THE UNIVERSITY OF MELBOURNE

Department of Economics

Working Paper Series

Tell Me Something I Don't Already Know: Informedness and External Validity in Information Programs

David P. Byrne, Andrea La Nauze, Leslie A.Martin

April 2014

Research Paper Number 1180

ISSN: 0819 2642
ISBN: 978 0 7340 4530 0

Tell Me Something I Don't Already Know: Informedness and External Validity in Information Programs

David P. Byrne* Andrea La Nauze* Leslie A. Martin*

April 29, 2014

Abstract

Information programs that leverage peer comparisons are used to encourage pro-social behavior in many contexts. We document how imperfect information generates heterogenous responses to treatments involving personalized feedback and peer comparisons. In our field experiment in retail electricity, we find that most households either overestimate or underestimate their relative energy consumption pre-treatment. Households that overestimated respond to new information by temporarily increasing electricity consumption, whereas households that underestimated take steps that lead to long term energy conservation. We explore the implications of these results for the external validity and design of information programs.

JEL Codes: C93, D12, D84, L94, Q41

---

*Department of Economics, University of Melbourne, 111 Barry Street, Carlton, Victoria 3053, Australia. Our email addresses are: `byrned@unimelb.edu.au`, `ala1@unimelb.edu.au`, and `leslie.martin@unimelb.edu.au`. We thank Billcap's Yann Burden and Lara Olsen for making the field experiment possible. We have received helpful comments from Koichiro Ito, David Rapson, and Tom Wilkening, as well as workshop participants at the 2013 U. Melbourne IO Workshop, 2013 Econometric Society Australasian Meetings, 2013 Behavioral Energy and Climate Conference, and the 2014 POWER conference. Funding from the University of Melbourne FBE Faculty Research Grant scheme and the Institute for a Broadband Enabled Society is appreciated. All errors, omissions, and views are our own.

1 Introduction

Information programs have long been used to encourage social and behavioral change. A range of development interventions affect behavior by providing information on the benefits of drinking water treatment (Kremer et. al 2010), hand washing (Luby et. al 2005), or on the costs of risky behavior, as with HIV risk prevalence (Dupas 2011). In the developed world, information programs have been shown, for example, to encourage retirement savings plan participation (Duflo and Saez 2003) and reduce individuals' levels of caloric intake (Bollinger, Leslie, and Sorensen 2011).

Interventions that further inform individuals about how their behavior compares with that of their peers have been shown to enhance voter turnout (Gerber and Rogers, 2009), rates of charitable giving (Frey and Meier 2004), retirement savings (Beshears et al. 2013) and water and energy conservation (Ferraro and Price 2013, Allcott 2011b). These non-pecuniary, nudge-based approaches to encouraging pro-social behavior are potentially a low-cost alternative to politically-sensitive regulation. As a result, they have garnered much attention from policymakers in these and other contexts.

A key issue confronting policymakers seeking to scale up these programs is the external validity of existing estimates of a program's impact. Do estimates from pilots or locally-run information treatments apply to larger and more diverse populations? Can population characteristics explain why similar programs yield different results? If so, what characteristics matter most? The extent to which information is new or surprising to individuals is likely to be of first order importance. At the extreme, providing information to perfectly informed individuals should have little effect, other than via increased salience.

Becker (1965) and Levitt and List (2007) provide frameworks that explain how information frictions can cause peer comparisons to produce heterogeneous treatment effects. In particular, individuals who heavily engage in pro-social behavior could decrease their levels of engagement when they learn that their peers are not matching their efforts. Such "boomerang effects" (Schultz et. al 2007) imply that individuals that overestimate their relative levels of pro-social behavior will potentially *decrease* their engagement if treated with an information program with peer comparisons.[1] In contrast, these models also predict that households that are minimally engaged in pro-social behavior will increase their participation when they learn that they fall short of the social norm.

Although the literature on information treatments is extensive, little is known about how unobserved, pre-treatment informedness actually shapes information program outcomes. There is virtually no evidence on how informedness affects the external validity of program effects, nor how it should be accounted for in program design. The latter issue is particularly relevant for policymakers looking to enhance pro-social behavior and avoid boomerang effects.

This paper sheds light on these issues using a field experiment that involves a novel two-step experimental design. We first run a baseline survey that elicits individuals' beliefs over their relative levels of pro-social behavior. We then randomly inform individuals about how their behavior actually compares with that of their peers, and track their levels of pro-social behavior over time. To our knowledge, we provide the first test of whether imperfect information over peers' behavior in fact generates heterogeneous effects in information programs. We further study a key implication from our test results, namely how the distribution of pre-treatment informedness affects the external validity of program effects.[2]

The context for our study is the retail electricity market and the pro-social behavior of interest is energy conservation. We are interested in information programs in this setting for two reasons. First, behavioral interventions that inform households about energy conservation strategies and provide peer comparisons over energy usage have been used extensively in

[1]Boomerang effects have been of concern in a number of interventions, including those that aim to promote energy conservation (Schultz et. al 2007 and Allcott 2011b), water conservation (Ferraro and Price 2013), and public goods contributions (Croson and Shang 2008, Chen et. al 2010).

[2]As such, the usefulness of our baseline survey for conducting such a test and studying related implications for external validity speaks to prescriptions from Duflo, Glennerster, and Kremer (2008) in development, who advocate using baseline surveys for identifying sources of treatment heterogeneity and evaluating external validity in social programs.

electricity markets. As Delmas et al. (2013) document, more than 150 informational field trials spanning more than 500,000 households have been run since 1975.

Second, the retail electricity market is particularly well-suited to studying informedness and its impact on the external validity and design of information programs. Treatment effect estimates from information programs in electricity markets are highly heterogeneous (Delmas et al. 2013). This heterogeneity provides *a priori* motivation to study external validity and the channels that generate heterogeneous treatment effects in our context. In addition, boomerang effects have been documented in electricity market information programs and are a central concern in their design (Schultz et. al 2007 and Allcott 2011b)[3]. Further, previous studies have found households exhibit a high degree of ignorance regarding energy consumption.[4] Finally, large volumes of pre- and post-treatment behavioral data can be collected at low cost.

Our paper delivers three main results. We first show that most households are generally uninformed about relative energy use: only one quarter of our sample correctly identify their relative consumption levels. Contrary to our expectations, households do not systematically underestimate their energy use. Rather, we find many think they are "average" users, and that incorrect high and low energy use households symmetrically overestimate and underestimate their positions in the energy use distribution.

We then test whether pre-treatment informedness generates heterogeneous treatment effects from our information program. The program is similar to previous information treat-

[3]Organizations like OPower also recognize the importance that boomerang effects play in information interventions in promoting energy conservation. The company attempts to mitigate their effect by simultaneously providing "injunctive norms", that is, reminders to all users that energy conservation is pro-social, and rewarding low users with smiley faces or other forms of praise. They also provide multiple reference points, for example, individual-level rankings within neighbors, so that almost everyone receives a message that, to some extent, says "you can still do better".

