The Persistence of **Intrinsic** and **Extrinsic Motivation**: Experimental Evidence from Energy Demand*

Koichiro Ito — Boston University and NBER

Takanori Ida — Kyoto University

Makoto Tanaka — GRIPS

This version: December 1, 2014

Abstract

Regulators often use moral suasion and monetary incentives to influence intrinsic and extrinsic motivations for economic activities. In our field experiment, we randomly assigned households to moral suasion and dynamic electricity pricing to stimulate energy conservation in peak demand hours. Using household-level consumption data for every 30 minutes, we find that moral suasion induced significant short-run effects, but the effects diminished quickly during repeated interventions. Monetary incentives produced larger and persistent effects, which resulted in habit formation after the final intervention. While welfare gains are substantial for each intervention, monetary incentives produce particularly large gains when we consider persistence.

*Ito: School of Management, Boston University, 595 Commonwealth Avenue, Boston, MA 02215, and NBER (e-mail: ito@bu.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). We thank Hunt Allcott, Masahiko Aoki, Severin Borenstein, Lucas Davis, Liran Einav, Meredith Fowlie, Takeo Hoshi, Caroline Hoxby, Katrina Jessoe, Paul Joskow, Chris Knittel, David Laibson, Michael Price, Matthew Rabin, Dave Rapson, Catherine Wolfram, Frank Wolak, and seminar participants at MIT, Harvard, Stanford, UC Berkeley, UC Davis, University of Tokyo, the NBER Summer Institute, the POWER Conference on Energy Research and Policy, and the RIETI 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 for their collaboration on this study. Ito thanks the Stanford Institute for Economic Policy Research for financial support.

1 Introduction

Recent economic theory emphasizes that both intrinsic and extrinsic motivations play important roles in economic decisions (Kreps, 1997; Bénabou and Tirole, 2003, 2006). For example, consider those who allocate their time and resources to prosocial behavior such as charitable giving and energy conservation. The decision can be driven by intrinsic motivation such as warm glow (Andreoni, 1989) or by extrinsic motivation such as monetary incentives. A growing number of policymakers recognize the importance of intrinsic and extrinsic motivations and attempt to design economic policies to influence these motivations for a variety of policy goals—increasing donations, promoting smoking cessation, motivating people to exercise, improving academic refereeing processes, and encouraging conservation of energy, water, and other scarce resources.[1]

A central question for economists and policymakers designing such policies is whether they can generate robust and persistent effects by appealing to intrinsic and extrinsic motivations (Gneezy, Meier and Rey-Biel, 2011). Many studies in psychology note that there are important differences between short-run and long-run psychological processes (Gneezy and List, 2006). For example, individuals may have *hot* versus *cold* decision-making; immediate reactions to an event, called *hot* decision-making, can differ from decision-making in the *cold* phase, which follows the hot phase (Loewenstein, 2005). Another example is potential *habit formation* that can be induced by policy interventions; short-run shocks to current consumption patterns may form a habit pertaining to consumption patterns in the subsequent time periods (Becker and Murphy, 1988). In theory, these factors could interact with intrinsic and extrinsic motivations quite differently. The resulting differences, if any, give rise to differences in welfare implications between policies targeting intrinsic motivation and extrinsic motivation. While the answer to this question is relevant to economic policies in many fields, we are not aware of existing studies that compare the persistence of these two policy interventions in a unified field experimental framework.

In this paper, we use a randomized field experiment to investigate whether policymakers can generate robust and persistent effects by appealing to intrinsic and extrinsic motivations. We

[1] A vast amount of literature provides empirical evidence of these policies, including improving academic refereeing (Chetty, Saez and Sándor, 2014), increasing blood donations (Lacetera, Macis and Slonim, 2012), charitable giving (Gneezy and Rustichini, 2000; Landry et al., 2006, 2010), energy conservation (Reiss and White, 2008), motivating people to exercise (Charness and Gneezy, 2009), smoking cessation (Volpp et al., 2009), and tax evasion (Dwenger et al., 2014).

consider two policy interventions that are most widely used by policymakers in practice. The first intervention is moral suasion, by which policymakers attempt to influence intrinsic motivation for prosocial behavior. The second intervention is a monetary incentive, by which policy makers attempt to influence extrinsic motivation according to the law of demand. Our main outcome variable is household-level electricity consumption for every 30 minutes. We began by randomly assigning households to one of three groups: 1) a moral suasion group, 2) a monetary incentive group, and 3) a control group. On peak demand days in summer and winter, we delivered day-ahead and same-day notifications about treatments. The moral suasion group received a message requesting voluntary energy conservation with no monetary incentives. The monetary incentive group received high electricity prices during peak demand hours. We repeated these interventions to analyze *hot* versus *cold* decision-making among the groups. The repeated interventions allowed us to estimate how treatment effects change between the first intervention and subsequent interventions. Moreover, we collected electricity usage data after the final interventions to investigate potential habit formation. Finally, we conducted a detailed follow-up survey to investigate the mechanism behind our findings.

We present several findings from the experiment. First, moral suasion induced short-run reductions in electricity usage, but the effect diminished quickly after repeated interventions. The moral suasion group showed a usage reduction of 8 percent for the first few treatment days. However, their usage became statistically indistinguishable from that of the control group for the rest of the interventions. Second, we found that monetary incentives created much larger and persistent effects. The monetary incentive group showed usage reductions of 17 percent. Moreover, the effect was persistent over repeated interventions. Third, we tested whether there was a spillover effect on consumption during nontreatment hours of the treatment days. The treatment groups had direct incentives to reduce their usage only during the treatment hours. However, the theory suggests that our treatment may incentivize consumers to reduce or increase usage during nontreatment hours. We found no spillover effects for the moral suasion group. In contrast, there were significant spillover effects, namely, usage reductions during the nontreatment hours, for the monetary incentive group. Fourth, we tested potential habit formation by estimating treatment effects using data collected after the experimental periods had ended. We found significant habit formation for the monetary incentive treatment only. After the end of the final interventions, the moral suasion group's usage was indistinguishable from that of the control group. On the other hand, the monetary incentive group continued to practice energy conservation even after we withdrew the incentives.

What drives the substantially different results for the moral suasion and monetary incentive groups? We investigate two potential channels. The first possibility is durable goods investments—households purchased energy-efficient appliances in response to the treatments. If such an effect was systematically large for the monetary incentive group, it could explain the persistent usage reductions. The second possibility is behavioral changes in lifestyles. Suppose that some customers had "bad habits" of inefficient energy use at home before we began our experiment. Our interventions may have acted as a trigger to change their lifestyles, thus helping them form good habits, namely, efficient energy use. If such an effect was systematically large for the monetary incentive group, it could explain the persistent usage reductions. Using follow-up survey data, we find no statistical evidence for the first hypothesis. In contrast, we find supporting evidence for the second hypothesis. Our data indicate that the monetary incentives induced behavioral changes in lifestyles—households in this group formed 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 stated survey responses, they provide suggestive evidence about the mechanisms behind our findings. The significant difference in the persistent effects between the treatment groups is unlikely to arise from durable goods investments. Instead, we may safely attribute them to the high electricity prices; customers who experience high electricity prices make lifestyle changes favoring energy conservation.

