Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand*

Koichiro Ito
University of Chicago and NBER

Takanori Ida
Kyoto University

Makoto Tanaka
GRIPS

This version: March 1, 2017

Abstract

Firms and governments often use moral suasion and economic incentives to influence intrinsic and extrinsic motivations for economic activities. To investigate persistence of such interventions, we randomly assign households to moral suasion and dynamic pricing that stimulate energy conservation in peak-demand hours. We find significant habituation and dishabituation for moral suasion—the treatment effect diminishes after repeated interventions but can be restored to the original level by a sufficient time interval between interventions. Economic incentives induce larger treatment effects, little habituation, and significant habit formation. Our results suggest moral suasion and economic incentives produce substantially different short-run and longer-run policy impacts.

---

*Ito: Harris School of Public Policy, University of Chicago, 1155 East 60th St., Chicago, IL 60637, and NBER (e-mail: ito@uchicago.edu). Ida: Graduate School of Economics, Kyoto University, Yoshida, Sakyo, Kyoto 606-8501, Japan (e-mail: ida@econ.kyoto-u.ac.jp). Tanaka: National Graduate Institute for Policy Studies, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8677, Japan (e-mail: mtanaka@grips.ac.jp). The authors would like to thank Tetsuya Kawamura for excellent research assistance. We thank Hunt Allcott, Masahiko Aoki, Severin Borenstein, Lucas Davis, Liran Einav, Meredith Fowlie, Takeo Hoshi, Caroline Hoxby, Katrina Jessoe, Paul Joskow, Ryan Kellogg, Chris Knittel, David Laibson, Michael Price, Steve Puller, Matthew Rabin, Dave Rapson, Catherine Wolfram, Frank Wolak, and seminar participants at MIT, Harvard, Stanford, UC Berkeley, UC Davis, University of Tokyo, Hitotsubashi University, the NBER Summer Institute, the AEA Annual Meeting, the POWER Conference on Energy Research and Policy, the RIETI, and the Centre for European Economic Research for helpful comments. We thank the Japanese Ministry of Economy, Trade and Industry, the prefecture of Kyoto, Kansai Electric Power Company, and Mitsubishi Heavy Industries, Ltd. for their collaboration on this study. We thank the New Energy Promotion Council for financial support. Ito thanks the Energy Institute at Haas and the Stanford Institute for Economic Policy Research for financial support.

1 Introduction

Firms and governments often use moral suasion and economic incentives to influence intrinsic and extrinsic motivations for a variety of economic activities. For example, such interventions are frequently used by regulators to promote energy conservation (Reiss and White, 2008), smoking cessation (Volpp et al., 2009), and tax compliance (Dwenger et al., 2016). Similar tools are widely used by firms and nonprofit organizations to stimulate academic refereeing (Chetty, Saez and Sándor, 2014), blood donations (Lacetera, Macis and Slonim, 2012), charitable giving (Landry et al., 2010), and exercise (Charness and Gneezy, 2009).

A central question for economists and policymakers designing such policies is whether appealing to intrinsic and extrinsic motivations can generate persistent effects on economic activities (Gneezy, Meier and Rey-Biel, 2011). In this study, we begin with existing theories in economics and psychology that provide three key predictions for an individual's dynamic response to a policy intervention—*habituation* (Thompson and Spencer, 1966), *dishabituation* (Rankin and Carew, 1988), and *habit formation* (Becker and Murphy, 1988). We characterize how these phenomena can interact with policy interventions that target intrinsic and extrinsic motivations.

We then empirically test our predictions in a field experiment in the context of electricity demand. The first treatment is moral suasion, by which policymakers attempt to influence intrinsic motivation to generate pro-social behavior.[1] The second treatment is an economic incentive, by which policymakers aim to affect extrinsic motivation based on standard demand theory. We randomly assign households to one of three groups: 1) a moral suasion group, 2) an economic incentive group, and 3) a control group. Using household-level electricity consumption data of 30-minute intervals, we examine how our treatments affect electricity usage in peak-demand hours in which the marginal cost of electricity is substantially higher than other hours.

The moral suasion group receives messages requesting voluntary energy conservation during the peak-demand hours—between 1 pm and 4 pm on summer treatment days, and between 5 pm and 8 pm on winter treatment days. The economic incentive group does not receive this message but is charged one of three high marginal prices for electricity in the same peak-demand hours. We design

[1] Moral suasion is widely used by regulators for energy conservation (Reiss and White, 2008; Costa and Gerard, 2015), air quality preservation (Cutter and Neidell, 2009), incentivizing workers (Dal Bó and Dal Bó, 2014), discouraging tax evasion (Dwenger et al., 2016), and law enforcement (Fellner, Sausgruber and Traxler, 2013).

the experiment to test three primary hypotheses. First, we repeat the treatments over many days to examine habituation of the two treatments. Second, we discontinue the intervention for a while and then restart it, which allows us to test if restarting a treatment generates dishabituation—a decreased response via habituation might be restored back to original response levels. Finally, we collect electricity usage data before, during, and after the interventions to investigate habit formation.

We present several findings from the experiment. First, moral suasion induces short-run reductions in electricity usage, but the effect diminishes quickly over repeated interventions, indicating strong habituation. The moral suasion group shows a usage reduction of 8 percent initially. However, their usage becomes statistically indistinguishable from that of the control group over further interventions. Second, we find that economic incentives create much larger and persistent effects. The economic incentive group shows usage reductions of 14 percent for the lowest critical peak price and usage reductions of 17 percent for the highest critical peak price. Moreover, the effect is much more persistent over repeated interventions, suggesting relatively little habituation compared to the case with moral suasion. Third, we examine dishabituation—whether habituated responses can be restored back to an original level. After the summer experiment, we purposely give households about a 3-months interval before we restart our intervention in the winter. We find that the habituated response to moral suasion is "reset" (i.e., recovers back to the original level) when we restart our intervention in the winter. Fourth, we test potential habit formation by estimating treatment effects using data collected after we withdraw the treatments. We find significant habit formation for the economic incentive group and no habit formation for the moral suasion group. After we withdraw the treatment, the moral suasion group's usage is indistinguishable from that of the control group. On the other hand, the economic incentive group continues to conserve energy even after the incentive is withdrawn.

What are the mechanisms behind the substantially different results for the moral suasion and economic incentive treatments? We investigate two potential mechanisms. The first possibility is investment in physical capital stock—households might have purchased energy-efficient appliances in response to treatment. If such an effect were systematically large for the economic incentive group, it could explain the persistent usage reductions, including weaker habituation and stronger habit formation. The second possibility is that the treatments might have induced new utilization habits for daily electricity use. Suppose that some households had "bad habits" of inefficient energy use before we begin our experiment. Our interventions might have triggered a lifestyle change, helping them form good habits, namely, efficient energy use. If this effect were systematically large for the economic incentive group, it could explain these households' persistent usage reduction. Using follow-up survey data, we find no statistical evidence that investment in physical capital stock can explain our experimental findings. By contrast, we obtain supporting evidence that the changes in utilization habit are a key mechanism behind the results. We provide evidence that the economic incentive treatment induces a new utilization habit for electricity use—households in the economic incentive group form a habit of efficient energy use for a variety of electric appliances, including air conditioners, heaters, computers, washers, and cleaners. Although these results are based on survey responses, they provide suggestive evidence about the mechanisms behind our main findings.

This study makes three primary contributions to the literature and has key implications for economic policy. First, our experiment is the first study to directly compare the habituation, dishabituation, and habit formation of moral suasion and economic incentives. A few recent studies examine a subset of these phenomena—mostly habit formation—for either pecuniary or non-pecuniary incentives. Habit formation is studied by Charness and Gneezy (2009) for monetary incentives on exercise, and Ferraro, Miranda and Price (2011) and Allcott and Rogers (2014) for non-monetary incentives on water and energy conservation. To our knowledge, the economics literature has not studied dishabituation. Allcott and Rogers (2014) document habituation in the effect of social comparison. Our contribution to this new literature is that we characterize commonly phrased *persistence* or *long-run effects* of a treatment by habituation, dishabituation, and habit formation based on theories in economics and psychology, and empirically test all of the three phenomena for monetary and non-monetary incentives in a unified field experiment.

The second contribution of this study is that we provide new evidence to the growing literature showing that consumers might not necessarily respond to marginal incentives (Borenstein, 2009; Kahn and Wolak, 2013; Ito, 2014; Copeland and Garratt, 2015; McRae and Meeks, 2016). A central question in this literature is whether providing transparent price information can induce consumers to respond to marginal incentives.[2] Our experiment deliberately gives consumers salient information

[2] Ito (2014) describes that a potential reason for his findings—residential electricity consumers in California do not

on hourly marginal prices via text messages and in-home displays, and then tests their responses to different marginal prices. While marginal prices in increasing block pricing and time-varying pricing are not directly comparable, our results suggest that consumers faced with transparent price information on time-varying pricing respond to marginal prices as standard economic theory predicts.

Third, our study contributes to the recent literature on energy and environmental economics studying the effects of pecuniary and non-pecuniary policies on the conservation of scarce resources. It has been challenging to separately identify the effects on monetary and non-monetary incentives on conservation because in non-experimental data, a variety of policies usually affect consumers simultaneously. For example, Reiss and White (2008) and Costa and Gerard (2015) examine energy conservation programs in California and Brazil, acknowledging that it is empirically challenging to separate the effects of voluntary conservation from those of other policies during their sampling periods. Our study addresses this problem by randomly assigning consumers to either moral suasion or economic incentives in a field experiment.[3]

Finally, our results provide important implications for economic policy. Moral suasion has become increasingly common when policymakers aim to promote pro-social behavior. In practice, whenever a country or state encounters an energy crisis, one of the most debated policy topics is whether regulators should use moral suasion or economic incentives to mitigate the crisis. Historically, moral suasion has been favored politically in many states and countries. Our results on habituation, dishabituation, and habit formation have three key policy implications. In the very short run, both moral suasion and economic incentives are likely to be useful ways to induce pro-social behavior. However, moral suasion is more likely to have quick habituation, which makes the policy ineffective when a policy is repeated over time. Our findings on dishabituation or the "reset effect" imply that moral suasion can become effective again when it has not been used for a while, although the impact is again likely to wane with repeated use of the policy. We highlight

respond to their marginal prices—is that information-acquiring costs are likely to be high for residential electricity consumers who have conventional monthly billing. A few recent studies conduct field experiments related to this question. Kahn and Wolak (2013) and McRae and Meeks (2016) find that information provision changes electricity consumers' responses to marginal prices. On the other hand, Chetty and Saez (2013) find that providing information about non-linear tax incentives does not systematically affect earnings on average, although these authors also find evidence of heterogeneous treatment effects among taxpayers.

[3] Other example of non-experimental studies documenting intrinsically motivated conservation include Cutter and Neidell (2009), who studies the "Spare the Air" program in California and Ferraro and Price (2013), who investigates water conservation.

these implications by providing back-of-envelope calculation of a policy evaluation in the online appendix.

2 Habituation, Dishabituation, and Habit Formation

Our empirical analysis aims to test the persistent effects of moral suasion and economic incentives. For this purpose, theories in economics and psychology provide useful guidance and characterize three key predictions about an individual's dynamic response to a stimulus.

Habituation is a theory formalized by several studies in psychology, including Thompson and Spencer (1966), Groves and Thompson (1970), and Rankin et al. (2009). It implies that repeated presentation of a stimulus might cause a decrease in reaction to the stimulus. For example, animals, such as cats, dogs, monkeys, and rats, strongly react to a stimulus (e.g., a loud sound in a laboratory experiment) when it is presented for the first time, but their responses often gradually wane when the same intervention is repeated over time. The opposite phenomenon is called *sensitization*, which implies that repeated presentation of a stimulus induces an increase in reaction to the stimulus. Because many economic policies in practice are implemented repeatedly over time, whether individuals exhibit habituation or sensitization is a central question for policy evaluation. Existing laboratory evidence in the psychology literature suggests that habituation is more ubiquitous than sensitization for most species. However, we are not aware of studies testing habituation and sensitization for moral suasion and financial incentives in a field experiment.

Another relevant key theory in psychology is *dishabituation*—declined responses, as a result of habituation, can be restored to an original level either by providing a new type of treatment, a stronger or weaker intensity of the same treatment, or the same treatment with a sufficient time interval between interventions. In particular, dishabituation obtained by the last approach—a proper time interval between interventions—is called *spontaneous recovery* of habituation. For policymakers, dishabituation is an important phenomenon because it could make a habituated policy impact effective again. Evidence of dishabituation is mostly from laboratory experiments in psychology, and there is little if any empirical evidence in the economics literature.

Finally, *habit formation* is a theory developed primarily in the recent economics literature (Becker and Murphy, 1988; Rozen, 2010). A short-run intervention might form a habit of consumption for the future, which influences the prolonged existence of policy impacts after the removal of the stimulus. The theory presented by Becker and Murphy (1988) suggests that past consumption forms consumption capital, which drives habit formation. In addition, the authors suggest that consumption capital depreciates. That is, treatment effects might continue to exist after the final intervention but decay over time. Empirical evidence of habit formation is documented in many instances, including the effect of monetary incentives on exercise (Charness and Gneezy, 2009) and non-monetary incentives on water and energy conservation (Ferraro, Miranda and Price, 2011; Allcott and Rogers, 2014). Our study is the first to test habit formation for moral suasion and compare it to habit formation for economic incentives.