[4]See, for example, Attari et al. (2010), Allcott (2011a), or Allcott (2012).

ments (for instance, Allcott 2011b) in that it provides households with energy saving tips and peer comparisons over relative energy use. Our pooled estimates are directly in-line with predictions from Becker (1965) and Levitt and List (2007): in our preferred specification, households that overestimate their relative consumption levels increase their use by four percent, indicating the presence of boomerang effects.[5] In contrast, households that underestimate their use decrease by two percent. We observe no change in energy use in response to treatment for households who were correct. The heterogeneous treatment effect results survive a number of robustness checks, including allowing "average user" responses to count as "I don't know", controls for potential mean reversion, and sample attrition bias corrections.

Expanding on these results, we study the evolution of these heterogeneous treatment effects over time. We uncover a new asymmetry in the persistence of these effects. Boomerang effects appear to be short-lived, perhaps salience-driven, phenomena: households that overestimate their relative energy use immediately increase their use in response to treatment, however exhibit negligible treatment effects from five months following treatment onward. In contrast, households that ex-ante underestimate their energy use see a steady decline in energy consumption through the entire treatment window. After learning they are in fact high energy users, these households appear to make permanent changes in behavior, possibly because of social learning through the information treatment's peer comparisons (e.g., Becker 1965), and/or in response to previously unfelt "moral costs" from the environmental impact of their high energy use (e.g., Levitt and List 2007).

Beyond their academic interest, these long-run findings also have policy relevance. They

[5]While our private partner was willing to elicit households' beliefs over relative energy use, they were primarily interested in program evaluation and only allow for one type of information treatment. As a result, we were restricted in our ability to discriminate between the Becker (1965) social learning and Levitt and List moral cost mechanisms through our experiment. See Ferraro and Price (2013) for a design that discriminates among these channels. They cannot, however, test for the effect of imperfect information over peer behavior in behavioral programs, which is fundamental to both of these mechanisms.

highlight the importance of looking beyond a 3 month window when evaluating energy conservation programs: rebound effects disappear and conservation effects gain in importance.

To complete the analysis, we quantify the extent to which sub-population informedness affects the external validity of information program treatment effects. Using Monte-Carlo simulations, we show pre-treatment informedness plays an important role quantitatively in determining treatment effects from information programs. Our results naturally raise questions about what type of households are likely to underestimate relative energy use. We show that higher income households that are more environmentally conscious are more likely to underestimate their relative energy use, and hence be more responsive to treatment.

Related literature

Our main contribution is to the cited body of research on information treatments that encourage pro-social behavior. Of all the treatments that we have reviewed, we have only found three that collect baseline information on informedness. Fitzsimons et al. (2012) and Davis et al. (2011) in the nutrition and sanitation literatures record knowledge both before and after information interventions. Both find their intervention caused a change in knowledge, however neither explore heterogeneous program effects as a function of baseline knowledge.

The closest example to our experimental design that we could find comes from a health intervention by Jalan and Somanathan (2008). The authors undertake a baseline survey to assess awareness of potential causes of diarrhoea. They find some evidence that pre-treatment behavior is correlated with awareness. They do not find that these households are more likely to respond to treatment. Our paper differs considerably from this study given our focus on: (1) behavioral programs that leverage peer comparisons; (2) testing whether information frictions generate heterogeneous treatment effects in such programs; and (3) quantifying the impact imperfect information has on the external validity of treatment effects.

We also build on prior studies of information interventions in electricity markets.[6] Our study and its focus on the determinants of heterogeneous treatment effects most closely re-

[6]See Delmas et al. (2013) for an excellent review of this extensive literature.

late to Allcott (2011b) and Costa and Kahn (2013). These authors respectively provide evidence of how differences in energy consumption (e.g., high vs. low users) and political ideology (e.g., Liberals vs. Conservatives) predict heterogeneous effects from information interventions. Neither are able to separately identify the role of imperfect information from other factors (e.g., environmental consciousness). Our paper thus makes a contribution by providing a direct test of how imperfect information leads to heterogeneous treatment effects in electricity market information programs.

Our study also relates to research on limited consumer awareness over energy consumption. This includes studies that find households are unaware of the relative energy use of different household appliances (Attari et. al 2010) and fuel economy across similar vehicles (Allcott 2011a and 2012). Related work by Jessoe and Rapson (2013) and Ito (2013) respectively imply that households are imperfectly informed about real-time energy use (which reduces price sensitivity), and that they respond more to average than marginal electricity prices. We provide the first empirical evidence on the degree and nature of household informedness over energy consumption relative to peers. As such, we complement previous work by identifying another information friction that affects household energy consumption behavior.

2 Experimental context, design, and mechanisms

The context for our study is the retail electricity market in the Australian state of Victoria. Approximately 70% of Victorians reside in Melbourne, a city with four million inhabitants. Although the climate is similar to that of San Francisco, Melbourne is known for having "four seasons in a day" with both warm and cold air masses from the Outback and the Antarctic affecting local weather. Retail residential electricity prices in the state are on average 25 cents US per kWh, which are high compared to the US (15 cents per kWh), but low compared to Europe (30 cents per kWh) (Mountain 2012).

The state has retail market competition, where electricity retailers compete for customers through pricing plans and marketing. Competition among retailers for customers is fierce: on average, one in five consumers switch between electricity retailers in a given year. Upstream electricity distributors, in contrast, are regulated geographic monopolists who pass on regulated network charges. In total, there are five distributors and sixteen retailers in the market within our sample period.

In June 2010, the state government mandated a statewide rollout of "smart" electricity meters, a policy that was expected to have large market impact. Unlike traditional electricity metering which provides energy use information at monthly (or even lower) frequencies, these new meters collect high frequency, half-hourly reads of household-level energy use. The mandate required all distributors to install smart meters for every residential and small business electricity customer in the state by December 2013. Retailers have no influence over where or when smart meters are provided to their customers. Our experiment coincides with this introduction of smart meters into Victorian homes.

2.1 Experiment design

We developed and implemented our information program with an electricity retailer between August 2012 and June 2013. The retailer is a relatively small player in the market with less than 30,000 customers nationwide. We also collaborated with a web-portal developer, Billcap (www.billcap.com). The retailer employed Billcap to provide its customers with online home energy reports that help customers make better-informed energy consumption decisions by visualizing a given household's smart meter data, providing energy saving tips, and presenting peer comparisons over energy use. Figure 1 depicts these various aspects of the home energy reports.

