Can Electricity Conservation in Student Housing Be Nudged? Evidence from National Tsing Hua University

This Version: March 22, 2022

Abstract

This paper applies a non-price-based conservation experiment in college dorms among 6,723 residents. We explore the effectiveness of behavioral intervention by emailing residents two types of social comparison based energy consumption feedback: quintile social comparison message and mean comparison message. A quintile social comparison message ranks residents' relative energy consumption in quintiles, while the mean comparison message compares residents' energy consumption with the average. Our results show that the effectiveness of these messages depends on the residents' baseline consumption pattern. The highest 20% of users reduce their energy consumption with a quintile comparison message, while the lowest 20% are more responsive to a mean comparison message. We also find that messages are more effective with undergraduate residents than graduate students. This paper provides empirical evidence of the conservation study in a university setting beyond the US context. Our results suggest that the feedback design should be context-specific.

1 Introduction

Climate change is one of the most vital issues facing the world. Governments and researchers have been increasingly applying nudge-style intervention strategies to reduce greenhouse gas emissions by changing people's energy conservation behaviors because of their political feasibility and cost-effectiveness. Many studies have shown that these non-price-based conservation programs effectively reduce energy consumption. For example, providing Home Energy Reports (HERs) to US residential customers can reduce energy consumption by 2% (Allcott, 2011). Similar HERs have been widely implemented in 9 countries (Allcott and Kessler 2019). Other researchers publicly display energy usage at the university dorms, and this approach leads to 17% $\sim$ 20% of energy savings (Delmas and Lessem 2014; Bruelisauer et al., 2018). Yet, HERs do not improve conservation behaviors in other settings: Myers and Souza's (2020) HERs experiment at UIUC dorms; Andor et al. (2020) HERs research in Germany. Among these studies, the discussion has shifted to designing more targeted messages for different subgroups of electricity users. Following our discussion, we conduct experiments in university dormitories to examine which type of information policy can be most effective. Specifically, we ask two research questions: (i) Which type of messages are more effective among people with different electricity consumption patterns? (ii) Does the number of messages sent affect conservation behaviors differently?

Our research is closest to the conservation experiments conducted in the university settings. Based on our understanding, this stream of literature could be broadly divided into two types: (1) the competition type and (2) the information-provision type. In the first type, researchers usually recruit a group of willing students to participate in an experiment framed as a competition with monetary or corresponding rewards (Brewer et al., 2011; Delmas and Lessem, 2014; Chen et al., 2021). In the second type, researchers only offer their own electricity usage information or comparison information to peers without more incentives (Dixon et al., 2015). The results from the previous studies show that both types have had significant conservation effects. But most of the current studies are through recruitment, and thus they may suffer from selection bias. For example, the recruited participants may be more aware of the environmental issues and thus more responsive in changing their energy usage behaviors. The external validity may also be questionable because many studies do not use large sample sizes. In terms of policy implications, hosting a competition at a university may not always be practical because universities may not be able to constantly allocate large monetary rewards for energy conservation all the time. As for the predicament, we are particularly interested in formulating an effective and efficient strategy based on the nudging theory. A slight push leads to beneficial behavior changes to reduce electricity consumption in a university dorm setting. We also overcome the issue of recruitment selection bias.

We implemented our experiment in the dormitories at National Tsing Hua University (NTHU) in Taiwan in the fall semester of 2019. The unique design of our study is that rather than recruiting voluntary respondents, we have access to all the residents' emails at NTHU. With the school's permission, we can send our designed messages through the central administrative system to all the students living in the university dorms. In the emails, we provide comparative feedback on students' electricity consumption. We design two types of framing. One is a quintile comparison where we inform residents of the quintile of their electricity consumption (*Pct group).* The other uses a mean comparison where residents only know whether their consumption is below or higher than the average (*Avg group*). Based on the framing design (*Pct group or Avg group*) and how many times the messages were sent (Once or twice), we divided our residents into five groups: the control group, the OnceQt group, the TwiceQt group, the OnceAvg group, and the TwiceQt group.

Our results show that the conservation effect is more significant when a higher user receives a quintile comparison. On the contrary, a lower user saves more when receiving a mean comparison message. This may be because a mean comparison, relative to quintile comparison, can avoid a possible boomerang effect—a typical concern in the literature where low users may increase energy consumption after receiving social comparison information. Besides, this type of comparative message significantly affects college students but not graduate students, probably because of the economic factor and the time spent in the dormitory. Furthermore, the number of times messages are sent is also critical–messages have to be sent more than once to be effective. The contributions of this study are: (i) the extra empirical evidence to the conservation study in the university setting beyond the US context, (ii) addressing the usual recruitment selection bias through our universal access to all residents' emails and (ii) the design of more refined messages to the target group to bring larger impacts.

The organization of this paper is as follows. The next section reviews the literature related to nudge theory and energy saving. Section 3 introduces our experiment's design, including the sample selection, experiment duration, and the design of a nudging strategy. A data description and estimation strategies will then be provided in Section 4. Section 5 describes the empirical results and discusses our findings compared with previous studies. The last section summarizes this paper.

2 Literature

"Nudge" theory has been popular in public policies and businesses. Nudge is a small change to get people's attention and this theory argues that a subtle change might lead to unanticipated beneficial outcomes of decision-making and behaviors (Sunstein & Reisch, 2017). The concept of nudging has been taken seriously in various governments, such as the United Kingdom, the United States, and Australia—they all have established "nudge units" to carry out behavioral insights in policymaking (Sunstein & Reisch, 2017).

Even if nudge theory is highly recognized and promoted, the effectiveness is not always guaranteed because of the variety of information provided to target groups. The energy efficiency label is one common example. Many governments hope to encourage households to purchase energy-efficient appliances by creating different labels. But the evidence shows that the information added to labels does not necessarily have any effect (Newell & Siikam¨aki, 2014). Such ineffective label-informed provision was also found by Allcott and Sweeny (2017) and Allcott and Knittel (2019). In addition, the same message may not have the same impact on different socioeconomic groups. For example, the environment and health-based information strategies are effective on families with children in achieving energy savings, but the same information is ineffective among other groups (Asensio & Delmas, 2015). A similar challenge exists when providing peers' energy usage. Energy conservation nudges are more acceptable among political liberals than political conservatives (Costa & Kahn, 2013). Therefore, designing a more targeted message for different subgroups is crucial to ensure effective nudging.

In energy-saving experiments, high and low users are the most common target group because they often respond differently to the same message differently due to their different energy usage habits. High-energy consumers often reduce electricity consumption when receiving comparative feedback. On the contrary, low-energy consumers reduce their consumption less than high-energy consumers (Allcott, 2011; Ferraro & Price, 2013), or even increase their consumption after uncovering their relative usages (Anderson et al., 2017; Bruelisauer et al., 2018)–the so-called boomerang effect (Schultz et al., 2007). The previous studies have often eliminated the boomerang effect by adding an injunctive norm (e.g., Schultz et al., 2007; Anderson et al., 2017; Bonan et al., 2020) that connotes the social approval or disapproval of a particular behavior and even constitutes a moral rule of a social group (Cialdini et al., 1991). This study intends to design effective messages for people with different initial energy usages.

Another consideration is the extent to which it drives significant behavioral change. Specifically, how many messages do we have to send for them to be effective, or how often do we have to send them? The literature is still inconclusive. No matter how often the messages are provided (daily, weekly, or monthly), some results show significant effects while some show insignificant effects (Fischer, 2008; Allcott, 2011; Asensio & Delmas, 2015; Asensio & Delmas, 2016; Meub et al., 2019; Myers & Souza, 2020). Yet the growing consensus is that people need sufficient time to adjust their electricity consumption habits to achieve a significant power-saving effect (Fischer, 2008; Allcott and Rogers, 2014; Anderson et al., 2017). One-shot information is not enough to encourage conservation behaviors (Fischer, 2008). Therefore, this study seeks to provide evidence to establish an effective nudge. Our study helps better understand how often and long it takes for messages to drive significant behavioral change. This insight is helpful for a nonprofit organization, as the public university in our sample, because implementing nudging policies can create a financial burden in the long run. A more fine-tuned nudging strategy can help organizations generate significant behavioral change more cost-effectively.

