

NBER WORKING PAPER SERIES

DOES INFORMATION FEEDBACK FROM IN-HOME DEVICES REDUCE ELECTRICITY
USE? EVIDENCE FROM A FIELD EXPERIMENT

Shahzeen Z. Attari
Gautam Gowrisankaran
Troy Simpson
Sabine M. Marx

Working Paper 20809
<http://www.nber.org/papers/w20809>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2014

This research was supported by funding from the National Science Foundation Cooperative Agreement (SES-0345840 and SES-0951516) awarded to the Center for Research on Environmental Decisions (CRED) at Columbia University. We thank Jonathan Rose Companies, Jonathan Rose, Nicole Sherwood, and Richard Nurse for access to the building and participants; ThinkEco Inc., Heidi Perry, Erika Diamond, Amanda Lurie, and Mei Shibata for installation coordination and access to plug-level electricity data for participants; Ken Gillingham, David H. Krantz, Derek Lemoine, Arik Levinson, Beth Shinn, Tom Wallsten, Elke Weber, and numerous seminar participants for helpful comments; and Galen Treuer, Andrew Wessbecher, and Cale Reeves for research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w20809.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Shahzeen Z. Attari, Gautam Gowrisankaran, Troy Simpson, and Sabine M. Marx. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Does Information Feedback from In-Home Devices Reduce Electricity Use? Evidence from a Field Experiment

Shahzeen Z. Attari, Gautam Gowrisankaran, Troy Simpson, and Sabine M. Marx

NBER Working Paper No. 20809

December 2014

JEL No. D03,D12,L94,Q30,Q40,Q54

ABSTRACT

There is limited evidence of behavioral changes resulting from electricity information feedback. Using a randomized control trial from a New York apartment building, we study long-term effects of information feedback from “Modlet” in-home devices, which provide near-real-time plug-level information. We find a 12–23% decrease in electricity use for treatment apartments, concentrated among individuals reporting higher willingness-to-pay for an energy monitoring system. Decrease in overall electricity use is similar among treatment apartments which received Modlets and those which declined Modlets, and does not specifically occur for outlets with Modlets. This decrease may be due to a Hawthorne or salience effect.

Shahzeen Z. Attari
School of Public and Environmental Affairs
Indiana University Bloomington
sattari@indiana.edu

Troy Simpson
Center for Research on Environmental Decisions
Columbia University
troyjsimpson@gmail.com

Gautam Gowrisankaran
Department of Economics
University of Arizona
P.O. Box 210108
Tucson, AZ 85721-0108
and HEC Montreal
and also NBER
gowrisankaran@eller.arizona.edu

Sabine M. Marx
Center for Research on Environmental Decisions
Columbia University
sm2234@columbia.edu

An online appendix is available at:
<http://www.nber.org/data-appendix/w20809>

1. Introduction

Carbon dioxide emissions from man-made sources are contributing to global climate change (IPCC, 2013) and residential electricity use accounts for about 14% of overall U.S. emissions.¹ Research has shown that improvements in efficiency and conservation can offer great potential to decrease an individual's carbon footprint.² Specific household behaviors, such as weatherizing one's home and changing thermostat settings, could reduce direct household emissions by 30% (Gardner & Stern, 2008). More importantly, researchers believe that these changes would cause only limited decreases in consumer well-being (Dietz, Gardner, Gilligan, Stern, & Vandenberg, 2009). There is evidence that people misperceive which behaviors are effective at reducing electricity use.³ For these reasons, it is important to investigate ways to facilitate sustained decreases in household electricity use that may help realize these potential savings.

The high potential value of reducing household electricity use has spurred significant innovation in devices that provide detailed information on electricity use. We study *in-home devices*, which provide energy feedback information in a residential setting.⁴ Some popular in-home devices include Google PowerMeter coupled with The Energy Detective (TED) 5000 by Energy, Inc., which provides household-level information; and Conserve by Belkin and Modlets by ThinkEco, Inc., which both provide plug-level feedback. Many global companies see these technologies as an important area for further investment. For instance, in 2014, Google purchased Nest, creator of the Nest Learning Thermostat, for \$3.2 billion.

¹ In 2012, U.S. residential electricity CO₂ emissions allocated by sales accounted for 760 million metric tons of the total 5,290 million metric tons emitted nationwide (EIA, 2013b). Energy-related carbon dioxide emissions increased from 2012 to 2013 by 2.5%, primarily due to colder weather and an increase in energy intensity (EPA, 2014).

² The classic paper by Pacala and Socolow (2004) proposes solutions to limiting atmospheric carbon dioxide to a concentration below double the preindustrial concentration of 280 parts per million. They consider energy efficiency, decarbonizing electricity, and biological storage in forests and soils and find that "improvements in efficiency and conservation probably offer the greatest potential" to lower carbon dioxide emissions.

³ For instance, people list curtailment actions such as turning off lights as being the most effective actions they can take to decrease energy use in comparison to efficiency actions such as replacing incandescent light bulbs in contrast to expert recommendations (Attari, DeKay, Davidson, & Bruine de Bruin, 2010; Gardner & Stern, 2008; Kempton, Harris, Keith, & Wehl, 1985).

⁴ There has also been extensive roll-outs of smart meters in the E.U. and U.S. (Faruqui, Harris, & Hledik, 2010). Like in-home devices, smart meters record electricity use at fine time intervals. Unlike in-home devices, smart meters also report use back to the electric utility.

Can real-time information feedback from in-home devices foster long-term electricity conservation? The evidence on whether these devices work in the field is limited and mixed. For instance, one randomized control trial (RCT) study investigated a hardware device coupled with the Google PowerMeter and found a 5.7% short-term reduction in electricity use, but the reduction did not persist past four weeks (Houde, Todd, Sudarshan, Flora, & Armel, 2013). Another RCT study found information in combination with prices to be very effective in affecting household electricity use (Jesoe & Rapson, 2014). Given the rush to develop better in-home devices and install smart meters, an evaluation of whether and in what ways in-home devices help with energy conservation is vital.

The purpose of this study is to design and analyze an RCT field experiment to understand whether near-real-time plug-level information feedback for households can foster long-term energy conservation. A number of authors view the disaggregation of information feedback at the appliance level as an important technique to foster energy savings (Armel, Gupta, Shrimali, & Albert, 2013; Ehrhardt-Martinez, Donnelly, & Laitner, 2010). Here, we examine preliminary versions of Modlets, the plug-level in-home device made by ThinkEco, Inc. Modlets and other plug-level devices allow households to track electricity use at the appliance level. To our knowledge, ours is the first RCT to investigate the effect of plug-level in-home devices. Specifically, our study asks: (1) does information related to how much electricity is being used by each household appliance encourage consumers to change their behavior? (2) does the potential behavior change persist over time? and (3) which consumer characteristics predict the largest behavioral changes from the feedback information (if any at all)?

Our study focuses on the possibility of reductions in residential electricity use that are independent of top-down hard regulations. Electricity demand experiments often focus on market based incentives such as price changes (Borenstein, 2005; Ito, 2014; Wolak, 2007). However, research has shown that people find voluntary action, or soft regulations (such as changes in status quo or defaults), more palatable than hard top-down regulations in helping them decrease fossil fuel use (Attari et al., 2009). Harder regulations may also foster strong negative reactions such as psychological reactance (Brehm, 1966). Finally, recent literature shows that people are affected by non-linear electricity pricing in ways other than predicted by rational economic

models with full information (Ito, 2014). For these reasons, we believe that information-based measures have the potential to facilitate behavior change and are of interest and value, particularly since we do not yet have an economy-wide market mechanism to regulate carbon emissions.

Our study design is as follows: study participants were residents of a newly opened “green” building in East Harlem, New York, NY, called the Tapestry Building. Importantly, the building had variation in income, with some units reserved for low- and middle-income tenants, and the rest leased at market rates. The field experiment, which was run by study authors, started in Dec. 2010. Apartments were randomized into two groups, where those in the treatment group were offered Modlet devices with the plug-level data made available for analysis by this study’s researchers. Apartments which were not in the treatment group were in an observation-only control group.

Our study principally uses four types of data: (1) data on apartment characteristics including an indicator for whether the apartment was in the treatment group and if so, its Modlet install date; (2) billed electricity data over a 30-month period; (3) data from a pre-treatment survey; and (4) detailed 15-minute level electricity use data from the Modlets. We performed the randomization separately by apartment income group and apartment size, assuring a balance of observable characteristics across treatment and control groups. Notably, our experiment used multiple in-home devices that were installed in several outlets in each apartment;⁵ we collected detailed disaggregated data at the plug-level and monthly electricity use data at the apartment-level; and we investigated a long time horizon.

Not all the participants who were designated to receive Modlets actually received them, due to complications with the implementation of the study as well as participants in the treatment arm preferring not to receive the devices. An intention-to-treat estimate reveals that being in the treatment arm led to a 23% decrease in billed electricity use. We also find that the decrease is persistent over the study period. We also have pre-treatment electricity use data on the treatment

⁵ The installation process required installers to enter the study participant’s apartment and label each plug with the device that was using that plug. Installations usually lasted between 30 and 60 minutes.

and control groups. While there is no significant difference between the groups prior to treatment, the control group had a somewhat higher electricity use ex-ante. To be conservative, we calculate a difference-in-difference treatment effect of 12%, which is not significant ($P=0.11$). We believe that the most accurate estimate likely lies somewhere between 12 and 23%. Our findings appear to be largest for designated middle-income apartments and not significant for low-income apartments. We find that a higher ex-ante willingness-to-pay for a home energy monitoring system predicts a larger drop in electricity use from treatment. However, we find no significant heterogeneity in treatment effects with respect to pro-environmental attitudes or numeracy.

Despite these findings on the difference in overall electricity use across the treatment and control groups, a closer look at the data reveals no significant impact of the Modlets themselves. We find this lack of significant effect in two ways. First, the treatment effects were virtually identical across the different subgroups within the treatment group which we examined, notably: (1) apartments that received Modlets, (2) apartments in the treatment group that did not receive Modlets, and (3) apartments that received Modlets and also the firmware update in April, 2012 that allowed them to view electricity use information on the web and also schedule plugs to turn off. Second, our individual plug data from the Modlets show that electricity use on monitored plugs was not significantly different in the 1-30 day period after Modlet installation from the longer-run electricity use, implying no significant decrease in use following Modlet installation. Additionally, only one of the participants who received Modlets scheduled their plugs using the technology, a main potential method of saving electricity.