The findings from the experiment provide important policy implications. Our results indicate that economic policies can have significantly different impacts depending on how policymakers design policy incentives in relation to intrinsic and extrinsic motivations of consumer behavior. To highlight this policy implication, we provide a welfare analysis in the context of electricity markets. One of the largest inefficiencies in electricity markets originates from the fact that retail electricity prices generally do not reflect the marginal cost of electricity—consumers pay time-invariant prices although the marginal cost of electricity tends to be extremely high in "critical peak hours", which are peak demand hours of peak demand days. This inefficiency has long been discussed by economists and forms a key policy issue in recent energy policies of many countries, particularly because the recent technological progress in "smart meters" is likely to address the problem (Wolak, 2011; Joskow, 2012; Joskow and Wolfram, 2012; Jessoe and Rapson, 2014). Our welfare analysis implies that the moral suasion policy can effect welfare gains for a very short run. However, the monetary incentive policy produces larger welfare improvement, particularly when we consider repeated interventions. Our analysis suggests that the welfare gain for the Japanese electricity market per summer is \$120 million for the monetary incentive policy and \$37 million for the moral suasion policy.

This paper contributes to two growing strands of the literature in recent economics studies. The first is the strand studying the effects of intrinsic and extrinsic motivations on a variety of economic activities. A central question unanswered in the literature is the persistence of intrinsic and extrinsic motivations and their implications for economic policies. To our knowledge, our study is the first field experiment that 1) randomly assigned two interventions appealing to intrinsic and extrinsic motivations, 2) conducted repeated interventions of the two treatments, and 3) observed high-frequency consumption data during and after the treatment interventions. This experimental design enables us to provide new empirical evidence on the research question.[2]

Our paper also builds on recent studies in energy and environmental economics that examine the effects of pecuniary and nonpecuniary incentives in the context of energy and environmental policies. A key challenge is how to empirically distinguish intrinsically motivated and extrinsically motivated behavior in nonexperimental data. This is challenging because typical environmental policies provide various types of incentives at once. Such environments make it difficult to identify the effect of each incentive, particularly when researchers estimate persistent effects. Our study used a randomized controlled trial to address this challenge, and its results provide key implications for designing effective energy and environmental policies.[3]

[2]Many previous papers examine whether extrinsic incentives crowd out or backfire intrinsic motivation. Examples include Frey and Oberholzer-Gee (1997) for 'Not In My Backyard' problem, Gneezy and Rustichini (2000); Landry et al. (2006) for charitable giving, Lacetera, Macis and Slonim (2012) for increasing blood donations, Dwenger et al. (2014) for tax evasion, and Chetty, Saez and Sándor (2014) for improving academic refereeing. Most of these studies focus on short-run effects, except for Landry et al. (2010), which analyzes dynamic effects by recording donor actions more than once. See Gneezy, Meier and Rey-Biel (2011) for a comprehensive survey. Other recent papers test habit formation by providing either a monetary incentive or a norm-based intervention. For example, Charness and Gneezy (2009) find that a monetary incentive induces habit formation for exercising at the gym. Allcott and Rogers (Forthcoming) find that informing electricity consumers about peers' consumption produces long-run conservation effects.

[3]Examples of nonexperimental studies documenting intrinsically motivated conservation include Reiss and White (2008) for the California electricity crisis, Cutter and Neidell (2009) for 'Spare the Air' program in California, and Gerard (2013) for the Brazilian electricity crisis. The authors of these papers acknowledge that it is challenging to identify voluntary conservation separately from the effects of other policies in effect during their sampling periods. Ferraro, Miranda and Price (2011) and Ferraro and Price (2013) are the first to use field experiments to examine

2 Experimental Design, Data, and Hypotheses

We used a field experiment with high-frequency data on household-level electricity usage to test several hypotheses. In this section, we describe our experimental design and data and then explain our treatments and hypotheses.

2.1 Experimental Design and Data

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

To invite as broad a set of households as possible, we provided generous participation rewards, which included free installations of an advanced meter and in-home display and a participation reward of 12,000 yen (approximately $120 in 2012). We contacted all 40,710 residential electricity customers in the area by mail. Of these, 1,659 customers confirmed their participation. We excluded students, customers who had a rooftop solar panel or other types of electricity self-generation mechanisms, and those without access to the internet. This process left us with 691 households. Similar to previous studies of electricity conservation (Wolak, 2006, 2011; Faruqui and Sergici, 2011; Jessoe and Rapson, 2014), our experiment was a randomized controlled trial (RCT) for self-selected participants, as opposed to a RCT for a purely randomly selected sample of the population. To explore the external validity of our sample, we collected data from a random sample of the population. We analyze observables between our sample and the random sample in the following section. We randomly assigned the selected 691 households to three groups: control (C), moral suasion (S), and monetary incentive (I).

Control Group (C): The 153 customers in this group received an advanced electricity meter, an in-home display, and the participation reward. Other than that, this group did not receive any treatment.[4]

intrinsically motivated conservation. Although their experiments do not have extrinsic treatments, they compare three types of nonpecurinary treatments for water conservation: information dissemination on behavioral and technological modifications, appeal for prosocial preferences, and provision of social comparisons.

[4]These advanced electricity meters are sometimes called "smart meters." Conventional electricity meters record usage monthly or bi-monthly. In contrast, advanced meters usually record usage at 15-, 30-, or 60-minute intervals.

Moral Suasion Group (S): 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.

Monetary Incentive Group (I): The 384 customers in this group received an advanced electricity meter, an in-home display, and the participation reward. In addition, this group received "monetary 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 focus on consumption data on weekdays, excluding holidays, from the summer of 2012 to the spring of 2013. We also conducted three surveys. The first survey was conducted prior to treatment assignment and collected demographic information. The second survey was conducted upon completion of the experiment to explore the mechanism behind our findings. The third survey was conducted for a random sample of households in the area to investigate the external validity of our sample.

[Table 1 about here]

Columns 1, 2, and 3 of Table 1 present the summary statistics of demographic variables and preexperiment consumption data by treatment group. A comparison across control and treatment groups indicates statistical balance in observables because of random assignment of the groups. Very little attrition occurred in each group. In total, 9 households (1.3 percent) dropped out from our sample because they moved residence. Because this small attribution occurred at approximately the same rate in each group, it does not significantly bias our estimates.

Column 4 shows the summary statistics for a random sample of the population in the area. We investigate 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 and standard errors in brackets. The differences are small and statistically insignificant for most variables. We find 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

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

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.

To obtain a sense of weather conditions in the experimental region, we compare the monthly average high and low temperatures between Kyoto prefecture in Japan and Washington D.C. 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^oF in each month, which implies that these two cities have quite similar weather conditions.

2.2 Treatments

A fundamental economic inefficiency in electricity markets is that electricity consumers generally do not pay prices that reflect the high marginal costs of electricity during peak demand hours (Borenstein, 2001). Policymakers usually consider two types of economic policies to address this inefficiency. The first policy instrument, which is most frequently used by many countries, is an appeal to intrinsic motivation by calling 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 create persistent effects on consumer behavior. To investigate this question, we design two treatments that reflect the two policies being used by regulators.

Our first treatment is "moral suasion," which intends to influence intrinsic motivation for energy conservation. After customers were assigned to the moral suasion group (S), they were educated about the need for energy conservation during critical peak demand hours because Japan was facing a serious electricity shortage after the Fukushima Daiichi Nuclear Disaster. We informed them about the provision of day-ahead and same-day messages for a "critical peak day" (treatment day) based on the day-ahead weather forecast. Note that they did not receive any extrinsic reward to conserve energy. We neither provided a monetary incentive nor did we disclose information about their conservation efforts to others. Therefore, we interpret the effect of this treatment on conservation, if any, as intrinsically motivated conservation as opposed to extrinsically motivated behavior.