3 Experiment Design

We conducted a randomized control trial (RCT) in the student dormitories of National Tsing Hua University. The delivered energy report creates an opportunity to reveal the relative electricity usage compared with neighbors. Our experiment focuses on exploring the subsequent behavioral change responding to the nudge. This section would expand on the experiment design.

Unlike most electricity experiments in dormitories, we cooperated with the administrative unit in which all residents were included and without the need to recruit. In total, there are 6,895 residents in the final sample. The residents are distributed among twenty buildings with 2,709 rooms[1]. There are four room types: single, twin, triple, and quadruple rooms [2].

Considering the University's academic schedule, the climate, and the electricity consumption habits in Taiwan, we conducted the experiment from September 12, 2019, to October 31, 2019. This timing follows the University's semester schedule–the fall semester in Taiwan starts in the middle of September and ends in early January. In addition, electricity consumption varies greatly depending on the season. Air conditioning is usually the main source of energy consumption. Heating is rarely used in Taiwan, so there is almost no heating in the dormitories. So our experiment is concentrated on the season when the air conditioning is used to observe the behavioral change. We determine that mid-September to October is the optimal time period because November in Taiwan has been relatively cool, and the frequency of using air-conditioning begins to reduce or even not in use. [3].

During the above period, electricity feedback messages were delivered bi-weekly with two versions: the quintile comparison and the mean comparison, by Email. The former portrays energy usage in quintiles: top 20% as high users, bottom 20% as low users, and the intermediate users as

[1]The nature of buildings could be classified according to student status and gender. For student status, thirteen of the dormitories are for undergraduates, three are graduate dormitories, and the remaining are mixed-use. In terms of gender classification, it includes six female-only buildings, nine male-only buildings, and five mixed-gender buildings. The overall building complex goes from the second to tenth floors, while most buildings have three or four stories.

[2]There are 248 single rooms, 1,522 twin rooms, 153 triple rooms, and 786 quadruple rooms

[3]Bruelisauer et al.'s (2018) experiment demonstrates that providing the feedback of high energy-use appliances could promote graduate students' energy conservation more effectively than providing a general energy usage feedback. This finding supports taking high-energy-consuming appliances into account.

21-40%, 41-60%, and 61-80% (see Figure 1). The latter depicts electricity consumption on average, so residents know whether their consumption is higher or lower than the average (see Figure 2). With this type of message, residents could only realize their relative position of the energy consumption among the student dormitory and the respective building. Facial expression in both versions helps recipients understand the information more quickly and clearly. The past studies also indicate that such a message could reveal social appropriateness for behavior and reduce the boomerang effect, especially among the low users (Schultz et al., 2007; Bonan et al., 2020). Additionally, a survey link was sent to residents with the electricity report to understand their intentions after receiving the message and the length of time spent in their room.

The energy report number is also a manipulated intervention in our experiment. With the same content of the messages, one group would receive the feedback twice while the other group would only receive the message once. This design seeks to examine whether the number of messages would affect the conservation behavior. For this reason, the length of the experiment is divided into three periods: baseline, Period I, and Period II. The treated information is only released in Periods I and II. The first three weeks are the baseline, while the duration of the follow-up periods is two weeks.

With the type and the number of messages, the residents in the twenty buildings are randomly assigned into five groups by floors in each building: the control group, the once-quintile group ($OnceQt$), the twice-quintile group($TwiceQt$), the once-average group($OnceAvg$), and the twice-average group($TwiceAvg$) (see Figure 3).[4]. This design intends to balance the distribution of the five groups across floors and buildings to reduce the possible bias caused by differences between student halls and floors. Table 1 shows the balance among the five groups by comparing the dormitory characteristics, student's backgrounds, and energy consumption. More details are provided in the following section.

4 Data and Methods

4.1 Data

To conduct an electricity experiment among all dormitories on campus, we must obtain the detailed resident list of all dorms and the electricity record of each room. We obtained detailed and reliable administrative-level residential data collected by the Division of Student Housing–a

[4] "Once-" and "twice-" represents the number of the energy reports they received.

unit responsible for managing and operating university dormitories through special permission. The administrative-level data comprises the student directory of all dorms, such as residents' department, gender, degree, grade, and email account. But, for the electricity records, the Office of General Affairs only holds the electricity usage per building. To observe residents' behavior, we need to collect room-level records by hand. The room-level electricity records can be observed from the Watt-hour meter installed outside each room. The Watt-hour meter only records air-conditioner usage charged by air-conditioner cards (mainly for 220V).[5] We hire eight part-time students to collect electricity records for each room every two to three weeks from September to October 2019.

Besides the student directory and usage data, we also conducted online surveys that we attached the link to the email to each resident in the treatment group. This survey helps us gather extra information on students' attitudes and behaviors toward electricity usage. For example, we ask them if they will reduce their usage when they receive the electricity consumption message. The detailed questions of the questionnaire can be found in Appendix A. To increase the response rate, the questionnaire contains only five questions. We also host a lottery to grant twenty-five respondents a NT$100 gift voucher.

We generate a rich individual-level data set by combining the student directory of the dormitories, the electricity records, and the online surveys. The data set contains various important variables. For example, the electricity usage is the main dependent variable defined as usage per person per week in baseline, Period I, and Period II. We also have information on the characteristics of rooms (i.e. the located building/floor of the rooms and numbers of residents per room), residents' characteristics (i.e. residents' gender, degree, grade, and college), and attitudes and behaviors related to energy conservation (see Appendix A). We exclude the rooms with incorrect electricity usage records, with no residents, and those that have students without email information. After this correction, the number of rooms decreased from 2,709 to 2,668 rooms, and the total sample size reduced from 6,895 residents to 6,723 residents.

[5]It is worth noting that what is charged via air-conditioner cards is different by buildings. For eight buildings, the air-conditioner cards are used to pay for the electricity usage of the air-conditioner, fans, and lighting. While the other twelve buildings only need to pay for the air-conditioner.

4.2 Descriptive Statistics

Table 1 shows the sample mean of the whole individual-level data, the sample mean by treatment status, and the results of the multivariate means test. Panel A of Table 1 presents the electricity usage of each period during the experiment. We can see that the mean usage in period II is the smallest, with only 2.227 kWh per person per week and the smallest standard deviation. The phenomenon can correspond to the trend of temperature during the experiment. As shown in Figure 4, the temperature in period II is stabler and lower than the baseline and period I. Panel B of Table 1 shows the characteristics of rooms. The mean located floor is 3.5, and the average number of residents in a room is around 3. Panel C of Table 1 describes residences' characteristics. In the sample, 58.2% are male. Most of the residents are undergraduate students, accounting for 70.3% of the sample. Most of the residents are first-year and second-year students. The last column of Table 1 shows the p-value of the test that tests whether the means across our control and four treatment groups are the same. As shown in the last column, all p-values are greater than 0.1. We fail to reject the null hypothesis that the means across five groups are the same, validating our randomization process. In terms of our main treatment effect, the last column of Panel A shows that after the first and second treatments, the average usage across five groups does not have significant differences. Our nudging policy may not have significant effects. This preliminary evidence, however, is insufficient to suggest the policy effect. We will use a panel DID regression model to evaluate the policy effect.

