RSF AWARD – FINAL NARRATIVE REPORT

Principal Investigator(s): Jessica Terman

RSF Award Number: 98-18-09

Fiscal Agent Name: George Mason University

Project Title: Behavioral Interventions in Energy Consumption

Award End Date: 30 June, 2019

Report Due Date: 28 August, 2019

1) Please provide a summary of the original project including main questions to be addressed and the data/methods to answer those questions. (300 words max).

   Does the employment of a loss aversion incentive (EEE) reduce consumption relative to other types of behavioral nudges and pecuniary incentives? Several other research questions were added as indicated below

   Methods and Analytical Approach
   This field experiment monitored plug-level dorm-room energy usage for 90 students, including a control and a treatment group. The primary hypothesis was that the incentive based on potential losses, manifested through an Energy Efficiency Escrow (EEE), will result in lower usage for the same potential savings offered by a pay-for-performance (P4P) program (see Figure 1).

   The unit of analysis are individual students, each of whom live in on-campus residence hall rooms. Wi-fi connected power strips were assigned to each student and transmitted the entire plug-load demand, with individual plug and total power recorded every second. A consulting energy company (GOEFER, Inc.) developed a user-app for students to control plug loads and to monitor their usage. At the beginning of the treatment period a new version of the GOEFER app added features to allow the control and treatment group to assess their financial incentives. Pre/post-survey data provided controls for sociodemographic differences and other factors.

   A combination of difference-in-differences methods and clustered OLS modeling of pre and post treatment periods was used for all analyses.

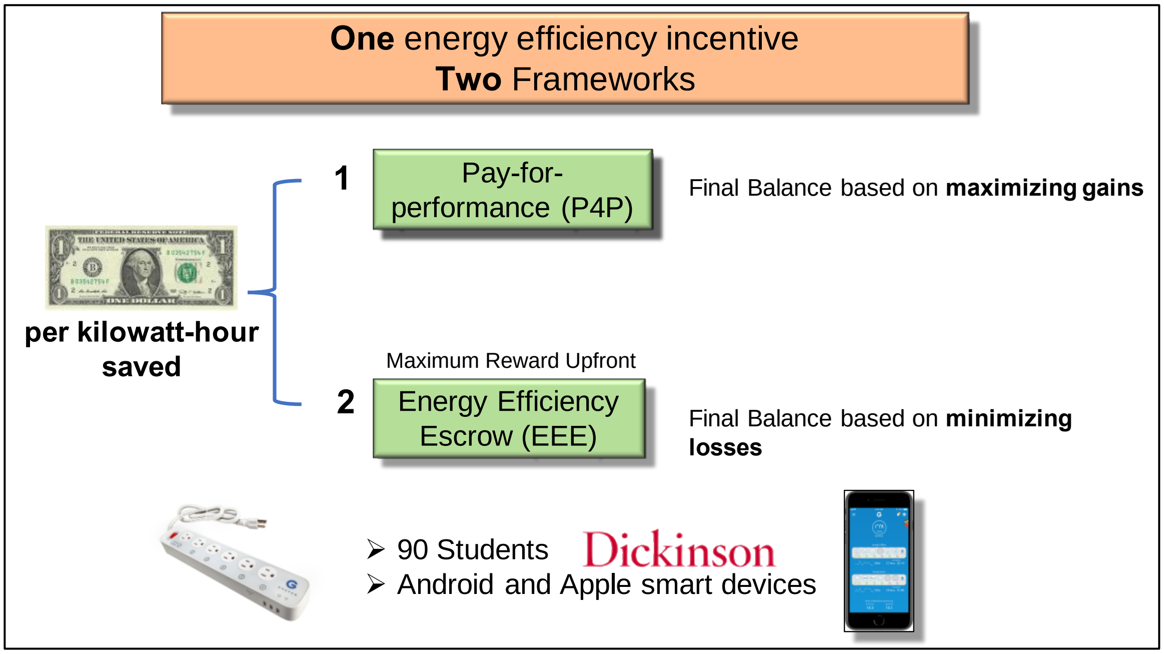


Figure 1: Overview of Field Experiment

One seminal difference of the EEE vs. P4P framework from most loss aversion studies is that the potential gains in a P4P program are notional and cannot be quantified a priori. For instance, at the beginning of the treatment period all of the participants in the treatment group (EEE) know exactly how much they can potentially lose, whereas in the control group (P4P) the potential gains are speculative. This is the first empirical study to compare how consumers respond when losses are explicitly revealed versus when potential gains are notional and implicit.

