Honor, Goal Setting, and Energy Conservation: Evidence from a Field Experiment in Student Dormitories

Qin, Botao and Xie, Siyuan and Xu, Chenyang

1 April 2024

Honor, Goal Setting, and Energy Conservation: Evidence from a Field Experiment in Student Dormitories[a]

Running title: Honor, goal setting, and energy conservation

Botao Qin

Jinhe Center for Economic Research, Xi'an Jiaotong Universiy, China

Siyuan Xie

Jinhe Center for Economic Research, Xi'an Jiaotong Universiy, China

Chenyang Xu

Department of Economics, University of Massachusetts Lowell

[a] *Corresponding author:* Botao Qin, Jinhe Center for Economic Research, Xi'an Jiaotong University, No 28 West Xianning Road, Xi'an, Shaanxi China. 710049. bqinecon@xjtu.edu.cn

HONOR, GOAL SETTING, AND ENERGY CONSERVATION: EVIDENCE FROM A FIELD EXPERIMENT IN STUDENT DORMITORIES

Abstract. Non-monetary incentives are increasingly being studied in encouraging energy conservation. In light of this, we conducted a natural field experiment in student dormitories to assess the effect of honor-based incentives and goal setting on electricity saving and the intrinsic motivation to save energy. Using a difference-in-difference model, we found that goal setting reduced the dormitories' electricity consumption by 15.93% on average compared to the control group. However, the honor-based incentives were not effective on average. In addition, the study found that both honor-based incentives and goal setting, on average, did not crowd out or crowd in the intrinsic motivation to save electricity in dormitories. The heterogeneity analysis showed that the more the dormitory values honor incentives, the more its intrinsic motivation was crowded in by honor incentives. We also found dormitory characteristics affect the crowding effect on the intrinsic motivation.

Keywords: Honor; Goal setting; Electricity use; Crowding effect

JEL classification: C93; D10; Q41

Acknowledgment: Funding for the project has been provided by China's Ministry of Education Humanities and Social Science Research Youth Fund. Grant Number: 19YJC790104.

Conflict of interest: None

1. Introduction

Literature from both economics and psychology suggests that behavioral interventions can be powerful environmental instruments. Non-monetary measures (e.g., public revelation, sending social-comparison messages, giving honors, and goal-setting) are relatively inexpensive compared to, for example, subsidies and do not infringe on people's liberty of choice as much as, for example, taxes. Consequently, both academics and practitioners view behavioral interventions as potential cost-effective complements for the traditional price instruments (Henry et al., 2019; Andor and Fels, 2018; Allcott, 2015; Allcott and Mullainathan, 2010). The literature has found that extrinsic motivators can either enhance or undermine intrinsic motivations, the so-called motivational "crowd-in" or "crowd-out" effects. However, the crowding effects are primarily examined when the extrinsic motivators are monetary instruments (see Rode et al. (2015) for a review). Less attention is paid to the crowding effects of non-monetary interventions. How does the intrinsic motivation and non-monetary incentives interact? This question is important, as the literature found that the crowding effects could be so large as to render the traditional price incentives counter-productive (Cardenas et al., 2000; Velez et al., 2010).

This paper takes up the tasks of examining both *the effectiveness* and *the crowding effects* of two types of non-monetary incentives in society—Honor and Goal setting—in inducing energy conservation. Honor is a widely used incentive tool in many societies. For example, most countries do not pay blood donors for their blood donations because doing so would crowd out the civic spirit of voluntary blood donations (Titmuss, 1970). Rather, it is common practice to award blood donors with blood donation certificates or symbolic souvenirs. Similarly, in China, the government grants "Civilized City" titles to cities that perform well in economy, environmental performance, and governance. How effective, though, are honor-based incentives in environmental protection?

And what impact do honor-based incentives have on people's intrinsic motivation to protect the environment? These are open questions that require more research.

Goal setting is often used as a commitment device to overcome present bias — people tend to delay the present tasks later. Goal setting could work because of reference-dependent preference: individuals evaluate their success or failure based on a predetermined level. Setting a goal increases pressure by conditioning satisfaction on a desired level of performance. Some studies have found that setting a goal can affect households' electricity consumption behavior. Harding and Hsiaw (2014), for instance, studied an energy-saving project in Illinois, USA, and found that residents who chose to join the goal-setting program saved an average of 4% of their electricity use. The savings were heterogeneous, with households that set realistic goals save more electricity than others. However, few studies we are aware of examine the crowding effects of goal setting. Would setting goals crowd in or crowd out people's intrinsic motivations?

To identify the effects of these two types of incentives, we designed a field experiment at Xian Jiaotong University in China in the context of energy savings in student dorms. Students' energy consumption in tertiary institutions accounts for a large portion of national energy use (Zhou et al., 2021; Ding et al., 2018). Of all energy consumption in university buildings, dormitory electricity use accounts for a significant proportion (Zhou et al., 2021). In addition, Chinese schools frequently use honor incentives to encourage students to achieve better grades and participate in extra-curricular activities. Goal setting is also often used in Chinese society. The central government sets goals in well-publicized Five-Year Plans. Parents and teachers frequently set goals for students in their semester plans. These nonmonetary incentives are thus not unfamiliar to students and have the potential to be applied in energy conservation.

The experiment selected two treatment groups—one of which was asked to set goals and the other was presented with honor-based incentives—and a control group. We use a three-stage design (with a pre-intervention, intervention, and post-intervention stage). This design allows us to separate the crowding effects from the effects of the nonmonetary incentives. For example, even if one finds lower energy consumption under interventions, this could reflect the effects of nonmonetary incentives or the crowd-in effects. If energy consumption is lower or higher under treatments than the control post-intervention after accounting for the difference of extractions in the pre-intervention stage, then we have evidence that the crowding effects are present.