We design our field experiment to test these three predictions—habituation, dishabituation, and habit formation—for moral suasion and economic incentives. In the next section, we begin by presenting our experimental design and data. We then describe hypotheses that we test in the experiment and explain why moral suasion and economic incentives might have different dynamic effects for the three phenomena.

3 Experimental Design, Data, and Hypotheses

3.1 Experimental Design and Data

We conducted the field experiment for households in the Keihanna area of Kyoto prefecture in Japan in the summer of 2012 and the winter of 2013. We implemented the experiment in collaboration with the Ministry of Economy, Trade and Industry, the prefecture of Kyoto, Kansai Electric Power Company, and Mitsubishi Heavy Industries, Ltd.

In order to invite as broad a set of households as possible, we provided generous participation rewards, including free installations of an advanced meter and in-home display in addition to a participation reward of 24,000 yen (approximately \$240 in 2012). We contacted all 40,710 residential electricity customers in the Keihanna area by mail.[4] Of these, 1,659 customers confirmed their participation. We excluded students, customers who had electricity self-generation devices, and those without access to the internet. This process left us with 691 households. Similar to previous field experiments in electricity demand (Wolak, 2006, 2011; Faruqui and Sergici, 2011; Jessoe and Rap-

[4]We include the English translation of the recruitment letter in the Appendix.

son, 2014), our experiment was a randomized controlled trial (RCT) for self-selected participants, as opposed to an RCT for a purely randomly selected sample of the population. Therefore, it is important to consider carefully the external validity of the experiment, although random assignment of treatments guarantees the internal validity of the experiment. To explore the external validity of our sample, we collected data from a random sample of the population in the corresponding geographical area. We analyzed the observables between our sample and the random sample, as outlined below.

We randomly assigned the 691 households to one of three groups: control (C), moral suasion (M), and economic incentive (E).

Control Group (C): The 153 customers in this group received an advanced electricity meter, an in-home display, and the participation reward. This group received no other treatment.

Moral Suasion Group (M): The 154 customers in this group received an advanced electricity meter, an in-home display, and the participation reward. In addition, this group received "moral suasion for energy conservation," which we describe below.

Economic Incentive Group (E): The 384 customers in this group received an advanced electricity meter, an in-home display, and the participation reward. In addition, this group received "economic incentives for energy conservation," which we describe below.[5]

The primary data for this study are high-frequency data on household electricity usage. Advanced electricity meters, often called "smart meters," were installed for all participating households, enabling us to collect household-level electricity usage at 30-minute intervals. We used consumption data from the summer of 2012 to the spring of 2013. In addition to the usage data, we collected data by three surveys. We conducted the first survey prior to treatment assignment and collected demographic information. Next, we conducted the second survey upon completion of the experiment to explore the mechanism behind our findings. Finally, we conducted the third survey for a random sample of 717 households in the area to investigate the external validity of our experimental sample.

Columns 1, 2, and 3 of Table 1 present the summary statistics of demographic variables and

[5]We assigned a relatively large number of participants to the economic incentive group to test the effects of different prices. If our sole objective were to compare the effects of the moral suasion and economic incentives, we would have assigned the same number of customers to each group in order to minimize the variance of the estimates (Duflo, Glennerster and Kremer, 2007).

pre-experiment consumption data by treatment group. A comparison across control and treatment groups indicates statistical balance in observables because of random assignment of the groups. Furthermore, very little attrition occurred in each group. In total, nine households (1.3 percent) were excluded from our sample because they moved residence. Because this small attrition occurred at approximately the same rate in each group, it is unlikely to bias our estimates significantly.

Column 4 shows the summary statistics for a random sample of the population in the corresponding geographical area. We investigated the external validity of our sample by comparing the mean for each observable variable between the random sample and our control group. Column 5 presents the differences in means and the standard errors of the differences in brackets. The differences are small and statistically insignificant for most variables. We found small but statistically significant differences at the 5 percent level for the age of buildings and household size. Note that there is still a possibility that unobservable characteristics can differ between the random sample and our experimental sample. However, the results in column 5 suggest that these two samples are statistically very similar, at least for the key observable variables for residential electricity demand.

This analysis investigated the external validity of our experimental sample to the population in the corresponding geographical area. However, there is another important external validity issue when a policy objective is to obtain estimates of the treatment effects for the population outside the experimental region. Allcott (2015) finds "site selection bias" when he compares the estimates of Opower's home-energy report among its 111 randomized controlled trials. The idea behind site selection bias is that cities that participated in a field experiment might be different from other cities in their observable and unobservable characteristics. Our experiment is not free from the possibility of site selection bias if households in our experimental city (Kyoto) are systematically different from those in other cities.

Two pieces of information are relevant to this point. First, we conducted similar dynamic pricing experiments for different research questions in three other locations in Japan—Yokohama, Toyota, and Kitakyushu—between 2012 and 2013. For each location, we found about 15–20 percent usage reductions in peak-hour electricity usage from similar dynamic electricity-pricing treatments. This finding implies that the estimates from the current study in Kyoto are not substantially different from those from other locations. Second, although Kyoto was the city that hosted the Kyoto Protocol on climate change in 1997, a general perception is that households in Kyoto are not necessarily more environmentally conscious than are households in other prefectures in Japan. For instance, Ida, Takemura and Sato (2015) conduct a survey for households in major cities in Japan and find that preferences for energy and environmental policies are not systematically different between average Japanese households and those in Kyoto.[6] While these two points provide suggestive evidence that site selection bias might not be severe in our case, we cannot provide definite evidence for this concern unless we are able to conduct the same experiment for a random sample of households in Japan. Results from our experiments, therefore, should be interpreted with this caution when they are applied to policies outside the experimental region.

To contextualize typical weather patterns in the experimental region, we compared the monthly average high and low temperatures between Kyoto prefecture in Japan (our experimental region) and Washington DC in the United States. We provide this comparison in Figure A.2 in the appendix. The average low and high temperatures are very similar between the two cities. The average high and low temperatures in the spring and summer months are almost identical between the two cities. In the fall and winter months, the average high and low temperatures are slightly higher in Kyoto, but the difference is less than 4^{o}F in each month, which implies that these two cities have quite similar weather conditions, which determine the major part of electricity consumed, namely, usage for cooling and heating.

3.2 Treatments

Electricity consumers generally do not pay prices that reflect the relatively high marginal costs of electricity during peak-demand hours. This mismatch is a fundamental economic inefficiency in electricity markets (Borenstein, 2002; Joskow, 2012). Policymakers usually consider two types of economic policies to address this inefficiency. The first policy instrument, which many countries use most frequently, is an appeal to intrinsic motivation by using moral suasion for voluntary energy conservation. The second policy instrument, which is motivated by the standard economic theory, is an appeal to extrinsic motivation by introducing dynamic pricing that reflects the time-varying marginal costs of electricity. An important question is whether these two types of policies can generate persistent effects on consumer behavior. To investigate this question, we designed two

[6] Ida, Takemura and Sato (2015) find that these preferences are systematically different between average Japanese households and those in Fukushima prefecture, where the Fukushima Daiichi nuclear disaster occurred in 2011.

treatments that reflect the two policies used by regulators in practice.

Our first treatment is "moral suasion," which is intended to influence intrinsic motivation for energy conservation (Kreps, 1997; Bénabou and Tirole, 2003, 2006).[7] After we assigned customers to the moral suasion group (M), we informed them that energy conservation was required in the critical peak-demand hours. The message sent to this group after the group assignment was "substantial energy conservation will be required for the society in 'critical peak-demand hours' on summer and winter peak-demand days, in which electricity supply will be very limited relative to demand." Note that customers in this group did not receive any monetary incentives to conserve energy.

Before the experimental period began, we informed customers of how they were going to receive the treatments. First, their treatment hours were predetermined—1 pm to 4 pm for the summer and 6 pm to 9 pm for the winter. These hours correspond to the system peak-demand hours in Japan. Second, we defined the treatment days as follows. A treatment day had to be a weekday in which the day-ahead maximum temperature forecast exceeded 31°C (88°F) for the summer and was lower than 14°C (57°F) for the winter.

To understand when and how consumers received the notifications of their treatments, consider an example treatment date, August 21. On the day before (August 20), the forecast maximum temperature for August 21 was reported to be above the cutoff level for treatment days. We delivered notifications to customers at 4 pm on August 20 by a text message to their in-home displays, cell phones, and computers. They were able to view the message on each device between 4 pm on August 20 and 4 pm on August 21. The text message sent to the moral suasion group was "Notice of Demand Response: In the following critical peak-demand hours, please reduce your electricity usage: 1 pm–4 pm on Tuesday, August 21."

Our second treatment is an "economic incentive," which was intended to influence extrinsic motivation for energy conservation. After we assigned customers to the economic incentive group (E), we informed them that they would be charged high electricity prices during the critical peak-demand hours on the critical peak-demand days. Precisely the same as we informed the moral suasion group, we told the economic incentive group that we would provide day-ahead notifications

[7] Note that the term, intrinsic motivation, is sometimes used slightly differently in economics and psychology. Our intervention of moral suasion came from an external authority but was aimed at influencing intrinsic motivation for energy conservation. This approach is similar to previous studies in economics such as Dwenger et al. (2016) that provided a shock to intrinsic motivations for tax compliance.

of critical peak days based on the day-ahead weather forecast.

Figure 1 shows the hourly price schedule for the economic incentive group. This price schedule is usually called variable critical peak pricing or critical peak pricing with variable peak prices, because it consists of variable marginal prices for the critical peak hours. On treatment days, the economic incentive group had a price increase of 40, 60, or 80 cents/kWh. Because the baseline price was approximately 25 cents/kWh, these price increases mean that the critical peak price was 65, 85, or 105 cents/kWh.[8] For example, at 4 pm on August 20, the economic incentive group received this message, "Notice of Demand Response: In the following critical peak-demand hours, you will be charged a very high electricity price, so please reduce your electricity usage: 1 pm–4 pm on Tuesday, August 21. The price will be 85 yen (+ 60 yen) per kWh." Participants were able to view this message on their in-home displays, cell phones, and computers between 4 pm on August 20 and 4 pm on August 21.

For a given treatment day, all customers in the economic incentive group had the same critical peak price. We randomized their prices across the treatment days—depending on the treatment day, the marginal price for the critical peak hours was different. We used stratified randomization to allocate the three critical peak prices (65, 85, and 105 cents/kWh) equally throughout the experiment period. First, we divided the treatment days into *treatment cycles*, which consisted of 3 treatment days. Then, we randomized the three critical peak prices within each cycle—each cycle included a treatment day with 65 cents/kWh, a treatment day with 85 cents/kWh, and a treatment day with 105 cents/kWh, in which we randomized the order of the three prices in each cycle. We repeated the interventions as long as the temperature forecasts met the threshold. In total, the treatment groups experienced 15 treatment days (five cycles) in the summer and 21 treatment days (seven cycles) in the winter.

As an example, consider two treatment cycles around August 21, which we used as an example of treatment days in the beginning of this subsection. The day-ahead forecasts for the maximum temperatures exceeded the threshold (31°C or equivalently 88°F) for August 17, 21, 22, 28, 29, and 31. Note that the treatment days were not necessarily consecutive because a treatment day had to be a weekday and its forecast maximum temperature had to be above the threshold. In

[8] Customers paid in Japanese yen, but we use US currency throughout the paper. 1 Japanese yen was approximately equivalent to 1 US cent (2012 exchange rate).

this example, we grouped August 17, 21, and 22 as a cycle, and August 28, 29, and 31 as another cycle. We then randomized the three critical peak prices in each cycle. As a result, customers in the economic incentive group had prices of 65, 105, 85, 85, 65, and 105 cents for these 6 treatment days.

We used this stratified randomization to minimize the correlation between the critical peak prices and temperatures. The minimum, average, and maximum of the daily maximum temperatures for the summer treatment days were 31.2°C, 33.9°C, and 36.5°C (88°F, 93°F, and 98°F). The minimum, average, and maximum of the daily maximum temperatures for the winter treatment days were 3.5°C, 7.8°C, and 11.4°C (38.3°F, 46.0°F, and 52.5°F). That is, while customers experienced hot temperatures for all summer treatment days and cold temperatures for all winter treatment days, there was some variation in the temperature across the treatment days. Using stratified randomization, we avoided the possibility of customers experiencing a certain critical peak price only on particularly hot or cold days. The resulting correlation between the temperatures for the treatment days and the critical peak prices was -0.06 for the summer and -0.05 for the winter.

3.3 Hypotheses

We tested four primary hypotheses using data from the field experiment. First, we asked whether standard economic theory could explain household responses to moral suasion and economic incentives. A stylized demand model for electricity usage predicts that 1) the economic incentive group lowers consumption in response to the marginal price of electricity according to a downward-sloping demand curve, and 2) the moral suasion group uses electricity in the same way as the control group, implying that moral suasion should not affect electricity usage. On the other hand, a demand model with warm glow, such as that of Andreoni (1989), predicts that moral suasion could alter usage levels, as described in our model in the online appendix. To test these hypotheses, we compared consumption between the control group, moral suasion group, and economic incentive group.

The second hypothesis is the habituation of treatment effect, as described in Section 2. We examined if the treatment effects of moral suasion and economic incentives stayed constant, decreased, or increased over repeated interventions. Standard demand theory predicts that usage should stay constant regardless of the number of interventions, given a certain marginal price of electricity. By contrast, the theory of habituation suggests that the treatment effects might habituate and thereby diminish over repeated interventions. In particular, Groves and Thompson (1970) and Rankin et al. (2009) present a theory and laboratory evidence that a less harmful treatment that does not involve an extrinsic reward is more likely to produce strong habituation. By contrast, a harmful treatment that involves an explicit reward is less likely to generate habituation. If households were to perceive moral suasion as a treatment that does not involve an explicit reward and is relatively less harmful than high electricity price, the moral suasion treatment might generate stronger habituation. Using the variation generated by random group assignment, we compared habituation between the moral suasion and economic incentive groups.