2) What are the main findings or conclusions of your funded research? Any unanticipated findings (positive or negative)?

Primary Research Questions and Findings:

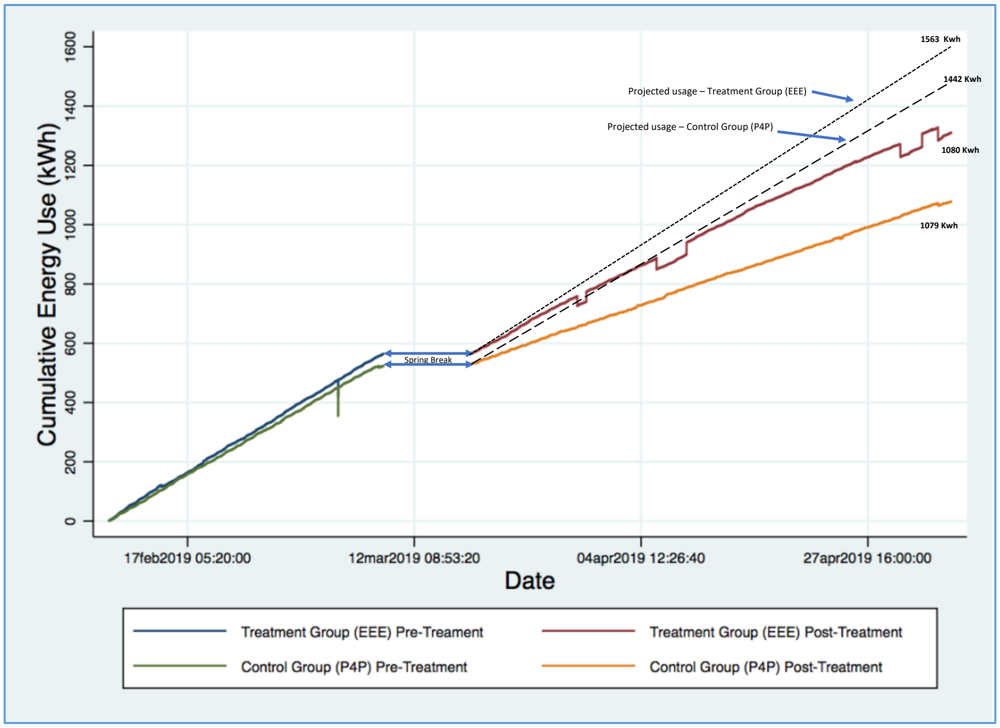
Research Question #1: Does the employment of a rate-based financial incentive reduce personal energy consumption?
Hypothesis: Hypothesis: A financial incentive will reduce personal energy consumption across multiple frameworks (EEE and P4P).

Average daily consumption across both groups (EEE and P4P) was reduced by 29.3 percent ($p < 0.001$). The hypothesis was confirmed.

Research Question #2: Does the employment of an energy efficiency escrow (EEE) reduce consumption relative to the same financial incentive presented as a potential gain, modeled by a P4P program?
Hypothesis: The EEE incentive will result in lower energy consumption relative to the same financial incentive framed as a P4P program.

Average daily reductions of 39.2 percent ($p < 0.05$) and 19.9 percent ($p < 0.05$) in the P4P and EEE groups, respectively, not only refuted the hypothesis, but it was statistically significant.