We used a phase-in design to randomly provide households access to home energy reports. Households were randomized into one of three waves of treatment: October 2012 (wave one), March 2013 (wave two) and June 2013 (wave three). Households in waves two and three act as a control for the wave one treatment; households in wave three are the control group for the wave two treatment.[7] Access to the reports was provided to treatment households through

[7]The third wave was a company-wide web-portal launch, implying that we do not have a control group after June 2013.

bi-weekly emails that encouraged households to freely access their personalized reports. Prior to the experiment, the retailer did not provide additional information to households based on their smart meter data; customers were only shown data aggregated at the level of their regular billing cycle (i.e. with monthly, bi-monthly or quarterly frequency). Our information treatment was thus expected to generate large changes in households' information sets regarding energy use.[8]

Beliefs survey

Before being offered or given access to home energy reports, we elicited households' beliefs over their relative energy use.[9] In particular, we asked households the following question:

> *Compared to energy use in Melbourne homes as large as yours, what statement best describes your household's monthly energy use?*
>
> a. *High (top 20%)*
>
> b. *Above average (top 40%)*
>
> c. *Average*
>
> d. *Below average (bottom 40%)*
>
> e. *Low (bottom 20%)*

We allocated households to one of five equal-sized quantiles of the energy use distribution, associating each quintile with an intuitive word-based descriptor. We intentionally included a middle, neutral response so that we could explicitly distinguish households that actively commit to classifying themselves as above or below average. We made the size comparison vague,

[8]Our retailer partner was also the first in the country to use Billcap's platform; households in our sample were unaware of Billcap and its services prior to August 2012. Over the course of our sample period, however, other retailers and distributors did launch similar services.

[9]For reference, we visually represent our phase-in experimental design and the timing of the beliefs surveys in Figure A.1 of the Appendix.

"as large as", to simplify the question, and then test two alternate interpretations: same number of bedrooms, and same number of occupants, two questions asked earlier in the survey.[10] Comparing a household belief to their actual use pre-treatment energy use quintile (conditional on home size) enables us to measure that households pre-treatment level of informedness.

2.2 Mechanisms for treatment effects

Why should the home energy reports have an effect on households' energy use? How should the effects vary depending on whether households over- or underestimate their relative energy use? Potential channels for information program effects have been discussed at length.[11] Here we summarize the relevant mechanisms, focusing on those that affect the information program treatment effects.

The reports' forecasting and budgeting tools and the energy savings tips can reduce energy use through at least three channels. The first channel appeals to the Becker (1965) household services model. In this model, households determine their utility-maximizing expenditure on energy and a composite good which, among other things, is affected by the utility derived from energy use (e.g., air conditioning). Home energy reports could affect consumption if households are not perfectly informed regarding their household production functions.[12] Providing

[10]Survey responses were collected and anonymized using the Billcap platform with email-based invitations. Respondents were also asked to provide information on household characteristics such as number of rooms, number of residents, and whether the house has gas appliances, is a detached dwelling or rented/owned. We sent out the survey to the retailer's existing customers prior to Billcap's October 2012 launch. We also surveyed all new customers with a smart meter who signed up with the retailer within our experimental window. These surveys were run shortly after customers signed up but before treatment.

[11]See, for example, Ferraro and Price (2013), Allcott (2011b), Levitt and List (2007).

[12]For example, a household may underestimate the electricity use of a clothes dryer or overestimate the use of a laptop or flat-screen TV. Attari et. al 2010 provide a number of examples.

households with the information necessary to undertake energy-saving activities and investments could therefore reduce utility-maximizing energy use.

Second, households may perceive home energy reports as a form of scrutiny over their existing energy use decisions, thereby creating a previously unfelt “moral cost” (Levitt and List, 2007) associated with the societal impact of their energy use, such as its environmental impact. Households may therefore reduce their energy use in response to the moral costs created by the treatment.

Third, the information treatment may affect households via a salience effect. Frequent email reminders may bring energy use to the forefront and encourage households to act with respect to their preferences.

The Becker (1965) model also predicts heterogeneous treatment effects from information programs that include peer comparisons. If households are uncertain about the nature of their household production functions, then peer comparisons can provide useful information. For example, suppose that a household thinks they are “average” (e.g., median) energy users and find out that they are actually use more energy than most of their peers. The family members may be unhappy to learn that they have a significantly higher electricity bills than their neighbors. They may also infer from this peer comparison that some of their practices or appliances are more electricity intensive than they believed. Conversely, a family that takes strict steps to limit electricity use could may be relieved to find out that most households use or pay more for electricity than they do. Learning about how their use compares to others may lead them to suspect that family members would be happier living under a less strict energy use regime. In either case, household responses could indicate a form of social learning, specificially, learning from peers about ones’ utility-maximizing electricity consumption level.

A Levitt and List (2007) moral cost story could also generate these patterns. If a household’s experienced moral cost depends on its beliefs over how its electricity consumption or associated environmental damage compares to that of peers, then high users who unknowingly underestimate their relative use will realize higher moral costs once informed about true their relative use. New information would also allow low users to relax upon learning that they had previously been over-estimating their relative consumption, giving them moral license (Harding and Rapson 2013) to consume electricity more freely.

Notice that both channels for heterogeneous effects rely on households exhibiting uninformedness/unawareness over how their energy use compares with that of their peers. Our beliefs survey, in combination with our information treatment, allows us to test of whether such information frictions indeed cause heterogeneous treatment effects. The treatment that we negotiated with our private partners does not, however, allow us to discriminate between the social learning and moral cost channels. We instead test empirically for the information frictions required for either of these channels to play a role.

3 Data

Within our sample window, the retailer had a total of 15,454 customers in the state with a smart meter. We drop all commercial customers, residential customers with solar panels, and large users who consume over 50 kWh/day.[13] We restrict our sample to customers with a smart meter who joined the electricity retailer up to and including March 2013 and track energy use up to June 2013. After imposing these sample restrictions our working sample has 8578 customers, with 765, 4548, and 3265 in treatment waves one, two, and three. Within this sample, 1188 households responded to the baseline survey on energy use beliefs.

The retailer provided us with half hourly smart meter interval data for all customers. We focus, however, on daily energy use so we aggregate the data to daily meter reads.[14] In addition, we were given the complete billing history for all customers. This data allow us to identify the exact date a household started with and left the retailer, which is important for our empirical specifications that deal with attrition below.

[13]We also removed any daily observation with less than 4 kWh/day consumption to avoid confounding from households who leave on holiday.

[14]On average, we have interval data for 190 days (S.D.=89) for a customer within the sample period. The full sample window is 337 days. Customers whose smart meters are exogenously switched on by upstream distributors within the sample window are eligible for treatment and are included in our sample.

We also collected demographic and elections data. Demographic information is collected at the lowest level of aggregation available, with 400 households living within a “Statistical Area Level 1” (SA1) census tract.[15] We further matched census tract centroids to nearest polling booth. Based on this matching we use the local share of votes for the federal Green party as a proxy for regional levels of environmental consciousness.[16]

3.1 Randomization checks

Tables 1 and 2 report descriptive statistics on pre-treatment energy use and household characteristics for our treatment and control groups. We observe balance and similar trends in energy use across the waves leading up to the October 2012 and March 2013 treatments. The households in each of the three waves also exhibit similar observable characteristics. Moreover, the summary statistics for our survey data in Table 2 largely align with figures from Tustin (2012), who reports that 70% of households in the country have air conditioning, and on average households have 2.6 residents and three bedrooms. Table A.1 in the Appendix further shows that survey respondents and non-respondents tend to have similar pre-treatment energy use and demographic characteristics. On the whole, this preliminary look at the data provides reassurance that randomization was successful, and that survey respondents are representative on observables.