Thus, our significant effects of treatment do not appear to be a result of the near-real-time information provided by the Modlets themselves. We postulate other explanations for our results, which cannot be disentangled in our current study. These include (a) the Hawthorne effect, where participants who know they are in a study may change their behavior based on the additional attention; and (b) that the extensive outreach to the study participants in the treatment arm for scheduling Modlet installs (by the experimenters, ThinkEco, Inc., and the building management office) made electricity use salient (for example, see Pallak, Cook, and Sullivan (1980)) which resulted in behavior changes for these participants. In either case, the effect would

have to apply to an apartment's overall electricity use behavior, and not for outlets with Modlets. After the study, in early 2014, our team contacted a random subset of participants with the aim of disentangling these competing explanations and understanding the results further. In general, participants did not recall many details about the study nor did they recall consciously changing their behavior if they were assigned to the treatment condition.

Our study builds on two literatures. First, a number of studies (including those mentioned above) analyze the effectiveness of in-home devices, though not on plug-level devices or with plug-level data. A recent review found a 7% reduction in electricity use (Faruqui, Sergici, & Sharif, 2010) though the research reviewed was largely not peer-reviewed nor based on RCTs, but rather based on utility pilot programs. Our study contributes to this literature by analyzing plug-level feedback devices, and by using a randomized design with a long follow-up. Second, recent research has used RCTs to examine the impact of information in fostering long-term decreases in energy use (Abrahamse, Steg, Vlek, & Rothengatter, 2005; Allcott & Rogers, 2012; Asensio & Delmas, 2014; Delmas, Fischlein, & Asensio, 2013; Ehrhardt-Martinez, et al., 2010; Kahn & Wolak, 2013). Notably, experiments by the company OPower that use social norms to inform consumers about residential electricity use have found small, but sustained drops in electricity use caused by social comparison information (Allcott & Rogers, 2012). OPower does not provide real-time electricity use data nor are its data disaggregated to the appliance level. Our study builds on this literature by considering the impact of plug-level in-home devices. Our use of near-real-time plug-level data allows us to better understand the sources of the changes in electricity use that we observe. Additionally, our finding that the electricity reductions do not stem from the plug-level information may also provide a cautionary note regarding interpreting the sources of electricity reduction in other studies.

The paper proceeds as follows: Section 2 describes the experiment and data. Section 3 outlines the model and estimation framework. Section 4 provides the results. Finally, Section 5 offers conclusions.

2. Experiment and data

2.1. Experimental setting

Our field experiment was designed to test the effectiveness of Modlets in reducing household electricity consumption. The participants in the study were residents of the Tapestry building (<http://www.tapestry124.com>).⁶ In Sep. 2010, Jonathan Rose Companies (JRC), co-developer of the Tapestry building, approached Columbia University's Center for Research on Environmental Decisions (CRED) with the idea of designing a study on electricity use in the building. The goal of the study was to evaluate how effective plug-level load monitoring devices could be in reducing energy use. JRC chose the Modlet devices for use in this building before approaching CRED.

The experiment was designed by the study authors who were at CRED at this time (Attari, Simpson, and Marx). The aim was to provide randomized evidence on whether plug-level devices could function effectively in the field. There is substantial evidence that people misperceive their electricity use (Attari, et al., 2010; Kempton & Montgomery, 1982). For this reason, there is the possibility, but not certainty, that detailed, plug-level information might help people make more informed choices, and thereby reduce their electricity use even in the absence of strong monetary incentives.

The Tapestry building was designed and constructed as a “green” building, and is certified to the Gold level by the U.S. Green Building Council's (USGBC) Leadership in Energy and Environmental Design (LEED) green building certification program. Many units in the building are reserved for households in two distinct income groups, low-income and middle-income as defined by the City of New York. The remaining units are market-rate apartments, which would typically rent to people in a higher income group than the reserved apartments. This variation allowed for a study that could evaluate the usefulness of information provided by plug-level devices across three distinct income levels. The Tapestry building provided a useful setting to explore household electricity use patterns, because each apartment was individually metered for electricity, allowing apartment-level electricity use data to be tracked and analyzed in conjunction with the plug-level data from the Modlets.

⁶ The online Appendix provides a map of the Tapestry building.

The building was newly opened at the time the study began, with first tenant lease dates beginning in Jun. 2010. Apartment units at Tapestry came supplied with energy-efficient appliances and fixtures, including ENERGY STAR refrigerators and energy-efficient hard-wired lighting in kitchens, living areas, and bathrooms. Microwaves, ENERGY STAR dishwashers, clothes washers, and electric condensation clothes dryers were also provided in some middle-income and market-rate apartment units. Communal washers and dryers were located in building common areas for units not supplied with those appliances. Stoves in residential units used natural gas, but Tapestry apartment units were not sub-metered for natural gas use.

Electrically powered pumps located in each apartment transported centrally-heated or -cooled water to the apartment for climate control. Tenants did not pay directly for the heating and cooling, but their electricity bills did include the electricity used to run the pumps in their apartments. The online Appendix includes a picture of the thermostat control system located in each apartment. Due to their configuration, the plug-level load monitors could not be used to individually track several appliances and fixtures in apartment units (such as the thermostat, hard-wired lighting, clothes washers and dryers, and some microwaves), but electricity used to run these appliances are included in residents' monthly electricity bills.

The main expected sources of electricity use in the Tapestry apartments are typical of many contemporary apartment buildings: space heating and cooling, refrigeration and freezing, laundry appliances, electronics, and lighting. Average residential electricity consumption per utility customer in New York State is 611 kWh per month (EIA, 2013a) or about 20 kWh per day. We anticipated that electricity use by Tapestry residents would be lower than average due to several factors: the overall efficiency of the building system, the fact that Tapestry is a dense apartment building, the presence of efficient owner-supplied fixtures and appliances, the lack of laundry machines in some units, and the presence of centrally-supported cooling and heating.⁷ Interpretation of our results should account for the fact that some electricity demand (such as cooling, common-area lighting, and clothes washing in some cases) and all non-electricity energy demand are not captured by apartment-level utility bills.

⁷ Levinson (2014) also finds that houses have lower electricity use in the first few years after they are built.

2.2. Modlets

Modlets were under development during our study and were not available for retail purchase. The product, which is now available through major retail sources, has undergone changes since the preliminary version used in our study. When operational, the ThinkEco, Inc. web platform allows users to view electricity use data for each plug in a given Modlet network, manually turn outlets off and on, and program automatic on/off schedules for specific plugs to avoid unwanted energy use in periods of inactivity. Modlet sales information not publicly available.⁸

Modlets plug into standard wall outlets and record energy use data for two plugs (see Figure 1a). They connect wirelessly to a USB key or special Ethernet Gateway using a ZigBee wireless protocol. When the USB key is not connected or the Ethernet Gateway is offline, the Modlet stores electricity data on the device itself for a short period of time until it can reconnect (see <http://thinkecoinc.com> for more details). If the USB is not connected in that period of time, the electricity use data will be lost. After downloading and installing software to connect to the Modlet, the user can access real-time data on device-specific energy use by logging onto a ThinkEco, Inc. website. The Modlet does not display electricity use data on the device itself. Figure 1b shows a screenshot of the mymodlet.com portal (with more screenshots in the online Appendix).⁹ Currently, the Modlets require a two-week baseline period during which normal energy use data is collected and observable to the user. Following the baseline period, users can create an active savings plan by scheduling devices to turn off.

2.3. Design of field experiment

Researchers at CRED set up a randomized field experiment study starting in Dec. 2010 (see Figure 2 for the project timeline). The first event of the study occurred in Dec. 2010, when CRED sponsored two resident outreach meetings with free pizza for attendees. At the meetings,

⁸ ThinkEco, Inc. and ConEd have recently installed 10,000 Modlets for air conditioners (see <http://www.coned.com/newsroom/news/pr20120426.asp> Accessed December 19, 2014). These devices are different from the Modlets used in our study.

⁹ The screenshots are taken from 2014 Modlet software which shows electricity use at the 1-minute level. The Modlets in our study showed use at the 15-minute level.

residents were informed that if they took part in the study, they would be randomly selected to be in either the treatment group or in the control group. Apartments in the treatment group would receive Modlets for free while apartments in the control group would not.

Under a protocol approved by the Columbia University Institutional Review Board, CRED distributed a resident pre-treatment survey (henceforth called pre-survey) and a consent form in individually addressed packets to all building residents in early Feb. 2012. The consent forms authorized the merging and analysis of apartment-specific electricity use and survey data in a manner that protected the confidentiality of study participants.

The goal of the pre-survey was to provide baseline evidence on the characteristics of study participants to evaluate how the effectiveness of Modlet treatments varied across individual characteristics. The survey asked for various information regarding individual characteristics, their knowledge of household electricity consumption, and their willingness to change their own behavior. Each participant was paid \$10 upon return of the pre-survey. The pre-survey is available in the online Appendix.

2.4. Randomization

The Tapestry building is a mixed-income 184 unit apartment building, consisting of 37 apartment units designated as “low-income” (LI), 55 units designated as “middle-income,” and 92 units designated as “market rate.”¹⁰ Apartment units range in size from 0-bedroom (studio) to 3-bedroom units across the various income designations. Low-income apartments were available to households earning no more than 40% to 50% of area median income, corresponding to a maximum of \$17,250 to \$44,450, depending on household size and the number of bedrooms in the leased apartment. Middle-income apartments were available to households earning no more than 175% of area median income, corresponding to a maximum of \$49,645 to \$138,600, depending on household size and number of bedrooms. Low-income households within the qualified income range could apply to the City of New York for subsidized housing. The city then determined housing assignments by lottery, implying that low-income Tapestry residents

¹⁰ The study excludes the one remaining unit, which was designated for use by the building superintendent.

did not choose Tapestry specifically. In contrast, middle-income residents would apply directly to JRC for the chance to reside in Tapestry, with the selection again based on lottery. Market rate units were priced according to current market conditions and did not have maximum household income limits.