The experiment runs for 12 weeks. We then employed the difference-in-difference (DID) identification strategy and found that compared to the control group, the goal-setting groups reduced their average electricity consumption by 15.93%. However, the effect of the honor-based incentives is not significant. We did not observe either crowd-in or crowd-out effects on the intrinsic motivation of dormitory energy conservation.

Our paper differs from previous literature in nudging energy conservation in two ways. First, ours are among the first studies that examine honor-based incentives and goal setting on people's *intrinsic motivation* to conserve energy. Secondly, although previous literature such as Harding and Hsiaw (2014), Lazaric and Toumi (2022), and Brandsma and Blasch (2019) has tested the effect of goal setting on energy conservation, the effect of honor-based incentives on energy conservation was rarely studied with experiments.

The next section reviews the literature and formulates the hypotheses. Section 3 describes the experimental design. The descriptive analysis and regression results follow in sections 4 and 5. Section 6 concludes.

2. **Literature review and hypotheses formulation**

People respond to non-monetary incentives, such as honor-based incentives, because they want to be perceived as civic-minded and responsible citizens (Benabou and Tirole, 2003; Bénabou and

Tirole, 2006). Ariely et al. (2009), for instance, found that people behaved prosocially in a charity-giving field experiment to boost their public image. Similarly, Delmas and Lessem (2014) found that public information about dormitories' conservation rating combined with private information about electricity use and social norms reduced electricity consumption by 20% but that private information alone did not work. Several other studies found that people engage in more environmentally conscious behavior when their actions are being observed by others (Barclay and Barker, 2020; Griskevicius et al., 2010; Sexton and Sexton, 2014; Ernest-Jones et al., 2011). In addition, Kraft-Todd et al. (2015) reviewed evidence in the field and found that social interventions based on observability are highly effective in promoting cooperation. We can speculate, then, that honor-based incentives can reduce residents' electricity consumption, which leads us to propose our first hypothesis:

Hypothesis 1: Honor-based incentives can reduce the electricity consumption of dormitories.

Goal setting is commonly used as a commitment device to encourage energy saving. The evidence on the effectiveness of commitment devices, however, is somewhat mixed. Review studies, such as those of Abrahamse and Steg (2013), Nisa et al. (2019), and Lokhorst et al. (2013), found commitment to be effective in promoting energy saving. However, Vesely et al. (2022) reviewed a wide range of field experiments and concluded that the effects of commitment-based interventions in energy conservation are almost zero. In addition, evidence from Harding and Hsiaw (2014), Lazaric and Toumi (2022), Liu et al. (2021), and Brandsma and Blasch (2019) suggests that setting a goal, especially a realistic goal, can reduce the electricity consumption of residents. Ishimura et al. (2024) has tested the effect of regional goal setting for waste reduction in Japan. They found it was effective and reduced waste output by 3.38 kg per capita per year. Therefore, we propose our second hypothesis:

Hypothesis 2: Goal setting can reduce electricity consumption in dormitories.

Previous studies have found that monetary incentives may crowd out people's intrinsic motivation (see Fehr and Falk (2002), Frey and Jegen (2001), and Bowles (2008) for a literature review). By contrast, non-monetary incentives, such as verbal and positive feedback, can crowd in people's intrinsic motivation. In addition, when incentives are perceived as supportive rather than controlling, they will crowd in people's intrinsic motivations (Deci, 1971; Deci et al., 1999; Lepper et al., 1973). The key is that these incentives provide individuals with the opportunity to satisfy their competency and self-determining needs (Hirst, 1988). Honor-based incentives provide positive feedback and are generally viewed as supportive instruments. Therefore, we propose that

Hypothesis 3: Honor-based incentives may crowd in the intrinsic motivation to conserve energy in dormitories.

The literature has shown that the effects of externally imposed goals on intrinsic motivation depend on the interests of the tasks. Assigning specific goals could undermine the intrinsic motivations of engaging in interesting tasks while enhancing the intrinsic motivations of undertaking boring tasks (Mossholder, 1980; Carroll Jr and Tosi, 1970; Locke and Bryan, 1967). Students may or may not view energy conservation as a boring task. This may depend on their environmental attitude. Thus we hypothesize that:

Hypothesis 4: Goal setting may crowd out or crowd in the intrinsic motivation to conserve energy in dormitories.

3. Design of experiments

We ran the experiment at the campus of Xi'an Jiaotong University, Xi'an, China. We targeted four dormitory buildings since they are the only buildings from which we can obtain electricity use data on both air-conditioners and lighting. One dormitory building was excluded due to a high resident turnover rate and too few residents. Among the three dormitory buildings (East-12,

Table I. Sample description

Group	Dormitory	Number of rooms
T1	East-12	77
T2	East-12	77
C	East-14	26
C	West-9	51

East-14, and West-9), we chose one (East-12) as the treatment group. The appendix shows the pictures of these buildings. The choice of treatment group is not random. East-12 building was chosen as the treatment group because it had enough dorms (154) so that we could have a large sample size for each treatment. Out of the 154 dorms, 40 have 2 students per dormitory and 1 dormitory has 1 student. The rest has 3 students per dorm. The other two buildings (East-14 and West-9) only have 26 and 51 dorms respectively and serve as control groups. All West-9 dorms have 3 students and all East-14 have 2 students. All student residents in these buildings are male. East-12 building includes only undergraduate students, whereas East-14 has only Ph.D. students and West-9 only Master's students. We realize this systematically different characteristic of the treatment and control groups may hinder our ability to estimate the causal impacts. However, as discussed later, we employ a difference-in-difference (DID) estimator, which gives us causal impacts as long as the parallel-trend assumption holds, i.e., it does not require the groups to be otherwise completely similar in properties as long as they show common trends (Fröhlich et al., 2019; Angrist and Pischke, 2014). The sample description is summarized in table I.

