. Figure 3 illustrates the average daily electricity use for the treated groups relative to control groups (1 on the vertical access) across both cost structures. Points below the horizontal line indicate daily averages in which the treated group uses less electricity on average compared to the baseline group.

FIGURE 3
Daily average electricity use for treated groups relative to control



To estimate the average treatment effects, we estimate the following panel model:

$$Y_{it} = \beta_0 + \beta_1 T_i + \beta_2 P_i + \beta_3 HDD_t + \beta_4 CDD_t + \beta_5 APT_i + \gamma \upsilon_i + \varepsilon_{it},$$

where Y is the normalized daily electricity usage, T indicates whether the unit is part of the *treated group* ($T = 1$ if the residents received messages and $T = 0$ if residents were in the control group), P denotes the *number of participants per unit* receiving messages, *HDD* and *CDD* are heating and cooling degree days and account for temperature variation.[10] The variable *APT* is the average daily usage by unit before the treatment begins (days 1—48) which controls for prior differences in usage independent of the assigned treatment. Variables in vector υ_i control for apartment characteristics (*bedrooms, square footage, age*), and γ is the corresponding vector of coefficient estimates. The model is estimated using data for the duration of the treatment period (days 49 through 272) using OLS and robust standard errors clustered at the unit level.[11]

We first run separate models for each cost structure to observe the influence of the nudge on each subsample. The first column of Table 4 reports the regression results over the treatment period for MC > 0, and the average treatment effect (ATE) is indicated by the shaded row. When MC > 0 we see an average reduction of 4.69% in daily electricity usage from the peer comparison information treatment. The average treatment effect in our study is on the higher end of other large-scale field experiments (e.g., Allcott 2011, Alcott 2015; Ferraro and Price 2013; Allcott and Kessler 2019; Henry et al. 2019; Jessoe et al. 2021), and is weakly significant using a one-tail hypothesis test (one-tail p = 0.096).

Usage increases significantly with the number of bedrooms, square footage of the apartment and the age of the unit. We also see a negative but insignificant effect of having more participants in the study residing in the same unit. Finally, we observe the pre-treatment usage control variable is significant and positive.

The second column of Table 4 shows an average treatment effect of positive 0.88% for non-ratepaying residents, though the estimate is statistically insignificant (one-tail p = 0.396). That we find the behavioral intervention does not lead to a significant treatment effect for non-ratepaying customers is consistent with Delmas and Lessem (2014), Crago et al. (2020) and Myers and Souza (2020). It is important to note that these studies, like ours, also have relatively small sample sizes and may suffer from limitations in statistical power. Specifically, Delmas and Lessem (2014) sample 66 rooms in university residence halls, Crago et al. (2020) sample 62 households in a single graduate housing complex and Myers and Souza (2020) sample fewer than 400 college residencies.

Our study makes a unique contribution to this literature by comparing two different cost structures within the same randomized control trial. When customers pay for electricity the magnitude of the average treatment effect in our study (–4.69%) is higher than the average treatment effects in other large-scale studies (e.g., Allcott 2015) but is only weakly significant. The average treatment effect of a positive 0.88% for those whose electricity is included in their rent is not significantly different from zero, suggesting that the behavioral nudge is less effective (or ineffective) in the absence of pecuniary motives to reduce consumption.

As illustrated in Figure 2, the average usage between ratepayers and non-ratepayers differs, so percentage changes due to the intervention should be interpreted accordingly. For example, in the first 48 days of the study the average daily use for MC > 0 units in the control group is 15.15 kWh, while MC = 0 units in the control group use 29.26 kWh on average. Therefore,

10. A degree day compares the mean (the average of the high and low) outdoor temperatures recorded for a location to a standard temperature, usually 65° Fahrenheit (F) in the United States (EIA.gov).

11. Participants in the peer-comparison treatments could opt-out of receiving messages. In total ten participant units opted out of receiving digital messages. Six of these were in the MC = 0 treatment and four in the MC > 0 treatment. In the MC = 0 treatment, one stopped receiving messages after January, four after March and one after April (the last text was sent in May). In the MC > 0 treatment, one unit stopped receiving messages after January, one after February, and two after March. We continue to receive usage data for these units even though they are no longer receiving text messages. The results in Tables 4 and 5 include data from these units which we interpret as 'intent to treat'.