In Feb. 2011, following the distribution and collection of consent forms and the pre-surveys, we divided the 156 at-the-time occupied apartments into cells based on the resident's income group and number of bedrooms. We then randomly assigned a fixed number of apartments within each income/bedroom cell into the treatment group. Specifically, we sorted all apartment units by the number of bedrooms (0, 1, 2, and 3 bedrooms, respectively), and then by the three income levels, in ascending order. This resulted in twelve cells. For each of the twelve cells, each unit was assigned a random number between 0 and 1. Then all units were sorted, within each cell, in ascending order by their random number. The first half of the units were assigned to the control group, and the second half of the units were assigned to the treatment group.¹¹ By randomizing within cells, we assure an adequate number of treatment and control observations in each cell. This list of unit conditions was distributed to the building management staff and ThinkEco, Inc. in late Feb. 2011. Table 1 shows the number of apartments in the study by income class, study participation, study assignment, and data availability. Of the 184 apartments in the building, 92 were leased at market rates with the remainder leased to either low-income or middle-income households.

After the pre-survey administration process was complete, the management staff and ThinkEco, Inc. began outreach to treatment group apartments to inform them of their opportunity to have Modlets installed in their apartments. Outreach was intensive and included placing fliers under the door and door hangers. Additionally the building manager called residents multiple times to schedule Modlet installation. Near the end of Apr. 2011, we randomized the 28 apartments that were previously unoccupied into treatment and control groups.

¹¹ In any case in which there was an odd number of units within a given category, the unit with the highest randomly assigned number was subsequently assigned a "tie-breaker" random number (also between 0 and 1). If that number was less than 0.5, the unit was assigned to the control group, and if the number was equal to or greater than 0.5, the unit was assigned to the treatment group.

After confronting challenges scheduling Modlet installations in a number of treatment group apartment units, in Jul. 2011, we decided to increase the treatment group size to help increase the odds of ultimately installing Modlets in a large enough number of apartments to collect meaningful data. To accomplish this, we randomly reassigned 30 of the 92 control group units to the treatment group. We applied the same random selection process described above to these new units. At the conclusion of this step, 122 units were in the treatment group and 62 units remained in the control group. Apartments randomly assigned to the treatment group – at the initial point as well as at subsequent points – comprise the *intention-to-treat* group.

Ultimately, there were two rounds of Modlet installations, both of which were coordinated by building management and ThinkEco, Inc. The first round of Modlet installations in 40 apartment units ran from Apr. 18, 2011 through Jul. 18, 2011, as shown in Figure 2. The second round of installations, which included 10 new installations and 11 reinstalls due to a firmware upgrade, ran between Apr. 15, 2012 and May 5, 2012. For the first round of Modlet installations, residents were informed of their placement in that condition on a rolling basis beginning in Mar. 2011. For the second round of Modlet installations, participants were informed of their new, or continued, placement in the treatment group on a rolling basis beginning in Apr. 2012. Only study participants with round 2 Modlets could view electricity use data on the web and schedule Modlets to turn plugs off. The ability to view electricity use data started on June 18, 2012, and the ability to schedule plugs started on July 16, 2012.

Overall, response to participating in the study was high, with 101 of the 184 apartments agreeing to participate in the study prior to the random assignment into treatment and control groups.¹² Response rates are similar across income class, ranging from 51% to 59% across the three income groups. Due to technical difficulties, we could not obtain ConEd electricity data for 13 of these apartments, leaving 88 apartments in the analysis sample. Of the apartments in the analysis sample, residents of 80 apartments completed parts of the pre-survey.

2.5. Data

¹² Some people were asked to participate in the study after the assignment to treatment groups. To avoid any potential concerns regarding sample selection, we discarded individuals for whom we could not verify the date at which they agreed to participate or for whom the assent date was after the date of randomization.

Our analysis principally merges data from four sources. First, we use apartment-level data on treatment assignment collected by CRED and data related to characteristics of residents collected by JRC. These data include the random assignment into treatment and control groups for each apartment, as well as information on whether the apartment was initially in the treatment group or added to the treatment group in 2012. These data include apartment characteristics such as the designated income level and square footage of the apartment. Finally, they also include the date that Modlets were installed in apartments in the treatment group.

Second, we use data on the ConEd billed electricity use. We obtained these data from the property management company, which uses a third-party online energy management and benchmarking tool operated by EnergyScoreCards, Inc., to access and track resident utility bills. The system uses resident account numbers, supplied by residents with their consent, to access and record energy use data from the local utility. The data are at the bill level, which generally corresponds to one apartment over one month. Each bill provides the total billed electricity use as well as the start and end date of the bill. We use this information to create a variable that measures the electricity use per day. We merge the apartment-level data with these data.

Third, we use pre-survey data that included a variety of information from each participant. From these data, we principally use three measures that we determined would be interesting to relate back to electricity use. First, we use pro-environmental attitudes, as measured by the New Ecological Paradigm (NEP), a 15-item instrument (Dunlap, Van Liere, Mertig, & Jones, 2000). We coded the 15 original responses (1 = *completely disagree*, 7 = *completely agree*) in the pro-environmental direction and then calculated an individual's NEP score as the mean value across the responses. Second, we use a numeracy assessment (Schwartz, Woloshin, Black, & Welch, 1997). This consisted of three non-multiple choice questions designed to test numerical aptitude. For example, "In the BIG BUCKS LOTTERY, the chance of winning a \$10 prize is 1%. What is your best guess about how many people would win a \$10 prize if 1000 people each buy a single ticket to BIG BUCKS?" Finally, we use the log of the reported willingness-to-pay for a home energy monitoring system with many features (such as the user can install the monitoring system themselves, the user can access the information from the web and phone, etc., see features

represented in Figure 3). Given that each adult resident of each apartment could fill out the survey, we have multiple responses per apartment in some cases. In these cases, we used the mean of the responses over the apartment. Twenty-one percent of our pre-survey participants reported that they would not be willing to pay anything for such a home energy monitoring system. To avoid taking the log of 0, we use the log of 1 plus the reported willingness-to-pay, measured in dollars.

Finally, we use data collected from the Modlet devices by ThinkEco, Inc. These data are at the level of the Modlet plug at a given time. Each Modlet plug records the time of day and date of use, the time period of observation, the plug function, and the total electricity consumed by this plug. As each outlet typically has two plugs, each will have two output streams reflecting two separate Modlet plugs. Generally, the time period of observation for a Modlet plug is a 15-minute interval (although newer Modlet devices measure electricity use in 1-minute intervals). Modlet installers inputted the plug function (i.e., the end-use device(s) connected to each plug) at the time of installation.

During the course of the experiment, we encountered some complications with the distribution and installation of Modlets. Specifically, not everyone who was in the intention-to-treat group received Modlets due either to residents not being interested in participating or difficulty contacting residents and/or scheduling device installations. In addition, in some cases where the Modlet was in fact installed, the date of Modlet installation was not recorded. Finally, recorded data from the Modlets are sporadic. The reasons for sporadic data recording, which are not easily discernable for any given apartment, may include device malfunction or failure of participants to plug the wireless USB key into their computers frequently enough (or at all) to upload data to the ThinkEco, Inc. cloud database. Thus, we define a separate *received Modlets* treatment group, which is a subset of the intention-to-treat group. Note that assignment to the received Modlets groups is potentially non-random, unlike assignment to the intention-to-treat group.

Table 2 provides details on the control and treatment groups of the analysis sample. From the 88 apartments in the analysis sample, 24 were in the control group and 64 in the treatment group. Among the 64 apartments in the intention-to-treat group, we have Modlet data from 32

apartments, which are the apartments in the received Modlets group. Only 15 of these 32 apartments in the received Modlet group received the upgraded Modlets in round 2. For the apartments which received Modlets, we have data on a mean of 14.8 Modlet plugs per apartment (recall that there are two plugs per Modlet device). As we would expect given our randomization, the NEP (pro-environmental attitude) and numeracy scores are very similar across the control and intention-to-treat group; they are also similar in the received Modlets group. The mean willingness-to-pay for an energy monitoring system is appreciably higher for the control group than for the intention-to-treat group, although this difference is not significant at the 5% level ($P=0.06$).¹³

Table 3 provides further analysis on the characteristics of the analysis sample by income. Overall, electricity use is relatively similar across the three income groups, ranging from 12.4 kWh/day for low-income apartments to 13.9 kWh/day for middle-income apartments. Apartment sizes are also similar across income groups, although slightly smaller for the low-income apartments. The two characteristics that are not similar across the groups are the mean numeracy score, which is higher for middle-income than for low-income apartments; and highest for market-rate apartments; and the willingness-to-pay for an energy monitoring system, which is lowest for low-income apartments. The overall Modlet use is similar across income groups. The use during different hours is also similar across income groups, with the heaviest use during the evening.

Table 4 provides characteristics of the Modlet data. We coded each Modlet plug into 10 functions based on the plug function text entered by the Modlet installers. In some cases, we coded a Modlet plug as having multiple functions. Hence, in Table 4, the number of Modlet plugs by function (in all rows but the last) adds up to more than the total (in the last row). Not surprisingly, the lowest electricity use is for plugs coded as “Empty.” The highest is for refrigerators, followed by “Routers, modems, and cable boxes,” and “Entertainment devices.” The Modlet plugs capture an average per-apartment electricity use of about 2 kWh/day in the 32 treated apartments. This is a small fraction (roughly 15%) of average household electricity use for apartments in our sample.

¹³ Our analysis uses logged willingness-to-pay. The difference in log willingness-to-pay across the control and intention-to-treat groups is not significant at even the 10% level ($P=0.22$).

3. Model and estimation framework

3.1. Model

We propose a very simple treatment effects model of apartment-level electricity use. In our model, electricity use in an apartment over a period of time is a function of external conditions (e.g., weather), preferences of the household, and treatment. Treatment reflects apartments which were offered Modlets, which we call the intention-to-treat. In further analysis, we split the intention-to-treat group into subgroups, based on having received Modlets and firmware upgrades.

In general terms, we can write electricity use as:

$$\text{Use}_{it} = f(\text{Conditions}_i, \text{Preferences}_i | \text{Treatment}_{it}). \quad (1)$$

In equation (1), i denotes an individual apartment; t denotes time; Use_{it} measures electricity use per day; Conditions_i measures the external conditions; Preferences_i measures household preferences; Treatment_{it} takes on values of 0 (for control) and 1 (for treatment); and $f(\cdot)$ is a random function of these variables.

We define the average treatment effect as the electricity use for the treatment group relative to the control group as:

$$\begin{aligned} \text{Treatment effect} = & E\left[f(\text{Conditions}_i, \text{Preferences}_i | \text{Treatment}_{it} = 1)\right] \\ & - E\left[f(\text{Conditions}_i, \text{Preferences}_i | \text{Treatment}_{it} = 0)\right], \end{aligned} \quad (2)$$

where the expectation $E[\cdot]$ in equation (2) is taken over the distribution of external conditions and preferences in our full sample.