There is some attrition in the treatment waves: on average, 3.7%, 6.5%, and 1.5% of households in waves one, two, and three drop out of the sample in a given month.[17] Such attrition

[15]The retailer used ArcGIS to match households to census tracts based on household street addresses (which were not made available to us).

[16]The Green party won 11.76% and 8.65% of all votes in the 2010 and 2013 general elections. This approach to using local party vote shares to reveal preferences pro-environmental programs/policies follows Costa and Kahn (2013).

[17]We formally test for differences in attrition rates between each of the waves in the Appendix. All of the differences are statistically significant. The particularly lower attrition rate in wave three partly reflects that a number of households in this wave sign-up with the retailer towards the latter part of our sample

behavior is an unavoidable econometric concern for our experiment because we study a competitive retail market.[18]

In Tables A.2 and A.3 of the Appendix we check whether attritors and non-attritors differ based on observable characteristics. We find some evidence that attritors tend to use more electricity, however no other significant differences arise between attritors and non-attritors within each treatment wave. All of our econometric specifications below include household fixed effects that account for any observed or unobserved time-invariant factors that generate differential attrition rates across our treatment waves. We also estimate models that allow us to quantify any attrition biases in our information program treatment effect estimates.[19]

Selection bias is another potential concern if certain types of customers non-randomly switch to our retailer within our sample window in response to the newly-offered Billcap home energy reports. However, given that our retailer was the first to use Billcap's services, and that Billcap was not promoted broadly, it is unlikely such non-random selection occurs.

---

window. As new customers, they are less likely to attrit as they become familiar with the company.

[18]This contrasts with information experiments from the U.S. in the context of regulated monopoly utilities where households have no choice over their electricity retailer. This aspect of our study is related to numerous studies from labor economics (for example, see Heckman and Smith 1999) that emphasize the need for non-experimental methods in evaluating social experiments such as labor market training programs, where individuals cannot be prevented from dropping out or pursuing alternative programs.

[19]In a complementary paper, Byrne et al. (2014), we focus specifically on empirically modelling customer churn. We use additional data on household interaction with the web interface to explore how some customers appreciate helpful reminders to conserve energy whereas others view those as spam, along with the implications for customer turnover and retailer incentives to provide personalized feedback. The development of these dynamic discrete choice models is well-beyond the scope of the current paper. Below we simply correct for any attrition biases in our estimates of the effect of our information program on energy use.

4 Findings

This section presents our experimental findings. We start by presenting results from our beliefs survey, highlighting the extent to which households over/under estimate their relative pre-treatment energy use. We then investigate how the treatment effects from our information program vary as a function of these prediction errors. We close by investigating a key implication from these treatment effect estimates, namely how the distribution of informedness over relative energy use affects the external validity of program effects.

4.1 Distribution of informedness over relative energy use

Our informedness measure compares households' beliefs to actual relative use levels. Specifically, we compare the quintile of the use distribution that households believe they are in to their actual pre-treatment average use quintile, conditional on the number of rooms in a given household's home. Table 3 presents contingency tables of beliefs and actual relative energy use in terms of counts (panel A) and percentages (panel B). If all households held correct pre-treatment beliefs, each cell along the diagonal would contain 20% of the observations and all other cells would be empty. We instead observe that only 25% of survey respondents are correct. Of those who are off by at least one quintile, about half overestimated their use (represented by the cells below the diagonal), whereas half underestimated their use (represented by the cells above the diagonal).

More than half of the households (57.5%) believe that their energy use is in the 40-60% quintile of the use distribution. That is, most people think they are "average" users. Of the many users (the remaining 42.5%) that take an active stand as to whether they are above or below average, 71% are wrong.

Figure 2 provides another view on the differences between households' beliefs and their actual relative energy use. For each household, we compute a prediction error as the number of quintiles its belief is above or below its actual quintiles in the energy use distribution, conditional on the number of bedrooms in a household's home. Panel A presents the error distribution for the entire sample; Panel B focuses on the sub-sample of households that believe they are above or below the middle quintile. The figures highlight the somewhat surprising result that the distribution of prediction errors is roughly symmetric. Similar percentages of respondents tend to underestimate (38%) and overestimate (37%) their relative energy use. These results contrast with our intuition prior to running the experiment that households would systematically underestimate their energy use (i.e. all think that they are better than average in conserving energy). Notice that the symmetry remains even after we drop respondents with 40-60% quintile/"average" pre-treatment beliefs.

Panels C and D present conditional distributions of prediction errors for different quintiles of the actual energy use distribution. Panel C presents distributions for households in the 20-40% and 60-80% quintiles of use. The peaks in the respective distributions are due to households systematically guessing they are average users; indeed 50% or more of the households within each group incorrectly guess they are in the average or 40-60% quintile. A similar pattern emerges in Panel D, where the the prediction errors are further accentuated.

4.2 Treatment effects from the information program

Graphical analysis

Figure 3 visualizes the distribution of average daily energy use across households and months, conditional on whether a given household in a given month was exposed to our information treatment.[20] We plot CDFs for these conditional distributions for four mutually exclusive subpopulations in the data: households who overestimate / underestimate / are correct about their relative energy use quintile (panels A, B, and C), and households who did not respond to the beliefs survey (panel D).

The leftward shift in the CDF in Panel B provides preliminary evidence that households who underestimate their relative energy use indeed reduce consumption once they are informed about their actual location in the energy use distribution. In contrast, we see mixed changes in

[20]The only transformation we make to these raw data is we de-mean average monthly household energy use by its sample average across all households for a given month. Doing so removes seasonal trends in energy use. The figures therefore depict distributions of the departures from average monthly energy use.

energy use at the bottom and top of the distribution in Panel A, similar to what we see in Panels C and D. This goes against predictions from Becker (1965) and Levitt and List (2007) that uninformed households who overestimate their relative energy use should increase consumption when provided with the new information contained in our treatment's peer comparisons (e.g. "boomerang effects" from Schultz et. al 2007).

Pooled treatment effects

To formally estimate the impact of providing home energy reports to households we use the following regression model :

$$Y_{it} = \beta_1 \times T_{it} + \mu_i + \eta_t + \epsilon_{it} \tag{1}$$

where Y_{it} is the logarithm of daily energy use (in kWh) for household i on date t, and T_{it} is a treatment dummy that equals one if household i was offered home energy reports on or before date t, and zero otherwise. The coefficient of interest, β_1, identifies the Intention-to-Treat Effect (ITT) from the information program. The μ_i and η_t terms correspond to household and date fixed effects which we include in all specifications. These control for time-invariant household characteristics that affect energy use (such has home size or environmental consciousness), as well as date-specific energy use shocks (such as particularly hot or cold days). For inference, we cluster our standard errors at the household level to account for within-household dependence in ϵ_{it}.