We informed customers about how they would be receiving 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. The treatment day had to be a week day in which the day-ahead maximum temperature forecast exceeded 31°C (88°F) for the summer and the day-ahead minimum temperature forecast was lower than 5°C (41°F) for the winter. Customers were told that they would receive day-ahead and same-day notices about the treatment day via a text message and a message on their in-home displays when the temperature forecasts met the criteria.

[Figure 1 about here]

Our second treatment is a "monetary incentive," which was intended to influence extrinsic motivation for energy conservation. After customers were assigned to the monetary incentive group (I), we informed them that they would receive higher electricity prices during the critical peak demand hours. We also told them that we would provide day-ahead and same-day messages for critical peak days based on the day-ahead weather forecast. Figure 1 shows the dynamic pricing schedule. There were four prices for the treatment hours. On treatment days, the monetary incentive group received one of the three critical peak prices: 65, 85, and 105 cents/kWh.[6] There was no price change for the control group and the moral suasion group, and they paid the baseline price, 25 cents/kWh. We used stratified randomization for the treatment days to decide which of the three critical peak prices would be allocated to the monetary incentive group. We defined *treatment cycles*, wherein each cycle consisted of three treatment days. We then randomized the three critical peak prices within the cycle. Using this stratified randomization, we minimized the correlation between the critical peak prices and the temperature to estimate whether consumers responded differently to each marginal price.[7] To analyze the persistence of the two treatments,

[6] Although participants paid in Japanese yen, we use U.S. cents in this figure. One Japanese yen was approximately equivalent to one U.S. cent (2012 exchange rate).

[7] 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 mimimum temperatures for the winter treatment days were -3.9°C, -1°C, and 4.6°C (25°F, 30°F, and 40°F). Because of the stratified randomization, 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.

we repeated the interventions as long as the temperature forecasts met the criteria. The treatment groups experienced 15 treatment days (5 cycles) in the summer and 21 treatment days (7 cycles) in the winter.

2.3 Hypotheses

We tested several hypotheses using the field experiment. The first hypothesis is based on the standard economic theory, which predicts that 1) the monetary incentive group lowers consumption in response to the changes in price based on the price elasticity of demand, and 2) the moral suasion group uses electricity in the same way as the control group. We compared consumption between the control group, moral suasion group, and monetary incentive group to test the hypotheses. In addition, we used the randomized critical peak prices to estimate whether consumers respond to the marginal price of electricity.

The second hypothesis is related to *hot* versus *cold* decision-making in the phycology literature (Loewenstein, 2005). Immediate reactions to an event, called *hot* decision-making, can differ from decision-making in the *cold* phase, which follows the hot phase. This is particularly important for energy policies because policymakers generally seek persistent responses from consumers. To test this possibility, we repeated our interventions. In total, consumers in the treatment group received treatments 15 times in the summer and 21 times in the winter. We tested whether 1) the monetary incentive was persistent through multiple treatment days and 2) the moral suasion was persistent through multiple treatment days.

The third hypothesis refers to potential habit formation that can be induced by the interventions (Becker and Murphy, 1988). To test this possibility, we withdrew the incentives after the final intervention but continued to collect high-frequency electricity consumption data. We asked whether 1) the moral suasion effect will disappear after the final intervention, so that the moral suasion group's usage will be indistinguishable from the control group's usage, and 2) the monetary incentive effect will disappear after we discontinued the treatment.

We test the three hypotheses in the following sections. We then investigate the mechanism behind our findings by analyzing follow-up surveys on durable goods investments and behavioral changes in energy-efficient lifestyles. Finally, we explore the policy implications by analyzing welfare gains from the two policies in the context of electricity markets.

3 Empirical Analysis and Results

In this section, we present the results of our field experiment. Recall that the treatment groups experienced several treatment days. We begin by including all treatment days in our regression, to estimate the overall treatment effects. We then explore their persistence, spillovers, and habit formation in the following sections.

3.1 Effects of Moral Suasion and Monetary Incentives

We begin by showing evidence from the raw data in Figure 2. It plots the mean 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. Half an 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 we observe small remaining differences for a few hours after the treatment hours. The figure provides visual evidence of the treatment effects for both treatment groups and suggests that the reductions seem to be larger for the monetary incentive group.[8]

Table 2 provides a formal econometric analysis with standard errors. We estimate the treatment effects by an OLS regression:

$$\ln x_{it} = \alpha S_{it} + \beta I_{it} + \theta_i + \lambda_t + \eta_{it}, \tag{1}$$

where $\ln x_{it}$ is the natural log of electricity usage for household i in a 30-minute interval t. S_{it} equals 1 if household i is in group S (the moral suasion group) and receives a treatment in t. Similarly, I_{it} equals 1 if household i is in group I (the monetary incentive group) and receives a treatment in t. We include household fixed effects θ_i and time fixed effects λ_t for each 30-minute interval to control for time-specific shocks such as weather. We cluster the standard errors at the household level to adjust for serial correlation. We include data from the preexperiment days and treatment

[8] In critical peak electricity pricing, the treatment (peak) hours are generally defined as the hours in which system-wise aggregate consumption reaches its peak. On typical summer days in Japan, the peak consumption hours occur during the night for residential customers, and therefore, they do not coincide with the system peak hours. On winter days in Japan, the residential peak hours coincide with the system peak hours.

days in this regression.[9] Recall that the treatment groups had 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 include only these hours to estimate the effects on the treatment hours. We examine potential spillover effects for nontreatment hours in the following section.

[Table 2 about here]

Column 1 shows that moral suasion causes a 3 percent reduction in peak-hour electricity usage for the summer 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.[10] This finding implies that the moral suasion policy can help to some extent when we consider the average effect over all treatment days. However, the reductions are much larger for the monetary incentive (17 percent). This result suggests that the monetary incentive policy can be used to substantially reduce costly electricity production from marginal power plants during critical peak hours.

An important question for the monetary incentive effect is whether 1) consumers responded to the changes in the marginal price or 2) they just reacted to a "pricing event." The two possibilities imply different policy implications because the former indicates that policymakers can use price as a tool to achieve certain levels of reductions. This question remains unanswered in the literature because most previous experiments use only one price for critical peak events.[11] We examine these hypotheses in column 2 by estimating the treatment effects by price. Recall that the marginal price for the control group was 25 cents/kWh. We randomize three prices for the critical peak price across our treatment days for the monetary incentive group.

Column 2 shows that, consistent with the prediction of the standard economic theory, the reduc-

[9] Treatment effects can have spillover effects on nontreatment days after the beginning of the experimental period. In this case, including nontreatment days (as control days) will underestimate the treatment effects. We, therefore, do not include nontreatment days in this regression. We estimate potential spillover effects in the following section.

[10] Reductions in peak hour electricity of 3 percent (our intrinsic effect) and 17 percent (our extrinsic effect) are large relative to other policies that aim for energy conservation. For example, home energy reports by Opower induce a reduction of about 1 percent for most of their treatment locations (Allcott and Rogers, Forthcoming). For a California conservation rebate program, which provided a 20 percent rebate if consumers reduced electricity usage by 20 percent, Ito (Forthcoming) realizes a reduction of about 5 percent only for customers who are likely to use air conditioners throughout summer. For a critical peak pricing experiment in Washington D.C., Wolak (2011) finds that consumption reduced by about 10 percent.