The dormitories comprising the treatment group were randomly divided into the honor-based incentive and the goal-setting group (77 dorms in each group). We received permission from the university's logistics energy department to conduct the experiment and obtained dormitory electricity consumption data from them. The experiment lasted 12 weeks and was divided into three stages (4 weeks for each stage). The first stage (the pre-treatment stage) ran from October 1st to October 31st and there was no intervention.

The second stage (the treatment stage) ran from November 1st to December 1st, during which the dormitories received their respective treatments. At the beginning of this stage, we announced a "Four-week Dormitory Energy Saving" activity for the dorms in the treatment building. We placed poster boards at the entrance of the building. On October 30th and 31st, we knocked door by door to inform the students of the activity and distribute different energy-saving leaflets to them. The contents of the flyer for the honor-based incentives group and the goal-setting group were different, as shown in the appendix. For the honor-based incentives group, the experimenter informed each dormitory that, after the event, the top 20% of dormitories with outstanding electricity-saving performance would be issued honorary certificates and that this outcome would be publicized. For the goal-setting group, the experimenter set the electricity-saving target at 15% of each dormitory's total electricity consumption in the first four weeks. The choice of this target was based on the research of Harding and Hsiaw (2014) and Lazaric and Toumi (2022). They suggest goals that are too high can be difficult to achieve, resulting in dormitories lacking sufficient motivation to save electricity; while targets that are too low will be too easy to reach and therefore not help save electricity. The experimenter then informed all dormitories of their first-stage electricity consumption. Experimenters also gave weekly feedback on the electricity consumption of each dormitory in the goal-setting group so that it could be compared with the electricity consumption in the first stage. At the end of the second stage, the experimenters informed each dormitory whether it had achieved the goal of saving electricity, with the dormitories that achieved their goals being awarded a USB flash drive engraved with "energy-saving dormitory".

The last stage (the post-treatment stage) ran from February 13th to March 14th. Due to the COVID-19 pandemic, students returned to their hometowns earlier than anticipated after the second stage, meaning the electricity consumption of the dormitories was close to zero. We therefore chose the new semester after students returned to campus as the third stage. In this stage, we stopped the respective treatments. By employing this three-stage design, we were able to identify the effect of the corresponding treatment on the intrinsic motivation of dormitory electricity consumption after the intervention was removed.

At the end of the third stage, the experimenters obtained weekly electricity consumption data from the three groups through the logistics energy department. In addition, the experimenters distributed corresponding questionnaires to each dormitory to collect data on their electricity use habits and electrical appliance usage. The timeline of the experiment is shown in the following table.

Table II. Experimental design

Group number	Intervention	First stage	Second stage	Third stage
T1	Honor	Week 1-4	Week 5-8	Week 9-12
T2	Goal setting	Week 1-4	Week 5-8	Week 9-12
C	None	Week 1-4	Week 5-8	Week 9-12

4. **Descriptive analysis**

Table III lists the electricity consumption for each week of the first stage for the control and treatment groups, including the mean, median, and standard deviation, respectively. Table IV reports the results of pairwise comparisons between groups. We can see that there are no significant differences between the groups' electricity consumption in the pre-intervention stage. In addition, we collected the dormitories' characteristics such as the number of residents, high-power electrical appliances, computers, and tablets, attitudes toward goals and feedback, the degree of importance attached to honor, and the electricity use habits. Table V shows the summary statistics of the characteristics. The variable AC days per month means the number of days the dormitory is getting used to using the air-conditioner. For AC habits and light habits, 1 means turning off the air conditioner at night when falling asleep and turning off the light when nobody is in the dormitory. We did a balance check of the characteristics among the treatment groups and the control group. We find there is no significant difference in electricity use habits such as turning off the air conditioner at night when falling asleep and turning off the light when nobody is in the dormitory. However, dormitories in the control group reported significantly fewer days of using air-conditioners per month. The dormitories are similar in the number of high-power electrical appliances as each dormitory is equipped with an air-conditioner, but the goal-setting group has relatively more computers and tablets than the control group. In addition, the treatment groups have relatively more residents than the control group. There is no significant difference between the two treatment groups in any of the dormitory characteristics, which confirms the randomization between the two groups. We later use a DID estimation model to identify the average treatment effect. It can allow for the pre-existing difference in characteristics between the treatment groups and the control group as long as the parallel trend assumption is met.

Table III. Weekly mean and median electricity consumption in the first stage (in kWh)

Group number	Week 1	Week 2	Week 3	Week 4
T1 (n=77)	23.1	19.6	17.5	16.2
	18.4	15.6	16.8	14.4
T2 (n=77)	23.9	20.6	18.5	17.1
	20.0	16.1	15.8	13.7
C (n=77)	23.0	19.5	17.9	16.2
	18.6	16.9	17.5	15.8

Notes: For each group, the upper and lower numbers are mean and median, respectively.

Table IV. Mann-Whitney test results

	First week	Second week	Third week	Fourth week
T1 and C	0.937	0.895	0.961	0.490
T2 and C	0.518	0.643	0.987	0.702
T1 and T2	0.531	0.450	0.460	0.367

Notes: The numbers in the table are p-value.