Our pooled treatment effect estimates are presented in Table 4. In the aggregate, we find no effect of home energy reports on energy use. As a check on whether this arises from studying noisy daily energy use data, we aggregate up our energy use data to the monthly level and re-estimate our model. The results, which are reported in column (2), yield similar results.[21]

Conservation effects may take time to emerge, depending on the rate at which households

[21] On the whole, these pooled results fall in line with a collection of previous well-designed information programs (as classified by Delmas et. al 2013) that collectively find negligible conservation effects.

learn about new strategies for conserving energy, or make investments in new energy-saving appliances. In contrast, if our information treatments only work through a short-lived salience effect, then we should expect an immediate response to treatment and subsequent "backsliding" to familiar energy use habits by individuals.[22]

We estimate time-varying ITTs by interacting months-since-treatment dummies with T_{it} in our baseline specification for equation (1). Figure 4 present the resulting time-varying treatment effect estimates and their corresponding 95% confidence intervals. The figures reveal that our information program delivers statistically significant and economically meaningful effects after seven months of treatment. At the seven month mark, energy use falls by 4.6 percent for households who were offered home energy reports.[23] It is interesting to note that despite the vastly different context of our study relative to the benchmark study of Allcott (2011b), we find an effect of similar magnitude. We also find a a similar pattern in the treatment effect over months since treatment.[24]

[22]See Allcott and Rogers (2012) for an extensive study into these long-run versus short-run effects from behavioral interventions in electricity markets.

[23]We have run a number of additional specifications that identify time-varying effects. These include models for daily energy use but with time-varying program effects by weeks since treatment, as well as models for weekly use with time-varying program effects by weeks since treatment. The results are very similar. We have also run auxiliary tests to check that our time-varying treatment effects are not simply due to cohort effects about households in the October 2013 and March 2013 treatment groups. These tests are available upon request.

[24]In particular, his ITTs steadily decline from 1% after seven months, falling to 3-4% after 20 months since treatment. The more rapid decline in our ITTs may reflect differences in the frequency and mode of delivery for the home energy reports in Allcott's and our study. In contrast to receiving monthly/weekly mail-based reports, the households in our sample can access their reports via the web at any time, which potentially hastens the rate at which households learn about ways to conserve energy.

Heterogeneous treatment effects among the informed and uninformed

Combining the data from the beliefs survey and information treatment, we now study whether the information program effects differ depending on whether households were ex-ante correct/incorrect about their relative energy use.[25] We use a similar difference-in-difference specification as above to estimate these heterogeneous treatment effects:

$$Y_{it} = \sum_{j=1}^{4} \beta_{3j} \times E_{ij} \times T_{it} + \mu_i + \eta_t + \epsilon_{it} \tag{2}$$

where E_{ij} is an indicator variable that equals one if treated household i belongs to group j, depending on whether the household underestimates ($j = 1$), is correct about ($j = 2$), or overestimates their relative energy use ($j = 3$), or whether they were treated but did not respond to the beliefs survey ($j = 4$).[26]

Columns 1 through 4 of Table 5 report our baseline heterogeneous treatment effects. They provide stark evidence that imperfect information over peers' behavior generates treatment effect heterogeneity, as predicted by Becker (1965) and Levitt and List (2007). Households who overestimate their relative energy use increase their consumption by seven percent once they are informed about their actual relative use. In contrast, households that underestimate their use decrease consumption by four to five percent. In other words, households move in the

[25]An assumption of note is that households' beliefs over relative use remain constant up until they date they are offered home energy reports. Under this assumption, we can compare specific treatment groups, say households treated in wave one who have just been told they overestimate their use, to control groups in waves two and three who later learn they have also being overestimate their use. Given the relatively short eight-month time horizon of our study, as well as the limited availability of other comparator websites for energy use available at the time, we believe this assumption is reasonable.

[26]In total, 864 households who were treated in the experiment replied to the survey. Comparing respondents use to households of similar size, we have 320 who overestimated, 210 who were correct, and 334 who underestimated.

direction of becoming they type of users that they thought they were. As expected, informed households (e.g., those who were correct about their relative use) do not exhibit any change in their consumption behavior.

The stability of the estimates across columns 1 through 4 imply these results are robust to how we construct peer groups, that is, whether we condition on the number of rooms, residents, or if we do not condition on household characteristics.[27]

The results are also robust to how we interpret "average user" responses, which are potential "don't know" or guess responses. In column 5 we present estimates dropping households who report they were "average" users.[28] We still find the boomerang effect among over-estimators. The point estimate for under-estimators is similar to our base results but no longer statistically different from zero.

Mean reversion is a potentially important confound for interpreting the heterogeneous treatment effects. The raw data shows that low users tend to increase their consumption over the sample period whereas high users tend to decrease use. Almost by definition, over-estimators are disproportionately lower users and under-estimators disproportionately higher users, so the estimates presented in columns 1 through 4 of Table 5 could embody both new information and differential trends.

We address potential mean reversion in three ways. First of all, we augment our models to include five sets of date fixed effects, one for each of the five quantiles in pre-treatment average energy use. This approach flexibly controls for differential trends among high and low energy users. The results, which are presented in column 6, yield similar conclusions as our

[27]We have also estimated specifications that allow for the intensity of the error, for example whether it matters if a household under/over estimated by one or two quantiles. We find no effect for households who underestimate their use by one or four quintiles. Those who underestimate their use by two or three quintiles reduce consumption by 7%. All households who overestimate their use increase their consumption.

[28]This regression also controls for mean reversion with use quintiles.

baseline results, with more modest point estimates. The magnitude of the increase (decrease) in energy use among households who overestimate (underestimate) their relative energy use falls however to five percent (three percent).

In an alternative specification, we include separate date fixed effects for each group of over-estimators, under-estimators, correct, and no-answer households. This approach is similar to allowing for separate time trends for each type of household, with no restriction on the form those trends may take. In this specification, however, the standard errors increase because we only have survey responses for a subset of households, whereas we know pre-treatment use quantiles for everyone. The point estimate for the boomerang effect is unchanged; the energy savings from under-estimators is smaller and no longer statistically different from zero.

Finally, we also show the estimates for each sub-group by average pre-treatment energy use quintile. Ours results are reported in Table 6. All of the coefficients shown come from a single regression. Although the point estimates are not statistically different from each other, we see that, of all the high users, it is only those who underestimated their use that respond to the information treatment in a statistically-significant way. Similarly, of all the low users, it is only those that thought that they were high users that respond to information by increasing use.

These regression results are also interesting in that we see that the treatment effect for those who underestimate is largely driven by those in the highest use quintile. Similarly, the effect on those who overestimate is driven by those whose pre-treatment actual use is in the smallest quintile. These findings are consistent with these groups having the most to learn from their peers in an environment with social learning, or realizing the highest moral costs or moral license in an environment with social norms.