TABLE 4
Average treatment effects by cost structure

	MC > 0	MC = 0
Treated Group (ATE)	–4.685	0.879
	(3.571)	(3.328)
	[0.191]	[0.792]
Number of Participants per Unit	–2.846	–2.918
	(3.408)	(1.967)
	[0.405]	[0.141]
Number of Bedrooms	12.86	11.43
	(4.858)	(2.068)
	[0.009]	[0.000]
Square Feet	0.0293	
	(0.012)	
	[0.016]	
Age of Unit	0.288	–1.089
	(0.170)	(1.090)
	[0.092]	[0.320]
Heat Degree Days	0.0514	0.0272
	(0.080)	(0.082)
	[0.524]	[0.739]
Cool Degree Days	–0.630	–0.515
	(0.565)	(0.578)
	[0.266]	[0.375]
Avg Pre-treatment Usage	2.835	1.022
	(0.280)	(0.272)
	[0.000]	[0.000]
Constant	5.525	54.66
	(10.779)	(19.430)
	[0.609]	[0.006]
Observations	52,192	26,880
Clusters	233	120
R^2	0.243	0.124

Notes: Standard errors (in parentheses) clustered by unit. In the MC = 0 treatments, square footage was excluded because square footage and bedroom count were extremely highly correlated (VIFs over 100). P-values for two-tail hypothesis tests in brackets.

using the pre-treatment period as a reference, a 4.69% reduction for ratepaying units is roughly a reduction of 0.71 kWh per day, and a 0.88% increase for non-ratepaying units is 0.26 kWh.

To improve the efficiency of the model we merge the data into a full panel and interact the cost structure dummy (MC = 0 is labeled as *MC0*) with the covariates and again examine the data through the duration of the treatment (day 49 to day 272). As a robustness check we exclude outliers (defined as daily usage ± 3 standard deviations from the mean, by treatment). Table 5 presents the results from both models.

Using the pooled data and before controlling for outliers, we find an average treatment effect of –5.40% when MC > 0, which is significant using a one-tail test ($p = 0.069$). The effect deceases to –4.51% when outliers are removed but remains weakly significant (one-tail $p = 0.091$). Therefore, we find that, on average, recipients of peer-comparison nudges reduce their usage between 4.51% and 5.40% relative to the control group when MC > 0. The interaction

TABLE 5
Average treatment effects using pooled data with interactions

	All data	Without outliers
Treated Group	–5.402	–4.510
	(3.625)	(3.368)
	[0.137]	[0.181]
Number of Participants per Unit	–4.365	–1.403
	(3.424)	(3.262)
	[0.203]	[0.667]
Number of bedrooms	23.94	17.56
	(3.107)	(2.762)
	[0.000]	[0.000]
Age of Unit	0.397	0.201
	(0.176)	(0.142)
	[0.025]	[0.158]
Heat Degree Days	0.0514	0.100
	(0.080)	(0.063)
	[0.523]	[0.114]
Cool Degree Days	–0.630	–1.052
	(0.565)	(0.462)
	[0.265]	[0.024]
Marginal Cost = 0 (MC0)	2.289	–4.237
	(22.211)	(18.781)
	[0.918]	[0.822]
MC0 x Treated Group	5.311	4.267
	(5.079)	(4.511)
	[0.296]	[0.345]
MC0 x Num Participants	2.276	0.181
	(3.972)	(3.692)
	[0.567]	[0.961]
MC0 x Bedrooms	–19.83	–11.25
	(3.721)	(3.045)
	[0.000]	[0.000]
MC0 x Age	–0.105	–0.481
	(1.194)	(0.925)
	[0.930]	[0.604]
MC0 x HDD	–0.0242	–0.0567
	(0.114)	(0.097)
	[0.833]	[0.559]
MC0 x CDD	0.116	0.746
	(0.807)	(0.677)
	[0.886]	[0.271]
Avg Pre-treatment Usage	2.392	2.037
	(0.221)	(0.204)
	[0.000]	[0.000]
Constant	11.87	28.18***
	(11.932)	(10.083)
	[0.321]	[0.005]
Observations	79,072	78,024
Clusters	353	353
R^2	0.206	0.153

Notes: Standard errors (in parentheses) clustered by unit. P-values for two-tail hypothesis tests in brackets.

between MC = 0 (labelled *MC0*) and the treatment (*MC0* x *Treated Group*) is positive in both models, which is consistent with our expectation that the nudge is less effective without financial incentives to lower electricity use. However, the effect is insignificant in both the full model (5.31, one-tail $p = 0.148$) and without outliers (4.27, one-tail $p = 0.173$). An *F*-test of whether the combined coefficients for *Treated Group* and *MC0 Treated Group* equal zero cannot be rejected (one-tail $p = 0.490$ and $p = 0.468$ for the models with and without outliers, respectively). In short, we find evidence that the behavioral nudge reduces electricity usage, but only for ratepaying customers.