Table 2 presents summary statistics of our two surveys. The first survey was sent to all four treatment groups (N=5,283), and the response rate was 28.13%. The second survey was sent with the second treatment information, and thus, only two treatment groups (N=2,590) received the second survey. The response rate of the second survey is 23.17%. We also asked them to fill out their usage ranking to see if their answer aligned with with our records. This question helps see whether residents have read the message carefully–they are actually "being treated." In the first survey, 92.2% correctly provided their usage ranking. In the second survey, 87.5% got the correct answer. Panel B in Table 2 compares the respondents' characteristics with the full sample. In our survey, females are more likely to respond to the questionnaires. The distribution of respondents' degrees similar to the full sample, while the proportion of master students is higher in the survey. Panel C shows the results of the survey questions. First, we want to know students' willingness to change their electricity behavior after receiving feedback. According to Table 2, around 40% of them would like to reduce their electricity usage. Second, over 86% of those surveyed reported that they care about climate change or global warming in daily life. Finally, the average time they spent in the dormitory (after deducting sleep time) is around 5 hours per day, and the mean days they stay in the dormitory per week is 5.685 days.[6]

4.3 Estimation Methods

Our estimation model is like a Panel Difference-in-Differences Method (DID), where we intend to evaluate the effect of the two proposed nudging policies. As pointed out in Section 3, we randomly assigned our rooms into five groups. One of them is the control group, while the other four are treatment groups that vary with feedback design and frequency. Because the experiment is for two months and we collect the electricity usage every two to three weeks (four times in total), we organize the data into a panel data set to help us control potential time-invariant individual effects, such as gender, grades, and other educational backgrounds. The general form of DID in the panel data framework is as follows:

$$Y_{it} = \alpha_i + \lambda_t + \gamma Post_t + \beta Treat_i + \delta(Treat_i \times Post_t) + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the outcome variable of resident i at time t. To be specific, in this study, Y_{it} is the electricity usage per person per week for resident i between the time point t and $t-1$. We divide the usage by the number of residents in their rooms because there may include one to four residents in each room. Moreover, because the frequency of collecting electricity records changes from every two to three weeks during the experiment, we transform the usage unit into a one-week unit. $Treat_i$ equals 1 if the entity i is assigned to the treatment group and 0 otherwise. $Post_t$ equals 1, representing the period after the policy, and 0 before the policy. The coefficient on the interaction term $Treat_i \times Post_t$ captures the causal treatment effect. The model also includes fixed effects, α_i, to control for resident fixed effects; and λ_t to control for time fixed effects. Finally, ε_{it} is the error term. In this main specification, δ is the parameter of interest, which measures the effects of electricity consumption induced by our nudging. We expect it to be negative if the policy is effective.

[6]For the selection options of the average time spent in the dormitory, 1.5 represents less than 3 hours, 5 is 4 to 6 hours, 8 is 7 to 9 hours, and 11 is more than 9 hours.

To further understand the effectiveness of the feedback, we designed a two by two experiment—four treatment groups varied by message type (quintile vs. average comparison) and message time (message sent once vs. twice) (see Figure 3). In the baseline (between $t-1$ and t), we did not inform all residents. The first treatment—message type treatment—was sent out at time t. Except for the control group, the treatment groups were provided with two types of information: one is the quintile ranking of an individual's baseline electricity usage (Pct group), and the other is the comparison with the average usage of the baseline electricity consumption (Avg group). Among all the treated groups, we randomly select some of them to receive the second treatment (the message time treatment), where they receive the second-time information at time $t+1$.

In the end, there are four treatment groups: once-quintile group ($OnceQt$), twice-quintile group ($TwiceQt$), once-average group ($OnceAvg$), and twice-average group ($TwiceAvg$). The $OnceQt$ group and $TwiceQt$ group received the quintile ranking of their baseline electricity usage, and the $OnceAvg$ group and $TwiceAvg$ group got the comparison with the baseline average electricity usage. At time $t+1$, only the $TwiceQt$ and $TwiceAvg$ group received the second-time information of their Period I electricity usage.

To estimate the treatment effects among different messaging formats, we combine some groups and reorganize the time period. We cut the time period of our experiment into two overlapping sections . As shown in Figure 3, the first section contains the baseline and Period I (from $t-1$, t, to $t+1$), and the second section includes Period I and Period II (from t, $t+1$, to $t+2$).

During the first section, only the first treatment was implemented. Therefore, our five groups can be shrunk into three groups. We classify the $OnceQt$ and $TwiceQt$ groups into one group, named $QtTreat$, as they receive the same quintile message treatment. Similarly, we put the $OnceAvg$ and $TwiceAvg$ groups together and name it $AvgTreat$ since they are exposed to the mean comparison message treatment. Our estimation is as follows:

$$\begin{aligned} Y_{it} = \alpha_i + \lambda_t + \gamma Post1_t + \beta_1 QtTreat_i + \delta_1(QtTreat_i \times Post1_t) \\ + \beta_2 AvgTreat_i + \delta_2(AvgTreat_i \times Post1_t) + \varepsilon_{it}, \end{aligned} \tag{2}$$

where $QtTreat_i$ equals 1 if the entity i is assigned to the $QtTreat$ group (including $OnceQt$ and $TwiceQt$, the quintile information group), and 0 otherwise. $AvgTreat_i$ equals 1 if the entity i is assigned to the $AvgTreat$ group (including $OnceAvg$ and $TwiceAvg$, the average information group), and 0 otherwise. The control group is the group without feedback. $Post1_t$ equals 1 representing Period I, and 0 is the baseline. The coefficient on the interaction term $QtTreat_i \times$

$Post1_t$ captures the treatment effect for the quintile message treatment. The coefficient on the interaction term $AvgTreat_i \times Post1_t$ captures the treatment effect for the mean comparison message treatment.

In the second section, our purpose is to evaluate the difference between the first-time and the second-time treatment effect. We drop our original control group and treat the $OnceQt$ group as the control group to compare with the $TwiceQt$ group. Since at time t, both the $OnceQt$ and $TwiceQt$ group got the same quintile comparison information. While at time $t + 1$, only the $TwiceQt$ group was exposed to the second mean comparison information. Following the same logic, we treat the $OnceAvg$ group as the control group of the $TwiceAvg$ group. By comparing the $TwiceQt$ with the $OnceQt$ group and the $TwiceAvg$ with the $OnceAvg$ group, we can examine whether the number of the experiments matters under different treatment designs. The separate two DID models are as follows:

$$Y_{it} = \alpha_i + \lambda_t + \gamma Post2_t + \beta TwiceQt_i + \delta(TwiceQt_i \times Post2)_{it} + \varepsilon_{it}, \tag{3}$$

$$Y_{it} = \alpha_i + \lambda_t + \gamma Post2_t + \beta TwiceAvg_i + \delta(TwiceAvg_i \times Post2)_{it} + \varepsilon_{it}, \tag{4}$$

where $TwiceQt_i$ equals 1 if the entity i gets the quintile information twice, and 0 if the entity i gets the quintile information once. $TwiceAvg_i$ equals 1 if the entity i gets the average information twice, and 0 if the entity i gets the average information once. $Post2_t$ equals 1 representing Period II, and 0 is Period I. The coefficient on the interaction term $TwiceQt_i \times Post2_t$ captures the differential treatment effect between the first-time and the second-time quintile message treatment, and the coefficient on the interaction term $TwiceAvg_i \times Post2_t$ captures the differential treatment effect between between the first-time and the second-time mean comparison message treatment.

5 Empirical Results

5.1 Overall Treatment Effects

The present research shows whether providing usage feedback may motivate residents to change their electricity behavior. We specifically randomize different types of messages and the number of times feedback was sent, and whether those treatment effects vary among different target groups. Table 3 summarizes the difference-in-differences regression results using equation 2 to equation 4 and presents each sample's mean usage in the "Mean" row. The top half of Table 3 presents the estimation results ($\hat{\delta_1}$ and $\hat{\delta_2}$) of equation 2. The estimated treatment effects of equation 3 and equation 4 are reported in the bottom left and the bottom right of Table 3. Panel A presents the results of the quintile message treatment compared with the control group, and Panel B summarizes the effects of the mean comparison message. We provide the estimation results with the full sample in column (1). As shown in column (1) in Table 3, the results reveal that neither the quintile message treatment nor the mean comparison message significantly affects residents' electricity behavior no matter how many times the feedback was sent. These results are much like those in Anderson et al. (2017), which found that normative messaging has no significant effect on energy conservation in the short run.

One possible explanation for no treatment effects is the context-dependency, which is also found by Andor et al. (2020). The electricity usage is relatively low in this study than the consumption level at the dormitory in the United States. As shown in Table 1, the average electricity usages during our experimental period are between 2.227–6.079 kWh per person per week, transforming to 8.908–24.316 kWh per month. In comparison, Delmas and Lessem (2014) show that the average electricity usage in the UCLA dorms is 198 kWh per person per month. Furthermore, the main source of energy consumption in our experiment in Taiwan comes from cooling, while heating and other appliance usages may be equivalently important in other countries. All the differences indicate that the treatment effects of feedback policy may be context-dependent.