Given that we have randomly assigned apartments into the treatment and control groups, Preferences_i will be independent from Treatment_{it} . If we further restrict our sample of treatment and control groups to have identical external conditions, then Conditions_i will also be orthogonal to Treatment_{it} . For instance, this orthogonality condition will be satisfied if we observe all treatment and control apartments over the exact same time periods. Thus, if this orthogonality condition holds, this model implies that we can estimate the treatment effect of receiving Modlets by taking the mean use in the received Modlets group and subtracting the mean use of the control group.

Our estimation adapts the basic model defined by equation (2) in three ways. First, individuals may drop out of the sample, due to a change in domicile, for instance. Thus, to help control for any non-random selection out of the sample, we include time dummies.

Second, treatment may vary over time due to a number of factors. For instance, treated individuals may learn slowly from the Modlets how to lower their electricity use. Alternately, they may experience a “Hawthorne effect,” where they lower electricity use due to the feeling of being watched, but this effect may diminish over time or may last until the study is completed (Schwartz, Fischhoff, Krishnamurti, & Sowell, 2013). In this case, we redefine treatment to indicate the length of time since the start of the intervention.

Finally, there may be heterogeneity in treatment across individuals. Our basic model allows for this possibility as it considers the expectation across individuals. Nonetheless, we may be interested in separately analyzing the treatment effect across different groups of residents. We can do this by stratifying our sample based on participant characteristics, for instance, by the designated income level of the apartment, and then examining the impact of treatment for each stratified sample.

3.2. Estimation framework

A central point of our analysis is to examine whether electricity use differs between the treatment and control groups. We estimate the treatment effect by regressing electricity use on treatment and month dummies:

$$Use_{it} = \beta \text{Treatment}_{it} + \alpha_t + \varepsilon_{it}, \quad (3)$$

where α_t are month dummies and ε_{it} reflects the deviation between electricity use and expected electricity use conditional on treatment and month dummies. The unit of observation for our regressions is an apartment observed over one electricity billing cycle, which is generally one month in length.

Given the random assignment of individuals to treatment or control groups in our study, Treatment_{it} is independent from ε_{it} provided that there is no differential selection across apartments out of the sample. Under these assumptions, our estimator of β will be an unbiased and consistent estimator of the mean treatment effect.

We now discuss some important details regarding our estimation. First, we observe several observations for each apartment in equation (3), i.e., we observe several bills per apartment during our sample period. These observations are likely not statistically independent and indeed our choice of time period for the unit of observation is somewhat arbitrary. Thus, we cluster our standard errors at the level of the apartment. By clustering standard errors at the level of the apartment, we ensure that our estimates are robust to the presence of correlated observations within an apartment.

Second, we observe pre-treatment electricity use data for both the control and treatment groups. We also examine pre-treatment use data to uncover whether there is differential electricity use between the treatment and control groups prior to the implementation of the treatment.

Third, we examine the heterogeneity of our treatment effects in three ways: (1) we include graphs that allow us to examine the extent to which the treatment effect is time varying; (2) we include regressions where we stratify treatment by income group of the rented apartment (low-income, middle-income, and market), and (3) we include regressions where we examine the interaction of treatment with pre-survey responses.

Finally, while our model concerns the intention-to-treat group, we also examine specifications where treatment denotes apartments that actually received Modlets and the 2012 firmware upgrade. We use the original control group for comparison here. It is possible that the decision to accept Modlets correlates with ε_{it} and hence we cannot view these estimates as fully causal. Nonetheless, we believe that we can evaluate the impact of Modlets versus other aspects of treatment by comparing regressions with the intention-to-treat definition of treatment to ones with receiving devices as the definition of treatment.

In addition to regressions based on equation (2), we also estimate regressions where we examine the plug-level impact of having Modlets using our Modlet plug-level data. These regressions allow us to understand more directly how quickly the reduction in electricity use occurs and where it occurs. The Modlet data allow us to answer two related questions. First, was electricity use for outlets with Modlets different in the days after installation than in the longer run? Second, which type of devices connected to the Modlets saw the biggest change in electricity use? For these regressions, we regress electricity use for a Modlet at the 15-minute level on plug fixed effects, date and hour indicators, and the number of days since the first Modlet observation for that plug. The date and hour indicators here are necessary because of the lack of a control group. A limitation is that we only examine the time since the Modlets were installed, but participants could start logging on to the website and scheduling plugs on particular dates (June 18 and July 16, 2012, respectively). Thus, this analysis, which includes date fixed-effects, only measures the effect of Modlet installation and not the effect of these other functionalities on plug-level use.

4. Results and Discussion

4.1. Results

We first analyze electricity use separately for the control and treatment groups. Table 5 presents these figures across the control and intention-to-treat groups, and also splits the intention-to-treat group into columns for apartments that received Modlets, that did not receive Modlets, and that received both Modlets and the 2012 firmware upgrades.

The first row measures the mean electricity use for each of the groups. The treatment group has a 3.0 kWh/day, or 19%, lower electricity use than does the intention-to-treat group, 12.5 kWh/day compared to 15.5 kWh/day. The mean electricity use in the received-Modlets group is very similar to the intention-to-treat group. Not surprisingly then, the mean electricity use among individuals in the intention-to-treat group who did not receive Modlets is also similar both to the intention-to-treat mean and received-Modlets mean. Finally, the group which received Modlets and the 2012 firmware upgrade also had a similar electricity use level.

In the period Jan. – Mar. 2011, the intention-to-treatment group had an electricity use of 13% less than the control group, of 15.3 kWh/day compared to 13.3 kWh/day, a difference that grows to 22% from Apr. 2011 onwards. Thus, this suggests that at least part of the difference in electricity use between control and treatment groups may be explained by statistical noise or other factors unrelated to the treatment. A difference-in-difference – of the drop in electricity use for the intention-to-treat group relative to the control group – then yields a treatment effect of – 9%.

Figure 4 summarizes the time-varying nature of electricity demand in the treatment and control groups. Figure 4a shows the mean use over time for the treatment group, as a whole and separately for the three groups in Table 5, and the control group, while Figure 4b shows the percentage difference in use between the treatment groups and the control group. The figure shows that there is a gap in electricity use between the control group and all treatments groups. Moreover, this gap increases starting in Apr. 2011 and does not diminish over time. In contrast, the treatment group and its three subgroups all have very similar monthly mean electricity use.

We next turn to a regression framework that further formalizes these results and also allows us to focus on heterogeneous responses, presented in Table 6. Each row of the table corresponds to a regression. For each regression, the table reports standard errors clustered at the apartment level. The table does not report month indicators (included in all specifications) or the coefficient on the pre-survey measure (included in the last three specifications) but reports all other coefficients.

Our base regression selects treatment as an indicator variable for apartments in the intention-to-treat group and at Apr. 2011 or after. We find that treatment caused a 3.5 kWh/day drop in electricity use relative to the control group, a figure that is statistically significant. This represents a 23% drop from the 15.5 kWh/day mean electricity use for the control group, reported in Table 5. In order to evaluate whether some of this treatment effect might be due to random differences in the groups prior to treatment, we next provide a specification with two indicators: the same treatment indicator as earlier and an indicator for the intention-to-treat group but prior to Apr. 2011. We find that the pre-treatment effect on the treatment is negative and -1.6 kWh/day (-10% using the same 15.5 baseline value), but not significant while the treatment effect remains unchanged. Thus, a difference-in-difference estimate here would yield a coefficient of -1.9 kWh/day (-12%) on treatment. Not reported in the table, the difference-in-difference lacks statistical significance ($P=0.11$).

As a rough idea of the potential cost-savings from these devices, note that Modlets currently cost \$45 and up.¹⁴ The mean number of Modlet plugs for apartments in our sample who received Modlets is 14.8 (see Table 2), with two Modlet plugs per Modlet, implying a total purchase cost of about \$333. The mean apartment in our sample spends about \$1,000 per year on electricity.¹⁵ Therefore, the cost buying 7.4 Modlets, accounts for roughly 33% of the total electricity cost per apartment in one year.

¹⁴ ThinkEco, Inc. website http://shop.thinkecoinc.com/products/home-starter-kit#.VHdUakv_gdU Accessed November 28, 2014.

¹⁵ This assumes a rate of 19.4 cents/kWh, which is the average price of electricity for New York state in September, 2014 (see http://www.eia.gov/electricity/monthly/epm_table_grapher.cfm?t=epmt_5_6_a Accessed November 28, 2014).

We next examine the impact of treatment using the received Modlets group. This reveals a similar estimate of -3.3 kWh/day. The impact of treatment among apartments which received Modlets and the 2012 firmware upgrade was also -3.3 kWh/day, though not statistically significant.

We then consider the variation in treatment across income groups using the intention-to-treat group. The largest impact of the treatment is on middle-income apartments where residents earn between \$50,000 and \$140,000. The coefficient on middle-income apartment is the only one that is significant and is the largest in magnitude. There is some evidence that treatment also reduces electricity use in market-level apartments, but it is not statistically significant. The coefficient on low-income apartments is the smallest, perhaps reflecting the fact that low-income residents were the only ones who did not necessarily choose to live in Tapestry and hence that the potential reductions they can actually achieve are the smallest.

Finally, Table 6 examines the variation in treatment across three pre-survey measures: NEP score (pro-environmental attitude), numeracy score, and willingness-to-pay for an energy monitoring system. For each measure, we estimate a specification where we include the interaction between treatment and the measure as a regressor, and also include the treatment indicator and the measure as two additional regressors. For both NEP and numeracy scores, we find that the interaction of treatment with the score is not significant.

In contrast, we find that the interaction of logged willingness-to-pay for an energy monitoring system and treatment is a significant negative predictor of electricity use. In other words, individuals who reported being willing to pay more for a monitoring system disproportionately lowered their electricity use following the treatment. The coefficient estimate implies that a doubling of willingness-to-pay causes a 0.54 kWh/day drop in electricity use from treatment. Although not reported in Table 6, the baseline coefficient on logged willingness-to-pay is 1.1 and statistically significant ($P = 0.001$). This implies that individuals who used more electricity also were, on average, the ones who were ex-ante willing to pay more money for a monitoring system.