In an ex-post analysis, we find that our statistical tests were underpowered largely because of much larger variation in usage in our sample compared to the estimates used in our ex-ante power analysis. The ex-ante power analysis was conducted using a historic estimate of variance provided by NRLP. Given the variance estimate, our expected sample size was sufficient to detect average treatment effects of at least two percent at lower levels of significance.[12] Ultimately, the usage in our sample was significantly more variable than predicted leading to low power. To illustrate, assume the true effect of the nudge is a two percent reduction in usage (as in previous studies). Given our estimated standard errors and using a one-tail test at a 5% significance level, the likelihood we correctly reject the null (of zero effect) is only 0.14. Therefore, there is a high likelihood of false negatives in our study.

4.2 Participant feedback

Shortly after the conclusion of the study, the residents that received text and email messages over the course of the year (i.e., those in the PC and PC0 treatments) were invited to complete a survey about receiving home energy reports. Residents were asked about their preferred method of receiving information (text vs email), whether the messages motivated them to change their behavior related to electricity usage, and how closely they read the message. We also asked questions to assess the recipients' feelings about the content of the messages. Of the 279 individual residents receiving messages, 83 completed the survey, which translates to a 30% response rate.[13] Though respondents are not randomly drawn from the study participants, results offer unique and important insights on how consumers react to information interventions like peer comparisons.

Roughly 52% of respondents preferred receiving the notices by text, 21% preferred email and 27% preferred to receive both. When asked how carefully they read the messages 88% answered either "very carefully" or "somewhat carefully", and 87% either "strongly agreed" or "agreed" that the messages motivated them to try and conserve electricity usage. That 87% of respondents agreed that the digital messages motivated them to try to reduce electricity consumption corresponds to our empirical findings. Whether or not they were able to translate that motivation into real energy savings is a function of many things, including how well informed the residents are of the ways they could reduce electricity consumption.

The messages raised different types of emotions among respondents (Table 6). A majority (57%) felt appreciative of the information, but 10% of respondents reported being annoyed by the messaging (and 11% reported feeling both appreciative and annoyed). Overall, 78% replied that receiving the notices made them feel better off compared to not getting messages. Therefore, for the respondents in our study, the behavioral intervention appears to have im-

12. The power analysis was reviewed as part of the grant application process.
13. We incentivized residents to complete the survey by selecting one respondent at random to receive a $50 Amazon gift card.

TABLE 6
Respondents' reported feelings regarding the messages (percent of survey respondents)

	Appreciative	Annoyed	Neither	Both
Some people feel either appreciative or annoyed when they receive the messages. Did you feel appreciative, annoyed, neither, or both?	57	10	22	11
	Inspired	Pressured	Neither	Both
Some people feel either inspired or pressured when they receive the messages. Did you feel inspired, pressured, neither, or both?	22	16	45	18
	Proud	Guilty	Neither	Both
Some people feel either proud or guilty when they receive the messages. Did you feel proud, guilty, neither, or both?	18	13	47	22
	Better off	Worse off	No effect	
Overall, do you feel that receiving the messages made you better off, worse off or had no effect?	78	0	22	

proved the perceived welfare of the residents. Our finding is consistent with Allcott and Kessler (2019) who show similar nudges (home energy reports) increase consumers' welfare.

5. CONCLUSION AND POLICY IMPLICATIONS

We set out to explore the effectiveness of a widely-used behavioral nudge—home energy reports with social comparison information—with and without monetary incentives to conserve energy. In addition to being able to study both cost structures in the same randomized control trial, another novelty of our study is the use of text messaging to deliver energy reports rather than traditional paper billings or emails. In order to send customized text messages to specific households, we collaborated on this study with a university-owned electric utility.

A handful of previously published studies have used similar behavioral interventions in samples of non-ratepaying households (Delmas and Lessem 2014; Crago et al. 2020; Myers and Souza 2020). Each of these studies find an insignificant effect on electricity usage and/or thermostat adjustments. The implication is that the behavioral nudge is ineffective in the absence of pecuniary motives to reduce electricity consumption. Our study adds to this discussion by assessing the intervention in both ratepaying and non-ratepaying residencies, allowing us to compare the relative effectiveness of the nudge in the same population. The existing studies with non-ratepaying customers cannot make this comparison. The magnitude of the average treatment effect we observe for ratepayers is on the higher end of large-scale randomized control trial studies starting with Allcott (2011). However, given the relatively small sample size and high variability in electricity use, the standard errors in our study are high leading to low statistical power. Therefore, the observed average treatment effect on this subsample of approximately a 4.69% reduction is only weakly significant. We find that the behavioral intervention motivates non-ratepaying units to use 0.88 more electricity than the control group, but this result is statistically equivalent to zero. However, as noted, we cannot rule out the possibility of a false negative given the limitation in power.