The third hypothesis is dishabituation—whether habituated responses can be restored back to an original level. After the summer experiment, we purposely gave households about a 3-month interval before we restarted the intervention in the winter. Laboratory experiments in psychology, such as those of Rankin and Carew (1988) and Phelps (2011), find that providing sufficient time intervals between interventions is one of the effective ways to obtain dishabituation for certain species. We know little about whether this method can produce a similar reset effect for economic activities, which we test by comparing the final intervention in the summer and the first intervention in the winter.

Fourth, we tested for habit formation using data from the post-intervention period. We withdrew our treatments after the final intervention but continued to collect high-frequency electricity consumption data. Standard economic theory predicts that electricity usage for the treatment groups should not differ from that for the control group once we withdrew our interventions. However, the theory of habit formation developed by Becker and Murphy (1988) suggests that short-run interventions of economic incentives and moral suasion could form a new consumption habit, which could affect future consumption. Empirically, the literature contains mixed evidence on what drives habit formation. We provide empirical evidence on habit formation both for moral suasion and economic incentives in our field experiment by comparing electricity usage between our groups in the post-experimental period.

In the next section, we test these four primary hypotheses. We then investigate the mechanism behind our findings by follow-up surveys on investment in physical capital stock and utilization habits for electricity usage.

4 Empirical Analysis and Results

We present our experimental results in this section. Recall that the treatment groups experienced many treatment days in the summer and winter. We included all treatment days in our regression, to show the overall treatment effects. We then explored habituation, dishabituation, spillover effects on non-treatment hours, and habit formation in the subsequent subsections.

4.1 Effects of Moral Suasion and Economic Incentives

We begin by showing evidence from the raw data in Figure 2, which plots the mean log electricity consumption for each group over 30-minute intervals on the summer treatment days. The figure indicates that usage in the pretreatment hours is essentially the same for all groups. About 1 hour before the treatment hours, usage for the treatment groups begins to drop relative to the control group. The reductions are consistent during the treatment hours (1 pm to 4 pm). Immediately after the end of the treatment hours, usage for the treatment groups returns to the control group's level, although there are small remaining differences for a few hours following treatment. The figure provides visual evidence of the treatment effects for both treatment groups and suggests that the reductions are larger for the economic incentive group.

Table 2 provides a formal econometric analysis with standard errors. We estimated the treatment effects by ordinary least squares (OLS) for a linear equation:

$$\ln x_{it} = \beta M_{it} + \gamma E_{it} + \theta_i + \lambda_t + \eta_{it}, \tag{1}$$

where $\ln x_{it}$ is the natural log of electricity usage for household i in a 30-minute interval t. We used the natural log of usage for the dependent variable so that we could interpret the treatment effects approximately in percentage terms. The treatment effects in exact percentage terms can be obtained by $\exp(\beta) - 1$ and $\exp(\gamma) - 1$.[9] In this paper, we report both the log points and exact percentage terms. M_{it} equals 1 if household i is in group M (the moral suasion group) and receives a treatment in t. Similarly, E_{it} equals 1 if household i is in group E (the economic incentive group) and receives a treatment in t. We included household fixed effects θ_i and time fixed effects λ_t

[9]Note that when an estimate (β or γ) is negative, its exact percentage term will be smaller than the corresponding log points in absolute terms. For example, when $\gamma = -0.167$, we obtain $\exp(\gamma) - 1 = -0.154$.

for each 30-minute interval to control for time-specific shocks, such as weather. We clustered the standard errors at the household level to adjust for serial correlation. We included data from the pre-experiment days and treatment days in this regression. Note that treatment effects can have spillover effects on non-treatment days after the beginning of the experimental period. In this case, including non-treatment days (as control days) will underestimate the treatment effects. Therefore, we excluded non-treatment days in this regression. Recall that the treatment groups have explicit incentives to reduce usage only during the treatment hours—1 pm to 4 pm for the summer and 6 pm to 9 pm for the winter. In this regression, we included only these hours to estimate the treatment effects on the treatment hours. We examined potential spillover effects for non-treatment hours in the following subsection.

Columns 1 and 3 of Table 2 show that moral suasion caused a reduction in peak-hour electricity usage by 0.031 log points (3.1 percent) for the summer treatment days and by 0.032 log points (3.2 percent) for the winter treatment days. A reduction in peak hour consumption by 3 percent is economically significant because the marginal cost of electricity is extremely high during critical peak hours.

This finding suggests that the moral suasion policy could have provided a meaningful impact on electricity usage when we considered the average effect across all treatment days.[10]

Nevertheless, the level of the reductions is much larger for the economic incentive treatment. The average treatment effect is 0.167 log points (15.4 percent) for the summer and 0.173 log points (15.9 percent) for the winter. They are statistically different from the effect of moral suasion at the 1 percent significance level. An important question regarding the economic incentive effect is whether consumers 1) merely reacted to "pricing events" and reduced their usage or 2) responded to the changes in marginal price and consumed electricity according to their demand curves. The two possibilities have different policy implications because the latter indicates that policymakers can use price as a tool to achieve certain levels of reductions in peak-hour electricity usage. The answer to this question remains unclear in previous field experiments on electricity demand. This is primarily because most experiments used only one price for critical peak events, and therefore,

[10] We included pre-experimental data to obtain household fixed effects, which improved the efficiency of our estimates. An alternative was to calculate a simple mean difference in log usage between the treatment and control groups during the treatment hours. With this approach, the moral suasion effect is -0.035 (0.016) for summer and -0.034 (0.023) for winter. The economic incentive effect is -0.170 (0.025) for summer and -0.177 (0.027) for winter.

could not disentangle the two possibilities. For example, Wolak (2006, 2011) have one treatment price in each of their two experiments. While Jessoe and Rapson (2014) have different treatment prices, the treatment hours and duration differed between the price regimes, which the authors acknowledge made it difficult to compare treatment effects across different prices. A few recent studies attempt to answer this question, but empirical evidence is limited. For instance, Carroll, Lyons and Denny (2014) estimate the effect of different electricity tariffs on electricity usage for residential customers in Ireland, finding no statistical evidence that customers respond to changes in tariff.

We tested these hypotheses in columns 2 and 4 of Table 2 by estimating a treatment effect for each marginal price. Recall that the baseline marginal price was 25 cents/kWh, and the economic incentive group experienced the three critical peak prices—65, 85, and 105 cents/kWh. The result is consistent with the prediction of the standard demand theory. Consumers reduced usage more in response to higher marginal prices. In the summer, the critical peak prices of 65, 85, and 105 cents/kWh produced reductions in usage by 0.151 log points (14 percent), 0.167 log points (15.4 percent), and 0.182 log points (16.6 percent). The winter results show a similar relationship between the prices and responses. For both seasons, we rejected the equality of the coefficients between 65 cents and 105 cents at the 5 percent statistical significance level—the p-value is 0.043 for the summer and 0.048 for the winter. Although we could not reject the equality of the coefficients between the middle price (85 cents) and other prices at the 5 percent statistical significance level, the point estimates indicate a monotonic relationship between price and response. This finding implies that households indeed responded to marginal prices when they were well informed about their time-varying marginal prices.

This result has an important implication for energy policy because regulators and utility companies often believe that electricity consumers do not respond to electricity prices at all, and therefore, a price-based policy is not a practical solution to mitigate problems on the retail side of electricity markets. In addition, our finding provides new evidence to the growing literature showing that consumers might not necessarily respond to marginal incentives (Borenstein, 2009; Kahn and Wolak, 2013; Ito, 2014; Copeland and Garratt, 2015; McRae and Meeks, 2016). Our result is in contrast to that of Ito (2014), who finds that electricity consumers in California do not respond to the marginal price of their non-linear price schedules but rather respond to the *average* price of their electricity bills. There are two main reasons why we found different evidence in our study. First, customers in our experiment had access to salient information about their real-time marginal price via in-home displays and text messages, whereas Californian customers in Ito's study received their price information only through their monthly bills. Although monthly electricity bills provide information on marginal price, such information is not usually transparent, and consumers receive it with a month lag. The different findings are consistent with the literature that emphasizes the importance of price salience (Chetty, Looney and Kroft, 2009; Finkelstein, 2009; Jessoe and Rapson, 2014). The second key difference is that our customers had a single marginal price within every hour, which varies only across hours, whereas the marginal price in Ito (2014) varies with each customer's cumulative monthly usage. Therefore, the different findings between the two studies could reflect the possibility that consumers are more likely to respond to time-varying marginal prices than marginal prices that vary with their cumulative usage during a month.

4.2 Habituation and Dishabituation

To investigate habituation of treatment effects, we repeated our interventions over 15 treatment days in the summer and 21 treatment days in the winter. Recall that we determined the treatment days by day-ahead weather forecasts, and therefore, they were not necessarily consecutive. As described in Section 3, we divided the 15 summer treatment days into five cycles, and the 21 winter treatment days into seven cycles so that each cycle has 3 treatment days with 65 cents, 85 cents, and 105 cents as the peak-hour prices. We examined the habituation of treatment effects by estimating OLS with treatment cycles $c = 1, .., 5$ for the summer and $c = 1, .., 7$ for the winter:

$$\ln x_{it} = \sum_{c \in C} (\beta_c M_{itc} + \gamma_c E_{itc}) + \theta_i + \lambda_t + \eta_{it}, \tag{2}$$

where β_c and γ_c are the effects of moral suasion and economic incentives for treatment cycle c. Our objective was to test how β_c changed over the repeated interventions.

Table 3 shows the results. Column 1 indicates that the moral suasion effect was statistically significant in the first cycle, but became insignificant in the remaining cycles for the summer. Although it reduced usage by 0.083 log points (8 percent) in the first cycle, this effect diminished quickly in the remaining interventions—the point estimate declined to 0.033 log points (3.2 percent) in the second cycle and nearly 0 in the further interventions. The p-values reported in the bottom of the table show that the treatment effect in the first cycle is statistically different from those in the remaining cycles at the 5 or 10 percent statistical significance level.

In column 3 of Table 3, we tested if the habituated moral suasion effect could be restored back to the original level by giving a sufficient time interval between interventions. In the first cycle of the winter, moral suasion induced a usage reduction by 0.083 log points (8 percent). This impact is identical to the effect found for the first cycle in the summer. It implies that while the moral suasion effect was habituated over repeated interventions in the summer, the impact was reset to the original level—dishabituation—when we restarted the treatment in the winter followed by a 3-month interval after the summer.

This finding is consistent with the theory and laboratory evidence of dishabituation and spontaneous recovery in the psychology literature (Thompson and Spencer, 1966; Rankin et al., 2009). Furthermore, the process of habituation found in our experiment resembles laboratory evidence for many types of species in previous psychology studies. Columns 1 and 3 of Table 3 suggest that the habituation process—how the treatment effect decayed—was similar between the summer and winter. The habituation emerged approximately as an exponential decay of response in both seasons. This exponential decay is consistent with the pattern of habituation found for a variety of species (Groves and Thompson, 1970; Rankin et al., 2009).[11]

Columns 2 and 4 of Table 3 tested habituation for economic incentives. Compared to moral suasion, our economic incentive treatment produced much more persistent effects and exhibited smaller habituation. In the summer experiment, the treatment effect was the largest in the second cycle (0.198 log points; 17.9 percent) and remained fairly stable until the fourth cycle. The effect in the fifth cycle was economically and statistically significant—a usage reduction of 0.127 log points (12 percent)—but the p-value for the difference in the treatment effects between the first and fifth cycles was 0.05. There are two potential reasons for this result. One possibility is that the response to the economic incentive indeed habituated when it came to the fifth cycle in the summer. Another potential reason is that all days in the fifth summer cycle happened to be in

[11] An alternative explanation for the decayed moral suasion effect was that households responded less as a result of a decrease in the strength or severity of the appeals. This is unlikely in our case because 1) the moral suasion message was identical across the repeated interventions and 2) the severity of the situation—limited electricity supply—was fairly consistent during the summer and during the winter.

September. Households in Japan are accustomed to use air conditioning between June and August, but much less in September. This tendency could partially explain why we find a smaller response in the fifth cycle: there was likely to be less discretionary consumption when households did not use much air conditioning. The winter results show similar patterns with even more stable effects of economic incentives across repeated interventions. The effect is the largest in the second cycle (0.205 log points; 18.5 percent) and is stable from the first to seventh cycles. We cannot reject the null hypothesis that the economic incentive effects are equivalent in the first and seventh cycles in the winter.

To compare the treatment effects across cycles visually, we plotted the estimation results in Figure 3. The figure shows reductions in usage by treatment cycles, and the interval bar shows 1 standard error for each treatment effect. Our findings on habituation and dishabituation have four key policy implications, particularly for policymakers aiming to generate persistent policy impacts over repeated interventions, as follows. 1) Both moral suasion and economic incentives are likely to produce sizable policy impacts in the short-run. 2) However, the effect of moral suasion is likely to habituate fast when the intervention is repeated over time. 3) The habituated response to moral suasion can recover back to an original level by providing a sufficient time interval between interventions. Finally, 4) the effect of economic incentives is much less likely to habituate than moral suasion is.