	Mean(Baseline Avg. daily use – Kwh)	Mean (Treatment Avg. daily use – Kwh)	St_Err	p_value
Control (P4P)	0.416	0.253	0.044	0.001
Treatment (EEE)	0.452	0.362	0.040	0.030



In summary, this research shows that when consumers are continuously informed about their usage and offered an incentive to reduce consumption, **those who are forced to cognitively calculate potential gains and are therefore forced to perform some form of risk management (P4P group), perform better than those who are explicitly and continuously aware of their loss exposure (EEE group)**.

Based on post-experiment surveys, it was observed that the effect of the EEE for those who did not understand how it was calculated was a significant barrier rather than an aid, or even a neutral factor, in conserving energy. This was true despite the fact that all persons in the EEE group still had the same tool (the energy history tab) that the P4P group had to monitor their usage.

Also, it was determined through post-treatment surveys that the P4P group, who had to estimate their gains based on their current usage patterns, received a reward greater than they expected (Pr=0.002 based on Chi-Squared estimator). Coupled with the fact that the P4P group conserved more energy, it supports the assertion that those participants “overshot” whatever conservation targets they may have set for themselves. It may also suggest that more individual utility can be gained with consumers who obtain a higher compensation than they were expecting at the end of the performance period. This is significant because customer satisfaction is an important element in any utility-run program. Whether this is sustainable over multiple treatment periods is left to be verified for further research.

Research Question #3: Do individuals with a higher potential reward for reducing energy, reflected in a higher baseline usage level, perform better relative than others?

Hypothesis: Individuals with a higher potential reward for conserving energy will reflect higher relative levels of energy reductions than those with lower potential rewards.

A clustered OLS model shows that as potential rewards increase, the relative percentage reduction in energy usage does increase. For every 1 Kwh of baseline usage, students reduced their consumption by 27.1 percent. This result is not especially surprising and has been observed over a handful of other DSM programs. Social norms tend to have a stronger effect on heavy energy consumers as well.

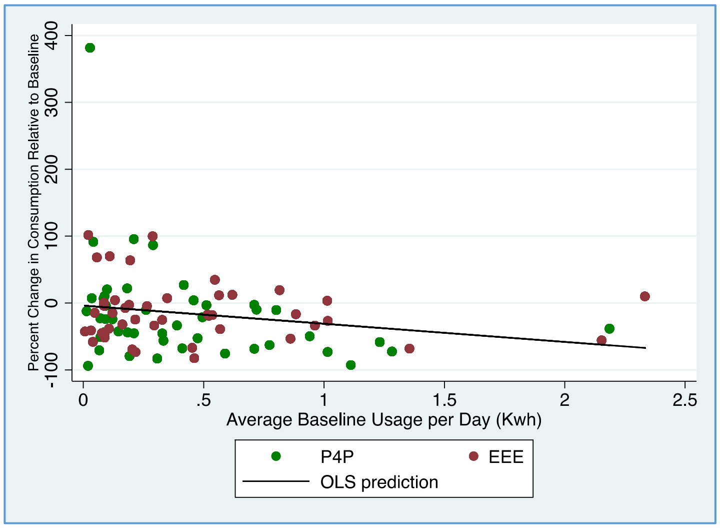


Figure 2: Energy Conservation Performance vs. Average Baseline Usage

One reason could be that heavier energy users tend to have a higher number of discretionary loads to moderate.

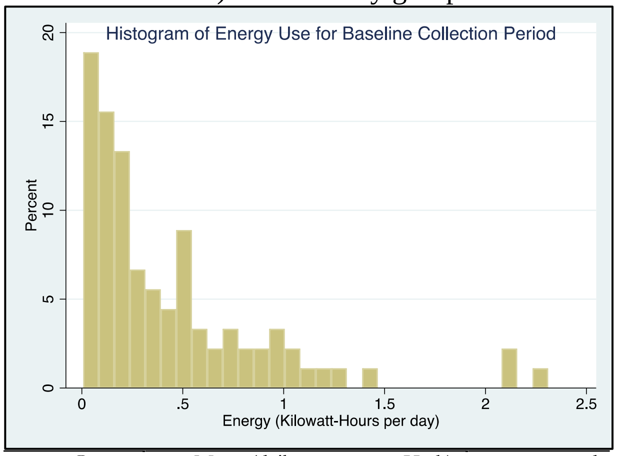
Research Question #3: Do levels of consumption reduction increase over time for the EEE group relative to the P4P group?
Hypothesis: The EEE group will show greater levels of consumption reduction than the P4P group as the incentive period approaches its conclusion.