5.2 Heterogeneous Treatment Effects

We next investigate whether the treatment effects vary by particular groups of residents. First, we focus on the residents with different levels of electricity use. We restrict the sample to the highest 20% and the lowest 20% users based on the baseline electricity consumption. The results in column (2) and column (3) in Table 3 show that a certain type of message is effective after providing residents the feedback twice. Among the top 20% of users, providing the quintile message twice motivates them significantly to reduce the weekly electricity usage by 0.887 kWh as compared to the first feedback. This effect is sizable, corresponding to 7.66% of their mean usage in Period I and 15.13% of the mean usage of the whole sample. In contrast, the quintile message is not effective among the lowest 20 % users, but the average-type feedback has a significant treatment effect. Providing the lowest 20% of users with the average-type feedback twice seems to decrease their usage by 0.201 kWh per person per week (although it is at a 0.1 significance level). This reduction is around 10.99% lower than the mean usage of the lowest 20% of electricity users in Period I.

The heterogeneous treatment effects arise in groups with different electricity usage levels, and the efficacy of the treatment relies on the number of times feedback was provided. The quintile information only affects residents with higher baseline electricity usage, while residents with lower baseline electricity usage are only affected by the mean comparison information. A possible explanation might be that the quintile information provides more precise position information than the average-type information. For the top 20% users, a more accurate message may alert them to the highest usage among their peers. As a comparison, the quintile information may not work well among the lowest 20% of users as this message may result in the boomerang effect that knowing their low usage induces people to use more (Schultz et al., 2007). Therefore, providing the average-type information to the lowest 20% of users may be more effective as it may "nudge" the users to save energy while the boomerang effect does not happen simutaneously.

We also examine heterogeneous treatment effects by other individual-specific characteristics: students' degree. Table 4 summarizes the subsample estimation results using equation 2 to equation 4 by students' degree. The results presented in Table 4 reveal that undergraduate students reduce their electricity consumption by 0.268 kWh per person per week in response to the second-time quintile-type information, which corresponds to 4.61% lower of their mean usage in Period I. The nudge policy, however, does not have a significant influence on masters and Ph.D. students. We suspect that these results may be driven by the economic factor and time spent in the dormitory. Because Undergraduate students are less likely to receive scholarships or teaching assistantships than graduate students, they are possibly more responsive to electricity usage information. Another possible explanation is that undergraduate students tend to spend more time in the dormitory than graduate students. We have some evident statistics through our online survey (see Figure 5, Figure 6, and Table 5)–undergraduate students stay significantly 0.3 more hours per day in the dorm than graduate students. The caveat is that the online survey is voluntary, so this sample is subject to selection bias. In all, our heterogeneity results suggest that the content of feedback should be designed based on the characteristics of the target group.

Apart from the heterogeneous effect, the subsample estimation represented in Table 3 and Table 4 also shows that the treatment should be implemented at least twice to drive the behavioral change significantly. This finding seems to align with Fischer (2008), which finds that one-shot information is too loose to create conservation behavior.

5.3 Cost-Effectiveness of the Treatment

Table 6 presents our experiment's cost-effectiveness and the cost-effectiveness of applying the nudging policy to the whole NTHU residents. According to our heterogeneous treatment analyses (see Table 3 and Table 4), the message treatment is effective among the highest and the lowest 20% baseline electricity users and among undergraduate residents after we provide the usage information twice. Focusing on the highest and the lowest users, we learn that providing the quintile comparison feedback twice to the highest 20% of users (252 residents in our treatment group) save 0.887 kWh per person per week. Providing the average-type feedback twice to the lowest 20% users save 0.201 kWh per person per week. To sum up, targeting at the highest and the lowest 20% users save 545.94 kWh usage during our study period (around one month).[7]

On the cost side, the main cost is the amount paid to collect electricity records of each room. We hire eight part-time students with a wage rate of NT$750 (US$24.25) per time (roughly three hours) to collect electricity records four times since NTHU does not have a smart meter.[8] This wage rate is NT$250 (US$8.08) per hour, which is above the market hourly minimum wage, NT$160 (US$5.17). Because the electricity feedback messages were delivered by email without extra cost, and thus, the total amount spent on the experiment is NT$24,000 (US$776.07).[9] Combining the estimated treatment effect with experiment costs, it costs NT$43.96 (US$1.42) to induce a reduction of 1 kWh electricity usage.[10] On the other hand, providing the second-time quintile comparison message to undergraduate residents (906 residents in our treatment group) saves 485.62 kWh electricity usage in total. Therefore, the cost-effectiveness of targeting undergraduate residents is NT$49.42 (US$1.60) per kilowatt-hour saved.[11]

If the nudging policy is promoted to the whole NTHU residents and we set the wage rate to the market hourly minimum wage, this policy will be even more cost-effective. Bringing this policy into practice, the implementation cost would be NT$11,520 (US$372.51).[12] If policymakers choose

[7] 545.94 kWh=(0.887kWh)*(252 residents)*(2 weeks)+(0.201kWh)*(246 residents)*(2 weeks).

[8] The exchange rate is 1 United States Dollar to 30.925 New Taiwan Dollar, the average exchange rate in 2019.

[9] NT$24,000=(8 part-time students)*(NT$250 hourly wage)*(3 hours per time)*(4 times). Besides the cost of collecting electricity records, we also conduct two online questionnaires to make sure receivers have read the content of the email. It includes twenty-five NT$100 (US$3.23) Family Mart Gift Vouchers which grants twenty-five respondents randomly in each survey. This gives a cost of NT$5,000 (US$161.68). But, bringing the nudging policy into practice in the future does not require online questionnaires. So we do not count the Gift Vouchers into our cost.

[10] NT$43.96=NT$24,000/545.94kWh.

[11] 485.62 kWh==(0.268kWh)*(906 residents)*(2 weeks). NT$49.42=NT$24,000/485.62kWh.

[12] The last electricity record is only to evaluate the treatment effect. Thus, implementing the policy only needs to collect the electricity record three times. NT$11,520= (8 part-time students)*(NT$160 hourly wage)*(3 hours per

to target the highest and the lowest 20% baseline electricity users (1,345 residents), they can save 2,927 kWh usage during our study period.[13] As a result, the cost of saving 1 kWh electricity usage is around NT$3.936 (US$0.127)(see Panel B in Table 6). To put the number into perspective, this number is less than the amount shown in Chen et al. (2021), who conducted a group contest at another major university in Taiwan. Their results show that providing both inter-ranking and intra-ranking information to all contestants costs NT$8 (US$0.259) to induce a reduction of 1 kWh. Furthermore, if our nudging policy is promoted to universities with a smart meter, the cost of saving 1 kWh electricity usage is much less than NT$3.936 (US$0.127).

6 Conclusion

We conducted a non-price-based conservation experiment in the dormitories at NTHU in Taiwan. We provided two types of comparative feedback on students' electricity consumption through emails. One message format is a quintile comparison (*Qt* group), and the other is a mean comparison (*Avg* group). In addition, we also experimented with the times the messages were sent (once or twice).

Our results show that no significant effect on changing overall residents' electricity behavior regardless of the information type and the times of the feedback. The treatment changes the conservation behavior significantly when breaking down the treatment effects by different groups of residents. The feedback is effective only after being provided twice. After providing the feedback twice, the highest 20% pf users reduced their electricity consumption to the message presented in quintile ranking (7.66%–a reduction compared with their Period I average usage); the lowest 20% pf users were responsive to the message presented in comparison with the average (10.99%–a reduction compared with their Period I average usage). The messages are more effective among undergraduate residents compared with graduate students. The back-of-the-envelope calculation suggests that this nudging policy along with the smart meter infrastructure is cost-effective for universities' electricity conservation.

This paper contributes to our existing knowledge of the conservation study in university settings beyond the US context. In addition, we conducted a randomized experiment on the entire university population with more refined message framing. From the policy perspective, our results show that

time)*(3 times).

[13] 2,972 kWh=(0.887kWh)*(1,345 residents)*(2 weeks)+(0.201kWh)*(1,345 residents)*(2 weeks).

the conservation message design should be context-specific. Messages should be designed differently for different target groups (ex: high vs. low users), and should be provided to residents more than once.