Figure 1 shows the changes in electricity consumption for all three groups at each stage. We can see that the electricity consumption trend of each group in the first stage was similar: as

Table V. Summary statistics and balance check of dormitory characteristics

	Honor (1)	Goal setting (2)	Control (3)	Balance check (p-value) (1) vs. (2)	(1) vs. (3)	(2) vs. (3)
High-power appliance	2.195	2.377	2.015	0.501	0.521	0.173
	(0.200)	(0.181)	(0.191)			
Computers and Tablets	3.961	4.208	3.803	0.130	0.383	0.013
	(0.126)	(0.102)	(0.128)			
AC days per month	11.909	12.091	7.879	0.861	0.000	0.000
	(0.769)	(0.694)	(0.804)			
AC habits	0.662	0.675	0.621	0.865	0.612	0.502
	(0.054)	(0.054)	(0.060)			
Light habits	0.636	0.688	0.727	0.499	0.249	0.613
	(0.055)	(0.053)	(0.055)			
Residents number per room	2.870	2.818	2.667	0.400	0.005	0.038
	(0.043)	(0.044)	(0.058)			
Number of dormitories	77	77	77			

Notes: In columns (1), (2), and (3), the upper and lower numbers are mean and standard deviation, respectively. The variable AC days per month means the number of days the dormitory is getting used to using the air-conditioner. For AC habits and light habits, 1 means turning off the air conditioner at night when falling asleep and turning off the light when nobody is in the dormitory.

temperature increased after the first week, the electricity consumption of all the groups declined. Then, in weeks 5 and 6 of the second stage, the electricity consumption of the control group began to rise, but the electricity consumption of the honor-based incentives group and goal-setting group continued to decline. Relative to the control group, the honor-based incentives and goal-setting groups both saw a decrease in electricity consumption, but by weeks 7 and 8, the trend of electricity consumption in each group became similar. Finally, in the third stage, the trend of electricity consumption across all the groups was similar, which suggests that, in the early stage of treatment, the treatments reduced energy consumption. However, this effect may not persist.

5. **Empirical model**

5.1. *Parallel trend test*

The identifying assumption for the DID method is that the counterfactual electricity consumption in the treatments and the baseline exhibit parallel trends. One often tests the assumption of parallel trends using a test of "pre-trends," i.e., the parallel trends before the treatment.

We adopt the approach implemented by both Kearney and Levine (2015) and Autor (2003). The approach involves regressing the outcome variable on the interaction between the treatment variable and the time dummies

$$Y_{it} = \beta_0 + \sum_{j=2}^{4} \beta_{j1} * D_i * pre_{-j} + \epsilon_{ij} \tag{1}$$

where Y_{it} is the electricity use for dormitory i at period t. pre_{-j} is the time dummy for the jth period before the treatment occurred, and pre_{-1} is the base period. D_i is a dummy variable for the treatment: if dormitory i is treated, the variable is 1, otherwise, it is 0. ϵ_{it} is the error term. The coefficient β_{j1} on the interaction term $D_i * pre_{-j}$ reflects the pre-intervention difference between the treatment group and the control group. If β_{j1} is not significantly different from 0, it implies that parallel trends hold for the pre-intervention stage. Fig 2 shows the regression coefficient and 95% confidence interval of the parallel trend test. We can see that there was no significant difference in the weekly electricity consumption across all three groups pre-treatment.

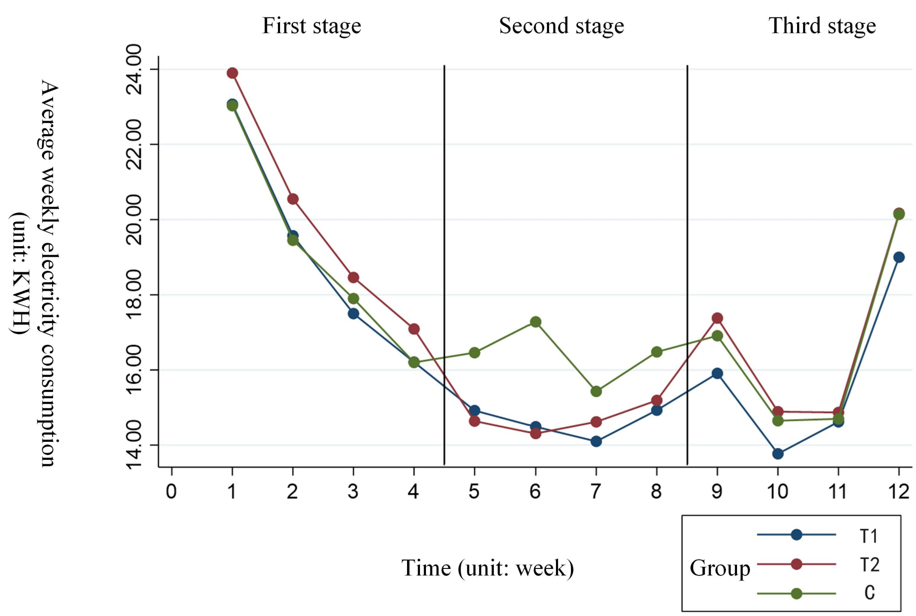

FIGURE 1. Average weekly electricity consumption of the treatment and control groups

5.2. *Average treatment and crowding effects*

We employ difference-in-differences (DID) estimators to identify the average treatment effect. The DID method allows us to control for observed and unobserved time-varying factors common to all groups and time-invariant characteristics that might be correlated to the treatments. This implies that the systematic difference in student body characteristics between the treatment and control groups will be controlled for. Our counterfactual is the amount of electricity that would have been consumed in the treatment group without the treatment. As shown before, we have evidence that this counterfactual follows the same trend as the control group during the treatment period.