With no effective coefficient in the quadratic or log modeling of the time variable the hypothesis can be refuted. Reductions were remarkably linear throughout the treatment period.

<u>Other important findings:</u>

Baseline Usage Period
A 28-day collection period was used to establish personal baselines for each student. Energy usage was archived every 15-minutes for all participants. Average daily usage was used in order to accommodate the few dropouts that occurred during the period. There were 2,688 (28*4*24) collection opportunities for this period. 85 of the 90 students experienced zero collection dropouts during the period. For any daily period where a full 96 (24 X 4) possible values were not collected on any student, their average for what was collected for a given day was normalized to a full 24-hour period. This allowed for more accurate comparisons between students and the inclusion of a higher percentage of panels.

Students were not told the parameters of the baseline collection period so as to minimize biases. A total energy usage of 1094 kWh was measured for the baseline period. Table 15 shows the distribution of energy usage – for the entire sample population (Mean = 0.434) and (Standard Deviation = 0.472) as well as by group.



Group	Mean (daily averages – Kwh)	sd
P4P	.416	.427
EEE	.451	.507
Combined	.434	.472

With average daily energy usage (Kwh) as the dependent variable, independent variables that showed increased usage were white and male students. Energy usage decreased with increasing age group, with seniors showing the lowest and only statistically significant academic year group. Financial assistance was not statistically significant. White and non-white differentiation showed significance with respect to Black or African American and Asian race groups using less energy.

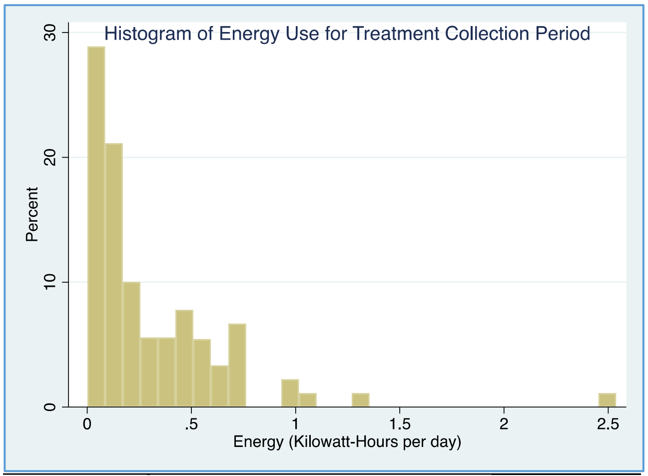
Students' estimation of their own energy consumption relative to other students (from pre-study survey) revealed interesting discrepancies. For instance, students who stated they used "a little more than average" consumed considerably more energy than all other users. Those who revealed their own use to be "much more than average" used less energy than all responses with the exception of those who stated they used "much less energy than average". All results were statistically significant.

Students' self-assessment of their energy knowledge correlated very well with baseline level usage, much better than did their estimation of personal energy usage. This is consistent with the theory of planned behavior (TPB) that suggest the main determining factors of behavioral intention are attitudes, which are influenced by knowledge and experience. Again, all results were statistically significant.

Lastly, it was not surprising that those students who requested multiple power strips used more energy. Although some students suggested that they needed additional strips to accommodate a larger space with limited outlets, that translated into more actual load.