One limitation of this study is the limited time period. We conducted the experiment between September and October due to the semester schedule and the electricity usage habit in Taiwan. Therefore, we can only observe an average behavioral change of a short time period rather than a long-term pattern. Further research is needed to trace the dynamics of behavioral change.

References

[1] Allcott, H. (2011). Social Norms and Energy Conservation. *Journal of Public Economics, 95*(9–10), 1082-1095.

[2] Allcott, H. & Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review, 104*(10),3003–37.

[3] Allcott, H. & Kessler, J. B. (2019). The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons. *American Economic Journal: Applied Economics, 11*(1): 236-76.

[4] Allcott, H. & Knittel, C. (2019). Are Consumers Poorly Informed about Fuel Economy? Evidence from Two Experiments. *American Economic Journal, 11*(1), 1–37.

[5] Allcott, H. & Sweeny, R. (2017). The Role of Sales Agents in Information Disclosure: Evidence from a Field Experiment. *Management Science, 63*, 21–39.

[6] Anderson, K., Song, K., Lee, S., Krupka, E., Lee, H., & Park, M. (2017). Longitudinal Analysis of Normative Energy Use Feedback on Dormitory Occupants. *Applied Energy, 189*, 623-639.

[7] Andor, M. A., Gerster, A., Peters, J., Schmidt, C. M. (2020). Social Norms and Energy Conservation Beyond the US. *Journal of Environmental Economics and Management, 103*, 102351.

[8] Asensio, O. I., & Delmas, M. A. (2015). Nonprice Incentives and Energy Conservation. *Proceedings of the National Academy of Sciences, 112*, 510-515.

[9] Asensio, O. I., & Delmas, M. A. (2016). The Dynamics of Behavior Change: Evidence from Energy Conservation. *Journal of Economic Behavior & Organization, 126*, 196-212.

[10] Brewer, R. S., Lee, G. E., & Johnson, P. M. (2011). *The Kukui Cup: A Dorm Energy Competition Focused on Sustainable Behavior Change and Energy Literacy.* 2011 44th Hawaii International Conference on System Sciences, Kauai, HI, 1-10.

[11] Chen, C. W., Chen, J.I., & Lin. M. J. (2021). Performance Feedback in a Group Contest: A Field Experiment on Electricity Conservation. *Journal of Economics & Management Strategy*, 1-25.

[12] Cialdini, R., Kallgren, C., & Reno, R. R. (1991). A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. *Advances in Experimental Social Psychology, 24*, 201-234.

[13] Costa, D. L. & Kahn, M. E. (2013). Energy Conservation 'Nudges' And Environmentalist Ideology: Evidence From A Randomized Residential Electricity Field Experiment. *Journal of the European Economic Association, European Economic Association, 11*(3), 680-702.

[14] Delmas, M. A., & Lessem, N. (2014). Saving Power to Conserve Your Reputation? The Effectiveness of Private Versus Public Information. *Journal of Environmental Economics and Management, 67*(3), 353-370.

[15] Dixon, G. N., Deline, M. B., McComas, K., Chambliss, L., & Hoffmann, M. (2015). Using Comparative Feedback to Influence Workplace Energy Conservation: A Case Study of a University Campaign. *Environment and Behavior, 47*(6), 667–693.

[16] Ferraro, P. J. & Price, M. K. (2013). Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. *The Review of Economics and Statistics, MIT Press, 95*(1), 64-73.

[17] Fischer, C. (2008). Feedback on Household Electricity Consumption: a Tool for Saving Energy? *Energy Efficiency, 1*(1), 79-104.

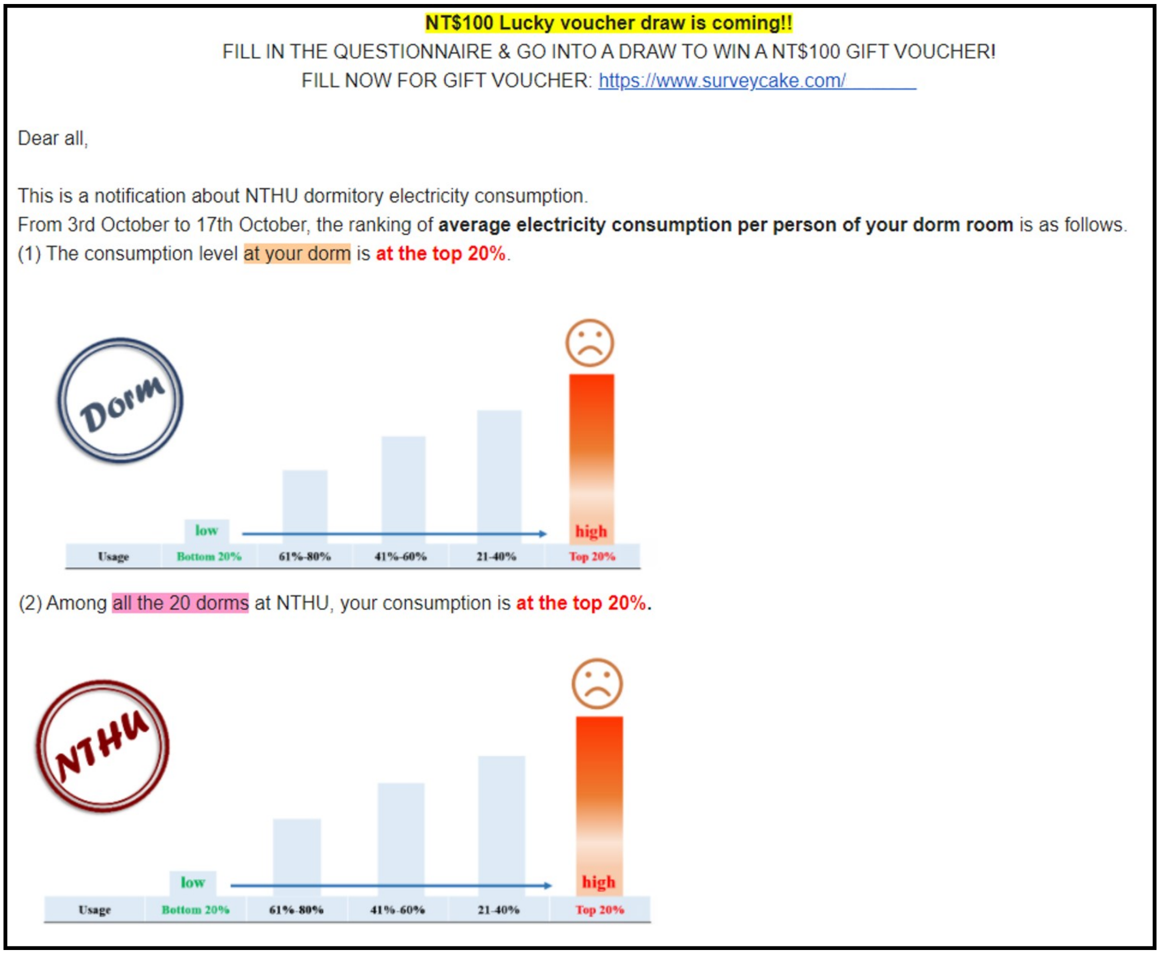
[18] Meub, L., Runst, P. & von der Leyen, K. (2019). *Can APPealing and More Informative Bills 'Nudge' Individuals into Conserving Electricity*? Center for European, Governance and Economic Development Research Discussion Papers 372, University of Goettingen, Department of Economics.

[19] Myers, E. & Souza, M.(2020). Social Comparison Nudges without Monetary Incentives: Evidence from Home Energy Reports. *Journal of Environmental Economics and Management, 101*, 102315, ISSN 0095-0696.

[20] Newell, R.G. & Siikamäki, J. (2014). Nudging Energy Efficiency Behavior: The Role of Information Labels. *Journal of the Association of Environmental and Resource Economists, 1*(4), 555-598.

[21] Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., & Griskevicius, V. (2007). The Constructive, Destructive and Reconstructive Power of Social Norms. *Psychological Science, 18*(5), 429–434.

[22] Sunstein, C. R., & Reisch, L. A. (Eds.) (2017). *The Economics of Nudge. Routledge.* Critical Concepts of Economics, Vol.. 4.



NT$100 Lucky voucher draw is coming!!

FILL IN THE QUESTIONNAIRE & GO INTO A DRAW TO WIN A NT$100 GIFT VOUCHER!