[11] For example, Wolak (2006, 2011) use one price for critical peak events for each experiment. Jessoe and Rapson (2014) include different critical peak hour prices, but the treatment hours and duration of the treatment differ across the prices. As the authors admit, this makes it difficult to compare treatment effects across prices.

tions monotonically increase when customers have higher marginal prices.[12] This finding implies that households indeed respond to time-varying marginal prices when they receive salient price information. While this result may be unsurprising to economists, it has a real policy implication because regulators and utility companies often believe that electricity consumers do not respond to electricity prices, and therefore, they cannot use a price-based policy. Our findings are in contrast to those of Ito (2014), who finds that electricity consumers in California do not respond to the marginal price of their nonlinear price schedules. One obvious difference in the environment is that customers in our experiment have access to salient price information via in-home displays and text messages, whereas Californian customers in Ito's study receive their price information only through their monthly bills. The difference in the findings is consistent with the literature that emphasizes the importance of price salience (Chetty, Looney and Kroft, 2009; Finkelstein, 2009; Jessoe and Rapson, 2014). The second important difference is that our customers had a single marginal price over an hour, which varies across hours, whereas the marginal prices in Ito's study vary with each customer's cumulative monthly usage. The different findings between the two studies, therefore, reflect the possibility that consumers are more likely to respond to time-varying marginal prices compared to marginal prices that vary with their cumulative usage. Finally, in columns 3 and 4, we find consistent results for the winter treatment days, when the treatment hours were between 6 pm and 9 pm.

3.2 Repeated Interventions and Persistence of Treatment Effects

The treatment groups experienced 15 treatment days in the summer and 21 treatment days in the winter. Our main objective of including many treatment days was to examine whether the treatment effects persisted across repeated interventions. This question is relevant to policy because policymakers, in reality, need to have repeated interventions to deal with multiple peak demand days for a given season. Because the critical peak event days in our study were based on the day-ahead weather forecasts, the treatment days were not necessarily consecutive. We, therefore, had 15 and 21 nonconsecutive days for the summer and the winter respectively. To analyze persistence, we divide the 15 summer treatment days into 5 cycles and the 21 winter treatment days into 7

[12]The coefficient estimate for 65 yen is statistically different from that for 105 yen at the 5 percent significance level. We cannot reject that the coefficient estimates for 65 yen and 85 yen are statistically equivalent at the 10 percent significance level.

cycles, in which each cycle includes 3 event days.[13] We estimate OLS regressions for treatment cycles $c = 1, .., 5$ for the summer and $c = 1, .., 7$ for the winter:

$$\ln x_{it} = \sum_{c \in C} (\alpha_c S_{itc} + \beta_c I_{itc}) + \theta_i + \lambda_t + \eta_{it}, \tag{2}$$

where α_c and β_c are the effects of moral suasion and monetary incentives for treatment cycle c respectively.

In Table 3, we find substantially different persistence between moral suasion and monetary incentives. The moral suasion effect is statistically significant only in the first cycle (the first three treatment days). This is consistent for both the summer and the winter experiments. The moral suasion effect in the first cycle causes a reduction of about 8 percent, which is economically significant for the critical peak hours. It is also larger than the effects of non-price interventions in the literature addressing social comparison of electricity usage, which usually induces an average usage reduction of 1 or 2 percent. This finding implies that moral suasion can be an effective policy to begin with. However, it is clear from the table that the effect decays quickly after the first cycle—it declines to 3 percent in the second cycle and diminishes to zero in the remaining cycles.

In contrast, we find that the monetary incentive produces considerably more persistent effects. For the summer, the effect is the largest in the second cycle and remains fairly stable between the first and third cycles. The effect becomes slightly lower in the fourth and fifth cycles. Nevertheless, the fifth cycle still shows an effect of about 9 percent, which statistically significant at the 1 percent significance level. The winter results show similar patterns with even more stable effects across repeated interventions. The effect is largest in the second cycle and is stable from the first to the seventh cycles. While we reject the null hypothesis that moral suasion effects are equivalent in the first and the seventh cycles, we cannot reject the null for the monetary incentive effects. Our findings suggest that the two policies are likely to have significantly different implications when policymakers intend to generate persistent effects across repeated interventions.

[Table 3 about here]

[13] As explained in the experimental design, the three critical peak prices (65, 85, and 105 yen/kWh) were randomized within the three days in a cycle.

3.3 Spillover Effects for Nontreatment Hours on Treatment Days

Customers in the treatment groups had a direct incentive to change their consumption only during the treatment hours. However, there are a few reasons for the existence of spillover effects for the nontreatment hours on the treatment days. First, households had the incentive to change their usage immediately before or after the treatment hours. For example, customers faced with a high critical peak price may increase their air conditioner usage right before the treatment hours to cool down or warm up their rooms. Similarly, they may 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 is high during peak hours, the marginal cost during shoulder hours is also likely to be high. The shift in consumption from peak hours to shoulder hours, if any, could lower the economic benefits of peak hour usage reductions.[14]

The second possibility is that consumers may shift their consumption to off-peak hours, which are hours outside peak hours and shoulder hours. In general, the marginal cost of electricity is lower in off-peak hours. Therefore, such shifts in consumption are likely to provide an economic benefit. The third possibility is that consumers may reduce their usage in all hours, including shoulder hours and off-peak hours. For example, consider that consumers bear a fixed cost of changing their lifestyle in terms of electricity usage (Wolak, 2011). When they are faced with a substantial increase in peak hour electricity price, they may change their lifestyle to be more energy efficient, which can have a similar spillover effects for other hours.

[Table 4 about here]

Table 4 provides the results of empirical tests for these possibilities. We estimate equation (1) by including data from different hours for each column. Column 1 shows the result for the treatment hours, which is equivalent to the results in Table 2. Column 2 shows the result for the shoulder hours, namely, three hours before and after the treatment hours. Finally, column 3 includes data for all 24 hours, showing the average treatment effect for all hours of the treatment days. For both treatments, we do not find shifts in consumption between the treatment hours and other hours.

[14] We could design a dynamic pricing schedule that includes relatively high prices for the shoulder hours to incentivize conservation, which could reflect the time-varying marginal costs more effectively. Similar to previous studies (Wolak, 2011; Jessoe and Rapson, 2014), our price changes focused on the critical peak hours only. This was mainly because we wanted to make the price change salient and easy to interpret for consumers.

Instead, we find usage reductions during the shoulder hours and all-day hours for the monetary incentive group. In contrast, we do not find such spillover effects for the moral suasion group. Their point estimates are negative but statistically insignificant. Our findings imply that moral suasion does not induce a spillover effect, whereas monetary incentives produce usage reductions by about 5 percent for the summer and by 3 percent for the winter in the nontreatment hours.

3.4 Habit Formation

In the previous section, we find that the two treatments produced different results in terms of the persistence and spillover effects on the nontreatment hours of the treatment days. These findings suggest the possibility that customers faced with extrinsic incentives may form a habit for energy-efficient lifestyles, which could be why we observe consistent reductions in consumption. To explore a potential habit formation effect, we collected usage data for the periods *after* the end of the final interventions. Customers did not receive any treatments for these periods. If our treatments did not form the habit, we should observe the same levels of consumption between the control and treatment groups.