Figure 5 shows a scatter plot of electricity use in the pre-treatment period against use by the same apartment in the treatment period. We split the apartments into treatment and control groups and also provide fitted regression lines for both groups. Overall, predicted electricity use in the treatment period for the treatment group is lower than for the control group for most values of pre-treatment electricity use in the sample. In particular, we find that there is a larger drop in predicted electricity use for the treatment group relative to the control group the larger the use in the pre-treatment period. In other words, apartments that used more electricity prior to the study were also able to decrease use more from treatment, which may be due to the fact that they have the greatest potential to decrease electricity use. This result is consistent with two findings from Table 6, when taken together, that (1) individuals with a higher electricity use had a higher willingness-to-pay for a home electricity monitoring system and (2) that individuals with a higher willingness-to-pay had a larger drop in electricity use from treatment.

In Table 7, we report plug-level electricity use following the installation of Modlets. We do not observe plug-level electricity use data from apartments without Modlets, since Modlets are necessary to observe this information. Since this specification does not directly include a control group, we evaluate how electricity use evolved at different plug types following the installation of Modlets. We report use levels for indicator variables covering the first month of Modlet use, with the omitted indicator being outlets that had a Modlet for more than a month. Thus, all coefficients can be interpreted as the electricity use for an outlet relative to the mean use for the same outlet for 31 and more days after the Modlet install. All our specifications include date and hour-of-day indicators and are clustered at the outlet level.

In none of the cases is the impact of use in the 15-30 day period after Modlet installation significantly different from the longer-run use. Excluding outlets marked as “empty,” only two of the 60 coefficients from 1 day after the Modlet were installed or later are significantly different from zero. The empty outlets had significantly negative coefficients for the 1-14-day periods, suggesting that, in many cases, individuals eventually plugged in a device into these outlets.

We find positive coefficients for the day of Modlet install in three categories: lights and lamps, refrigerators, and other devices, implying higher electricity use on the day of install than in the longer-run period. Though statistically significant, these effects are small. For instance, the 0.03 kW effect on “lights and lamps” corresponds to an extra 30W average per plug following the period at which the Modlets started recording. This might just reflect extra use during or after the install, for instance due to some residents staying home from work on the day of the install.

Because participants in the treatment group who received Modlets in the first round of installations never had access to the ThinkEco, Inc. web interface before the devices stopped functioning, we know that none of them viewed Modlet electricity data online. ThinkEco, Inc. also reported to us that no Tapestry users used Modlets to schedule any of their plugs to automatically go off until Dec. 2012, when only one participant created an online schedule. We do not know how many participants in the treatment group accessed Modlet electricity use data on the web after it was made available to them.

Figure 3 reports statistics on pre-survey responses on the importance of different features of monitoring systems. The pre-survey elicited the importance of these features with a five-point scale with “dislike very much” as 1 and “like very much” as 5. The three most desired features, in declining importance, were: (1) saving up to 20% on your electricity bill; (2) monitoring savings in terms of dollars; and (3) checking if each device was functioning correctly. The three least important features were, in increasing importance: (1) the ability to play an online game with neighbors or friends to compare electricity savings; (2) comparison of electricity savings to acres of forests planted; and (3) comparison of electricity use to neighbors or friends.¹⁶ To summarize, users reported wanting their monitoring system to report practical information that would help them save money rather than provide behavioral or motivational information.

At the end of the study, we conducted a brief exit telephone interview, which we gave to a random set of study participants from the intention-to-treat, received-Modlets, and control group

¹⁶ Although participants in our study do not care to compare their electricity use to others, researchers find that such social norms matter. For example, Allcott & Rogers (2012) found that OPower’s social norm cues (such as comparing participants’ electricity use to their most efficient neighbor) significantly decreased electricity use. Additionally, other studies have found that providing information on injunctive and descriptive norms does effectively change behavior (Goldstein, Cialdini, & Griskevicius, 2008).

($N = 22$), after the study was completed (Jan. 2014). The interview protocol can be found in the online Appendix. The interview was designed to understand why there were little to no differences between the intention-to-treat and received-Modlets groups, and what caused the differences between these two groups and the control group. In general, participants who were interviewed scarcely remembered any details of the study. Some participants who had received Modlets claimed to not use them, for example one participant stated “they are still hooked up, but I don't use them at all.” Overall, the exit interview implies that people did not actively remember Modlets, or the study, as helping them reduce their electricity use.

4.2. Discussion

Overall, our findings are that electricity use decreased by 1.9 – 3.5 kWh / day, or 12–23%, in the intention-to-treat group relative to the control group, depending on whether we control for the difference relative to the pre-treatment data or not. We find that study participants who reported being willing to pay more for an energy monitoring system were disproportionately the ones who saw a reduction in electricity use, indicating that they may have been more open to finding ways or had more options to reduce their electricity use.

Despite the findings of an overall drop in electricity use, we see that apartments that were in the treatment arm but did not receive Modlet experienced a similar drop in electricity use to those that received Modlets. Moreover, we found no effect of a decrease in electricity use in outlets with Modlets following the installation. Finally, no treated individuals logged on to the ThinkEco, Inc. website to schedule plugs to automatically switch off, except for one participant later in the study. This is one of the potential energy savings features of the Modlets. Thus, it does not appear that providing the type of real-time information showcased in our study effectively motivated behavior change.

What could explain our seemingly contradictory results, to the extent that they do not reflect random variation? We propose two possible explanations. First, it is possible that the outreach given to the treatment group caused these participants to use less electricity. Participants in the intention-to-treat group were given door hangers and other reminders to schedule Modlet

installations, which may have made electricity savings more salient and spurred them to learn about ways to reduce their electricity use. Second, it is possible that our results are due to a Hawthorne effect (Adair, 1984), where study participants change their behavior not because of the treatment but rather because they knew that investigators were paying attention to their electricity use reduction. The Hawthorne effect has been shown to decrease electricity use by 2.7% per month over the long run for extremely non-invasive communications, with the effects disappearing as soon as the participants know that the study is over (Schwartz, et al., 2013).¹⁷ For both explanations, the effects may be larger in our study given that our treatment was fairly intrusive, involving home installations, door hangers, and face-to-face interactions. Also, the reduction in electricity use would have to occur for activities not related to the Modlet plugs, for example changing settings on the HVAC system, and changing the use of hard-wired lighting and non-monitored plugs.

Why were the Modlets not more effective here? We hypothesize that, despite their innovative design, the preliminary version of the devices used in the study may have required more active engagement than study participants were willing to provide in order to reduce energy use. For instance, using Modlets required plugging in a USB key to one's computer periodically (unless it is left plugged in at all times, which is unlikely in the case of laptop use). It required logging in after the baseline period to actually schedule devices to switch off. Finally, the Modlets used in our study were still undergoing development and needed firmware upgrades and hardware exchanges midway.

While our study only investigates plug-level in-home devices, our results relate to work on a different type of in-home device, the programmable thermostat. Overall, programmable thermostats have failed to save the energy that they were designed to save. As noted on the EPA ENERGY STAR labeling site: "The ENERGY STAR specification for programmable thermostats was suspended on December 31, 2009 and the ENERGY STAR label is no longer available for this category. While EPA recognizes the potential for programmable thermostats to save significant amounts of energy, there continue to be questions concerning the net energy

¹⁷ Schwartz et al. (2013) used a large field experiment where participants only received postcards notifying and reminding them that they were in a study of home electricity use. The authors found evidence of a reduction in electricity use that vanished when the study was over.

savings and environmental benefits achieved under the previous ENERGY STAR programmable thermostat specification.” The question remains as to why programmable thermostats failed to save energy. A likely reason is that consumers did not program these thermostats and that manufacturers did not set effective default settings that made them useful in the absence of substantial consumer effort. More generally, Levinson (2014) finds that building construction codes meant to encourage energy efficiency did not save energy, while Fowlie *et al.* (2014) find only modest savings from subsidies for weatherization.

Our study has several limitations. First, due to the study design, the Modlets used in our study were preliminary versions where the users could not connect to the web or schedule devices immediately. In addition, there are (non-statistically-significant) pre-treatment differences between the control and treatment groups even though we had random assignment, reflecting differences between these groups due to our limited sample size. Given that all the apartment units were in one building, there could have been spillovers across treated apartments or between treated and control apartments. We also can only track residents according to the activity of their utility accounts, so if tenants move, we lose track of them. Also, we cannot track energy used for central heating or cooling (though we do track the electricity used to pump hot or cold air into each apartment). We do not track energy use for cooking, as stoves are natural gas and Tapestry apartment units are not sub-metered for gas, nor do we capture building common area energy use. Participants who initially received Modlets in round one were not given access to the web interface and scheduling, which may have deterred them from interacting with the devices immediately after installations.

5. Conclusion

This paper reports on an RCT field experiment implemented by study investigators to test the usefulness of in-home devices in reducing electricity consumption. To our knowledge, this study comprises the first RCT field experiment of plug-level in-home devices that provide real-time feedback. We find that electricity use in the treatment arm dropped but that this was likely not due to residents meaningfully interacting with the Modlets. We find heterogeneity in the treatment effects, with a larger drop in electricity use for individuals who were ex-ante willing to

pay more for a home energy monitoring system, indicating that they may have been more open to finding ways to reduce their electricity use.

Even though the existing literature has found that real-time information feedback at the household level has the potential to decrease electricity use, our RCT does not support this hypothesis. Instead we find that the possible reductions in electricity use achieved by study participants in our treatment group might be due to other factors, such as salience and a Hawthorne effect. We believe that our findings of a significant reduction in electricity use are driven by factors other than the information from Modlets serves as a caution for other studies that have found effects from electricity conservation programs regarding the sources of any reductions that they are finding. The use of detailed data, such as plug-level data in our case, can help in more accurately pinpointing the sources of reductions.

Why were these in-home devices not effective in the intended way? As with programmable thermostats, using these devices requires a significant investment in time, effort, and motivation. We believe that the user interfaces of in-home devices will have to be modified for them to be useful for the average user. For instance, it may be helpful to limit the user entry of information to one time. It may also be helpful to simplify the programming, eliminate the physical burden of repeatedly connecting a USB key to a computer, set appropriate default rules for the devices, and provide direct feedback mechanisms that do not require accessing a web-based interface.

Finally, we believe that more RCT experiments such as ours are needed to evaluate if newer in-home devices can work more effectively in terms of actually creating electricity savings from real-time information feedback coupled with other forms of motivational cues. We hope our work helps facilitate additional research in this area.