Looking at the combined panel of data from the start of treatment, we find that the nudge caused ratepaying residents to reduce their electricity use by 5.4% relative to the untreated residents, and this effect is weakly significant. The effect of the nudge almost entirely disappears when used on non-ratepaying customers. This evidence suggests the importance of financial motives underpinning the success of peer-comparison nudges.

We point out a few limitations of our study. First, legal restrictions on NRLP's ability to send text messages to residents without consent (not specific to this research, but in general) limited the number of participants in our study, resulting in lower power for our results. Second, as in similar studies, we could not observe whether or not participants read their text messages or control access to other information. Third, the student population we are sampling from exhibits more variability in electricity usage than traditional households, as some students spend time away from their apartment (e.g., visit home on weekends) while others have guests visit their apartment (e.g., frequent parties). Fourth, we cannot rule out that the baseline average difference in usage between rate structures is impacted by self-selection of high users into non-ratepaying units.

Our findings offer relevant insights on the role of monetary incentives in the behavioral responses to information nudges. Further, estimates highlight the important role that marginal pricing plays in regulating household electricity usage, an important result considering the significant number of non-ratepaying households. Based on the 2015 and 2020 Residential Energy Consumption Surveys (U.S. Energy Information Administration), approximately 5–6 percent of all U.S. households have some or all of their electricity use included in their rent or condo fees. The impact of flat pricing schemes has led many states to require new residential units to be individually metered, but policymakers should consider processes to transition households away from existing non-ratepayer arrangements.[14]

ACKNOWLEDGEMENTS

We thank the Energy and Environment Program at the Alfred P. Sloan Foundation for generously funding this research. The project was made possible through a collaboration with the Appalachian Energy Center and New River Light and Power at Appalachian State University in Boone, NC. We thank Jason Hoyle, Ed Miller, Janet Miller, Diana Wilcox and many research assistants for their help on this project. We thank Charles Holt for his comments and suggestions, and three anonymous referees who provided helpful insights.

References

Allcott, H. 2011. "Social Norms and Energy Conservation." *Journal of Public Economics* 95(9–10):1082–95. https://doi.org/10.1016/j.jpubeco.2011.03.003.

Allcott, H. 2015. "Site Selection Bias in Program Evaluation." *Quarterly Journal of Economics* 130(3): 1117–1165. https://doi.org/10.1093/qje/qjv015.

Allcott, H. and S. Mullainathan. 2010. "Behavior and Energy Policy." *Science* 327(5970): 1204–1205. https://doi.org/10.1126/science.1180775.

14. https://www.ncsl.org/research/energy/utility-submetering.aspx

Allcott, H. and T. Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review* 104(10): 3003–3037. https://doi.org/10.1257/aer.104.10.3003.

Allcott, H., and J.B. Kessler. 2019. The Welfare Effects of Nudges: A Case Study of Energy use Social Comparisons. *American Economic Journal: Applied Economics*, *11*(1), 236–76. https://doi.org/10.1257/app.20170328.

Ayers, I., S. Raseman and A. Shih. 2013. "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage." *Journal of Law, Economics and Organization* 29(5): 992–1022. https://doi.org/10.1093/jleo/ews020.

Benabou, R. and J. Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70(3): 489–520. https://doi.org/10.1111/1467-937X.00253.

Benabou, R. and J. Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96: 1652–1678. https://doi.org/10.1257/aer.96.5.1652.

Brulisauer, M., L. Goette, Z. Jiang, J. Schmitz and R. Schubert. 2020. "Appliance-specific Feedback and Social Comparisons: Evidence from a Field Experiment on Energy Conservation." *Energy Policy* 145. https://doi.org/10.1016/j.enpol.2020.111742.

Buckley, P. 2020. "Prices, Information and Nudges for Residential Electricity Conservation: A Meta-Analysis." *Ecological Economics* 172. https://doi.org/10.1016/j.ecolecon.2020.106635.

Costa, D.L. and M.E. Kahn. 2013. "Energy Conservation "Nudges" and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment." *Journal of the European Economic Association* 11(3): 680–702. https://doi.org/10.1111/jeea.12011.

Crago, C. L., J.M. Spraggon and E. Hunter. 2020. "Motivating Non-ratepaying Households with Feedback and Social Nudges: A Cautionary Tale." *Energy Policy* 145 (2020). https://doi.org/10.1016/j.enpol.2020.111764.