[Table 5 about here]

Table 5 provides the results of the empirical tests for habit formation. For both summer and winter, we employ usage data for the three-month period after the end of the final treatment to examine if usage levels differ among the control and treatment groups. We find that the moral suasion group's usage is not statistically different from the control group's. In contrast, households that received the extrinsic treatment continued to have lower consumption by about 10 percent relative to the control group's consumption. Unfortunately, we cannot analyze if this habit formation lasted beyond the three-month period for each season because customers received treatments after these periods. However, the robust findings for the summer and winter provide evidence that the extrinsic treatment created habit formation, which lasted for at least three months after the final intervention.

3.5 Heterogeneity

In this section, we explore potential heterogeneity in the treatment effects. We interact the treatment dummies with two policy-relevant household characteristics—household income and average electricity usage. Policymakers are often concerned about how these policy instruments affect higher- and lower-income households differently. In addition, it is important to know how high and low consumers of electricity respond to these policies when policymakers target a certain level of aggregate usage reductions. We collected household income from the preexperiment survey. We use electricity usage data from the preexperiment period to calculate the household-level average daily electricity use in kWh per day. We rescale the unit of the income variable to hundred thousand dollars.

Table 6 shows the estimation results for equation (1), which includes the interaction terms for household income and usage. Although we find weak evidence for the moral suasion effect being larger for higher income households, the estimates are not statistically significant. We find a consistent relationship between extrinsic treatment and income; the extrinsic treatment 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) in column 2 implies that an increase in household income by $10,000 would be associated with a 1.26 percentage point increase for the coefficient for the extrinsic treatment dummy variable (i.e., a 1.26 percentage point decrease in the treatment effect). For usage, we find that the extrinsic effect is larger for high users in the summer, but we do not find such effects in the winter.

[Table 6 about here]

3.6 Mechanisms Behind the Effects of the Two Treatments

We find significant differences in the persistence of the two treatments, namely, moral suasion and monetary incentives. Moral suasion was effective only for the first few treatment days and did not induce a persistent effect. Monetary incentives, in contrast, produced strong persistent effects on energy conservation. To investigate the mechanisms behind the findings, we conducted a detailed follow-up survey. We examine two potential channels. The first possibility is that the treatments may have induced durable goods investments—households in the treatment groups may have purchased energy-efficient appliances in response to the treatments. If this effect was systematically large for the monetary incentive group, it could explain why we observed persistent usage reductions for this group. The second possibility is that the treatments may have induced behavioral changes in lifestyles. Suppose that customers had "bad habits" of using energy inefficiently. It is possible that experiencing high electricity prices may have triggered a change in their lifestyles, encouraging them to become more energy efficient.

[Table 7 about here]

We begin by testing the first possibility in Table 7. We asked customers if they purchased energy-efficient appliances since the start of the experiment. We estimate 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, therefore, gives the ratio of control group customers who purchase an energy-efficient appliance. The coefficients for the treatment dummy variables indicate a percentage point increase in case a group purchases an energy-efficient appliance.

The moral suasion increased the purchase of air conditioners by 8 percentage points, whereas the monetary incentive increased the same by 9 percentage points. These estimates suggest that customers in the two treatment groups had similar significant increases in purchasing energy-efficient air conditioners compared to the control group. We do not find statistically significant effects for other products. These results suggest that durable goods investments are unlikely to explain why we find significant differences in the persistent effects between the two treatments.[16]

[Table 8 about here]

Table 8 explores the second potential channel, namely, behavioral changes in lifestyles. After the experimental period, we asked customers two questions. The first question inquired about their efforts toward adopting an energy-efficient lifestyle. Customers evaluated their lifestyles in terms

[15]For robustness checks, we also run probit and logit models. Our results are the same as those of the linear probability model.

[16]An important caveat from the evidence about durable goods investments is that the customers in our experiments were aware about the time span of their experimental period, namely, one year. For example, more consumers find durable goods investments economical if they know that economic incentives will be available for a longer time period.

of energy efficiency on a scale of 1 (lowest) to 5 (highest). We regress this score on the dummy variable for each treatment group and a constant term. Column 1 implies that the monetary incentive increased this score by 0.4 from the baseline level of 3.03. We find a slight positive effect for the moral suasion group, but it is statistically insignificant.

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 estimate a linear probability model, in which the dependent variable is binary choice, and the independent variables include dummy variables for each treatment group. The model also contains a constant term. For each appliance, we find that monetary incentives had a statistically significant effect by 8 to 15 percentage points. In contrast, moral suasion did not have statistically significant effects on the energy-efficient use of each appliance.

Although these results are based on stated survey responses, they provide suggestive evidence about the mechanisms behind our findings. The significant differences in the persistent effects between the treatment groups are unlikely to originate from durable goods investments. Instead, they are likely to originate from the fact that customers who experience high electricity prices make lifestyle-related changes in terms of energy efficiency.

4 Welfare Implications

When pursuing a variety of policy goals, policymakers can design policies to influence intrinsic or 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.

4.1 Conceptual Framework

We introduce a simple conceptual framework for a model of electricity consumers to guide our welfare analysis. When there is no need for or effort made toward energy conservation, a 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 a consumer is requested to reduce power usage in some hours, he/she may voluntarily decrease his/her 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 monetary terms, Pg, is added to the numeraire y, which totals $Y = y + Pg$. Alternatively, $Y = M - Px$ from the budget constraint of a consumer with income M.

We assume that utility is additively separable into utility $u(x)$ from consuming electricity, utility $v(M - Px)$ from numeraire consumption, and utility $\phi(g;\theta)$ from conservation of electricity. We further assume that each utility function is increasing ($u' > 0$, $v' > 0$, and $\phi' > 0$) and concave ($u'' < 0$, $v'' < 0$, and $\phi'' \leq 0$) in each variable. 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 $\phi'_\theta = \frac{\partial \phi'}{\partial \theta} < 0$. 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 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 contributions by other consumers or a disutility component of social pressure.

The BAU consumption level $\bar{x}$ without any treatment can be expressed as $\bar{x} = \arg\max\{u(x) + v(M - Px)\}$ without utility from electricity conservation. A consumer in the monetary incentive group receives only a price signal and simply adjusts his/her consumption, such that $u' - Pv' = 0$, responding to the price changes. A consumer in the moral suasion group receives a conservation request without monetary incentives. When he/she receives an appeal calling for conservation, he/she maximizes the following overall utility function:

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

This problem can be rewritten as follows:

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

Let x^* denote the optimal solution for the maximization problem (4), namely, the optimal consumption level under a conservation request. Note that x^* satisfies $u' - Pv' - \phi' = 0$.

The effect of repeated interventions on voluntary conservation can be easily derived by differentiating the first order condition for the optimization problem (4).[17] Simple calculation yields

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

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, repeated interventions decrease voluntary conservation of electricity. This result is consistent with our empirical findings. In the next section, we use this conceptual framework to highlight the welfare implications of our empirical findings.

4.2 Welfare Gains from the Two Policies

We highlight the welfare implications of the two types of policy instruments, policies appealing for intrinsic motivation and those providing extrinsic motivation, in the context of the electricity market in Japan. Recall that the fundamental inefficiency in most 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 85 yen/kWh.[18] In the baseline case, consumers paid 25 yen/kWh for their electricity, the average residential electricity price in Japan in 2012.

We consider two policies. The first policy is our monetary incentive treatment. We consider that consumers availing of this policy pay the price that equals the marginal cost, which is 85

[17] Total differentiation of the first order condition for (4) gives $(u'' + P^2 v'' + \phi'')dx - (Pv'' + \phi'' \bar{x}_M)dM - (v' - Px^* v'' + \phi'' \bar{x}_P)dP - \phi_{\theta}' d\theta = 0$. Thus, we have $x_{\theta}^* = \frac{\phi_{\theta}'}{u'' + P^2 v'' + \phi''} > 0$ with $dM = dP = 0$.