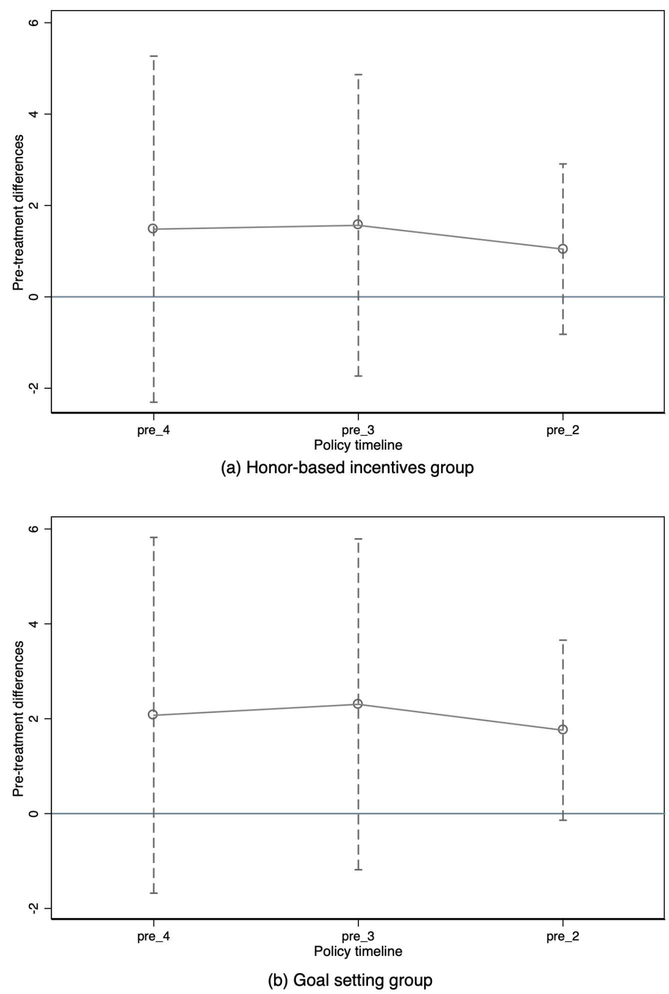

FIGURE 2. Parallel trend test coefficient and confidence interval

We used the following DID model to analyze the treatment effect.

$$Y_{it} = \beta_0 + \beta_1 \boldsymbol{D_i} + \beta_2 S_2 + \beta_3 S_3 + \beta_4 \boldsymbol{D_i} S_2 + \beta_5 \boldsymbol{D_i} S_3 + v_i + \epsilon_{it} \tag{2}$$

where Y_{it} is the electricity use for dormitory i at period t. $\mathbf{D}$ is a vector of dummy variables that indicate whether the dormitory is in one of the treatment groups or not. S_2 and S_3 indicate whether the dormitory is in the second (treatment stage) or third stage (post-treatment stage). v_i captures the dormitory fixed effect. ϵ_{it} is the error term. β_4 captures the treatment effect and β_5 captures the crowding effect. The estimated results of the DID model with both random effect and fixed effect are shown in Table VI.

Table VI. Estimates of the DID model

	(1) Random effect	(2) Fixed effect
$S_2 D_{honor}$	-1.750*	-1.750*
	0.949	(0.949)
$S_2 D_{goal}$	-2.580***	-2.580***
	0.974	(0.974)
$S_3 D_{honor}$	-0.443	-0.443
	1.138	(1.137)
$S_3 D_{goal}$	-0.811	-0.811
	1.230	(1.229)
S_2	-2.729***	-2.729***
	0.727	(0.727)
S_3	-2.287***	-2.287***
	0.804	(0.803)
D_{honor}	-0.055	
	1.655	
D_{goal}	0.858	
	1.589	
Constant	19.142***	19.41***
	1.042	(0.321)
Number of observations	3696	3696
R-squared	0.039	0.039
Number of dormitory	231	231

Notes: Robust standard errors are in parentheses. $*p < 0.1 **p < 0.05 ***p < 0.01$

We perform a Hausman test and the results fail to reject the null that there are no systematic differences between the two models' estimates. Using the fixed effect model, we see the honor-based incentives reduced dormitories' electricity use by 1.75 kWh on average but the effect is only significant at the 10% level. The goal setting reduced dormitories' electricity use by 2.580 kWh on average at the 1% significance level. The average reduction of the energy consumption for the honor-based incentives and goal-setting are 10.80% and 15.93% respectively compared to the baseline level. Thus our evidence supports hypothesis 2 that goal setting can reduce dormitory electricity consumption. However, honor-based incentives do not seem to have significant impacts and we reject hypothesis 1. Comparatively, Allcott (2011) found that regularly sending home energy reports with social comparison information helped to reduce household electricity use by an average of 2.0%. Moreover, Magali and Doctori (2010) conducted a meta-analysis to assess the impact of different types of information interventions on energy-saving behavior. They found that the average treatment effect of information-based strategies was 7.4%. Although the goal setting in this study exhibits a greater treatment effect on electricity saving than that seen in previous studies, we caution that the intervention only lasts for one month and is shorter than previous studies.

The coefficients of S_3D_{honor} and S_3D_{goal} which capture the crowding effect of the two interventions are negative but insignificant. Therefore, there seems to be no crowding effect for both honor-based and goal-setting incentives. We thus reject hypotheses 3 and 4.

5.3. *Heterogeneous treatment and crowding effects*

Although the average treatment effect of honor-based incentives is not significant, this effect might be significant in certain subpopulations. We thus examine the impacts of dormitory characteristics and electricity use habits on the average treatment effects. We collected the data on dormitory characteristics and electricity use habits of each dormitory through an online questionnaire post-experiment. In addition, for the dormitories of the honor incentive group and the control group, we distributed a questionnaire about the attitude towards honor; for dormitories in the goal-setting treatment, we distributed questionnaires about their views on setting goals. These questionnaires all included 5 questions, each with three options (negative, moderate, and positive, to which we assigned scores 0, 1, and 2 respectively). The higher the score, the more honor and goal-setting the dormitory values.

To explore the heterogeneous treatment effect, we modified the previous DID model to include variables such as dormitory characteristics and students' electricity use habits

$$Y_{it} = \beta_0 + \beta_1 D_i + \beta_2 S_i + \beta_3 \boldsymbol{X} + \beta_4 D_i S_i + \beta_5 D_i \boldsymbol{X} + \beta_6 S_i \boldsymbol{X} + \beta_7 D_i S_i \boldsymbol{X} + v_i + \epsilon_{it} \quad (3)$$

where $\boldsymbol{X}$ is the vector variable of dormitory characteristics such as the number of permanent residents, the number of high-power electrical appliances, the number of computers and tablets, attitudes toward goals and feedback, the degree of importance attached to honor, and the electricity use habits. All these variables have been demeaned to mitigate the concern of multicollinearity due to the presence of multiple interaction terms (Iacobucci et al., 2016). Coefficient β_7 measures the effects of the dormitory characteristics on the treatment effect and the crowding effect. Table VII shows the results of the influence of each characteristic on the treatment effect and the crowding effect.