Delmas, M.A. and N. Lessem. 2014. "Saving Power to Conserve your Reputation? The Effectiveness of Private vs. Public Information." *Journal of Environmental Economics and Management* 67(3): 353–370. https://doi.org/10.1016/j.jeem.2013.12.009.

Ferraro, P.J. and M.K. Price. 2013. "Using Nonpecuniary Strategies to Influence Behavior: Evidence form a Large-Scale Field Experiment." *Review of Economics and Statistics* 95(1): 64–73. https://doi.org/10.1162/REST_a_00344.

Henry, M.L., P.J. Ferraro and A. Kontoleon. 2019. "The Behavioral Effect of Electronic Home Energy Reports: Evidence from a Randomized Field Trial in the United States." *Energy Policy* 132: 1256–1261. https://doi.org/10.1016/j.enpol.2019.06.039.

Ioannidis, J.P.A., Stanley, T. D. and H. Doucouliagos. 2017. "The Power of Bias in Economics Research." *The Economic Journal* 127(605): F236–F265. https://doi.org/10.1111/ecoj.12461.

Ito, K., T. Ida and M. Tanaka. 2018. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy* 1: 240–67. https://doi.org/10.1257/pol.20160093.

Jessoe, K. and D. Rapson. 2014. "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review* 104(4): 1417–38. https://doi.org/10.1257/aer.104.4.1417.

Jessoe, K., G.E. Lade, F. Loge and E. Spang. 2021. "Spillovers from Behavioral Interventions: Experimental Evidence from Water and Energy Use." *Journal of the Association of Environmental and Resource Economists* 8(2). https://doi.org/10.1086/711025.

Levinson, A. and S. Niemann. 2004. "Energy Use by Apartment Tenants when Landlords Pay for Utilities." *Resource and Energy Economics* 26(1): 51–75. https://doi.org/10.1016/S0928-7655(03)00047-2.

Levitt, S. and J. List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *Journal of Economic Perspectives* 21(2): 153–174. https://doi.org/10.1257/jep.21.2.153.

Myers, E. and M. Souza. 2020. "Social Comparison Nudges without Monetary Incentives: Evidence from Home Energy Reports." *Journal of Environmental Economics and Management* 101. https://doi.org/10.1016/j.jeem.2020.102315.

Munley, V.G., L.W. Taylor and J.P. Formby. 1990. "Electricity Demand in Multi-Family, Renter-Occupied Residences." *Southern Economic Journal* 57(1): 178–194. https://doi.org/10.2307/1060488.

Nolan, J.M., P.W. Schultz, R.B. Cialdini, J. Goldstein and V. Griskevicius. 2008. "Normative Social Influence is Underdetected." *Personality and Social Psychology Bulletin* 34: 913–923. https://doi.org/10.1177/0146167208316691.

Pellerano, J.A., M.K. Price, S.L. Puller and G.E. Sanchez. 2017. "Do Extrinsic Incentives Undermine Social Norms? Evidence from a Field Experiment in Energy Conservation." *Environmental and Resource Economics* 67(3): 413–428. https://doi.org/10.1007/s10640-016-0094-3.

Royal, A. and G. Rustamov. 2018. "Do Small Pecuniary Incentives Motivate Residential Peak Energy Reductions? Experimental Evidence." 50(57). https://doi.org/10.1080/00036846.2018.1489508.

Schultz, P.W., J.M. Nolan, R.B. Cialdini, N.J. Goldstein and V. Griskevicus. 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science* 18(5): 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x .

US Census. 2013. American Housing Survey 1997–2013.

U.S. Energy Information Administration (EIA). 2019. "Demand-side Management Programs Save Energy and Reduce Peak Demand," Today in Energy, March 29, 2019.

IAEE

International Association for
ENERGY ECONOMICS

Membership in the International Association for Energy Economics is open to anyone worldwide who has an interest in the fields of energy or energy economics. Our membership consists of those working in both the public and private sectors including government, academic and commercial. Our current member base consists of 3900+ members in over 110 nations, with 28 nations having local affiliate organization.

We are an independent, non-profit, global membership organization for business, government, academic and other professionals concerned with energy and related issues in the international community. We advance the knowledge, understanding and application of economics across all aspects of energy and foster communication amongst energy concerned professionals.

We are proud of our membership benefit offerings, which include access to a rich library of energy economics related publications and proceedings as well as a robust line-up of webinars, podcasts and conferences. Learn more about the benefits of membership at: https://www.iaee.org/en/membership/benefits.aspx

In addition to traditional membership, we offer student and institutional memberships.