Post-Treatment Results

A 49-day treatment period was used to collect energy use information after the start of the incentive. Again, energy usage was taken every 15-minutes for all participants and normalized to a 24-hour period. This amounted to 4705 observations per student. Table 1 shows the distribution of energy usage – for the entire sample population (Mean = 0.307) and (Standard Deviation = 0.364).



Group	Mean (daily averages - Kwh)	sd
P4P	.252	.258
EEE	.362	.442
Combined	.307	.364

The total expected energy usage for the treatment period was calculated using the baseline period usage (1094 kWh) and multiplied by the fraction of days in each period (49/28) = 1915 kWh. The total energy used for the treatment period was 1354 kWh for a reduction of 561 kWh, or 29.3 percent.

Total student payouts for the incentive amounted to $679. This puts the effective payout per-kWh-saved at $679/561(kWh saved) = $1.21, since individuals were not penalized if they exceeded their baseline usage. Another way of stating this is to normalize the savings per dollar invested, thus showing an equivalent 29.3 percent/1.21 = **24.2 percent reduction in energy usage for every $1 of program rate-based incentive**. This 5.2 percent delta can be regarded as the cost of allowing users who do not take advantage of the incentive to revert back to the mean of their normal usage. These users must be compensated for by energy savers. Figure 3 shows a scatterplot of how the population varied their treatment period usage compared to their baseline energy usage. All those in green increased energy usage (28 users) relative to their baseline and represent consumption for which energy conservationists (62 users) had to overcome in order to reach the 29.3 percent reduction.

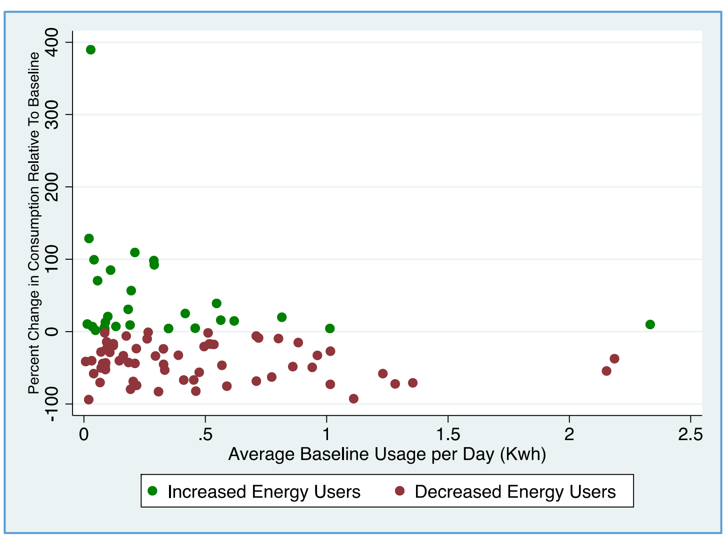


Figure 3: Comparison of Increased Usage vs. Decreased Treatment Usage by Baseline Usage

Other Findings:

Usage Feedback and Monitoring

The existence of feedback has an effect on behavior, even if only for the short term. This research contributes some new information regarding frequency of usage feedback and performance. Most usage feedback programs push information to users, through reports and social media reminders. While some utilities do allow users to log into a web portal and view their usage, the barrier to feedback is still rather high. However, the smart device-enabled GOEFER app is a pull device that allows users to check usage virtually anywhere and at any time. Based on the results of the post-survey question "How often did you check the GOEFER app to monitor your usage"

A clustered OLS model was used to determine that those who checked their app daily conserved more energy relative to their baseline than those that checked weekly or monthly, but conserved less than those who checked less frequently than monthly. However, none of those results were statistically significant.

Usage Migration

Because students had some power loads that were migratory, such as laptops and cellphones (fixed loads were required to remain in a GOEFER power strip) the post-survey asked students "when you consciously reduced your energy consumption in your room, how often did you simply use the same power in a different location?". This gives some assessment if energy migration might account for some of the measured reductions. In other words, how much did migration affect their usage behavior as well as measured energy usage?