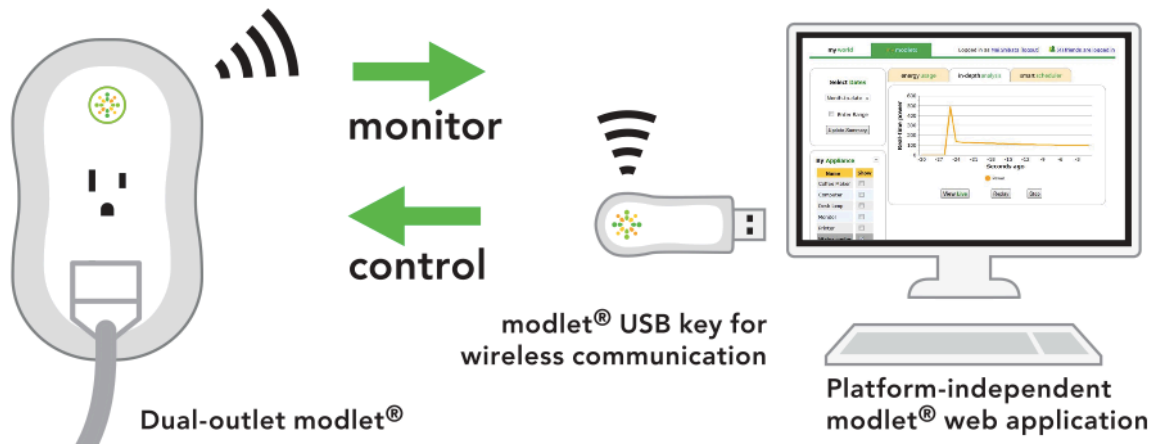
References

- Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25(3), 273-291.
- Adair, J. G. (1984). The Hawthorne effect: A reconsideration of the methodological artifact. *Journal of Applied Psychology*, 69(2), 334.

- Allcott, H., & Rogers, T. (2012). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation: National Bureau of Economic Research.
- Armel, K. C., Gupta, A., Shrimali, G., & Albert, A. (2013). Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy*, 52, 213-234.
- Asensio, O. I., & Delmas, M. A. (2014). The dynamics of information framing: The case of energy conservation behavior *Mimeo.*: University of California, Los Angeles.
- Attari, S. Z., DeKay, M. L., Davidson, C. I., & Bruine de Bruin, W. (2010). Perceptions of energy consumption and savings. *Proceedings of the National Academy of Sciences*, 107(37), 16054-16059.
- Attari, S. Z., Schoen, M., Davidson, C. I., DeKay, M. L., Bruine de Bruin, W., Dawes, R. M., & Small, M. J. (2009). Preferences for change: Do individuals prefer voluntary actions, soft regulations, or hard regulations to decrease fossil fuel consumption? *Ecological Economics*, 68(6), 1701-1710.
- Borenstein, S. (2005). The long-run efficiency of real-time electricity pricing. *The Energy Journal*, 93-116.
- Brehm, J. W. (1966). *A Theory of Psychological Reactance*. New York: Academic Press.
- Delmas, M. A., Fischlein, M., & Asensio, O. I. (2013). Information Strategies and Energy Conservation Behavior: A Meta-analysis of Experimental Studies from 1975-2012.
- Dietz, T., Gardner, G., Gilligan, J., Stern, P. C., & Vandenbergh, M. (2009). Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. *Proceedings of the National Academy of Sciences*, 106(4), 18452-18456.
- Dunlap, R., Van Liere, K., Mertig, A., & Jones, R. (2000). New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: A revised NEP scale. *Journal of Social Issues*, 56(3), 425-442.
- Ehrhardt-Martinez, K., Donnelly, K. A., & Laitner, S. (2010). Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities (Vol. E105): American Council for an Energy-Efficient Economy Washington, DC.
- EIA. (2013a). How much electricity does an American home use? Table 5A. Residential average monthly bill by Census Division, and State 2011. Retrieved 10/17/13
<http://www.eia.gov/tools/faqs/faq.cfm?id=97&t=3>
- EIA. (2013b). U.S. Energy related carbon dioxide emissions, 2012.
- EPA. (2014). U.S. energy-related carbon dioxide emissions, 2013 Retrieved November, 28th, 2014, from <http://www.eia.gov/environment/emissions/carbon/>
- Faruqui, A., Harris, D., & Hledik, R. (2010). Unlocking the, 53 billion savings from smart meters in the EU: How increasing the adoption of dynamic tariffs could make or break the EU's smart grid investment. *Energy Policy*, 38(10), 6222-6231.
- Faruqui, A., Sergici, S., & Sharif, A. (2010). The impact of informational feedback on energy consumption—A survey of the experimental evidence. *Energy*, 35(4), 1598-1608.
- Fowlie, M., Greenstone, M., & Wolfram, C. (2014). *Do energy efficiency investments deliver? Evidence from the weatherization assistance program*. Mimeo. University of California, Berkeley.
- Gardner, G., & Stern, P. (2008). The short list: The most effective actions U.S. households can take to curb climate change. *Environment* 50(5), 12-24.

- Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, 35(3), 472-482.
- Houde, S., Todd, A., Sudarshan, A., Flora, J. A., & Armel, K. C. (2013). Real-time feedback and electricity consumption: A field experiment assessing the potential for savings and persistence. *The Energy Journal*, 34(1), 87-102.
- IPCC. (2013). Climate Change 2013: The Physical Science Basis. *Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Retrieved October 6th, 2013, from <http://ipcc.ch/report/ar5/wg1/> - .UIHC5hZTDbU
- Ito, K. (2014). Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *American Economic Review*, 104(2), 537-563.
- Jessoe, K., & Rapson, D. (2014). Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use. *American Economic Review*, 104(4), 1417-1438.
- Kahn, M. E., & Wolak, F. A. (2013). Using Information to Improve the Effectiveness of Nonlinear Pricing: Evidence from a Field Experiment. *Mimeo. University of California, Los Angeles*.
- Kempton, W., Harris, C. K., Keith, J. G., & Wehl, J. S. (1985). Do Consumers Know" What Works" in Energy Conservation? *Marriage & Family Review*, 9(1), 115-133.
- Kempton, W., & Montgomery, L. (1982). Folk quantification of energy. *Energy*, 7(10), 817-827.
- Levinson, A. (2014). *How much do building energy codes really save? Evidence from California*. Mimeo. Georgetown University.
- Pallak, M., Cook, D., & Sullivan, J. (1980). Commitment and energy conservation. *Applied Social Psychology Annual*, 1, 235-253.
- Schwartz, D., Fischhoff, B., Krishnamurti, T., & Sowell, F. (2013). The Hawthorne effect and energy awareness. *Proceedings of the National Academy of Sciences*.
- Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The Role of Numeracy in Understanding the Benefit of Screening Mammography. *Annals of Internal Medicine*, 127(11), 966-972.
- Wolak, F. A. (2007). Residential customer response to real-time pricing: The anaheim critical peak pricing experiment. *Center for the Study of Energy Markets*.

Figure 1: Information on Modlets
(a) Image of dual-outlet Modlet system used in the Tapestry building



Source: ThinkEco, Inc.