[18] We conduct the same analysis based on different assumptions for the marginal cost (65 and 105 yen/kWh) in the Appendix.

yen/kWh. The second policy is our moral suasion treatment. Consumers offered this policy pay the baseline price but receive shocks to their intrinsic motivation for conservation in the form of a request for voluntary conservation.

Consider a quasi-linear utility function for equation (4). To be consistent with our empirical estimation for electricity demand, we characterize the electricity demand by $\ln x = a + \alpha D + \epsilon \ln p$, where D equals 1 if consumers receive the moral suasion treatment, p is the electricity price, and ϵ is the price elasticity. Parameters α and ϵ are obtained from our experiment.[19] The inverse demand is defined by $p(x) = x^{\frac{1}{\epsilon}} \cdot (\exp(a) \cdot \exp(\alpha D))^{-\frac{1}{\epsilon}}$.

The baseline consumption is $\bar{x} = \exp(a) \cdot 25^{\epsilon}$. When consumers receive the monetary incentive, the usage becomes $x_e = \exp(a) \cdot 85^{\epsilon}$. The efficiency gain is characterized by $\int_{x_e}^{\bar{x}} (c - p(x))dx$, the area between the marginal cost c and the demand curve x in the range between x_e and $\bar{x}$. We begin by calculating this efficiency gain for the Japanese electricity market. The summer peak hour electricity consumption from residential customers is about 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 9 about here]

Column 1 of Table 9 shows the efficiency gain from the monetary incentive policy. With the short-run policy, the total efficiency gain for the 3 treatment days is \$25.44 million. We then calculate the welfare gains for the repeated policy with 15 treatment days based on the estimated parameters (β) from our experimental findings for the repeated interventions. Because the responses to the monetary incentive treatments (β) do not decay much, more treatment days provide further efficiency gains. With 15 treatment days, the efficiency gain is \$118.88 million. The difference between the short-run and repeated polices is \$93.44 million and statistically significant. These results suggest that 1) the monetary incentive policy can provide substantial efficiency gains for

[19] Recall that we estimate α (the effect of the moral suasion) and β (the effect of monetary incentives). We use β for the case with treatment price 85 yen kWh to calculate the price elasticity $\epsilon = \beta / \ln(85/25)$.

the electricity market, and 2) repeated interventions can obtain further gains when there are many critical peak demand days, during which time the marginal cost of electricity becomes very high.

When consumers receive moral suasion, the usage becomes $x_i = \exp(a) \cdot \exp \alpha \cdot 25^{\epsilon}$. The efficiency gain is characterized by $\int_{x_i}^{\bar{x}} (c - p(x))dx$, which is shown in Column 2. With the short-run treatment, the efficiency gain is \$17.28 million, which is lower than the gain with the monetary incentive treatment, but it still has a meaningful magnitude for the market. Because the moral 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), we consider that consumers who receive a conservation request change their usage from $\bar{x}$ to x_i because they feel a pleasure such as warm glow or self-satisfaction from behaving prosocially. In this case, consumers obtain a surplus from their consumption $g = \bar{x} - x_i$.[20] If we assume a linear functional form for $\phi(\cdot)$, we can calculate the exact welfare gain from warm glow by using $\int_{x_i}^{\bar{x}} \phi(x)dx = \phi \cdot (\bar{x} - x_i)$. We can calculate the lower bound for the welfare gain without imposing this assumption. For a general form of $\phi(x)$, the lower bound is $\int_{x_i}^{\bar{x}} (p(x) - 25)dx$. Intuitively, consumers have to attain at least this lower bound gain, which equals the loss in their consumer surplus from deviating their ex ante optimal consumption $\bar{x}$ to the new consumption x_i. We provide the sum of the efficiency gain and warm glow in the last column of Table 9. The results suggest that if there is a positive gain from warm glow, the total welfare gains from the moral suasion policy can be close to the gains from the monetary incentive policy. However, this is not the case for the repeated intervention, in which the welfare gain is much larger for the monetary incentive policy even if we incorporate potential gains from warm glow. Finally, these results suggest that while, in theory, welfare gains can also originate from the warm glow effect, in our case at least, the major welfare gains originated from the efficiency gains from having consumers face effective prices that are closer to the actual marginal cost of electricity

[20] Note that our experiment was not able to identify if the conservation from the moral suasion group originated from such a positive motive or other motives. For example, another possibility is that consumers reduced usage because they felt social pressure (DellaVigna, List and Malmendier, 2012) or the need for obedience. In such cases, it is possible that consumers may lose a surplus when they receive a conservation request. Although we think that in our experimental setting, consumers were more likely to conserve because of their prosocial motives, our experiment cannot empirically distinguish these motives. We, therefore, provide the welfare changes from the efficiency gain and warm glow separately in this analysis and interpret the gain from warm glow with this caution.

for the critical peak hours.

5 Conclusion

We used a randomized field experiment to study the persistence of moral suasion and monetary incentives and its welfare implications. Using high-frequency electricity usage data at the household level, we found that moral suasion induced a usage reduction of 8 percent in the short run, which is economically and statistically significant for improving economic inefficiency in electricity markets. However, the effect diminished quickly after repeated interventions. In contrast, the monetary incentive produced larger and persistent consumption reductions of 17 percent. The monetary incentive also resulted in habit formation after the final intervention. Our follow-up survey data indicated that most of the persistent changes were likely to originate from behavioral changes in lifestyle rather than durable goods investments. Finally, in the welfare analysis, we highlighted that both policy instruments would create substantial welfare gains to electricity markets, though the welfare gains from the monetary incentive dominate those from moral suasion when we consider persistence for repeated interventions.

References

Allcott, Hunt, and Todd Rogers. Forthcoming. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review.*

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

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. 2001. "The Trouble With Electricity Markets (and some solutions)." Regulation2point0 Working paper 394.

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

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.

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.

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. 2014. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *Working Paper.*

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.

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.

Frey, Bruno S., and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out." *The American Economic Review*, 87(4): 746–755.

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, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." *The Quarterly Journal of Economics*, 115(3): 791–810.

Gneezy, Uri, and John A List. 2006. "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments." *Econometrica*, 74(5): 1365–1384.

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.

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. Forthcoming. "Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program." *American Economic Journal: Economic Policy*.

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.

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*, 1197–1207.

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. 2006. "Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment." *The Quarterly Journal of Economics*, 121(2): 747–782.

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.

Loewenstein, George. 2005. "Hot-cold empathy gaps and medical decision making." *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, 24(4 Suppl): S49–56.