The remaining robustness checks in columns 8 through 10 of Table 5 show that our results are largely robust to using data only for survey respondents, aggregating to monthly energy use, and correcting for attrition bias. The column 8 results, which involve a substantial reduction in sample size, highlight how the identification and accuracy of the underestimating households depends on the inclusion of the non-survey respondents in the control group. To reflect the timing of decisions to change retailers, we estimate treatment effects accounting for attrition at the monthly level. For the attrition correction we use a parametric inverse probability weight estimator, where we allow for attrition at date t to depend on a household's pricing plan, the number of a months it has been with the retailer, and lagged energy consumption.

We then estimate time-varying heterogeneous treatment effects among households that overestimate/are correct about/underestimate their relative use. In constructing our peer groups, we continue to condition on the number of rooms in a home. We again present treatment effect estimates at monthly frequencies. The results, which are presented in Figure 4, reveal key insights for the literature on information treatments. Panel A shows that households that overestimated their energy use (e.g., low users who thought they were high users) exhibit a temporary boomerang effect in the first four months of treatment that subsequently dissipates over time. In contrast, Panel B shows that households that underestimated (e.g., high users that thought they were low users) reduce their electricity consumption gradually and consistently over time. Panels C and D show that informed households and non-survey respondents do not exhibit significant treatment effects at any point within our sample window.

Taken together, these figures are encouraging for policymakers looking to promote energy conservation through peer comparisons. In our sample boomerang effects that harm energy conservation efforts largely appear to be shortly-lived salience effects. Underestimating households, on the other hand, appear to permanently change their behavior in response to the new information, potentially by learning from their peers' behavior and/or responding to previously unfelt moral costs from using relatively large amounts of energy. As the latter effect starts to dominate, we should expect energy conservation effects to emerge over time in the population as we found in Figure 4, and as others have documented (Allcott 2011b and Allcott and Rogers 2012).[29]

[29]These patterns in treatment effects for households that over/underestimate energy use are robust to correcting for attrition, though the estimates are somewhat noisier under these corrections.

4.3 Informedness and external validity

To conclude our analysis, we use our novel treatment effect estimates to highlight the pivotal role that population (un)informedness plays in determining information program effects. This issue of external validity is perhaps the most important implication stemming from our heterogeneous treatment effects. It is particularly relevant for policymakers concerned with scaling up programs on the basis of evidence from locally-run or pilot experiments.

We use Monte-Carlo simulations for our investigation. Specifically, we create Monte Carlo samples by re-sampling from our original dataset with nine different frequencies of households who under and over estimated their consumption. We take 500 samples of each of these nine different frequencies. In constructing the samples, we keep constant both the control group and the subsample that did not answer the survey. We also hold fixed the ratio of total under and over estimators to those that stated their true consumption quantile. We then re-estimate pooled and time-varying treatment effects for each Monte Carlo sample. We then estimate pooled and time varying effects for each of these 4500 samples, bootstrapping standard errors.

Our simulation results in Table 7 quantify the rate at which conservation effects from the information program rise with the fraction of the population that underestimate their energy use. Panel A does so using the pooled treatment effects from Table 4 (e.g., the program effect estimate most previous studies tend to focus on), while Panel B uses the time-varying treatment effects from Figure 5. The point estimates for the pooled effects range from -0.002 (not significant) for a sample with very few households who overestimate to 0.010 (highly significant) with a high proportion of overestimating households. The time varying estimates display a similar pattern. The point estimates all increase with the proportion of the same who overestimate. These time varying estimates also provide further evidence that the boomerang effect for the over-estimators is temporary, by seven months after treatment, the average effect of treatment is four percent regardless of the proportion of over-estimators in the sample.

The above results beg the following question: *who* are the underestimating households that are most likely to exhibit long-run conservation effects from information programs? Given that most households think they are average, and in light of the estimates by usage level and informedness in Table 6, high energy users clearly stand out as a particularly sensitive group to behavioral nudges with peer comparisons.[30]

To examine other factors that might explain household informedness, we run auxiliary multinomial logistic regressions that correlate census-block household demographics and voting records with whether a household overestimates, is correct about, or underestimates their energy usage. We report the coefficient estimates from these regressions in Table 8.[31] The main takeaway is that Green voters, particularly those with higher incomes, systematically underestimate their energy use (i.e. think they conserve more energy relative to their peers than they actually do). Although this result may not be surprising, it does contradict an equally plausible finding that Greens are more informed about their use because they care more about conservation. This result offers an alternative, information-based interpretation of the findings from Costa and Kahn (2013). These authors interpret Liberal votes' relatively large responses to a similar information program as stemming from political ideology. Alternately, the responses that they observe could be driven – or at least amplified – by a systematic tendency to underestimate carbon footprints. Liberals may respond the most because they are learning the most.

5 Conclusions

In this paper, we used a field experiment to study the impacts of an information program with peer comparisons. We found that the effects differ depending on what households knew

[30]It is worth noting that we strictly identify an informational channel for this effect. This contrasts with interpretation of heterogeneous treatment effects in previous studies on energy and water conservation (Allcott 2011b, Ferraro and Price 2013) where it is argued that in addition to the role information frictions potentially play in generating larger conservation effects among high energy/water users, that these types may be more responsive to peer comparisons because it is easier or less costly for them to find ways to reduce their energy/water usage.

[31]We estimate similar multinomial logits at the household-level using home characteristics data from our pre-treatment survey. For the interested reader, we present these estimates in the Table A.5 of the Appendix.

pre-treatment about how their behavior compared to that of their peers. In the context of the retail electricity market, we found that households who overestimated their relative energy use exhibited a temporary boomerang effect. Their consumption increased post-treatment for a few months before returning to long-run levels. In contrast, households who underestimated their relative use permanently reduced their consumption after becoming informed about their position in the energy use distribution.

These heterogeneous treatment effects have implications for the external validity of information treatments, and hence the scope for their broader application as a policy tool. We demonstrated that the same information program can have very different average effects, depending on the distribution of pre-treatment informedness. While this finding may seem obvious ex-post, no one has previously empirically identified the effects of pre-treatment informedness on the effects of behavioral interventions with peer comparisons. Rather, this information-based channel for treatment effect heterogeneity has been assumed to exist. By showing that estimated effects differ by prior knowledge, we provide the first direct test of this channel. In doing so we have, at the very least, validated previous researchers' interpretations of information program effects. Our findings should also guide future attempts to uncover mechanisms through which information treatments affect behavior.

From a program evaluation perspective, we effectively experimented with a standard experimental design for information programs by first running a pre-treatment beliefs survey. We found the survey to be extremely useful for uncovering an important source of unobserved heterogeneity – informedness about peer behavior – and thus understanding the channels through which program effects arise.

Depending on the application, households may feel harmed from revealing their beliefs exante and being told they were wrong ex-post. These ethics concerns will be minimized, however, when there is a significant lag between survey collection and information treatment or when there is low social pressure to know the correct answer to the question. Keeping these potential ethics concerns in mind, we hope that our study will motivate researchers to collect data on baseline beliefs prior to implementing information treatments.