(b) Screenshot of the mymodlet.com portal

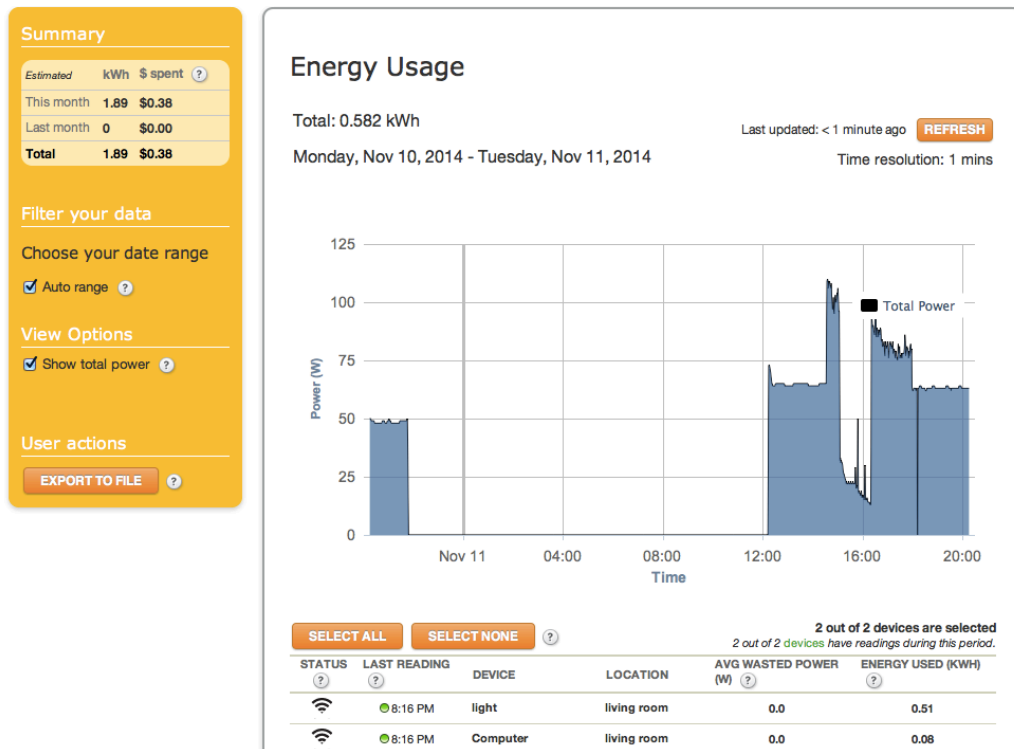


Figure 2: Project timeline and overview

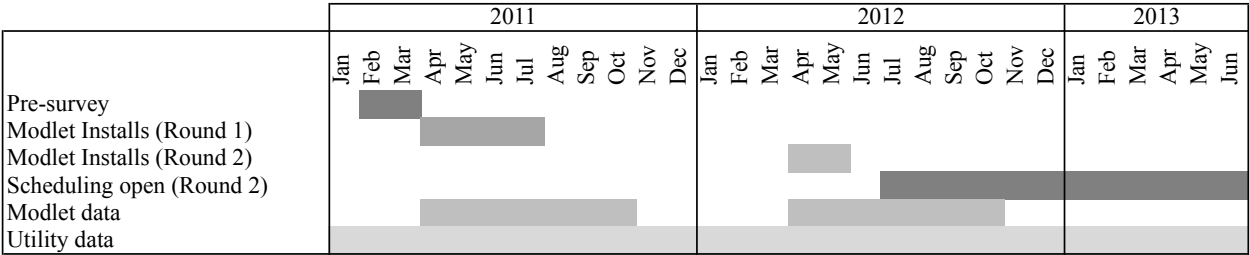
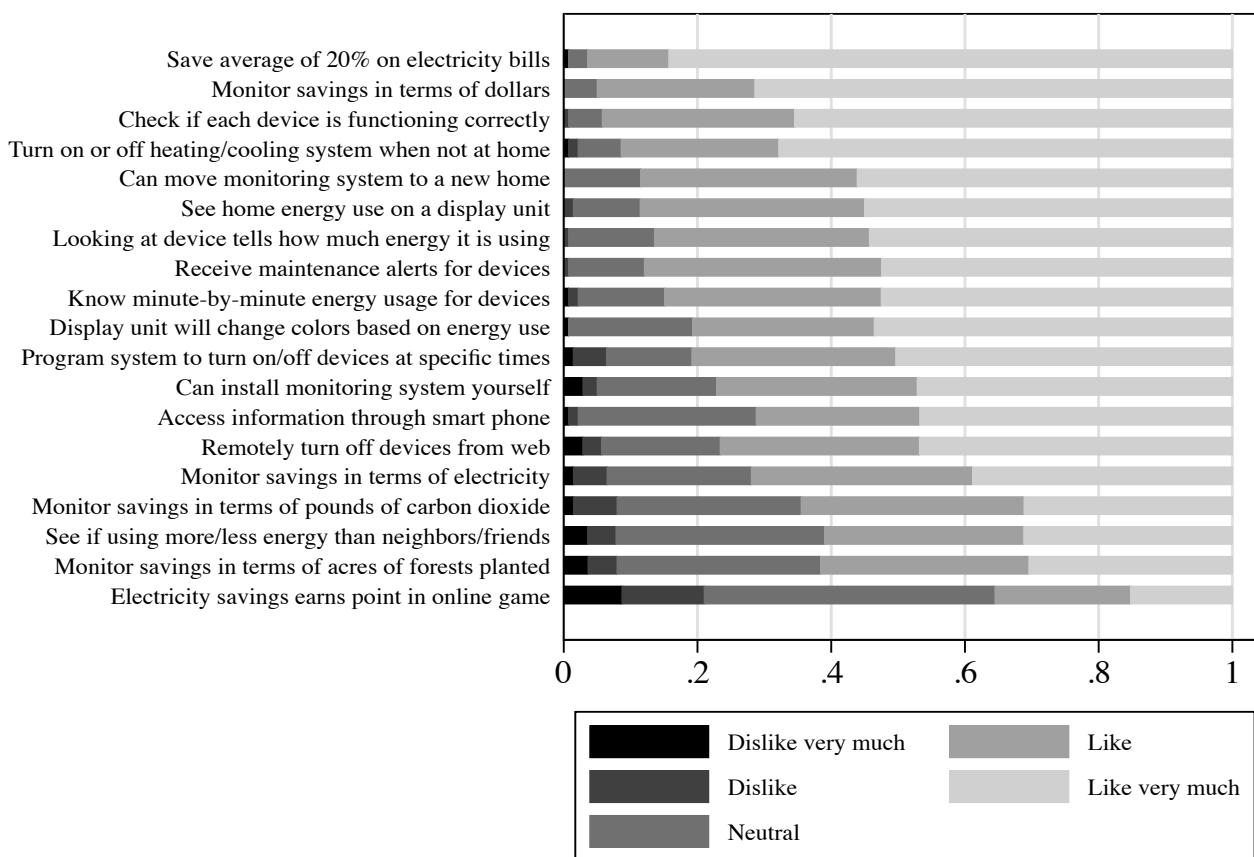
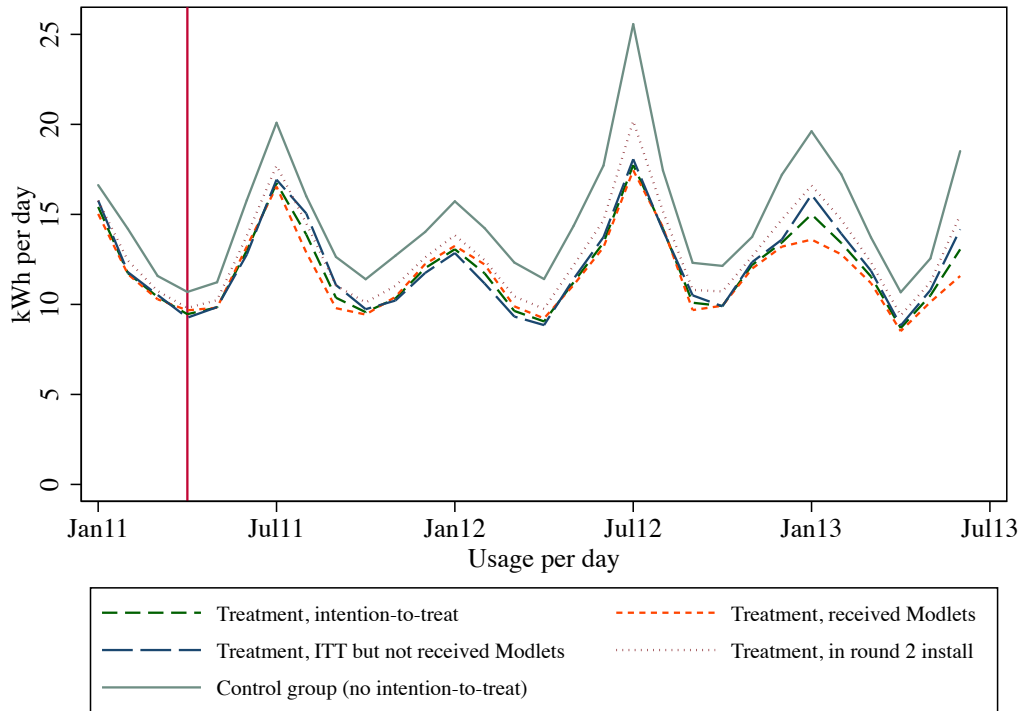


Figure 3: Preferences for energy monitoring systems characteristics



Note: this figure presents statistics of the survey responses for the 138-141 individuals who responded to these questions in the pre-survey. This sample differs from the analysis sample because not all individuals who responded to the survey are in the analysis sample and the analysis sample is at the apartment level while this figure is at the individual level.

Figure 4: Electricity use for treatment and control groups
(a) Treatment groups and control group use



(b) Treatment groups use relative to control group use

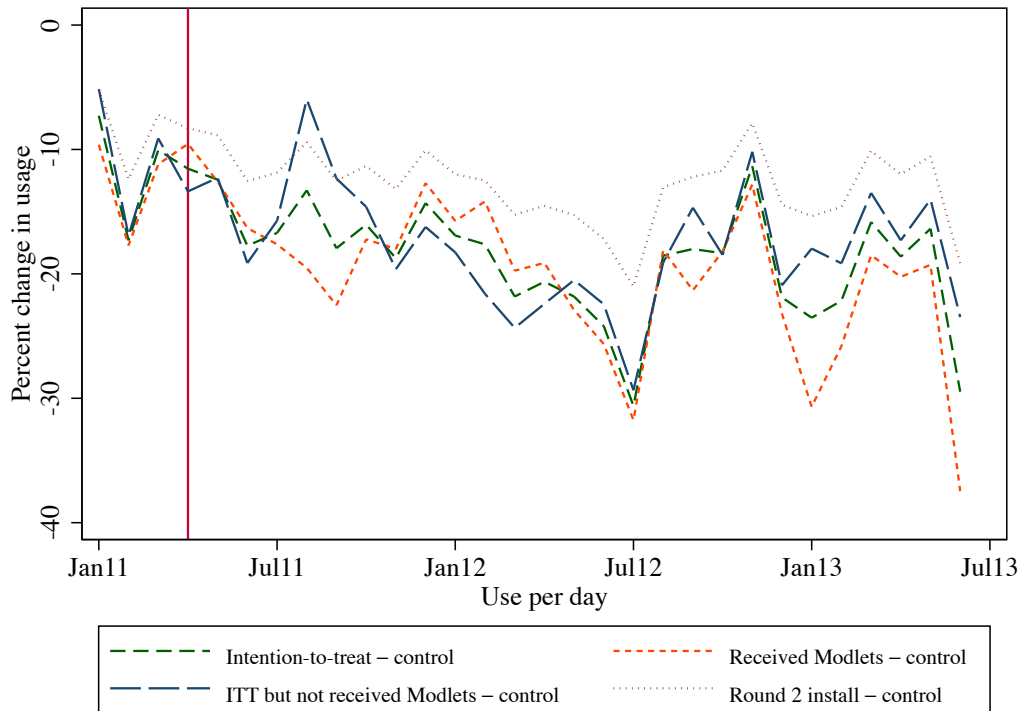


Figure 5: Pre-treatment and treatment periods electricity use by group

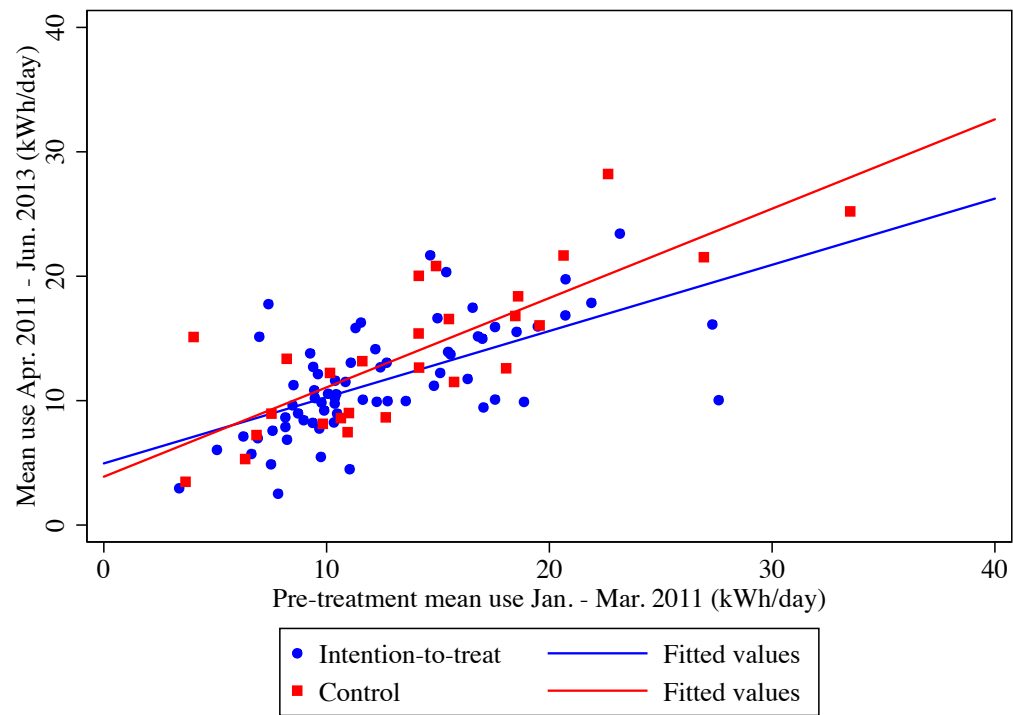


Table 1: Characteristics of population and analysis sample

Sample	Low income	Middle income	Market rate	Total
Apartments in Tapestry building	37	55	92	184
Apartments with consent form received prior to randomization	22	28	51	101
Apartments with some ConEd data	22	30	50	102
Apartments in analysis sample (consent forms in time and some ConEd data)	18	26	44	88
Apartments in analysis sample with some pre-survey data	17	21	42	80