It is useful to compare the policy impacts of our treatments to those found in previous studies. One way to compare the magnitude of the economic incentive effect is to calculate the implied price elasticity. The implied price elasticity for each cycle is between -0.104 and -0.162 for the summer and between -0.137 and -0.167 for the winter. Combining all cycles, the average price elasticities and standard errors are -0.136 (0.017) for the summer and -0.141 (0.018) for the winter. These estimates are close to the price elasticity estimates for residential electricity customers in the literature. For example, Wolak (2011) and Ito (2014) find that the price elasticity for residential customers in Washington DC and California are around -0.1. Because our study and these two studies examined different samples that had different price schedules, it is notable to observe that the estimated price elasticities for residential customers are similar among these studies.

Another useful exercise is to compare our moral suasion effect—about 8 percent reductions in peak-hour electricity usage for the first cycle and 3 percent average reductions for the entire cycles—

to the effects of other non-monetary incentives on energy conservation. Reiss and White (2008) find that public appeals during the California electricity crisis provided about a 7 percent reduction in electricity usage for residential consumers in San Diego.[12] Allcott and Rogers (2014) find that sending a report showing peer comparison of energy usage induced about a 1 percent reduction in electricity usage for residential consumers in the United States. Finally, Schwartz et al. (2013), after sending consumers postcards that stated their electricity usage was being observed, find that the treated customers had a 2.7 percent reduction in electricity usage relative to customers who received nothing.[13]

When interpreting the size of the treatment effect, we want to emphasize again that our households were not a random sample of the population. While observable characteristics are not statistically different between the experimental sample and a random sample of the population in Table 1, there can be unobservable factors that are different between the two groups. For example, our sample households could be interested in smart meter technologies and potentially more willing to respond to our treatments. If that is the case, the treatment effect for our sample can be larger than that for the population. We showed that our estimated price elasticity is similar to that found in previous studies in residential electricity demand, and that our moral suasion effect is close to the effect of similar interventions conducted in the past. This comparison provided suggestive evidence that our experimental sample is unlikely to be substantially different from the population in terms of responses to treatment. However, this is merely suggestive evidence, and we emphasize that our experiment cannot be entirely free from external validity issues.

4.3 Spillover Effects for Non-treatment Hours on Treatment Days

We specifically targeted the treatments in our experiments at electricity usage during the treatment hours—1 pm to 4 pm for the summer and 6 pm to 9 pm for the winter. For a few reasons, however, the treatments can generate spillover effects for electricity usage in the non-treatment hours on the treatment days. First, households might change their usage immediately before or after the

[12] During the California electricity crisis in 2000–2001, there was an economic incentive (a spike in electricity prices) right before the conservation campaign period. Therefore, there is a possibility that part of the 7 percent effect might include a potential persistent effect of the economic incentive.

[13] In our experiment, both the control and treatment groups received smart meters and in-home displays. Therefore, the effect of being observed was likely to be captured by the control group. However, if there was a Hawthorne effect of getting a message about moral suasion or price changes, such an effect was not fully captured by the control group and included in our treatment effects.

treatment hours. For example, those who had a high critical peak price for the treatment hours could increase their air conditioner usage immediately before the treatment hours. Such pre-cooling or pre-heating could be rational given the high critical peak prices for the treatment hours. Similarly, households might increase their air conditioner usage immediately after the treatment hours. Hours immediately before and after the peak hours are called "shoulder hours." In general, when the marginal cost of electricity supply is high during peak-demand hours, the marginal cost is also likely to be relatively high during shoulder hours. Therefore, if customers were to increase their usage in the shoulder hours, this could attenuate the economic benefits of interventions focused on peak-demand hours.[14] This is also an important question for environmental policy because increases in usage in non-treatment hours could increase the total emissions from electricity generation (Holland and Mansur, 2008).

The second possibility is that consumers might shift their consumption to off-peak hours, which are hours outside peak hours and shoulder hours. In most electricity markets, the marginal cost of electricity is much lower in off-peak hours relative to peak-demand hours. Therefore, such consumption shifting is still likely to provide a meaningful economic benefit. Finally, the third possibility is that consumers could reduce their usage in all hours, including shoulder hours and off-peak hours. For example, consider that there is a fixed cost for consumers to change their lifestyles in terms of electricity usage (Wolak, 2011). When consumers face a substantial increase in peak-hour electricity prices, they might change their lifestyles as a whole to be more energy efficient. In such cases, it is possible that customers lower their electricity usage in all hours when they face interventions that are targeted primarily at peak-demand hours.

Table 4 provides the results of empirical tests for these possibilities. We estimated equation (1) by including data for different hours on the treatment days for each column. Column 1 shows the results for the treatment hours, which are equivalent to the results in Table 2. Column 2 shows the results for the shoulder hours, that is, 3 hours before and after the treatment hours. Finally, column 3 includes data for other non-treatment hours on the treatment days. For both the moral suasion group and the economic incentive group, we find no increase in consumption either in the shoulder hours or in other off-peak hours. Instead, we found usage reductions for the economic

[14] To address this concern, policymakers can design dynamic pricing schedules that include relatively high prices for shoulder hours in addition to high prices for peak-demand hours, which could reflect the time-varying marginal costs of electricity supply more effectively.

incentive group during the non-treatment hours. For example, we found usage reductions of 0.060 log points (5.8 percent) for the shoulder hours and by 0.022 log points (2.2 percent) for the off-peak hours in the summer experiment. The findings for the winter experiment are consistent with those for the summer experiment. By contrast, we found no such spillover effects for the moral suasion group. This group's usage in the shoulder hours and off-peak hours is statistically indistinguishable from the control group's usage. These results imply that the economic incentives in our experiment motivated customers to lower their usage in both the non-treatment hours and the treatment hours.

4.4 Habit Formation

In the previous subsections, we found that moral suasion and economic incentives produced substantially different results in terms of habituation, dishabituation, and spillover effects during non-treatment hours. In this subsection, we tested for habit formation (Becker and Murphy, 1988)—customers faced with our interventions might have formed a habit for efficient use of energy. To test for habit formation, we collected data for the periods *after* the treatment was withdrawn. During this post-intervention period, households did not receive any treatment. We expected that unless our treatment induced habit formation, there should be the same levels of consumption between the control group and each treatment group during this period.

Table 5 shows the results. We examined usage data during the 3-month post-experimental period for both the summer and winter experiments, wherein customers did not receive any treatment. We tested whether usage in peak-demand hours during this period differed between the control and each treatment group. The table shows that the moral suasion group's usage is not statistically different from that of the control group. By contrast, the economic incentive group's consumption is statistically different from that of the control group as well as the moral suasion group. This result implies that consumers who received our economic incentives continued to have lower consumption of 0.077 log points (7.4 percent) after the summer experiment and 0.069 log points (6.7 percent) after the winter experiment.

This finding has an important policy implication because the existence of habit formation could offer additional policy impacts for post-intervention periods. In the literature, habit formation has been studied largely for "bad" habits. Our finding is among recent studies that investigate the formation of "good" habits. For example, Charness and Gneezy (2009) find that providing monetary incentive for exercising induced a habit formation of exercise for college students on campus. In addition, a few studies explore potential habit formation for the treatment effect of peer comparison in the context of water and energy conservation. For instance, Ferraro, Miranda and Price (2011) and Allcott and Rogers (2014) find that providing information about neighbors' water usage and electricity usage, respectively, induced conservation effects that lasted even after the intervention was withdrawn. Our finding is consistent with the empirical evidence in Ito (2015), which finds that residential electricity customers in California exhibited similar habit formation of electricity consumption in response to a short-run economic incentive of a subsidy program. Why do we observe habit-formation effects only for the economic incentive group and not for the moral suasion group? We explore the mechanism behind these findings in the final subsection of our empirical analyses. Before we proceed to analyzing the mechanism, in the next subsection, we investigate potential heterogeneity in the treatment effects.

4.5 Mechanisms behind the Treatment Effects

We found significant differences in the effects of moral suasion and economic incentives in their overall impacts, habituation, spillovers to non-treatment hours, and habit formation. Moral suasion was effective only for the first few treatment days and habituated quickly over repeated interventions. By contrast, economic incentives produced strong persistent effects on energy conservation. To investigate the mechanisms behind these findings, we conducted a detailed follow-up survey. We examined two potential mechanisms.

The first possibility is that the treatments might have induced investment in physical capital stock—households might have purchased energy-efficient appliances in response to the treatments. If this effect were systematically large for the economic incentive group, it could explain why we observed less habituation and stronger habit formation for this group. The second possibility is that the treatments might have induced new utilization habits for daily electricity use. Suppose that customers had "bad habits" of inefficient or wasteful energy use. It is possible that experiencing high electricity prices triggered a change in their utilization habits, encouraging them to use electricity more efficiently.

Table 6 tests the first possibility—investment in physical capital stock. We asked customers if they purchased energy-efficient appliances since the start of the experiment. We estimated a linear probability model, which includes a binary choice dependent variable, dummy variables for the treatment groups as independent variables, and a constant term.[15] The constant term gives the ratio of customers in the control group who each purchased an energy-efficient appliance. The coefficients for the treatment dummy variables represent percentage-point increases for the treatment groups.

The table suggests that moral suasion increased the purchase of energy-efficient air conditioners by 8 percentage points, whereas economic incentives increased it by 9 percentage points. The estimates suggest that customers in the two treatment groups had similar statistically significant increases in purchasing energy-efficient air conditioners compared to the control group. We found no statistically significant effects for other products for both treatment groups. The results suggest that investment in physical capital stock is unlikely to explain why we found significant differences in the persistent effects between moral suasion and economic incentives.[16]

Table 7 explores the second potential mechanism—behavioral changes in utilization habits. After the experimental period, we asked customers two questions related to this point. The first question inquired about their efforts toward adopting an energy-efficient lifestyle. Customers evaluated their lifestyles in terms of energy efficiency on a scale of 1 (lowest) to 5 (highest). We regressed this score on the dummy variable for each treatment group and a constant term. Column 1 implies that the economic incentive increased this score by 0.4 from the baseline level of 3.03. We found a slightly positive effect for the moral suasion group, but it is statistically insignificant. The difference between the coefficients for moral suasion and economic incentives (0.13 and 0.40) is statistically significant at the 1 percent significance level.

We then asked consumers whether they were using each electric appliance in an energy-efficient way. We asked this question for air conditioners, electric heaters, personal computers, washers, and vacuum cleaners. We estimated a linear probability model, in which the dependent variable is binary choice, and the independent variables include dummy variables for each treatment group. In addition, the model contains a constant term. For each appliance, we found that the economic incentive had a statistically significant effect of 8 to 15 percentage points. By contrast, moral

[15] In addition, for robustness checks, we run probit and logit models and found the same results.

[16] An important caveat of this analysis is that customers knew that they were going to receive the treatments only during the experimental period. If customers had experienced their treatments for a longer time, more consumers might have found investment in physical capital stock economical.

suasion had no statistically significant effects on the energy-efficient use of each appliance.

Although stated survey responses inform these results, they provide suggestive evidence about the mechanisms behind our findings. Investment in physical capital stock is unlikely to explain the significant differences in the persistent effects between the two treatments. Instead, experiencing high electricity prices is likely to trigger a change in utilization habit for electricity use, encouraging customers to use electricity more efficiently.

5 Conclusion

In this study, we used a randomized controlled trial in residential electricity demand to investigate whether moral suasion and economic incentives generated persistent impacts on economic activities. Using high-frequency electricity-usage data at the household level, we found that moral suasion induced a short-run reduction in peak-hour electricity usage by 8 percent. This short-run effect is economically significant for improving economic efficiency in electricity markets. However, the response to moral suasion habituated quickly—the treatment effect diminished toward zero when we repeated the intervention over time. In addition, we found evidence of dishabituation—the habituated response could be restored back to its original level when we provided a sufficiently long time between interventions. However, the recovered response once again habituated when we repeated the intervention frequently.

We compared the impact of moral suasion to that of a more conventional policy tool—economic incentives for consumers to reduce energy usage during peak-demand hours. The economic incentive treatment produced usage reductions of 14 percent for the lowest treatment price and usage reductions of 17 percent for the highest treatment price. Moreover, compared to the impact of moral suasion, the effect was more persistent over repeated interventions, suggesting little habituation. In addition, the economic incentive induced habit formation—households continued to conserve electricity even after the treatment was withdrawn. Our follow-up survey data indicated that most of the persistent changes were likely to originate from changes in utilization habit of electricity usage rather than investment in physical capital stock.

As described in the experimental design section, an important limitation of our experiment is that subjects were not a random sample of the population. We showed that observable characteristics are statistically indistinguishable between our experimental sample and a random sample of the population. We also presented that our estimated price elasticity is similar to that found in previous experiments, and that our moral suasion effect is close to the effect of similar interventions conducted in the past, which provided suggestive evidence that our experimental sample is unlikely to be substantially different from the population in terms of responses to treatment. However, this is merely suggestive evidence, and we emphasize that our experiment cannot be entirely free from external validity issues.

Our results suggest that economic policy can induce significantly different welfare effects depending on whether policymakers use moral suasion or economic incentives to promote pro-social behavior. To illustrate this implication, we conduct a welfare analysis in the online appendix. One of the largest inefficiencies in electricity markets is that retail electricity prices usually do not reflect the marginal cost of electricity—consumers pay time-invariant prices while the marginal cost of electricity is time-variant (Wolak, 2011; Joskow, 2012; Joskow and Wolfram, 2012; Jessoe and Rapson, 2014). Our analysis shows that moral suasion can provide significant welfare gains in the short run, but such gains are likely to diminish rapidly when the intervention is repeated over time. By contrast, economic incentives produce greater welfare improvement, particularly when long-run effects, such as habituation and habit formation, are taken into account.[17]

References

Allcott, Hunt. 2015. "Site Selection Bias in Program Evaluation." *The Quarterly Journal of Economics*, 130(3): 1117–1165.