FILL NOW FOR GIFT VOUCHER: https://www.surveycake.com/

Dear all,

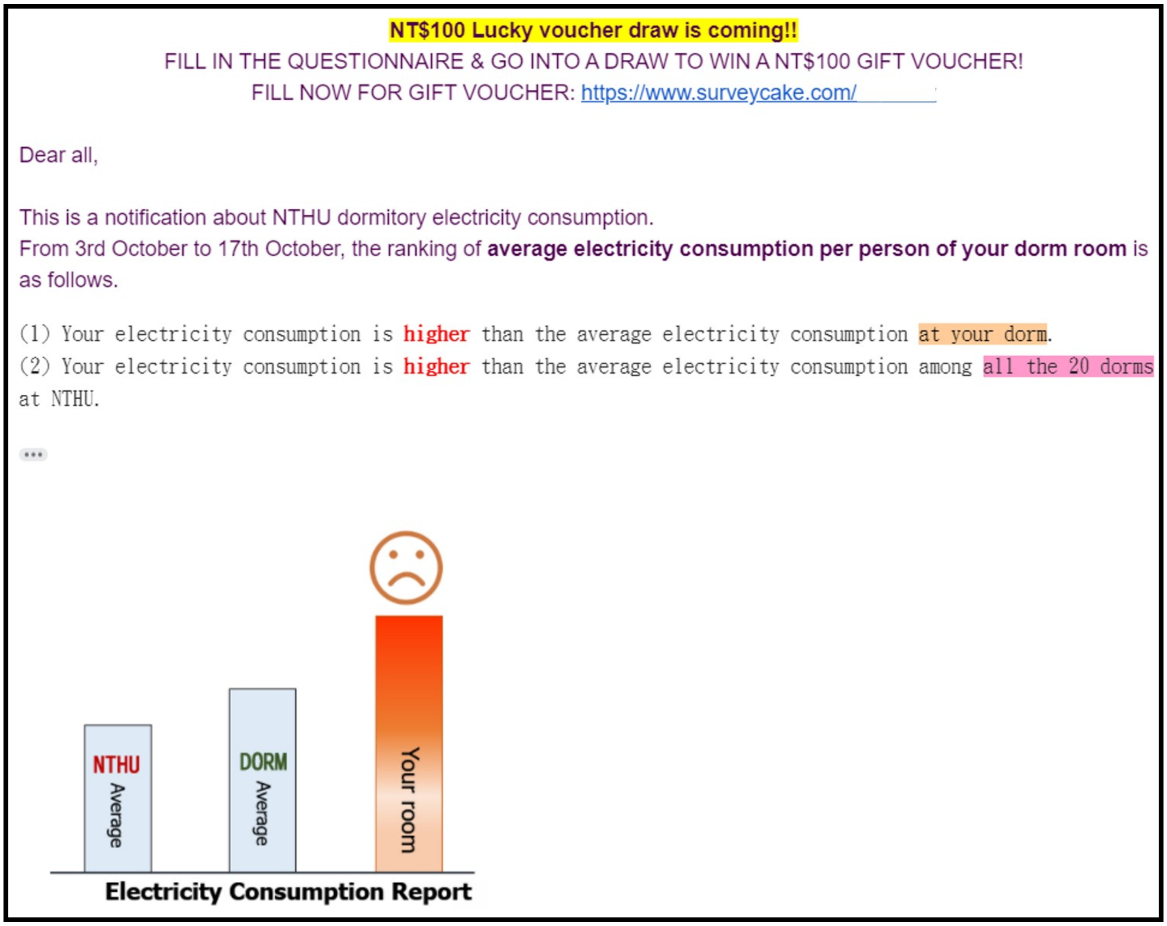
This is a notification about NTHU dormitory electricity consumption.

From 3rd October to 17th October, the ranking of **average electricity consumption per person of your dorm room** is as follows.

(1) The consumption level at your dorm is **at the top 20%**.

(2) Among all the 20 dorms at NTHU, your consumption is **at the top 20%.**

Figure 1: Electricity Feedback – Quintile Information



NT$100 Lucky voucher draw is coming!!

FILL IN THE QUESTIONNAIRE & GO INTO A DRAW TO WIN A NT$100 GIFT VOUCHER!

FILL NOW FOR GIFT VOUCHER: https://www.surveycake.com/

Dear all,

This is a notification about NTHU dormitory electricity consumption.
From 3rd October to 17th October, the ranking of **average electricity consumption per person of your dorm room** is as follows.

(1) Your electricity consumption is **higher** than the average electricity consumption at your dorm.
(2) Your electricity consumption is **higher** than the average electricity consumption among all the 20 dorms at NTHU.

...