We first examine the heterogeneous treatment effect by dormitory characteristics. In table VII, the first row of each model shows the treatment effects. Models 1 and 2 show that the more high-power electrical appliances, computers, and tablets the dormitories own, the larger the energy-saving effect from the two treatments. This makes intuitive sense since the more appliances students have, the more room for energy conservation. Models 3 to 5 show the heterogeneous treatment effect of the dormitories' electricity-using habits. We see that the more days the dormitories use an air-conditioner and the more likely they turn off the air-conditioner after falling asleep, the larger the energy-saving effect from the two treatments. These findings also make sense since if students didn't use air-conditioners that often, there was less room for energy conservation. In addition, students who turned off air-conditioners after falling asleep might be already environmentally concerned and thus were more responsive to the cue provided by the incentives. The temperature and the habit of turning the light off when nobody is present have no impact on the treatment effect. This might also make sense since light is not as energy-consuming as air conditioners. Model 7 shows the more the dormitory values honor, the better the electricity-saving effect of honor incentives.

Next, we examine the heterogeneous crowding effect. Although section 5.2 shows an insignificant crowding effect in general, the crowding effects might still exist within certain sub-populations. In table VII, the second row of each model illustrates the impacts of different characteristics on the crowding effects. In model 1, the coefficient of the triple interaction term in the goal-setting group is significantly negative, implying the more high-power-electrical appliances in the dormitory, the more likely there is a crowding-in effect due to the goal-setting incentives. This might be because more appliances are more likely to help students form energy conservation habits. Therefore, even after the incentives are ceased, students with more appliances are still more likely to conserve energy. Model 2 shows the more computers and tablets in the dormitory, the more likely there is a crowd-in effect due to both forms of incentives. Models 3 and 5 show the more days dormitories use air conditioners, or the dormitories with the habit of turning off the air conditioner at night, the students' intrinsic motivations were more likely to be crowded in by both forms of incentives. As mentioned before, previous literature found that the crowd-in effects of goal-setting are more likely to happen when the tasks are not boring. If the students already have the habit of turning off the air conditioner at night, this might not be a boring task to complete and hence the possibility of crowd-in effects. Models 4 and 6 found that whether lights were turned off when no one was present and the temperature had no effect on the intrinsic motivation of energy saving in dormitories. Model 7 shows the more students value honor, the stronger the crowd-in effects of honor incentives.

TABLE VII. Heterogeneous treatment and crowding effects by dormitory characteristics

	Honor-based incentives group	Goal setting group
Model 1		
$Numberapp_i S_2 D$	-1.893***	-2.312***
	(0.732)	0.756
$Numberapp_i S_3 D$	-0.989	-2.383***
	0.753	0.739
Model 2		
$Computer_i S_2 D$	-1.989**	-1.674*
	(0.891)	(0.866)
$Computer_i S_3 D$	-2.192**	-3.782***
	1.023	1.052
Model 3		
$Airconditioner_i S_2 D$	-3.968*	-9.269***
	(2.254)	(2.115)
$Airconditioner_i S_3 D$	-6.703**	-9.101***
	(2.649)	(2.558)
Model 4		
$lightoff_i S_2 D$	1.451	-1.509
	(2.804)	(2.818)
$lightoff_i S_3 D$	-2.899	-3.267
	(2.750)	(2.918)
Model 5		
$ACoff_i S_2 D$	-0.784***	-0.860***
	(0.187)	(0.201)
$ACoff_i S_3 D$	-0.450**	-0.753***
	(0.189)	(0.185)
Model 6		
$Temperature_i S_2 D$	-0.173	-0.273
	(0.719)	(0.701)
$Temperature_i S_3 D$	-0.0320	-0.163
	(0.705)	(0.697)
Model 7		
$Honor_i S_2 D$	-2.245***	
	0.294	
$Honor_i S_3 D$	-2.351***	
	(0.398)	
Observations	2288	2288

Notes: Robust standard errors are in parentheses. $*p < 0.1 **p < 0.05 ***p < 0.01$

5.4. *Robustness test*

We used a double-randomized placebo test to determine whether unobserved features affected the results. We randomly sampled 3000 times for the treatment stage S_2 and the interaction item $\boldsymbol{D_i} S_2$ for the treatment stage, respectively. The kernel densities of the estimated coefficients for the two treatments are shown in Figure 3. We can see that the estimated coefficients followed normal distributions and had a mean of 0. Therefore, the placebo test was valid.