Allcott, Hunt, and Judd B. Kessler. 2015. "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons." National Bureau of Economic Research NBER Working Paper 21671.

Allcott, Hunt, and Todd Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review*, 104(10): 3003–37.

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97(6): 1447–58.

[17]A key assumption in our welfare analysis is that consumers gain positive utility (e.g. warm glow) from their responses to non-pecuniary interventions. This assumption can be violated if such interventions were to create negative utility (e.g. social pressure). This challenge suggests that an important future research question is how to disentangle such possibilities to provide precise analysis for welfare implications of non-pecuniary interventions. See DellaVigna, List and Malmendier (2012) and Allcott and Kessler (2015) for more discussions on this topic.

Becker, Gary S., and Kevin M. Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy*, 96(4): 675–700.

Bénabou, Roland, and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *The Review of Economic Studies*, 70(3): 489–520.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–1678.

Borenstein, Severin. 2002. "The Trouble With Electricity Markets: Understanding California's Restructuring Disaster." *Journal of Economic Perspectives*, 16(1): 191–211.

Borenstein, Severin. 2009. "To What Electricity Price Do Consumers Respond? Residential Demand Elasticity Under Increasing-Block Pricing." *Working Paper*.

Carroll, James, Seán Lyons, and Eleanor Denny. 2014. "Reducing household electricity demand through smart metering: The role of improved information about energy saving." *Energy Economics*, 45: 234–243.

Charness, Gary, and Uri Gneezy. 2009. "Incentives to Exercise." *Econometrica*, 77(3): 909–931.

Chetty, Raj, and Emmanuel Saez. 2013. "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients." *American Economic Journal: Applied Economics*, 5(1): 1–31.

Chetty, Raj, Emmanuel Saez, and László Sándor. 2014. "What Policies Increase Prosocial Behavior? An Experiment with Referees at the Journal of Public Economics." *Journal of Economic Perspectives*, 28(3): 169–188.

Chetty, R., A. Looney, and K. Kroft. 2009. "Salience and taxation: Theory and evidence." *The American Economic Review*, 99(4): 1145–1177.

Copeland, Adam M., and Rodney J. Garratt. 2015. "Nonlinear Pricing with Competition: The Market for Settling Payments." *FRB of New York Staff Report*, 737.

Costa, Francisco, and François Gerard. 2015. "Hysteresis and the Social Cost of Corrective Policies: Evidence from a Temporary Energy Saving Program." Working Paper.

Cutter, W. Bowman, and Matthew Neidell. 2009. "Voluntary information programs and environmental regulation: Evidence from 'Spare the Air'." *Journal of Environmental Economics and Management*, 58(3): 253–265.

Dal Bó, Ernesto, and Pedro Dal Bó. 2014. ""Do the right thing:" The effects of moral suasion on cooperation." *Journal of Public Economics*, 117: 28–38.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier. 2012. "Testing for Altruism and Social Pressure in Charitable Giving." *The Quarterly Journal of Economics*, 127(1): 1–56.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Chapter 61 Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*. Vol. Volume 4, , ed. T. Paul Schultz and John A. Strauss, 3895–3962. Elsevier.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *American Economic Journal: Economic Policy*, 8(3): 203–232.

Faruqui, Ahmad, and Sanem Sergici. 2011. "Dynamic pricing of electricity in the mid-Atlantic region: econometric results from the Baltimore gas and electric company experiment." *Journal of Regulatory Economics*, 40(1): 82–109.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler. 2013. "Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information." *Journal of the European Economic Association*, 11(3): 634–660.

Ferraro, Paul J., and Michael K. Price. 2013. "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment." *Review of Economics and Statistics*, 95(1): 64–73.

Ferraro, Paul J, Juan Jose Miranda, and Michael K Price. 2011. "The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment." *American Economic Review*, 101(3): 318–322.

Finkelstein, Amy. 2009. "E-ZTAX: Tax Salience and Tax Rates." *The Quarterly Journal of Economics*, 124(3): 969–1010.

Gerard, Francois. 2013. "The Impact and Persistence of Ambitious Energy Conservation Programs: Evidence from the 2001 Brazilian Electricity Crisis." Social Science Research Network SSRN Scholarly Paper ID 2097195, Rochester, NY.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives*, 25(4): 191–210.

Groves, P. M., and R. F. Thompson. 1970. "Habituation: a dual-process theory." *Psychological Review*, 77(5): 419–450.

Holland, Stephen P., and Erin T. Mansur. 2008. "Is Real-Time Pricing Green? The Environmental Impacts of Electricity Demand Variance." *Review of Economics and Statistics*, 90(3): 550–561.

Ida, Takanori, Kosuke Takemura, and Masayuki Sato. 2015. "Inner conflict between nuclear power generation and electricity rates: A Japanese case study." *Energy Economics*, 48: 61–69.

Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review*, 104(2): 537–63.

Ito, Koichiro. 2015. "Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program." *American Economic Journal: Economic Policy*, 7(3): 209–237.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review*, 104(4): 1417–38.

Joskow, Paul L. 2012. "Creating a Smarter U.S. Electricity Grid." *Journal of Economic Perspectives*, 26(1): 29–48.

Joskow, Paul L., and Catherine D. Wolfram. 2012. "Dynamic Pricing of Electricity." *American Economic Review*, 102(3): 381–85.

Kahn, Matthew E., and Frank A. Wolak. 2013. "Using Information to Improve the Effectiveness of Nonlinear Pricing: Evidence from a Field Experiment : Final Report." Stanford University Working Paper.

Kingma, Bruce Robert. 1989. "An accurate measurement of the crowd-out effect, income effect, and price effect for charitable contributions." *The Journal of Political Economy*, 97(5): 1197–1207.

Kotchen, Matthew J. 2006. "Green markets and private provision of public goods." *Journal of Political Economy*, 114(4): 816–834.

Kotchen, Matthew J., and Michael R. Moore. 2007. "Private provision of environmental public goods: Household participation in green-electricity programs." *Journal of Environmental Economics and Management*, 53(1): 1–16.

Kreps, David M. 1997. "Intrinsic Motivation and Extrinsic Incentives." *The American Economic Review*, 87(2): 359–364.

Lacetera, Nicola, Mario Macis, and Robert Slonim. 2012. "Will There Be Blood? Incentives and Displacement Effects in Pro-Social Behavior." *American Economic Journal: Economic Policy*, 3(4): 186–223.

Landry, Craig E, Andreas Lange, John A List, Michael K Price, and Nicholas G Rupp. 2010. "Is a Donor in Hand Better than Two in the Bush? Evidence from a Natural Field Experiment." *American Economic Review*, 100(3): 958–983.

McRae, Shaun, and Robyn Meeks. 2016. "Price Perceptions and Electricity Demand with Nonlinear Prices." University of Michigan Working Paper.

Phelps, Brady I. 2011. "Dishabituation." In *Encyclopedia of Child Behavior and Development.* , ed. Sam Goldstein and Jack A. Naglieri, 509–512. Springer US. DOI: 10.1007/978-0-387-79061-9_867.

Rankin, Catharine H., Thomas Abrams, Robert J. Barry, Seema Bhatnagar, David Clayton, John Colombo, Gianluca Coppola, Mark A. Geyer, David L. Glanzman, Stephen Marsland, Frances McSweeney, Donald A. Wilson, Chun-Fang Wu, and Richard F. Thompson. 2009. "Habituation Revisited: An Updated and Revised Description of the Behavioral Characteristics of Habituation." *Neurobiology of learning and memory*, 92(2): 135–138.

Rankin, C. H., and T. J. Carew. 1988. "Dishabituation and sensitization emerge as separate processes during development in Aplysia." *The Journal of Neuroscience*, 8(1): 197–211.

Reiss, Peter C., and Matthew W. White. 2008. "What changes energy consumption? Prices and public pressures." *RAND Journal of Economics*, 39(3): 636–663.

Rozen, Kareen. 2010. "Foundations of Intrinsic Habit Formation." *Econometrica*, 78(4): 1341–1373.

Schwartz, Daniel, Baruch Fischhoff, Tamar Krishnamurti, and Fallaw Sowell. 2013. "The Hawthorne effect and energy awareness." *Proceedings of the National Academy of Sciences*, 110(38): 15242–15246.

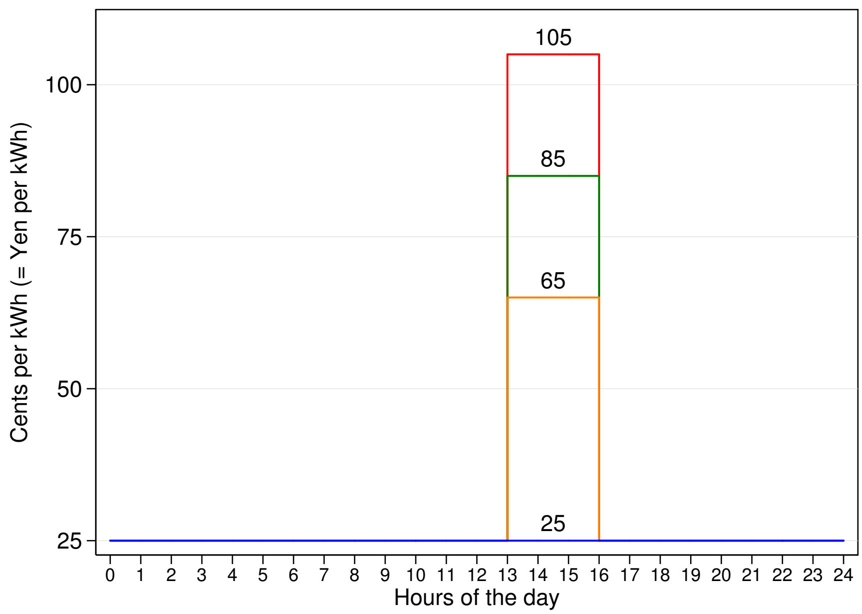
Thompson, R. F., and W. A. Spencer. 1966. "Habituation: a model phenomenon for the study of neuronal substrates of behavior." *Psychological Review*, 73(1): 16–43.

Volpp, Kevin G., Andrea B. Troxel, Mark V. Pauly, Henry A. Glick, Andrea Puig, David A. Asch, Robert Galvin, Jingsan Zhu, Fei Wan, Jill DeGuzman, Elizabeth Corbett, Janet Weiner, and Janet Audrain-McGovern. 2009. "A Randomized, Controlled Trial of Financial Incentives for Smoking Cessation." *New England Journal of Medicine*, 360(7): 699–709.

Wolak, Frank A. 2006. "Residential customer response to real-time pricing: the Anaheim critical-peak pricing experiment." *Working Paper.*

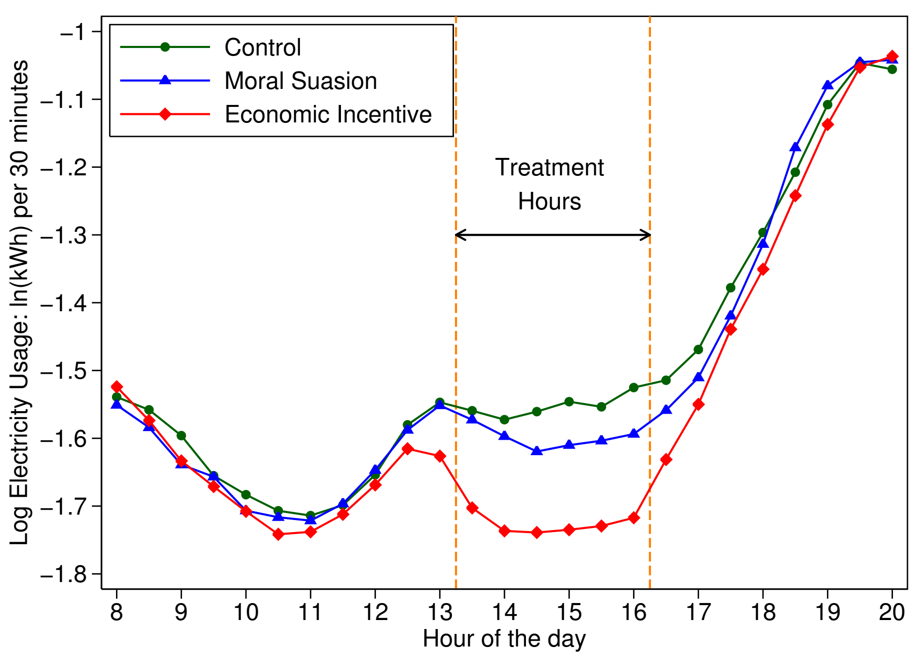
Wolak, Frank A. 2011. "Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment." *The American Economic Review*, 101(3): 83–87.

Figure 1: Economic Incentives: Dynamic Electricity Pricing



Notes: This figure shows the dynamic electricity pricing schedule for the economic incentive group and the baseline price (25 cents/kWh). Although our participants paid in Japanese yen, we use US. currency throughout the paper. One Japanese yen was approximately equivalent to one US. cent in 2012.