Figure 2: Electricity Feedback – Average Information

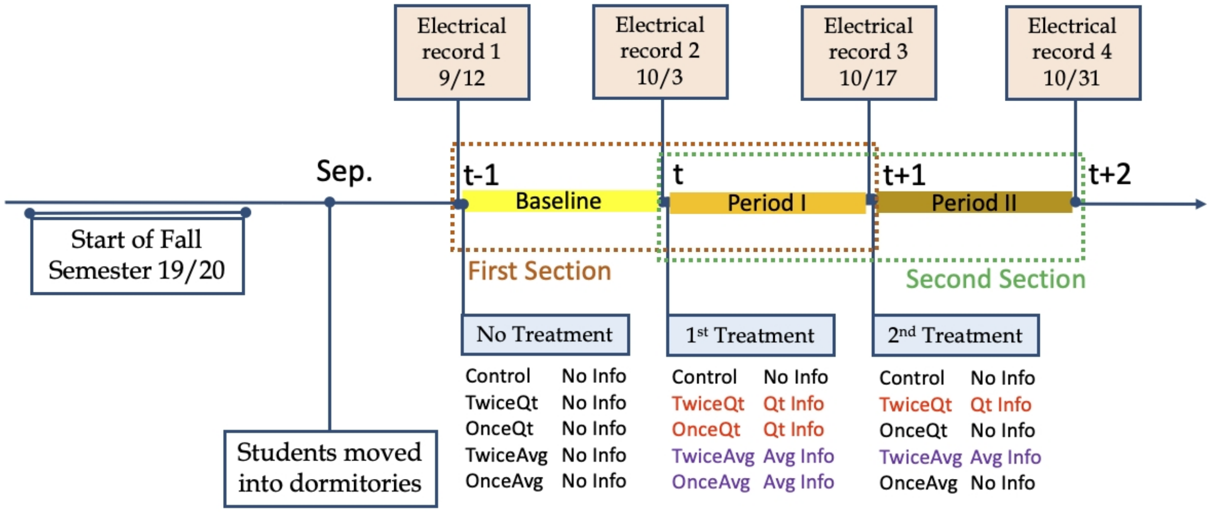


Figure 3: Model Design

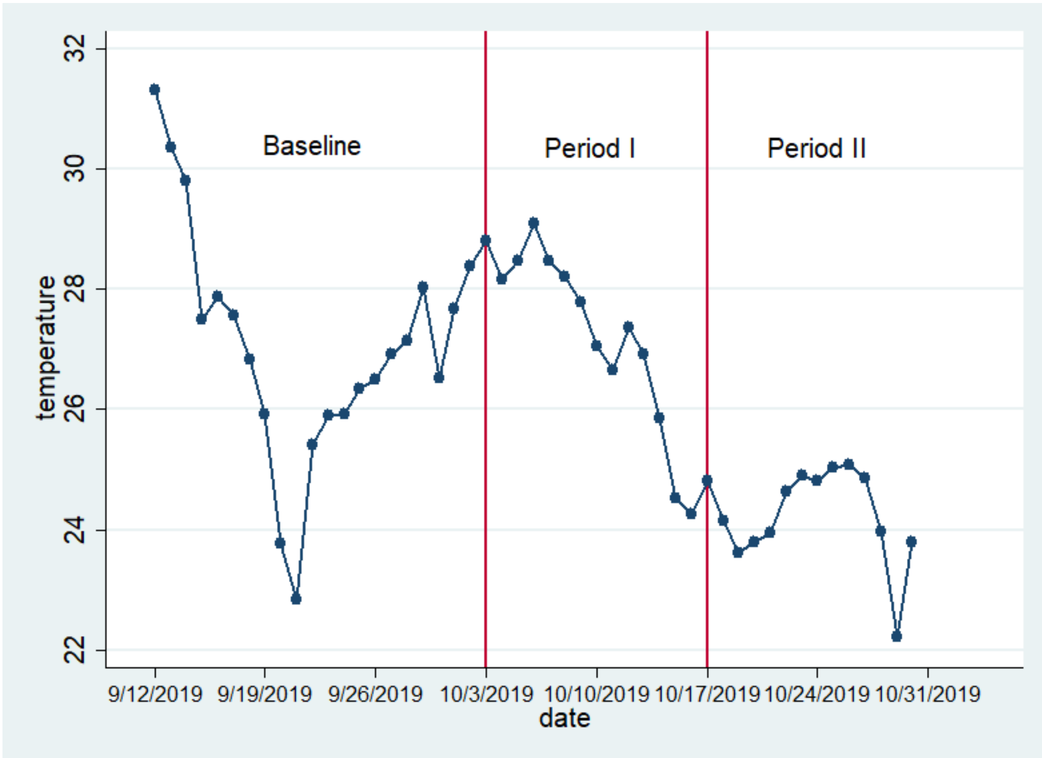


Figure 4: Temperature

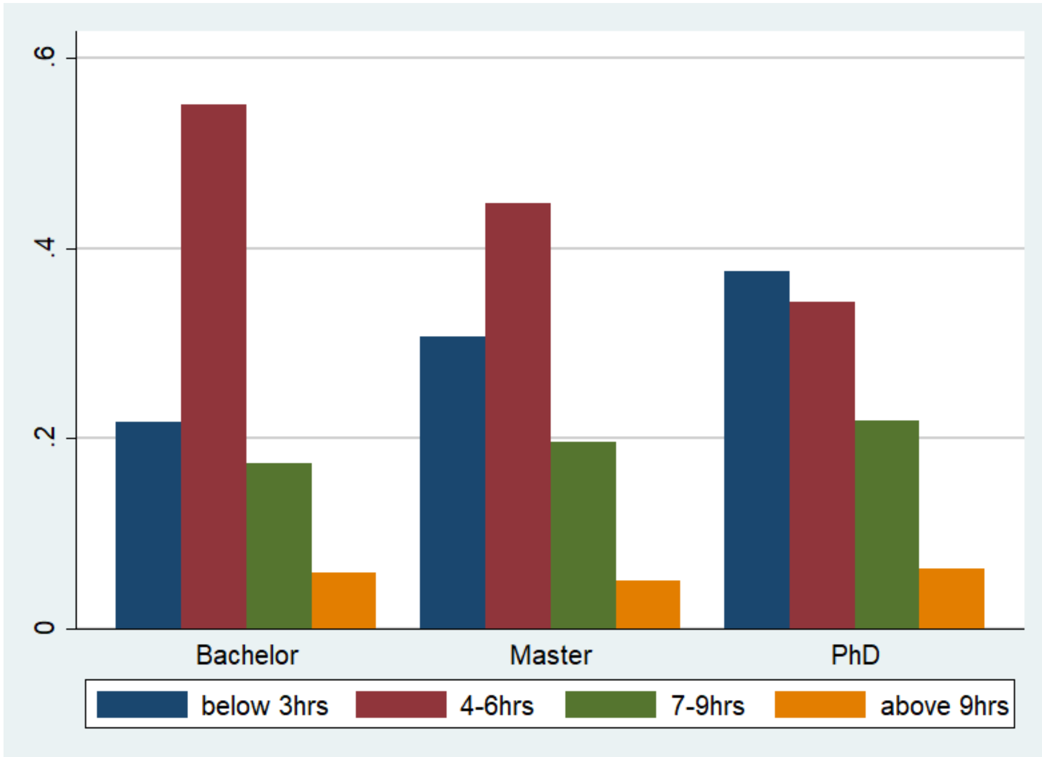


Figure 5: Hours Staying in the Dormitory on Average

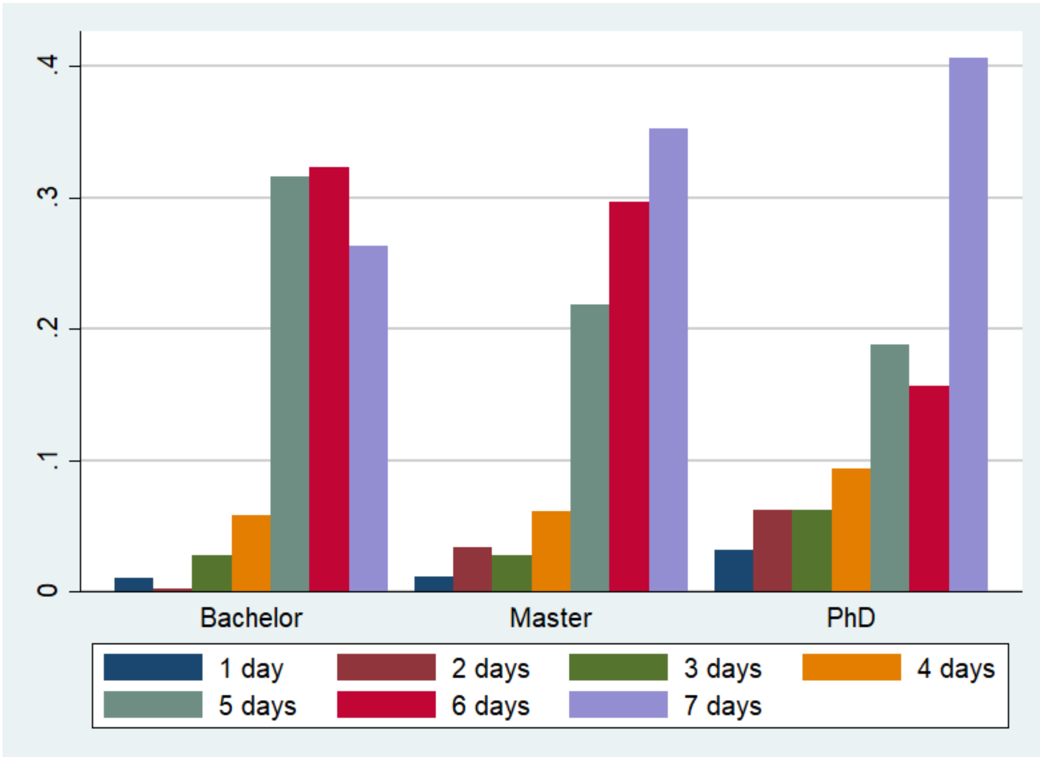


Figure 6: Days Staying in the Dormitory on Average

Table 1: Summary Statistics at the Individual Level by Treatment Status

	All	(1) Control	(2) TwiceQt	(3) OnceQt	(4) TwiceAvg	(5) OnceAvg	Multivariate means test
Variable	N=6,723	N=1,440	N=1,290	N=1,309	N=1,300	N=1,384	Prob>F
Panel A. Electricity Usage							
usage in baseline	5.646	5.697	5.737	5.487	5.783	5.530	0.298
	(4.301)	(4.282)	(4.339)	(4.051)	(4.457)	(4.366)	
usage in period I	6.079	6.179	6.137	5.965	6.119	5.992	0.719
	(4.808)	(4.806)	(4.942)	(4.447)	(4.874)	(4.951)	
usage in period II	2.227	2.347	2.155	2.167	2.325	2.133	0.142
	(2.868)	(3.024)	(2.809)	(2.604)	(3.120)	(2.743)	
Panel B. Characteristics of Rooms							
located building ID	10.564	10.480	10.658	10.568	10.670	10.462	0.857
	(6.172)	(6.105)	(6.209)	(6.195)	(6.222)	(6.143)	
located floor	3.485	3.466	3.458	3.517	3.477	3.507	0.926
	(1.994)	(1.998)	(1.985)	(2.003)	(1.991)	(1.993)	
# studs per room	2.931	2.950	2.930	2.912	2.948	2.915	0.803
	(1.025)	(1.027)	(1.024)	(1.012)	(1.024)	(1.038)	
Panel C. Residents' characteristics							
male	0.582	0.580	0.583	0.581	0.592	0.573	0.899
	(0.493)	(0.494)	(0.493)	(0.494)	(0.492)	(0.495)	
bachelor	0.703	0.722	0.702	0.691	0.694	0.704	0.401
	(0.457)	(0.448)	(0.457)	(0.462)	(0.461)	(0.457)	
master	0.243	0.235	0.242	0.250	0.253	0.236	0.748
	(0.429)	(0.424)	(0.428)	(0.433)	(0.435)	(0.425)	
phd	0.053	0.042	0.055	0.059	0.052	0.059	0.215
	(0.224)	(0.200)	(0.228)	(0.235)	(0.221)	(0.236)	
college	4.319	4.316	4.299	4.300	4.372	4.307	0.932
	(2.379)	(2.381)	(2.387)	(2.420)	(2.420)	(2.293)	

Notes: The usage variables report weekly usage (kWh) per person. Standard deviations are shown in parentheses. Located building ID is the building ID of twenty dormitories. The columns (1), (2), (3), (4) and (5) display the sample means for the five groups. The multivariate means test tests whether the means across our control and four treatment groups are the same.