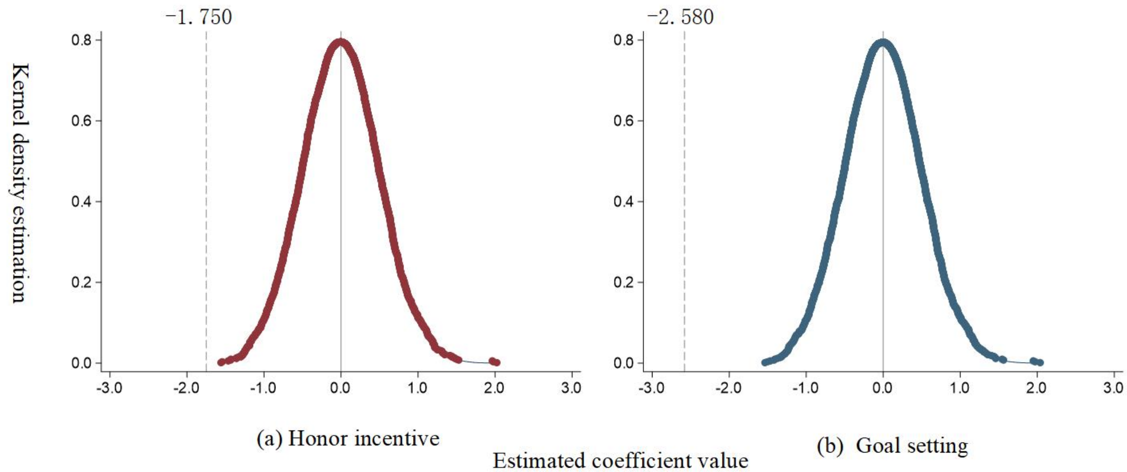

FIGURE 3. Estimation coefficient kernel density distribution

6. **Conclusion and policy implications**

Although price instruments in encouraging energy savings are useful policy tools, they also suffer from certain limitations, such as the political difficulty of raising electricity prices and the possible crowd-out effect on people's intrinsic motivation. In contrast, non-monetary incentives, such as sending out social comparison messages, providing information feedback, and setting goals, have the promise to be both effective and politically palatable. This study designed a natural field experiment in the context of energy saving in student dormitories to assess the treatment and crowding effects of two commonly used non-monetary instruments — honor-based incentives and goal setting — on dormitory electricity conservation. By estimating a DID model, we found that goal setting significantly reduced electricity consumption in dormitories and the average magnitude was 15.93% compared to the control group. Honor-based incentives, however, are shown to be not effective in encouraging energy saving. This implies that giving honors alone may not provide enough incentives to encourage energy savings. The study also found that both honor-based incentives and goal setting in general do not exhibit crowding effects on students' intrinsic motivation to conserve energy. The heterogeneity analysis found that the more the dormitory values honor incentives, the more its intrinsic motivation was crowded in by honor incentives. In addition, the more the number of computers and tablets in the dormitory, the more days air conditioners are used, or the habit of turning off the air conditioner at night, the more likely there existed crowd-in effects of both forms of incentives. Finally, the more high-power-electrical appliances in the dormitory, the more likely there existed crowd-in effects of the goal setting.

The findings of this study support the use of goal setting combined with electricity use feedback as an instrument for electricity conservation for policy makers. Policy makers can invite participants to set a moderately electricity-saving goal such as 15% and then send them electricity use feedback periodically to let them know the progress towards the goal. A symbolic award for those who achieve the goal is also helpful for the success of the goal setting. As for the honor-based incentives, policy makers need to target those specific subgroups such as those households that value honor more, have more high-powered electrical appliances, or have the habit of turning off the air-conditioner after falling asleep. Although we do not find a crowd-out/in effect on average, our heterogeneity analysis shows both the goal-setting and honor-based incentives can crowd-in the intrinsic motivation of the above subgroups to save electricity. Policy makers can target the subgroups to create the habit of electricity saving.

We admit there are several caveats of our study. First, we carried out the study in student dormitories, but the electricity consumption of student dormitories is less than that of ordinary households. Whether the effectiveness can extend to household settings, large-scale field experiments might be needed to test the effects of these two incentives. Moreover, our experimental design externally imposed the goal on the participants. One might expect self-determined goals to be less intrusive and thus more likely to lead to a crowd-in effect than externally imposed goals. Future research could experimentally examine the differential impacts of these two types of goals.

References

Abrahamse, W. and Steg, L. (2013). Social influence approaches to encourage resource conservation: A meta-analysis. *Global Environmental Change*, 23(6):1773–1785.

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

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics*, 130(3):1117–1165.

Allcott, H. and Mullainathan, S. (2010). Behavior and energy policy. *Science*, 327(5970):1204–1205.

Andor, M. A. and Fels, K. M. (2018). Behavioral economics and energy conservation–a systematic review of non-price interventions and their causal effects. *Ecological economics*, 148:178–210.

Angrist, J. D. and Pischke, J.-S. (2014). *Mastering'metrics: The path from cause to effect*. Princeton university press.

Ariely, D., Bracha, A., and Meier, S. (2009). Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99(1):544–55.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics*, 21(1):1–42.

Barclay, P. and Barker, J. L. (2020). Greener than thou: people who protect the environment are more cooperative, compete to be environmental, and benefit from reputation. *Journal of Environmental Psychology*, 72:101441.

Benabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3):489–520.

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American economic review*, 96(5):1652–1678.

Bowles, S. (2008). Policies designed for self-interested citizens may undermine "the moral sentiments": Evidence from economic experiments. *Science*, 320(5883):1605–1609.

Brandsma, J. S. and Blasch, J. E. (2019). One for all?–the impact of different types of energy feedback and goal setting on individuals' motivation to conserve electricity. *Energy policy*, 135:110992.

Cardenas, J. C., Stranlund, J., and Willis, C. (2000). Local environmental control and institutional crowding-out. *World development*, 28(10):1719–1733.

Carroll Jr, S. J. and Tosi, H. L. (1970). Goal characteristics and personality factors in a management-by-objectives program. *Administrative Science Quarterly*, pages 295–305.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *J of Personality & Social Psychology*, 18(1):105–115.

Deci, E. L., Koestner, R., and Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6):692–700.

Delmas, M. A. and 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.

Ding, Y., Zhang, Z., Zhang, Q., Lv, W., Yang, Z., and Zhu, N. (2018). Benchmark analysis of electricity consumption for complex campus buildings in china. *Applied Thermal Engineering*, 131:428–436.

Ernest-Jones, M., Nettle, D., and Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior*, 32(3):172–178.

Fehr, E. and Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, 46(4):687–724.

Frey, B. S. and Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, 15(5):589–611.

Fröhlich, M., Frölich, M., and Sperlich, S. (2019). *Impact evaluation*. Cambridge University Press.