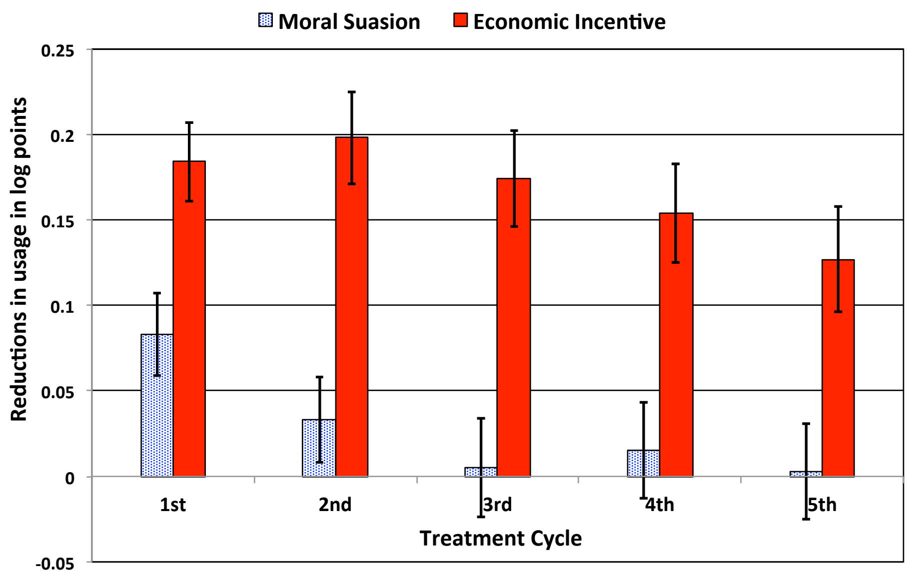
Figure 2: Effects of Moral Suasion and Economic Incentives on Electricity Usage



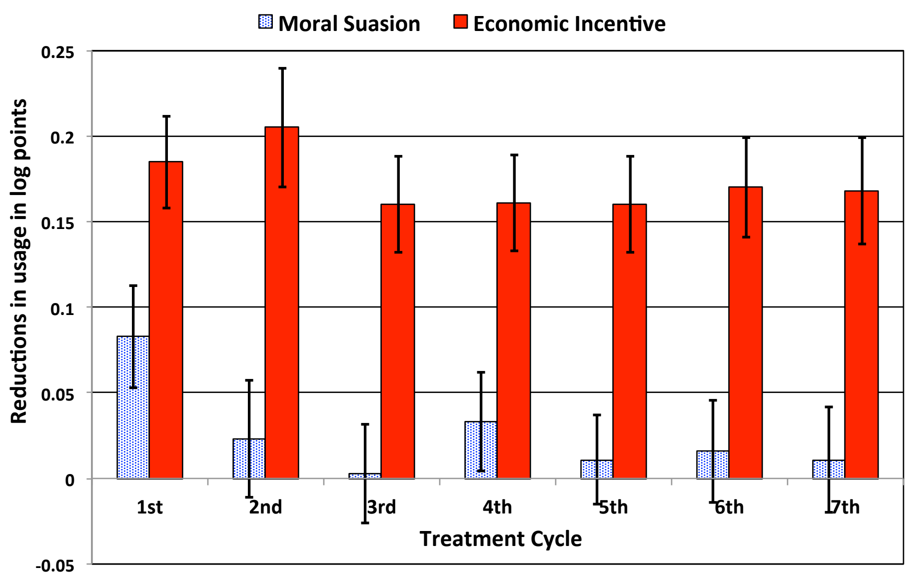
Notes: This figure shows the mean of log electricity usage (kWh) for 30-minute intervals by treatment groups for the summer treatment days. We calculate the mean log usage using data from all treatment days in the summer.

Figure 3: Treatment Effects by Treatment Cycles

Panel A: Summer Experiment



Panel B: Winter Experiment



Notes: This figure shows the treatment effects of moral suasion and economic incentives by treatment cycles in terms of the reductions in electricity usage in log points. The estimates are obtained from the estimation results in Table 3. The interval bars show one standard error of the treatment effect. In the estimation, we include household fixed effects and time fixed effects for each 30-minute interval. The standard errors are clustered at the household level to adjust for serial correlation. Each cycle includes three treatment days. There were 15 treatment days in the summer and 21 treatment days in the winter.

Table 1: Summary Statistics

	Sample in the field experiment			Random sample of population	Difference between sample and population
	Moral suasion (M)	Economic incentive (E)	Control group (C)	(P)	
Electricity use (kWh/day)	15.14	15.76	15.92	16.23	-0.45
	(6.91)	(8.49)	(8.47)	(7.97)	[0.62]
Household income (1,000USD)	66.74	66.59	67.06	66.83	-1.69
	(31.49)	(31.34)	(31.01)	(41.81)	[3.68]
Square meter of the house	121.49	113.08	122.15	125.90	-8.95
	(57.54)	(46.92)	(46.52)	(59.65)	[4.66]
Number of AC	3.46	3.50	3.68	3.95	-0.43
	(1.93)	(1.67)	(1.64)	(1.71)	[0.23]
Mean age of the household	42.26	42.22	40.31	41.91	-0.11
	(17.67)	(19.07)	(17.38)	(16.76)	[0.48]
Age of building (years)	13.83	13.39	13.12	15.05	-1.62
	(8.25)	(7.54)	(8.20)	(8.11)	[0.28]
Household Size	3.21	3.14	3.32	2.98	0.22
	(1.18)	(1.23)	(1.25)	(1.41)	[0.09]

Notes: The first three columns show the sample mean and standard deviation of observables by treatment group. Because of the random assignment, the observables are balanced across the groups. Column 4 shows the mean and standard deviation of observables for a random sample of the population in the experimental area (randomly sampled 717 households). Column 5 presents the difference in the means between the sample for the field experiment and the random sample of the population. Standard deviations are in parentheses in columns 1 to 4, and standard errors are in brackets in column 5.

Table 2: Effects of Moral Suasion and Economic Incentives on Electricity Usage

	Summer		Winter	
	(1)	(2)	(3)	(4)
Moral suasion	-0.031	-0.031	-0.032	-0.032
	(0.014)	(0.014)	(0.020)	(0.020)
Economic incentive	-0.167		-0.173	
	(0.021)		(0.022)	
Economic incentive (price = 65)		-0.151		-0.163
		(0.022)		(0.024)
Economic incentive (price = 85)		-0.167		-0.164
		(0.023)		(0.023)
Economic incentive (price = 105)		-0.182		-0.189
		(0.024)		(0.024)
Observations	123106	123106	244891	244891

Notes: This table shows the estimation results for equation (1) for the treatment hours. The dependent variable is the log of household-level 30-minute interval electricity consumption. We include household fixed effects and time fixed effects for each 30-minute interval. The standard errors are clustered at the household level to adjust for serial correlation. The difference between the coefficients for 65 and 105 cents is statistically significant at the 5 percent level. The implied price elasticity estimates are -0.136 (0.017) for the summer and -0.141 (0.018) for the winter.

Table 3: Repeated Interventions: Habituation and Dishabituation of Treatment Effects

	Summer		Winter	
	Moral Suasion (β_c)	Economic Incentive (γ_c)	Moral Suasion (β_c)	Economic Incentive (γ_c)
1st cycle	-0.083	-0.184	-0.083	-0.185
	(0.024)	(0.023)	(0.030)	(0.027)
2nd cycle	-0.033	-0.198	-0.023	-0.205
	(0.025)	(0.027)	(0.034)	(0.035)
3rd cycle	-0.005	-0.174	0.003	-0.160
	(0.029)	(0.028)	(0.029)	(0.028)
4th cycle	-0.015	-0.154	-0.033	-0.161
	(0.028)	(0.029)	(0.029)	(0.028)
5th cycle	-0.003	-0.127	-0.011	-0.160
	(0.028)	(0.031)	(0.026)	(0.028)
6th cycle			-0.016	-0.170
			(0.030)	(0.029)
7th cycle			-0.011	-0.168
			(0.031)	(0.031)
p-values of the differences in the treatment effects relative to the effects in the 1st cycle				
2nd cycle	0.075	0.474	0.124	0.522
3rd cycle	0.024	0.678	0.026	0.394
4th cycle	0.054	0.120	0.194	0.428
5th cycle	0.030	0.050	0.041	0.409
6th cycle			0.080	0.626
7th cycle			0.069	0.608

Notes: This table shows the estimation results for equation (2). The dependent variable is the log of household-level 30-minute interval electricity consumption. We include household fixed effects and time fixed effects for each 30-minute interval. The standard errors are clustered at the household level to adjust for serial correlation. Each cycle includes three treatment days. There were 15 treatment days in the summer and 21 treatment days in the winter. The treatment days were not necessarily consecutive.

Table 4: Spillover Effects for Non-treatment Hours on Treatment Days

	Summer			Winter		
	Treatment Hours (1pm-4pm)	Shoulder Hours (10am-1pm, 4pm-7pm)	Other Hours	Treatment Hours (6pm-9pm)	Shoulder Hours (3pm-6pm, 9pm-12pm)	Other Hours
	(1)	(2)	(3)	(4)	(5)	(6)
Moral suasion	-0.031	-0.010	-0.008	-0.032	-0.009	-0.007
	(0.014)	(0.010)	(0.005)	(0.020)	(0.015)	(0.012)
Economic incentive	-0.167	-0.060	-0.022	-0.173	-0.034	-0.007
	(0.021)	(0.015)	(0.010)	(0.022)	(0.017)	(0.014)
Observations	123106	248621	634387	244891	482902	1182574

Notes: This table shows the estimation results for equation (1) for the treatment hours and other hours on the treatment days. The shoulder hours are three hours before and after the treatment hours. Columns 3 and 6 include non-treatment hours except for the shoulder hours. The dependent variable is the log of household-level 30-minute interval electricity consumption. We include household fixed effects and time fixed effects for each 30-minute interval. The standard errors are clustered at the household level to adjust for serial correlation.

Table 5: Habit Formation After the Treatments Were Withdrawn

	After Summer Experiment	After Winter Experiment
	(1)	(2)
Moral suasion	0.006	0.021
	(0.028)	(0.026)
Economic incentive	-0.077	-0.069
	(0.034)	(0.022)
Observations	426770	478605

Notes: This table shows the estimation results for equation (1) for the three-month period after we withdrew our treatments. Column 1 shows the result for usage in peak demand hours (1pm to 4 pm) after the summer experiment. Column 2 shows the result for usage in peak demand hours (6pm to 9 pm) after the winter experiment. The dependent variable is the log of household-level 30-minute interval electricity consumption. We include household fixed effects and time fixed effects for each 30-minute interval. The standard errors are clustered at the household level to adjust for serial correlation.

Table 6: Treatment Effects on Investments in Physical Capital Stock

	Dependent variable: binary choice				
	Room AC	Refrigerator	Washer	Electric fan	Light bulb
	(1)	(2)	(3)	(4)	(5)
Moral suasion	0.08	0.01	0.01	-0.00	0.03
	(0.04)	(0.03)	(0.03)	(0.05)	(0.05)
Economic incentive	0.09	-0.01	0.01	-0.01	-0.03
	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)
Constant	0.06	0.08	0.05	0.23	0.29
	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)
Observations	640	640	640	640	640

Notes: We asked customers if they purchased an energy-efficient appliance since the experiment started. We estimate a linear probability model, with a binary choice dependent variable, dummy variables for the two treatment groups as independent variables, and a constant term. The constant term, therefore, provides the ratio of control customers who purchased an energy-efficient appliance. The coefficients for the group dummy variables provide a percentage point increase for the group. The robust standard errors are in parentheses. We had 51 customers who did not respond to this question. However, the number of non-response is balanced across the three groups.

Table 7: Treatment Effects on Utilization Habits

	Energy-efficient lifestyle	Energy-efficient use of appliances (Dependent variable: binary choice)				
	(Degree: 1 to 5)	AC	Heater	PC	Washer	Cleaner
	(1)	(2)	(3)	(4)	(5)	(6)
Moral suasion	0.13	-0.00	0.08	0.01	-0.03	-0.03
	(0.08)	(0.06)	(0.06)	(0.05)	(0.04)	(0.04)
Economic incentive	0.40	0.13	0.15	0.09	0.08	0.12
	(0.07)	(0.05)	(0.05)	(0.04)	(0.03)	(0.04)
Constant	3.03	0.61	0.53	0.11	0.08	0.07
	(0.06)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Observations	626	626	626	626	626	626

Notes: After the experimental period, we asked customers two questions. The first question was whether they were trying to have an energy-efficient lifestyle. Customers rated their lifestyles on a scale of 1 (lowest) to 5 (highest) in terms of energy efficiency of their lifestyles. We regress this score on the dummy variable for each treatment group and a constant term. Second, we asked consumers whether they were using each of the following electric appliances in an energy-efficient way: air conditioners, electric heaters, personal computers, washers, and vacuum cleaners. We estimate a linear probability model, which includes a binary choice dependent variable, dummy variables for the two treatment groups as the independent variables, and a constant term. The robust standard errors are in parentheses. We had 65 customers who did not respond to this question. However, the number of non-response is balanced across the three groups.

Online Appendices (Not for Publication)

Online Appendix A: Welfare Implications

When pursuing a variety of policy goals, policymakers can design policies to influence intrinsic and extrinsic motivations. Our empirical findings suggest that these two policy instruments are likely to have different policy implications, particularly when we consider persistence. In this section, we highlight such policy implications by analyzing the welfare gains from the two policies in the context of electricity markets.

Conceptual Framework

We introduce a simple conceptual framework for a model of electricity consumers to guide our welfare analysis. When consumers receive no treatment, each consumer uses electricity $\bar{x}$ at a given power price P, where $\bar{x}$ can be regarded as a "business as usual" (BAU) consumption level. When they receive moral suasion for conserving energy, they may voluntarily decrease their consumption from $\bar{x}$ to x. Voluntary conservation of electricity, g, is then expressed as the difference between $\bar{x}$ and x. The saved amount in economic terms, Pg, is added to the numeraire y, which totals $Y = y + Pg$. Alternatively, $Y = I - Px$ from the budget constraint of a consumer with income I.