Table 2: Characteristics of control and treatment groups for analysis sample

Treatment	Control group	Intention-to-treat	Received Modlets	Received Modlets and round 2 upgrade
Description of treatment	Randomly selected to not receive Modlets	Randomly selected to receive Modlets	Apartments receiving Modlets	Apartments receiving Modlets and 2012 upgrade
Number of low income apartments	4	14	5	2
Number of middle income apartments	7	19	13	8
Number of market rate apartments	13	31	14	5
Total number of apartments	24	64	32	15
Mean number of Modlet plugs per apartment	0 (0)	7.4 (8.7)	14.8 (6.3)	15.6 (6.5)
Mean NEP score (pro-environmental attitudes)	4.9 (0.5)	4.8 (0.7)	4.9 (0.7)	4.8 (0.6)
Mean numeracy score	2.0 (0.9)	2.1 (0.8)	2.1 (0.8)	2.1 (0.9)
Mean willingness-to-pay for an energy monitoring system (\$)	136 (142)	76 (119)	75 (75)	80 (69)

Note: standard deviations in parentheses.

Table 3: Characteristics of analysis sample by apartment income

Apartment type	Low income	Middle income	Market
Mean electricity use (kWh/day)	12.4 (6.3)	13.9 (6.9)	13.7 (7.1)
Mean area of apartment (square feet)	748 (149)	762 (230)	807 (184)
Mean NEP score	4.6 (0.7)	4.9 (0.6)	4.8 (0.6)
Mean numeracy score	1.3 (0.8)	2.0 (0.8)	2.5 (0.7)
Mean willingness-to-pay for an energy monitoring system (\$)	35 (45)	106 (84)	104 (139)
For apartments known to have received Modlets:			
Mean number of Modlet plugs	15.2 (6.5)	17.4 (5.6)	13.5 (6.4)
Mean overall electricity use per Modlet plug (kWh/day)	0.09 (0.02)	0.08 (0.02)	0.10 (0.07)
Mean electricity use per Modlet plug during morning hours 6-9AM (kWh/day)	0.07 (0.02)	0.07 (0.02)	0.09 (0.06)
Mean electricity use per Modlet plug during daytime hours 10AM-4PM (kWh/day)	0.09 (0.03)	0.08 (0.03)	0.10 (0.07)
Mean electricity use per Modlet plug during evening hours 5PM-10PM (kWh/day)	0.13 (0.04)	0.10 (0.03)	0.13 (0.08)
Mean electricity use per Modlet plug during nighttime hours 11PM-5AM (kWh/day)	0.08 (0.01)	0.07 (0.02)	0.10 (0.06)

Table 4: Characteristics of Modlet data

Modlet function	Number of Modlet plugs	Number of 15 minute observations	Mean electricity use (kWh/day)
Empty	33	188,029	0.001 (0.016)
Lights and lamps	137	648,766	0.03 (0.13)
Entertainment devices	159	774,720	0.17 (0.31)
Home office equipment	156	719,832	0.07 (0.20)
Chargers	91	370,934	0.02 (0.12)
Routers, modems, and cable boxes	96	399,977	0.23 (0.31)
Clocks	19	76,477	0.03 (0.06)
Air purifiers, fans, and heaters	21	106,597	0.03 (0.08)
Refrigerators	65	376,589	0.27 (0.32)
Other devices	82	493,034	0.05 (0.14)
All functions	671	3,340,831	0.10 (0.23)

Note: each Modlet generally includes two Modlet plugs. Data are recorded separately at the Modlet plug level. Some Modlet plugs used for multiple functions. Standard deviations in parentheses.

Table 5: Electricity use by treatment and control groups in analysis sample

Treatment assignment	Control group (no intention to treat)	Intention-to-treat treatment	Received Modlets treatment	Intention-to-treat, but did not receive Modlets	Received Modlets and round 2 upgrade treatment
Electricity use	15.5 (8.2)	12.5 (6.0)	12.6 (5.9)	12.4 (6.0)	12.8 (6.2)
Pre-treatment electricity use, Jan.–Mar. 2011	15.3 (8.5)	13.3 (7.1)	13.0 (6.9)	13.6 (7.3)	12.8 (5.9)
Treatment period electricity use, Apr. 2011–	15.6 (8.1)	12.2 (5.4)	12.5 (5.5)	12.0 (5.4)	12.7 (6.3)
Electricity use before round 2 upgrade, Jan. 2011–Mar. 2012	--	--	--	--	13.4 (6.9)
Electricity use after round 2 upgrade, Apr. 2012–	--	--	--	--	13.2 (6.0)
Mean daily electricity use before first Modlet data	--	--	13.1 (7.0)	--	--
Mean daily electricity use after first Modlet data	--	--	12.3 (5.1)	--	--

Note: “before first Modlet data” is defined by the last day of the electricity bill being before the first Modlet data for that apartment. Standard deviations in parentheses. Electricity use is measured in mean kWh/day.

Table 6: Treatment effects

Treatment measured by:	Effect of treatment on electricity use	Pre-treatment effect on treated	Interactions with pre-treatment survey	N	R ²
Intention-to-treat: Randomly selected to receive Modlets and Apr. 2011 and after	-3.5** (1.4)	--	--	2,195	0.20
Intention-to-treat: with regressor for pre-treatment effect on treated	-3.5** (1.4)	-1.6 (1.6)	--	2,195	0.20
Received Modlets treatment: Apartment received Modlets and after first Modlet observation	-3.3** (1.4)	--	--	1,417	0.21
Received Modlets and round 2 upgrade treatment	-3.3 (2.2)		--	777	0.20
Intention-to-treat: low income apartments only	-0.95 (2.1)	--	--	592	0.24
Intention-to-treat: medium income apartments only	-5.8* (3.1)	--	--	612	0.27
Intention-to-treat: market apartments only	-3.0 (1.9)	--	--	991	0.20
Intention-to-treat with interaction for NEP score and treatment	-11.8* (5.9)	--	1.8 (1.2)	2,008	0.20
Intention-to-treat with interaction for numeracy score and treatment	-5.7** (2.5)	--	1.2 (0.9)	2,008	0.21
Intention-to-treat with interaction for logged monitoring system willingness-to-pay	-0.16 (1.4)		-0.78** (0.38)	2,008	0.26

Note: each row corresponds to one regression. For the received Modlet specifications, the control group includes only apartments which were randomized to have no intention to treat. All regressions also include month indicators; the last three also include the pre-survey measure. Standard errors are clustered at the apartment level. * and ** denote significance at the 10% and 5% levels respectively. Electricity use is measured in kWh/day.

Table 7: Mean electricity use at Modlet plug level

Effect of Modlet plug by function on electricity use (kWh/day):	Day of install	1 day after	2 days after	3-4 days after	5-7 days after	8-14 days after	15-30 days after
Empty	−0.007 (0.004)	−0.008* (0.005)	−0.009* (0.005)	−0.008* (0.005)	−0.007* (0.004)	−0.004** (0.002)	−0.0003 (0.0018)
Lights and lamps	0.03*** (0.015)	−0.005 (0.006)	−0.006 (0.01)	−0.007 (0.006)	0.002 (0.005)	0.0006 (0.004)	−0.003 (0.003)
Entertainment devices	0.004 (0.025)	0.008 (0.018)	−0.001 (0.019)	0.004 (0.017)	0.005 (0.017)	0.006 (0.015)	−0.008 (0.008)
Home office equipment	0.02 (0.015)	−0.010 (0.010)	−0.002 (0.01)	−0.016* (0.009)	−0.01 (0.01)	−0.003 (0.009)	−0.004 (0.005)
Chargers	0.014 (0.013)	−0.009 (0.008)	−0.008 (0.011)	−0.03** (0.011)	−0.006 (0.005)	−0.003 (0.005)	−0.004 (0.003)
Routers, modems, and cable boxes	0.035 (0.031)	0.013 (0.024)	−0.01 (0.026)	0.016 (0.024)	0.014 (0.026)	0.025 (0.022)	−0.006 (0.012)
Clocks	0.007 (0.02)	0.003 (0.016)	0.003 (0.019)	−0.01 (0.01)	−0.01 (0.01)	−0.005 (0.009)	0.004 (0.007)
Air purifiers, fans, and heaters	0.11 (0.07)	0.04 (0.03)	0.025 (0.03)	0.02 (0.027)	0.016 (0.02)	0.000 (0.013)	0.003 (0.01)
Refrigerators	0.04* (0.025)	0.026 (0.022)	0.019 (0.02)	0.015 (0.022)	0.016 (0.02)	0.02 (0.02)	0.02 (0.01)
Other devices	0.018* (0.010)	0.008 (0.008)	0.010 (0.008)	0.006 (0.007)	0.003 (0.006)	0.002 (0.006)	0.0003 (0.004)
All functions	0.020** (0.008)	0.004 (0.006)	0.002 (0.006)	−0.0008 (0.006)	0.001 (0.005)	0.002 (0.005)	−0.002 (0.003)

Note: each row corresponds to one regression. Different rows include different samples of outlets based on reported Modlet function. Omitted category is >30 days. All regressions also include date and hour-of-day indicators. Standard errors are clustered at the apartment level. *, ** and *** denote significance at the 10%, 5% and 1% levels respectively. Electricity use is measured in kWh/day.