Griskevicius, V., Tybur, J. M., and Van den Bergh, B. (2010). Going green to be seen: status, reputation, and conspicuous conservation. *Journal of personality and social psychology*, 98(3):392.

Harding, M. and Hsiaw, A. (2014). Goal setting and energy conservation. *Journal of Economic Behavior & Organization*, 107:209–227.

Henry, M. L., Ferraro, P. J., and Kontoleon, A. (2019). The behavioural effect of electronic home energy reports: Evidence from a randomised field trial in the united states. *Energy Policy*, 132:1256–1261.

Hirst, M. K. (1988). Intrinsic motivation as influenced by task interdependence and goal setting. *Journal of Applied Psychology*, 73(1):96.

Iacobucci, D., Schneider, M. J., Popovich, D. L., and Bakamitsos, G. A. (2016). Mean centering helps alleviate "micro" but not "macro" multicollinearity. *Behavior research methods*, 48:1308–1317.

Ishimura, Y., Shinkuma, T., Takeuchi, K., and Hosoda, E. (2024). The effects of regional goal setting on household waste. *Ecological Economics*, 215:108015.

Kearney, M. S. and Levine, P. B. (2015). Media influences on social outcomes: The impact of mtv's 16 and pregnant on teen childbearing. *American Economic Review*, 105(12):3597–3632.

Kraft-Todd, G., Yoeli, E., Bhanot, S., and Rand, D. (2015). Promoting cooperation in the field. *Current Opinion in Behavioral Sciences*, 3:96–101.

Lazaric, N. and Toumi, M. (2022). Reducing consumption of electricity: A field experiment in monaco with boosts and goal setting. *Ecological Economics*, 191:107231.

Lepper, M. R., Greene, D., and Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality &*

Social Psychology, 28(1):129–137.

Liu, Y., Kua, H., and Lu, Y. (2021). Spillover effects from energy conservation goal-setting: A field intervention study. *Resources, Conservation and Recycling*, 170:105570.

Locke, E. A. and Bryan, J. F. (1967). Performance goals as determinants of level of performance and boredom. *Journal of Applied Psychology*, 51(2):120.

Lokhorst, A. M., Werner, C., Staats, H., van Dijk, E., and Gale, J. L. (2013). Commitment and behavior change: A meta-analysis and critical review of commitment-making strategies in environmental research. *Environment and Behavior*, 45(1):3–34.

Magali, D. and Doctori, B. V. (2010). Measuring corporate environmental performance: the trade-offs of sustainability ratings. *Business Strategy and the Environment*, 19(4):245–260.

Mossholder, K. W. (1980). Effects of externally mediated goal setting on intrinsic motivation: A laboratory experiment. *Journal of Applied psychology*, 65(2):202.

Nisa, C. F., Bélanger, J. J., Schumpe, B. M., and Faller, D. G. (2019). Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. *Nature Communications*, 10(1):4545. Number: 1 Publisher: Nature Publishing Group.

Rode, J., Gómez-Baggethun, E., and Krause, T. (2015). Motivation crowding by economic incentives in conservation policy: A review of the empirical evidence. *Ecological Economics*, 117:270–282.

Sexton, S. E. and Sexton, A. L. (2014). Conspicuous conservation: The prius halo and willingness to pay for environmental bona fides. *Journal of Environmental Economics and Management*, 67(3):303–317.

Titmuss, R. M. (1970). The gift relationship. from human blood to social policy. *Medical Care*, 11(3):261–263.

Velez, M. A., Murphy, J. J., and Stranlund, J. K. (2010). Centralized and decentralized management of local common pool resources in the developing world: experimental evidence from fishing communities in colombia. *Economic Inquiry*, 48(2):254–265.

Vesely, S., Klöckner, C. A., Carrus, G., Tiberio, L., Caffaro, F., Biresselioglu, M. E., Kollmann, A. C., and Sinea, A. C. (2022). Norms, prices, and commitment: A comprehensive overview of field experiments in the energy domain and treatment effect moderators. *Frontiers in Psychology*, 13:967318.

Zhou, Y., Sun, L., Hu, X., and Ma, L. (2021). Clustering and statistical analyses of electricity consumption for university dormitories: A case study from china. *Energy and Buildings*, 245:110862.

Appendix

Figure 4. Treatment dormitories

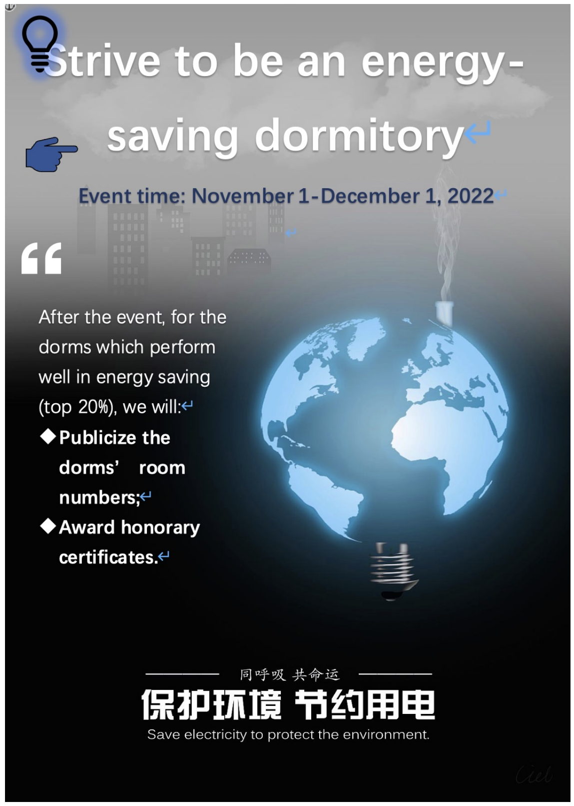

FIGURE 5. Honor-based-incentives group's electricity saving flyer

FIGURE 6. Electricity saving flyer for goal-setting group