We assume that utility is additively separable into three components. The first term, $u(x)$, denotes utility from consuming electricity, which is assumed to be increasing and concave ($u' > 0$ and $u'' < 0$). The second term, $v(I - Px)$, is utility from numeraire consumption, which is assumed to be increasing and weakly concave ($v' > 0$ and $v'' \leq 0$). Lastly, we consider utility from conservation of electricity, $\phi(g; \theta)$. This is also assumed to be increasing and weakly concave ($\phi_g = \frac{\partial \phi}{\partial g} > 0$ and $\phi_{gg} = \frac{\partial^2 \phi}{\partial g^2} \leq 0$). The utility term $\phi(g; \theta)$ may represent a warm glow component, which is a type of impure altruism, as discussed by Andreoni (1989). Let θ be a parameter that represents the frequency of interventions. We assume that utility and marginal utility of electricity conservation are decreasing in the frequency of interventions, that is, $\phi_\theta = \frac{\partial \phi}{\partial \theta} < 0$ and $\phi_{g\theta} = \frac{\partial^2 \phi}{\partial \theta \partial g} < 0$. The subscript notation denotes a partial derivative.

Andreoni (1989) and Kingma (1989) argue that there are several competing theoretical models of charitable contributions. In the case of pure altruism (pure public good), consumers may care about the total contributions to voluntary conservation. Moreover, consumers may take account of the utility cost (disutilily) of social pressure for not contributing or only contributing a small amount toward voluntary conservation, as illustrated by DellaVigna, List and Malmendier (2012). It is not our primary focus to compare these competing models, but note that we can extend our simple model to incorporate other potential mechanisms behind contributions to voluntary conservation.[18]

The BAU consumption level, $\bar{x}$, in the absence of treatment can be expressed by $\bar{x} = \arg\max\{u(x) + v(I - Px)\}$. Consumers in the economic incentive group have a price change and simply adjust their consumption such that $u' - Pv' = 0$ responding to the price changes. Consumers in the moral suasion group receive moral suasion without economic incentives. When they receive moral suasion, they maximize the following overall utility function:

$$\begin{aligned} &\max_{x,g}\ u(x) + v(I - Px) + \phi(g; \theta) \\ &\text{s.t. } g = \bar{x} - x. \end{aligned} \tag{3}$$

[18] For example, Kotchen (2006); Kotchen and Moore (2007) consider different participation mechanisms for environmental public goods and show how they relate to existing theory on either *pure* or *impure* public goods.

This problem can be rewritten as follows:

$$\max_{x} \ u(x) + v(I - Px) + \phi(\bar{x} - x; \theta). \tag{4}$$

Let x^* denote the optimal solution for the maximization problem (4), namely, the optimal consumption level under moral suasion. Note that x^* satisfies $u' - Pv' - \phi_g = 0$.[19]

The effect of repeated interventions on voluntary conservation can be easily derived by differentiating the first order condition for the optimization problem (4).[20] Simple calculation yields

$$g^*_\theta = \bar{x}_\theta - x^*_\theta = -x^*_\theta = -\frac{\phi_{g\theta}}{u'' + P^2 v'' + \phi_{gg}} < 0. \tag{5}$$

The optimal consumption of electricity is increasing in θ, that is, $x^*_\theta > 0$, while the BAU consumption level $\bar{x}$ is not affected by θ, that is, $\bar{x}_\theta = 0$. Therefore, the model suggests that repeated interventions may decrease voluntary conservation of electricity, which is consistent with our empirical findings. In the next subsection, we use this conceptual framework to highlight the welfare implications of our empirical findings.

Welfare Gains from the Two Policies

We examine the welfare implications of two policy instruments that are intended to reduce energy usage during peak demand hours: 1) moral suasion and 2) economic incentives. Recall that the fundamental inefficiency in electricity markets is that consumers do not pay time-varying prices for electricity. Thus, they do not have an incentive to use less energy when the marginal cost becomes very high during peak demand hours. We begin with the assumption that the marginal cost of electricity for the critical peak hours is 65 cents/kWh, which was the peak wholesale price in the Japanese wholesale electricity market, the Japan Electric Power Exchange, during our experimental period. For a few reasons, this number is likely to be a lower bound for the social marginal cost of electricity in the critical peak hours in Japan during the period.[21] Therefore, we provide the same analysis for different assumptions on the marginal cost of electricity supply (85 and 105 cents/kWh) in Tables A.2 and A.3 in the Appendix. Different assumptions on the marginal cost do not change the qualitative results of our welfare analysis, although the welfare gains are larger when we consider

[19] Alternatively, we may consider social pressure instead of warm glow. The utility maximization problem may be represented as $\max_x \ u(x) + v(I - Px) - \phi(x; \theta)$, where $\phi(x; \theta)$ can be interpreted as a utility cost of social pressure for not contributing to conservation. This argument is in line with those in DellaVigna, List and Malmendier (2012) and Gerard (2013). Note that x^* satisfies $u' - Pv' - \phi_x = 0$. Thus, if the functional form of $\phi(\,\cdot\,; \theta)$ is the same for both warm glow and social pressure, we obtain the same results in a marginal sense.

[20] Total differentiation of the first order condition for (4) gives $(u'' + P^2 v'' + \phi_{gg})dx - (Pv'' + \phi_{gg}\bar{x}_I)dI - (v' - Pxv'' + \phi_{gg}\bar{x}_P)dP - \phi_{g\theta}d\theta = 0$. Thus, we have $x^*_\theta = \frac{\phi_{g\theta}}{u''+P^2v''+\phi_{gg}} > 0$ with $dI = dP = 0$.

[21] The Japanese electricity market was only partially deregulated during our experimental period. As a result, not all electricity was traded in the centralized Japanese wholesale market. Regulators knew that most of the marginal power plants supplying electricity for peak demand hours were owned by vertically integrated local monopoly power companies, whose electricity was usually not sold in the centralized wholesale market. During our experimental period, these power companies needed to run their old and inefficient power plants to meet unexpected supply shortages after the Fukushima Daiichi Nuclear Disaster. This is one of the reasons why our assumption of 65 cents/kWh is likely to be a lower bound for the marginal cost. In addition, regulators avoided system-wide blackouts by forcing manufacturing firms to stop operating during peak demand hours. If this cost is considered to be a marginal cost for peak hour electricity, the marginal cost can be much higher than the wholesale electricity price in this partially deregulated market. Finally, the wholesale price did not include environmental externalities from electricity generation, the cost of which is likely to underestimate the social marginal cost of electricity.

a higher marginal cost of electricity supply.[22]

We consider two policies as well as a baseline case with no policy intervention. In the baseline case, consumers pay 25 cents/kWh for their electricity usage, the average residential electricity price in Japan in 2012. The first policy is our economic incentive treatment. We consider that consumers with this policy pay the price that equals the marginal cost, which is 65 cents/kWh. The second policy is our moral suasion treatment. Consumers with this policy pay the baseline price but receive moral suasion for energy conservation.

Consider a quasi-linear utility function for equation (4). To be consistent with the empirical estimation for electricity demand from our field experiment, we characterize the electricity demand by $\ln x = a + \beta D + \epsilon \ln p$, where D equals 1 if consumers receive the moral suasion treatment, p is the electricity price, and ϵ is the price elasticity. We obtain parameters a, β and ϵ from our field experiment.[23] The inverse demand is defined by $p(x) = [x/(\exp(a) \cdot \exp(\beta D))]^{1/\epsilon}$.

The baseline consumption is $\bar{x} = \exp(a) \cdot 25^{\epsilon}$. When consumers receive the economic incentive, the usage becomes $x_e = \exp(a) \cdot 65^{\epsilon}$. The efficiency gain is characterized by $\int_{x_e}^{\bar{x}} (c - p(x))dx$, the area between the marginal cost c and the inverse demand curve $p(x)$ in the range between x_e and $\bar{x}$. We begin by calculating this efficiency gain for the Japanese electricity market. For a typical summer peak hour, electricity consumption from residential customers is 46,800 MWh. An important assumption in this welfare calculation is that residential customers in Japan respond in the same manner to these two policies as the consumers in our experimental households. We consider two scenarios. In the first scenario, we provide the policy for a short run only, by having only 3 treatment days. In the second scenario, we offer the treatment repeatedly for a total of 15 treatment days. This comparison is consistent with our empirical analyses in the previous section, from which we obtain necessary parameters for our welfare calculation.

[Table A.1 about here]

Column 1 of Table A.1 shows the efficiency gain from the economic incentive policy. With the short-run policy, the total efficiency gain for the three treatment days is \$16.84 million. We then calculate the welfare gains for the repeated policy with 15 treatment days based on the estimated parameters from our experimental findings for the repeated interventions. Because the responses to the economic incentive treatments (γ) do not decay much, more treatment days provide further efficiency gains. With 15 treatment days, the efficiency gain is \$76.55 million. The difference between the short-run and repeated polices is \$59.71 million and statistically significant. These results suggest that 1) the economic incentive policy can provide substantial efficiency gains for the electricity market, and 2) repeated interventions can obtain further gains when there are many critical peak demand days, during which the marginal cost of electricity becomes very high.

When consumers receive moral suasion, the usage can be characterized by $x^* = \exp(a) \cdot \exp(\beta) \cdot 25^{\epsilon}$. The efficiency gain is $\int_{x^*}^{\bar{x}} (c - p(x))dx$, which we calculate in Column 2. With the short-run treatment, the efficiency gain is \$11.37 million, which is lower than the gain from the economic incentive treatment, but it still has a meaningful magnitude for the market. Because the moral

[22] Another reason why our welfare calculation is likely to provide a lower bound is that it does not consider long-run avoidable investment costs for generation capacity. According to Kansai Electric Power Company, their long-run avoidable cost for a 600 MW thermal plant is \$150,000/MW per year, assuming that the payment period is 10 years and the discount rate is 4%. The maximum total electricity load from residential customers in Japan is 46,800MW, which implies that our economic incentive policy would induce a reduction in the maximum load by 7,198 MW ($= 46,800 \cdot 0.1538$). Therefore, a back of envelop calculation of the long-run avoidable cost from the economic incentive policy is \$1,080 million ($= 7,198 \cdot 150,000$) per year, which is significantly larger than the welfare gains in Table A.1, which does not consider long-run avoidable investment costs for generation capacity.

[23] Recall that we estimated β (the effect of the moral suasion) and γ (the effect of economic incentives) in our field experiment. We use γ for the case with treatment price 65 cents/kWh to calculate the price elasticity $\epsilon = \gamma/\ln(65/25)$.

suasion effect decays, the efficiency gain does not increase much with repeated interventions. We cannot reject the null that the efficiency gain from the moral suasion treatment is the same for the short-run policy and repeated policy.

When consumers receive moral suasion, there is one more channel through which the welfare can be changed. In our model in equation (4), consumers who receive moral suasion would change their usage from $\bar{x}$ to x^* because they feel warm glow or self-satisfaction from behaving prosocially. In this case, consumers obtain a surplus from their conservation $g = \bar{x} - x^*$. Note that consumers do not necessarily gain a surplus if we consider different models that could explain their motives. For example, consumers may reduce usage because they feel social pressure (DellaVigna, List and Malmendier, 2012) or obedience for authorities. In such cases, it is possible that consumers may lose a surplus when receiving moral suasion. Given our experimental setting, the primary motive for our consumers was more likely to be warm glow. However, we cannot completely exclude the possibility that our households may have lost a surplus or gained no surplus when receiving moral suasion. Therefore, we provide the welfare change from the efficiency gain and that from (potential) warm glow separately in the table and interpret the gain from warm glow with this caution. Recall that the inverse demand is $p(x) = [x/(\exp(a) \cdot \exp(\beta D))]^{1/\epsilon}$. The surplus from warm glow is, therefore, obtained by $\int_{x^*}^{\bar{x}} \left([x/\exp(a)]^{1/\epsilon} - [x/(\exp(a) \cdot \exp(\beta))]^{1/\epsilon}\right) dx$, in which parameters a, β, and ϵ are obtained from the field experiment.

We provide the sum of the efficiency gain and warm glow in the last column of Table A.1. The results suggest that if we take account of a positive gain from warm glow, the total welfare gains from the moral suasion policy can be close to the gains from the economic incentive policy in the short-run. However, this is not the case for the repeated intervention, in which the welfare gain is much larger for the economic incentive policy even if we incorporate potential gains from warm glow. Finally, these results suggest that while in theory welfare gains can arise from the warm glow effect in theory, the major welfare gains in our context arise from the efficiency gains—from letting consumers pay prices that reflect the actual marginal cost of electricity during the critical peak hours.

Online Appendix B: Additional Tables and Figures

Table A.1: Welfare Gains from the Two Policies (Assumption on Marginal Cost = 65 cents/kWh)

	Economic Incentive	Moral Suasion	
	Efficiency Gain ($M)	Efficiency Gain ($M)	Efficiency Gain + Warm Glow ($M)
Short-Run Treatments (3 days)	16.84	11.37	15.02
	(1.99)	(2.55)	(4.62)
Repeated Treatments (15 days)	76.55	24.40	27.32
	(9.04)	(9.92)	(12.38)

Notes: This table shows the estimated welfare gains per season from the two policies in our field experiment. We use 46,800 kWh as the peak hour residential electricity consumption in the Japanese electricity market for the baseline case, which does not refer to either of our policies. We use 65 cents/kWh as the marginal cost of electricity for these critical peak hours. In the Appendix, we provide the same analyses for different assumptions of the marginal cost of electricity.