Table 2: Summary Statistics of Two Online Surveys

	Full Sample	First Survey	Second Survey
Panel A. Objects and Sample			
objects	N=6,723	N=5,283	N=2,590
response sample		N=1,486	N=600
response rate		28.13%	23.17%
error detecting rate		3.8%	13.5%
Panel B. Respondents' Characteristics			
male	0.582	0.489	0.475
	(0.493)	(0.500)	(0.500)
bachelor	0.703	0.688	0.655
	(0.457)	(0.463)	(0.476)
master	0.243	0.264	0.293
	(0.429)	(0.441)	(0.456)
phd	0.053	0.046	0.050
	(0.224)	(0.210)	(0.218)
Panel C. Survey Questions			
intention to reduce usage		0.409	0.393
		(0.492)	(0.489)
care global warm		0.861	
		(0.346)	
hours in the room			5.012
			(2.609)
days in the room			5.685
			(1.232)

Notes: Standard deviations are shown in parentheses. Error detecting rate is a proportion of respondents whose answer of their usage ranking is inconsistent with our record to the whole respondents. "Intention to reduce usage" and "care global warm" are dummy variables, 1 for yes and 0 for not. "Hours in the room" is the average hours that they spend in the dormitory per day after deducting sleep time. "Days in the room" is the average days that they spend in the dormitory per week.

Table 3: Impact of Quintile and Average Information on Usage by Electricity Usage

Panel A. Info: Quintile				Panel B. Info: Average			
	(1)	(2)	(3)		(1)	(2)	(3)
	Whole	Highest 20%	Lowest 20%		Whole	Highest 20%	Lowest 20%
QtTreat×Post1	-0.043	-0.081	-0.137	AvgTreat×Post1	-0.082	0.013	-0.086
	(0.068)	(0.201)	(0.106)		(0.068)	(0.214)	(0.099)
Mean	5.863	11.585	1.829	Mean	5.863	11.585	1.829
N	13,446	2,544	2,596	N	13,446	2,544	2,596
TwiceQt×Post2	-0.184	-0.887**	-0.079	TwiceAvg×Post2	0.065	0.583	-0.201*
	(0.141)	(0.440)	(0.143)		(0.138)	(0.438)	(0.116)
Mean	4.106	8.082	1.211	Mean	4.140	8.590	1.485
N	5,198	992	978	N	5,368	1,008	1,052
Time fixed effects	Yes	Yes	Yes	Time fixed effects	Yes	Yes	Yes
Entity fixed effects	Yes	Yes	Yes	Entity fixed effects	Yes	Yes	Yes

Notes: The outcome variable is the usage (kWh) per person per week during each period. Standard errors in parentheses are clustered at the individual level. * $p < 0.1$, **$p < 0.05$, ***$p<0.01$. Residents are classified into "Highest 20%" and "Lowest 20%" subsample based on their levels of electricity usage in the Baseline. QtTreat includes the OnceQt group and TwiceQt group, the quintile comparison information group. AvgTreat includes the OnceAvg group and TwiceAvg group, the average information group. QtTreat×Post1 and AvgTreat×Post1 report the treatment effects from equation 2. TwiceQt×Post2 and TwiceAvg×Post2 report the treatment effects from equation 3 and equation 4 respectively. "Mean" refers to the mean usage of each sample during each period.

Table 4: Impact of Quintile and Average Information on Usage by Students' Degree

Info: Quintile				Info: Average			
	(1)	(2)	(3)		(1)	(2)	(3)
	Bachelor	Master	PhD		Bachelor	Master	PhD
QtTreat×Post1	-0.121	0.157	0.134	AvgTreat×Post1	-0.093	-0.025	-0.082
	(0.077)	(0.149)	(0.398)		(0.076)	(0.150)	(0.406)
Mean	5.810	6.047	5.731	Mean	5.810	6.047	5.731
N	9,452	3,266	714	N	9,452	3,266	714
TwiceQt×Post2	-0.268*	0.149	-0.619	TwiceAvg×Post2	0.085	0.020	0.157
	(0.161)	(0.313)	(0.695)		(0.153)	(0.315)	(0.711)
Mean	4.169	4.024	3.731	Mean	4.103	4.367	3.576
N	3,620	1,278	296	N	3,752	1,312	298
Time fixed effects	Yes	Yes	Yes	Time fixed effects	Yes	Yes	Yes
Entity fixed effects	Yes	Yes	Yes	Entity fixed effects	Yes	Yes	Yes

Notes: The outcome variable is the usage (kWh) per person per week during each period. Standard errors in parentheses are clustered at the individual level. * $p < 0.1$, **$p < 0.05$, ***$p<0.01$. QtTreat includes the OnceQt group and TwiceQt group, the quintile comparison information group. AvgTreat includes the OnceAvg group and TwiceAvg group, the average information group. QtTreat×Post1 and AvgTreat×Post1 report the treatment effects from equation 2. TwiceQt×Post2 and TwiceAvg×Post2 report the treatment effects from equation 3 and equation 4 respectively. "Mean" refers to the mean usage of each sample during each period.

Table 5: Hours and Days Staying in the Dormitory on Average by Degrees of Programs

	Undergraduates	Graduates	
	Mean	Mean	Diff.
times in the room	5.111	4.799	-0.313*
days in the room	5.687	5.692	0.005
	N=396	N=211	

Table 6: Cost Effectiveness

Grouping Basis	Usage		Degree
Target Group	Highest 20%	Lowest 20%	Bachelor
In our experiment			
Information type	Quintile	Average	Quintile
Second TE (kwh/pre week)	-0.887	-0.201	-0.268
Treated residents	252	246	906
Treated weeks	2	2	2
Total saving kWh	-447.048	-98.892	-485.616
Cost (NT$)	24,000		24,000
Cost-Effectiveness (NT$/kwh)	43.961		49.422
Apply it to whole target sample			
Whole residents	1,345	1,345	4,726
Treated weeks	2	2	2
Total saving kWh	-2386.03	-540.69	-2533.136
Cost (NT$)	11,520		11,520
Cost-Effectiveness (NT$/kwh)	3.936		4.548

Notes: The exchange rate is 1 United States Dollar to 30.925 New Taiwan Dollar, the average exchange rate in 2019.

Appendices

A Online Questionnaires

Q1. Your electricity consumption is

(quintile group)

a. top 20% (higher usage)

b. 21-40%

c. 41-60%

d. 61-80%

e. bottom 20% (lower usage)

(average group)

a. higher than the average dorm consumption

b. lower than the average dorm consumption,

Q2. When you received the electricity consumption message, you will

a. try to reduce the usage

b. live as usual

Q3. Do you care about the issue about climate change or global warming in daily life?

a. Of course! I am very concerned about this kind of information!

b. Oh, I am not very concerned about such issues.

Q4. How long do you spend in the dormitory on average (after deducting sleep time)?

a. less than 3 hours

b. 4 to 6 hours

c. 7 to 9 hours

d. More than 9 hours

Q5. How many days do you spend in the dormitory per week on average?

a. 1 day; b. 2 days; c. 3 days; d. 4 days; e. 5 days; f. 6 days; g. 7 days