The post treatment results analysis suggests no discernible effect between the different survey response. Those responding "never", "rarely", and "often" had similar levels of usage reduction, while those responding "sometimes" actually increased usage. None of those four were statistically significant. The only significant result was from a single respondent who responded "I never altered the way I used energy", who increased energy usage throughout the treatment period.

Thus, assuming students answered the survey question honestly, and understanding that each student may have a different estimate of words such as "often" or "sometimes", there does not appear to be a correlation between energy behavior and levels of reported energy migration.

Import of Financial Incentive on Behavior

In response to the question "How significant was the financial incentive in changing your energy use?", 28.9 percent stated that the financial incentive was "not at all significant" in reducing energy use., while only 13.3 percent stated that it was "very significant". For the majority of the sampled population (57.8 percent), the financial incentive appears to be part of several factors involved in changing behavior.

Two important findings, 1) students who thought the financial incentive was "very significant" in changing their energy use reduced their average daily use significantly, and 2) those same students had significantly higher levels of baseline use than others. Again, this reinforces the importance of how higher potential gains effectively motivated students to change behavior, both in their stated and revealed preferences.

Gender Differences in Post-Treatment

Gender was the most significant sociodemographic factor in the study. It was significant in the outcome of the baseline collection period. Table 1 reveals several interesting observations:

Table 1: Paired T Test Results Pre-Post Treatment by Gender

	Baseline Mean (Avg. Kwh/day)	Post-Treatment Mean (avg. Kwh/day)	St_Err	p_value
Female	0.378	0.252	0.033	0.001
Male	0.580	0.444	0.064	0.044
	Percent Reduction relative to baseline			
Female	-14.000		7.881	0.081
Male	-20.096		11.366	0.090

Individual males reduced their energy usage by an average of 20.1 percent ($p < 0.1$) and individual females by 14.0 percent ($p < 0.1$). Yet, overall, male and female energy reduction was 23.4 percent and 33.3 percent, respectively. This result suggests that high level female baseline users resulted in a lot of the overall post-treatment mean reductions for females.

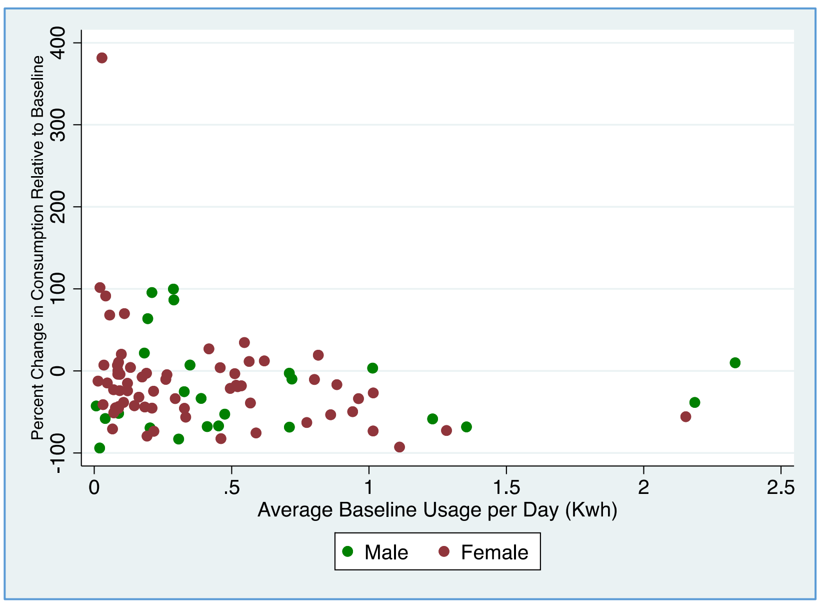


Figure 4: Energy Conservation Performance by Gender

This result is not surprising in view of two previous finding, 1) females had lower levels of baseline energy use, and 2) higher baseline levels showed higher percent reductions. Remember that overall energy savings was much higher due to the higher levels of reductions amongst heavier baseline users. Figure 4 shows this gender difference well.