Table A.2: Welfare Gains from the Two Policies (When Marginal Cost = 85 cents/kWh)

	Economic Incentive	Moral Suasion	
	Efficiency Gain ($M)	Efficiency Gain ($M)	Efficiency Gain + Warm Glow ($M)
Short-Run Treatments (3 days)	26.15	17.38	22.11
	(3.08)	(3.98)	(6.69)
Repeated Treatments (15 days)	118.88	36.91	40.65
	(14.03)	(15.13)	(18.29)

Notes: This table shows the estimated welfare gains per season from the two policies in our field experiment. We use 46,800 kWh as the peak hour residential electricity consumption in the Japanese electricity market for the baseline case, which does not refer to either of our policies. For this table, we use 85 cents/kWh as the marginal cost of electricity for these critical peak hours. In the Appendix, we provide the same analyses for different assumptions of the marginal cost of electricity.

Table A.3: Welfare Gains from the Two Policies (When Marginal Cost = 105 cents/kWh)

	Economic Incentive	Moral Suasion	
	Efficiency Gain ($M)	Efficiency Gain ($M)	Efficiency Gain + Warm Glow ($M)
Short-Run Treatments (3 days)	35.78	23.51	29.10
	(4.21)	(5.47)	(8.70)
Repeated Treatments (15 days)	162.65	49.50	53.89
	(19.18)	(20.41)	(24.12)

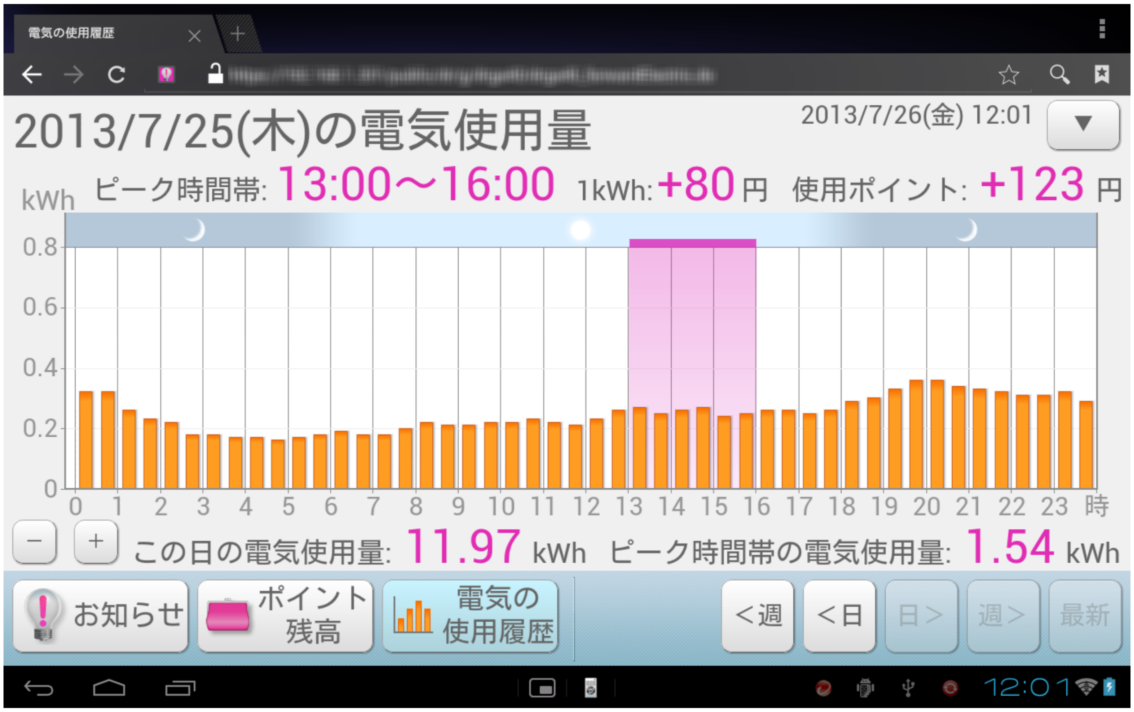
Notes: This table shows the estimated welfare gains per season from the two policies in our field experiment. We use 46,800 kWh as the peak hour residential electricity consumption in the Japanese electricity market for the baseline case, which does not refer to either of our policies. For this table, we use 105 cents/kWh as the marginal cost of electricity for these critical peak hours. In the Appendix, we provide the same analyses for different assumptions of the marginal cost of electricity.

Table A.4: Heterogeneity in the Treatment Effects

	Summer			Winter		
	(1)	(2)	(3)	(4)	(5)	(6)
Moral suasion	-0.044	-0.045	-0.045	-0.034	-0.034	-0.034
	(0.014)	(0.014)	(0.014)	(0.022)	(0.022)	(0.022)
Economic incentive	-0.168	-0.178	-0.178	-0.178	-0.177	-0.177
	(0.022)	(0.023)	(0.023)	(0.023)	(0.024)	(0.023)
Moral suasion × Income	-0.052		-0.054	-0.002		-0.003
	(0.029)		(0.030)	(0.040)		(0.040)
Economic incentive × Income	0.119		0.126	0.108		0.100
	(0.051)		(0.050)	(0.046)		(0.046)
Moral suasion × Usage		0.058	0.069		0.005	0.007
		(0.089)	(0.089)		(0.117)	(0.119)
Economic incentive × Usage		-0.516	-0.531		0.138	0.072
		(0.178)	(0.171)		(0.117)	(0.117)
Observations	105107	105107	105107	205357	205357	205357

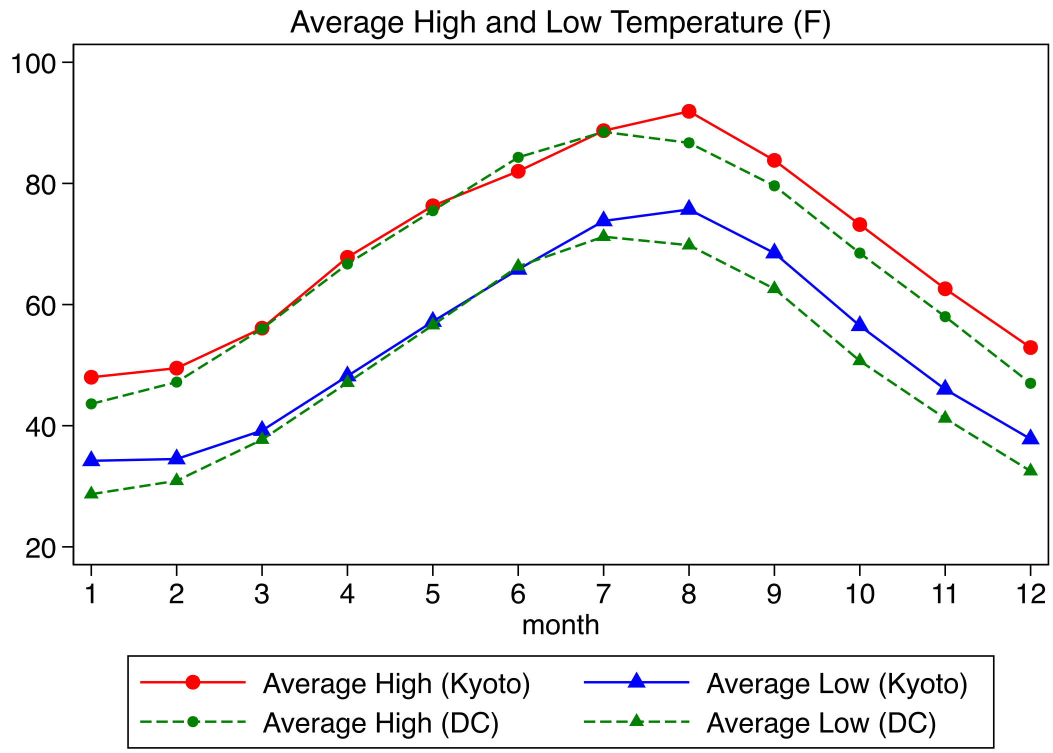
Notes: Table A.4 shows the estimation results for regressions that include the interaction terms for household income and usage. Columns 1 and 4 include the interaction terms for household income, columns 2 and 5 include the interaction terms for pre-experiment usage levels, and columns 3 and 6 include both interaction terms. Although we found weak evidence for the moral suasion effect being larger for higher-income households, the estimates are not statistically significant. We found a consistent relationship between economic incentives and income—the economic incentive effect is lower for higher-income households compared to lower-income households. Note that our dependent variable is the log of electricity usage, and the treatment variables are dummy variables. Therefore, for example, the coefficient (0.126 log points) in column 2 implies that an increase in household income by $10,000 would be associated with a 0.0126 log-point increase for the coefficient for the economic incentive dummy variable (i.e., a 0.0126 log-point *decrease* in the treatment effect).

Figure A.1: Information Provided by an In-Home Display



Notes: This figure shows an example screenshot of the in-home displays that were installed for both the control and the treatment consumers in the experiment. On the top of the figure, it shows "Electricity usage for July 25, 2013. Peak hours: 13:00 to 16:00. The price increase is +80 yen per kWh." The figure in the middle shows usage in kWh for each 30-minute interval from hour 0 to hour 24 of the day. The shaded area shows the peak hours, which are from 13:00 to 16:00. On the bottom of the figure, it shows "The daily electricity usage is 11.97 kWh. Usage for the peak hours is 1.54 kWh."

Figure A.2: Average High and Low Temperatures in Kyoto, Japan and Washington DC, United States



Notes: This figure compares the average high and low temperatures (°F) in Kyoto, Japan and Washington DC, United States.

Online Appendix B: Materials from the Field Experiment

Invitation Letter (Translated in English)

The Keihanna Eco-City Next-Generation Energy/Community System Demonstration Project

Questionnaire for Assessing Interest in Participating in the Smart Power Usage Program

The Keihanna Eco-City Next-Generation Energy/Community System Demonstration Project Promotion Council created with the support of the Ministry of Economy, Trade and Industry, Keihanna Science City's Next-Generation Energy/Community System Demonstration Project consists of a variety of initiatives designed to create a leading low-carbon community in Japan. As part of this project, we have recently started a power usage demonstration program. As part of this program, we request several households to adopt an energy-saving but easily sustainable lifestyle. We have created this questionnaire to assess the interest of Keihanna Science City residents participating in the program. Please take some time to read and complete this questionnaire. Thank you for your cooperation.

Points to Note before Filling Out the Questionnaire

● Respondents who agree to participate in the program and receive an at-home program briefing will be rewarded with a 1,000-yen prepaid card.

● Read the program overview (on the other side of this sheet) before responding to the questionnaire (separate sheet).

● Place the completed questionnaire in the prepaid return envelope provided and mail it before February 13 (Mon).

■ Questionnaire participants

This questionnaire was distributed by Japan Post's Yu-Mail designated delivery area service after selecting survey areas from among the districts of Keihanna Science City (Kyotanabe, Kizugawa, and Seika).

■ Terms of privacy for personal information

Personal information obtained using this questionnaire will be rigorously managed by the questionnaire administrator, Mitsubishi Heavy Industries. It will be used only to implement the Smart Power Usage Program and for no other purposes. If information about your electric power agreement, facilities, or usage is required for the program, Mitsubishi Heavy Industries will request the information from the Kansai Electric Power Company, and the Kansai Electric Power Company will provide Mitsubishi Heavy Industries with the information requested about your electric power agreement, facilities, or usage.

The Keihanna Eco-City Next-Generation Energy/Community System Demonstration Project Promotion Council

Members: Kyoto Prefecture, City of Kyotanabe, City of Kizugawa, Town of Seika, Public Foundation of Kansai Research Institute, Kansai Electric Power Company, Mitsubishi Electric Corporation, Mitsubishi Heavy Industries, other private-sector companies

For inquiries about the questionnaire, contact the questionnaire administrator organization below.

Questionnaire administrator organization: Regional Futures Research Center

Staff members: Horibe, Yoshiura, Tabuchi

Tel. (toll-free): 0120-79-7711 (9:30~17:00, except weekends and holidays)

Please continue to the program overview on the other side.

Program Overview

Three aspects of smart power usage

● The program will use modern telecommunications technology to create smarter and more streamlined power use by equipping households to moderate their power usage volume and adopt energy-saving habits.

✓ Awareness of energy-saving timing

✓ Visibility of power wastage

✓ Advice from other households

Program Description

● Participating households will engage in some of the following activities.

– The activities will vary depending on households and will be set at random according to the needs of the survey.

–Participants will not incur any cost as a result of taking part in these activities.

Activity 1 Setting variable power charges

● We will provide simulated power charges that vary in time slots of rising power demand.

● You will work on moderating your power usage as much as possible in time slots of high power charges (about 2 or 3 hours during the day).

Activity 2 Providing information on power usage

● We will provide a system enabling participating households to check their power usage every hour.

● You will check your power usage in each time slot and devise ways to minimize wasteful power usage.

Activity 3 Providing energy-saving advice

● After analyzing your power usage, we will advise you on areas such as power-usage methods and replacing appliances (in 2013 summer).

● You will follow the above advice to reduce wasteful power use.

You will not be pressured into replacing appliances.

Program Period, Rewards

● We plan to conduct the program from July 2012 through the end of 2014.

● Participants will receive a small reward (in addition to the reward for completing this questionnaire).