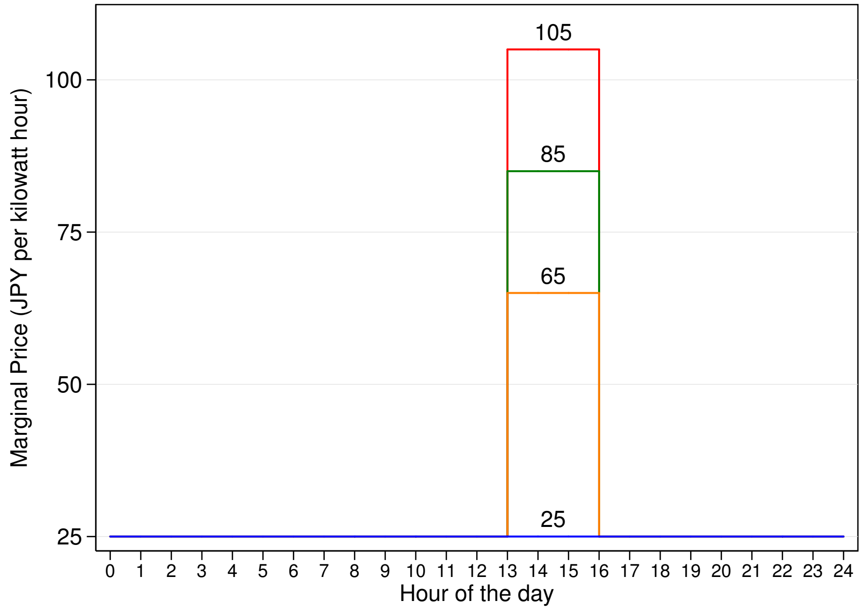
Reiss, Peter C., and Matthew W. White. 2008. "What changes energy consumption? Prices and public pressures." *RAND Journal of Economics*, 39(3): 636–663.

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. PMID: 19213683.

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

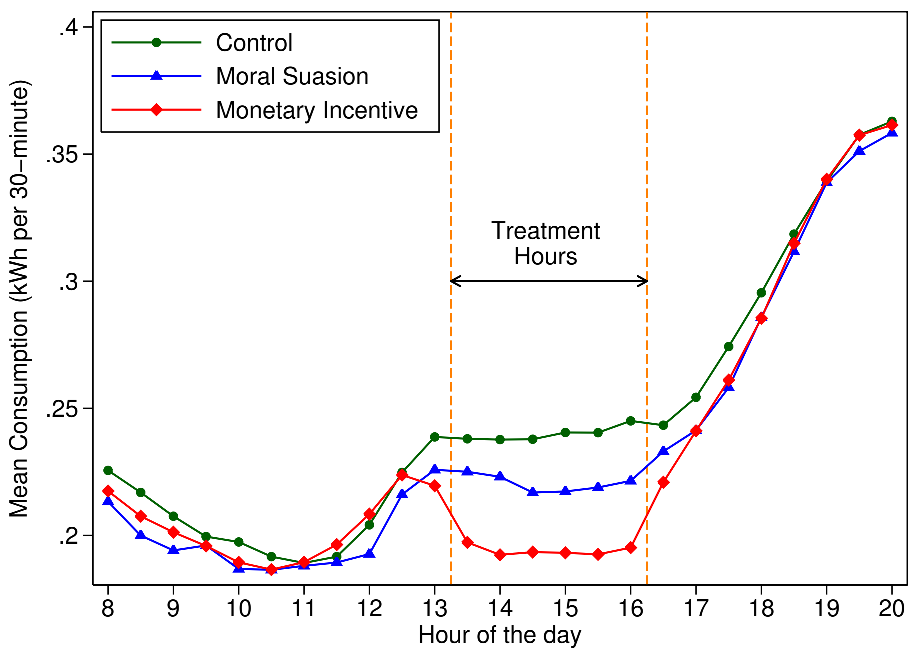
Wolak, Frank A. 2011. "Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment." *The American Economic Review*, 101(3): 83–87.

Figure 1: Monetary Incentives: Dynamic Electricity Pricing



Notes: This figure shows the dynamic electricity pricing schedule for the monetary incentive group and the control price (25 cents/kWh). Although participants paid in Japanese yen, we use U.S. cents in this figure. One Japanese yen was approximately equivalent to one U.S. cent in 2012.

Figure 2: Effects of Moral Suasion and Monetary Incentives on Electricity Consumption



Notes: This figure shows the mean of electricity usage in 30-minute intervals for the control, intrinsic, and extrinsic groups for the summer treatment days. We calculate the mean usage using data from all treatment days in the summer.

Table 1: Summary Statistics

	Sample in the Field Experiment			Random	Difference
	Moral Suasion (S)	Monetary Incentive (I)	Control Group (C)	Sample of Population (P)	[S.E.] (C)−(P)
Electricity use (kWh/day)	15.14	15.76	15.92	16.23	-0.31
	(6.91)	(8.49)	(8.47)	(7.97)	[0.86]
Household income (1,000USD)	66.74	66.59	67.06	66.83	0.22
	(31.49)	(31.34)	(31.01)	(41.81)	[3.93]
Square meter of the house	121.49	113.08	122.15	125.90	-3.75
	(57.54)	(46.92)	(46.52)	(59.65)	[5.41]
Number of AC	3.46	3.50	3.68	3.95	-0.28
	(1.93)	(1.67)	(1.64)	(1.71)	[0.16]
Mean age of the household	42.26	42.22	40.31	41.91	-1.60
	(17.67)	(19.07)	(17.38)	(16.76)	[1.56]
Age of building (years)	13.83	13.39	13.12	15.05	-1.92
	(8.25)	(7.54)	(8.20)	(8.11)	[0.75]
Household Size	3.21	3.14	3.32	2.98	0.34
	(1.18)	(1.23)	(1.25)	(1.41)	[0.13]

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 three groups. Column 4 shows the mean and standard deviation of observables for a random sample of the population in the area of our experiment. We collected the data to investigate the external validity of our sample. Column 5 presents the difference in the means between the field experiment's control group and the random sample. 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 Monetary Incentives on Electricity Consumption

	Summer		Winter	
	(1)	(2)	(3)	(4)
Moral suasion	-0.031	-0.031	-0.032	-0.032
	(0.014)	(0.014)	(0.020)	(0.020)
Monetary incentive	-0.167		-0.173	
	(0.021)		(0.022)	
Monetary incentive (price = 65)		-0.151		-0.163
		(0.022)		(0.024)
Monetary incentive (price = 85)		-0.167		-0.164
		(0.023)		(0.023)
Monetary 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 and Persistence of Treatment Effects

	Summer		Winter	
	Moral Suasion (α_c)	Monetary Incentive (β_c)	Moral Suasion (α_c)	Monetary Incentive (β_c)
1st cycle	-0.083	-0.179	-0.083	-0.185
	(0.024)	(0.023)	(0.030)	(0.027)
2nd cycle	-0.033	-0.196	-0.023	-0.205
	(0.025)	(0.027)	(0.034)	(0.035)
3rd cycle	-0.005	-0.158	0.003	-0.160
	(0.029)	(0.028)	(0.029)	(0.028)
4th cycle	-0.015	-0.128	-0.033	-0.161
	(0.028)	(0.029)	(0.029)	(0.028)
5th cycle	-0.003	-0.091	-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)

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 Nontreatment Hours on Treatment Days

	Summer			Winter		
	Treatment Hours (1pm-4pm)	Shoulder Hours (10am-1pm, 1pm-4pm)	All Day	Treatment Hours (6pm-9pm)	Shoulder Hours (3pm-6pm, 9pm-12pm)	All Day
	(1)	(2)	(3)	(4)	(5)	(6)
Moral suasion	-0.031	-0.010	-0.011	-0.032	-0.008	-0.010
	(0.014)	(0.008)	(0.006)	(0.020)	(0.014)	(0.012)
Monetary incentive	-0.167	-0.054	-0.049	-0.173	-0.029	-0.033
	(0.021)	(0.014)	(0.010)	(0.022)	(0.016)	(0.014)
Observations	123106	374285	1006114	244891	637124	1910367

Notes: This table shows the estimation results for equation (1) for the treatment hours and other hours on the treatment days. 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
	Treatment Hours [1pm-4pm]	Treatment Hours [6pm-9pm]
	(1)	(2)
Moral suasion	-0.034	-0.036
	(0.049)	(0.050)
Monetary incentive	-0.089	-0.097
	(0.040)	(0.045)
Observations	214385	128198

Notes: This table shows the estimation results for equation (1) for the period without treatments (i.e., after the treatments were withdrawn) for both treatment groups. 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: Heterogeneity in the Treatment Effects

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

Notes: This table shows the estimation results for equation (1) with the interaction terms of the treatment dummies with demographic variables. The income variable is in hundred thousands U.S. dollars. The usage variable is the average daily electricity usage in the preexperimental period. 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 7: Durable Goods Investments in Response to Moral Suasion and Monetary Incentives

	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)
Monetary 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.