References

ALLCOTT, H. (2011a): "Consumers' Perceptions and Misperceptions of Energy Costs," *American Economic Review*, 101, 98–104.

——— (2011b): "Social norms and energy conservation," *Journal of Public Economics*, 95, 1082–1095.

——— (2012): "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market," Tech. rep., Working Paper, New York University (January). Available at: https://files. nyu. edu/ha32/public/research. html.

ALLCOTT, H. AND T. ROGERS (2012): "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation," Tech. rep., National Bureau of Economic Research.

ATTARI, S. Z., M. L. DEKAY, C. I. DAVIDSON, AND W. B. DE BRUIN (2010): "Public perceptions of energy consumption and savings," *Proceedings of the National Academy of Sciences*, 107, 16054–16059.

BECKER, G. (1965): "A Theory on the Allocation of Time," *Economic Journal*, 75, 493–517.

BESHEARS, J., J. J. CHOI, D. LAIBSON, B. C. MADRIAN, AND K. L. MILKMAN (2013): "The Effect of Providing Peer Information on Retirement Savings Decisions," Mimeo, Yale.

BOLLINGER, B., P. LESLIE, AND A. SORENSEN (2011): "Calorie posting in chain restaurants," *American Economic Journal: Economic Policy*, 3, 91–128.

BYRNE, D., A. L. NAUZE, AND L. MARTIN (2014): "Helpful tool or spam? Personalized information and customer churn," *University of Melbourne working paper.*

CHEN, Y., F. M. HARPER, J. KONSTAN, AND S. X. LI (2010): "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *American Economic Review*, 100, 1358–1398.

COSTA, D. L. AND M. E. KAHN (2013): “Energy conservation “nudges” and environmentalist ideology: Evidence from a randomized residential electricity field experiment,” *Journal of the European Economic Association*, 11, 680–702.

CROSON, R. AND J. Y. SHANG (2008): “The Impact of Downward Social Information on Contribution Decisions,” *Experimental Economics*, 11, 221–233.

DAVIS, J., A. J. PICKERING, K. ROGERS, S. MAMUYA, AND A. B. BOEHM (2011): “The effects of informational interventions on household water management, hygiene behaviors, stored drinking water quality, and hand contamination in peri-urban Tanzania,” *The American journal of tropical medicine and hygiene*, 84, 184.

DELMAS, M. A., M. FISCHLEIN, AND O. I. ASENSIO (2013): “Information Strategies and Energy Conservation Behavior: A Meta-analysis of Experimental Studies from 1975-2012,” *UC Center for Energy and Environmental Economics Working Paper Series.*

DUFLO, E., R. GLENNERSTER, AND M. KREMER. (2008): “Using Randomization in Development Economics Research: A Toolkit,” *Handbook of Development Economics*, 4, 3895–3962.

DUFLO, E. AND E. SAEZ (2003): “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment,” *Quarterly Journal of Economics*, 118, 815–842.

DUPAS, P. (2011): “Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya,” *American Economic Journal: Applied Economics*, 3, 1–34.

FERRARO, P. AND J. MIRANDA (2013): “Heterogeneous Treatment Effects and Mechanisms in Information-Based Environmental Policies: Evidence from a Large-Scale Field Experiment,” *Resource and Energy Economics*, 35, 356–379.

FERRARO, P. AND M. PRICE (2013): “Using non-pecuniary strategies to influence behavior: Evidence from a large scale field experiment,” *Review of Economics and Statistics*, 95, 64–73.

FITZSIMONS, E., B. MALDE, A. MESNARD, AND M. VERA-HERNÁNDEZ (2012): "Household responses to information on child nutrition: experimental evidence from Malawi," Tech. rep., CEPR Discussion Paper No. DP8915.

HARDING, M. AND D. RAPSON (2013): "Do Voluntary Carbon Offsets Induce Energy Rebound? A Conservationist's Dilemma," Tech. rep., Manuscript: Stanford.

HECKMAN, J. J. AND J. A. SMITH (1999): "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies," *Economic Journal*, 109, 313–348.

ITO, K. (2013): "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing," *American Economic Review*, forthcoming.

JALAN, J. AND E. SOMANATHAN (2008): "The importance of being informed: Experimental evidence on demand for environmental quality," *Journal of Development Economics*, 87, 14–28.

JESSOE, K. AND D. RAPSON (2013): "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use," *American Economic Review*, forthcoming.

LEVITT, S. D. AND J. A. LIST (2007): "What Do Field Experiments Measuring Social Preferences Reveal About The Real World?" *Journal of Economic Perspectives*, 21, 153–174.

LUBY, S. P., M. AGBOATWALLA, D. R. FEIKIN, J. PAINTER, W. BILLHIMER, A. ALTAF, AND R. M. HOEKSTRA (2005): "Effect of Handwashing on Child Health: A Randomised Controlled Trial," *Lancet*, 366, 225–233.

MOUNTAIN, B. (2012): "Electricity Prices in Australia: An International Comparison," Tech. rep., Carbon and Energy Markets CME.

SCHULTZ, P., J. NOLAN, R. CIALDINI, N. GOLDSTEIN, AND V. GRISKEVICIUS (2007): "The constructive, destructive, and reconstructive power of social norms," *Psychological Science*, 18, 429–434.

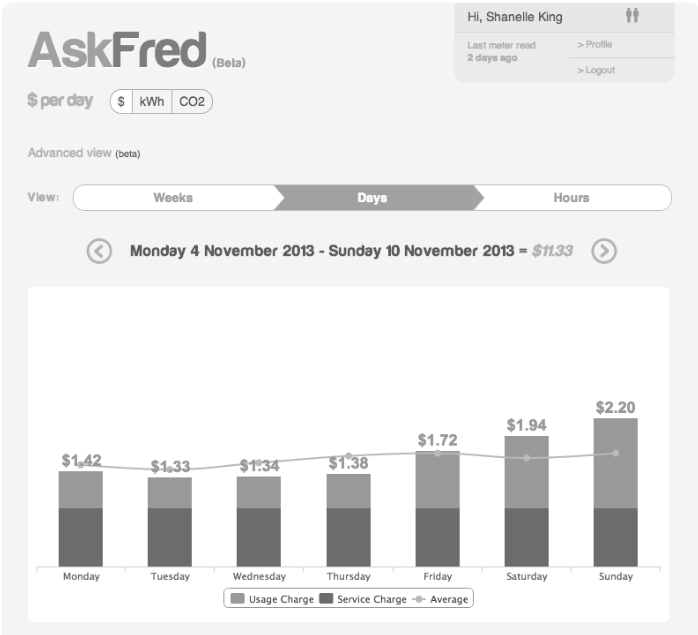
TUSTIN, J. (2012): "Electricity bill benchmarks for residential customers; Special report for the Australian Energy Regulator," Tech. rep., ACIL Tasman.