Table 8: Behavioral Changes in Response to Moral Suasion and Monetary Incentives

	Energy-efficient lifestyle (Degree: 1 to 5)	Energy-efficient use of appliances (Dependent variable: binary choice)				
		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)
Monetary 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). 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.

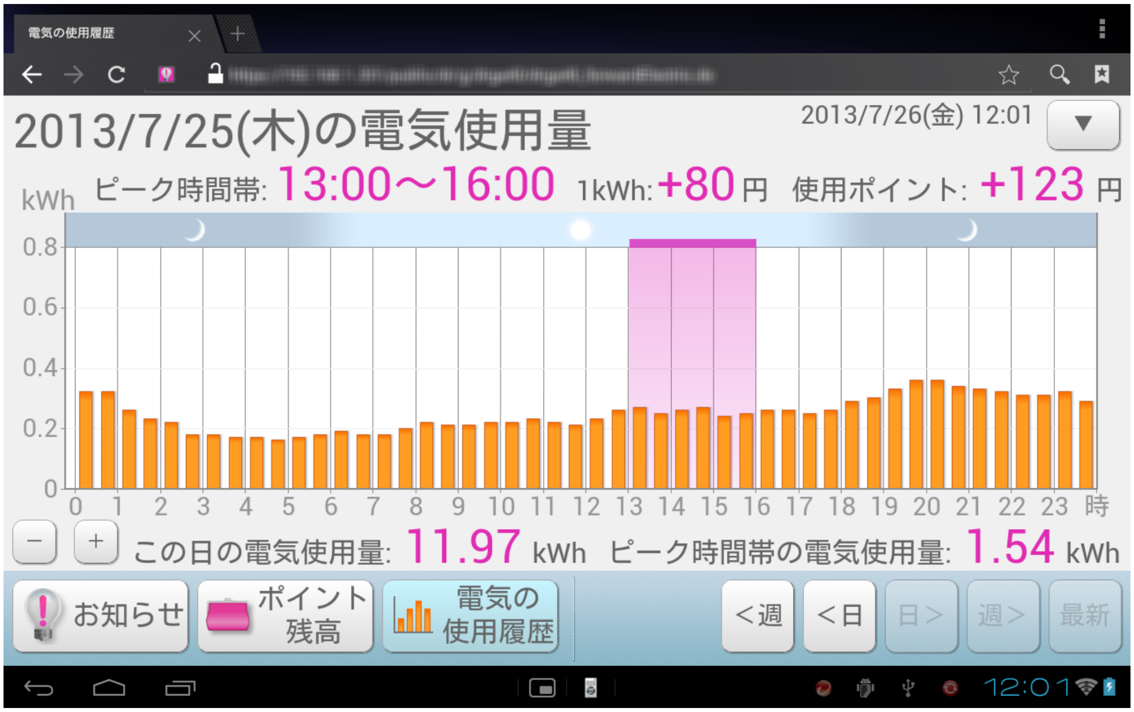
Table 9: Welfare Gains from the Two Policies (When Marginal Cost = 85 yen/kWh)

	Monetary Incentive	Moral Suasion	
	Efficiency Gain ($M)	Efficiency Gain ($M)	Efficiency Gain + Warm Glow ($M)
Short-Run Treatments (3 days)	25.44	17.28	22.16
	(3.00)	(3.92)	(6.71)
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. We use 85 yen/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.

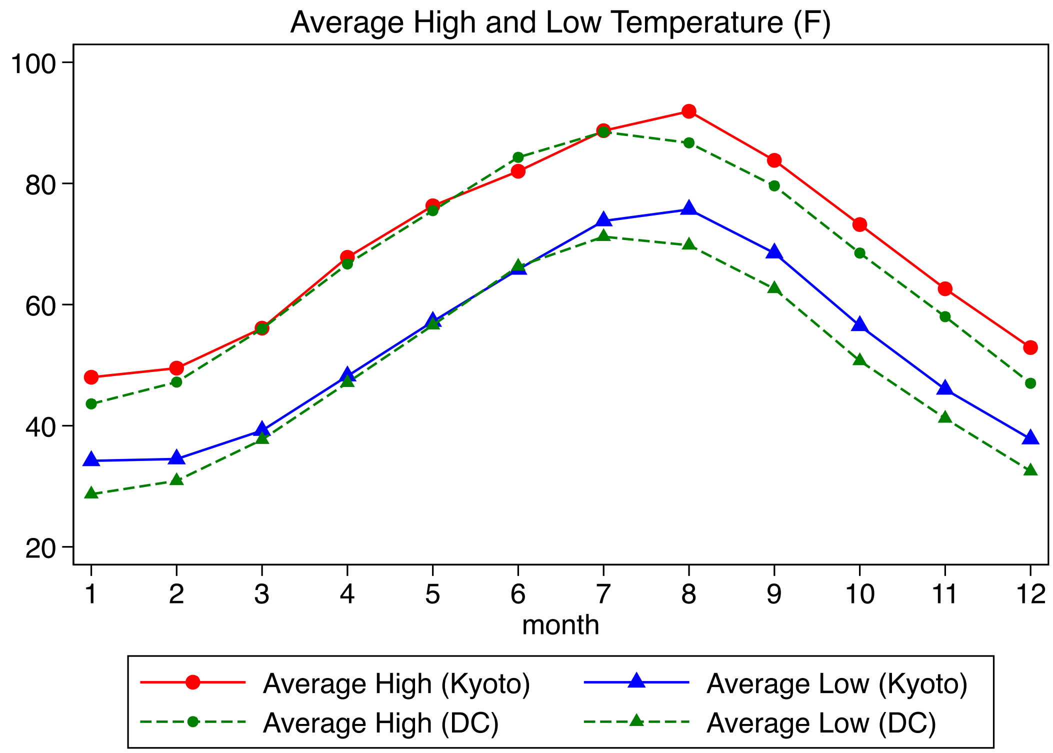
Online Appendix (Not For Publication)

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.

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



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

Table A.1: Welfare Gains from the Two Policies (When Marginal Cost = 65 yen/kWh)

	Monetary Incentive	Moral Suasion	
	Efficiency Gain ($M)	Efficiency Gain ($M)	Efficiency Gain + Warm Glow ($M)
Short-Run Treatments (3 days)	16.38	11.29	15.07
	(1.94)	(2.50)	(4.64)
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. For this table, we use 65 yen/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 = 105 yen/kWh)

	Monetary Incentive	Moral Suasion	
	Efficiency Gain ($M)	Efficiency Gain ($M)	Efficiency Gain + Warm Glow ($M)
Short-Run Treatments (3 days)	34.81	23.38	29.16
	(4.10)	(5.39)	(8.72)
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 yen/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.