A Figures

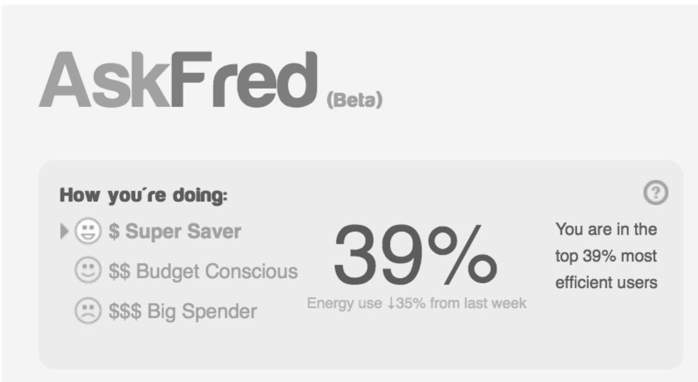
Figure 1: TREATMENT: ONLINE HOME ENERGY REPORTS

Front Page Summary

Daily Electricity Use Cost Summary



Peer comparisons



Energy savings tips

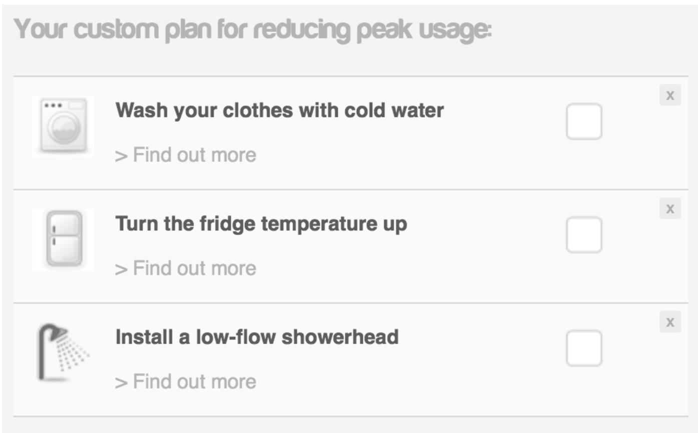


Figure 2: DISTRIBUTION OF PREDICTION ERRORS ABOUT RELATIVE ENERGY USE

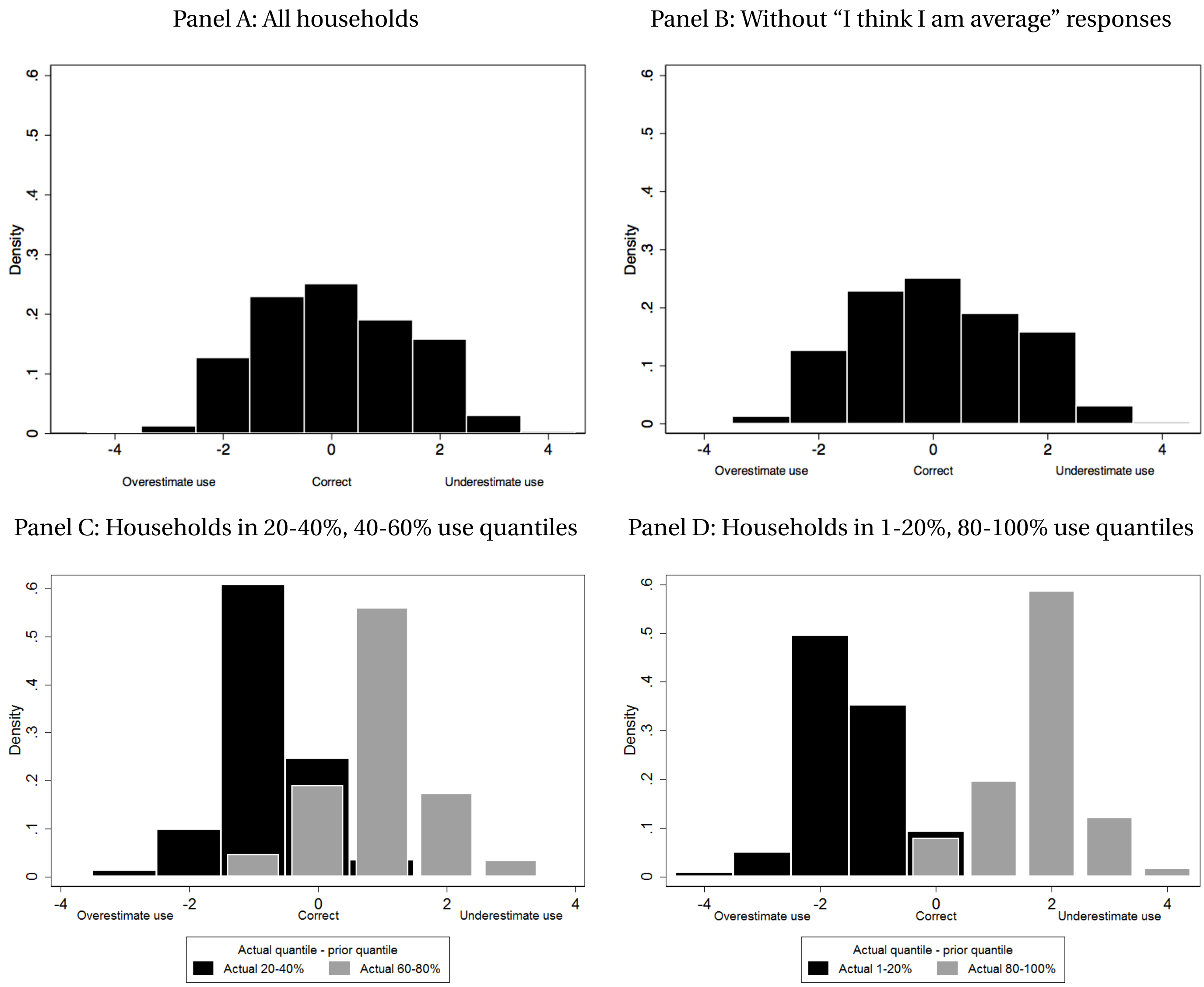
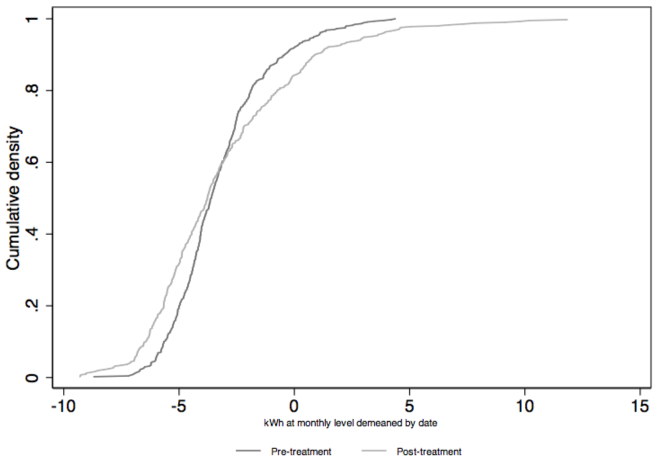
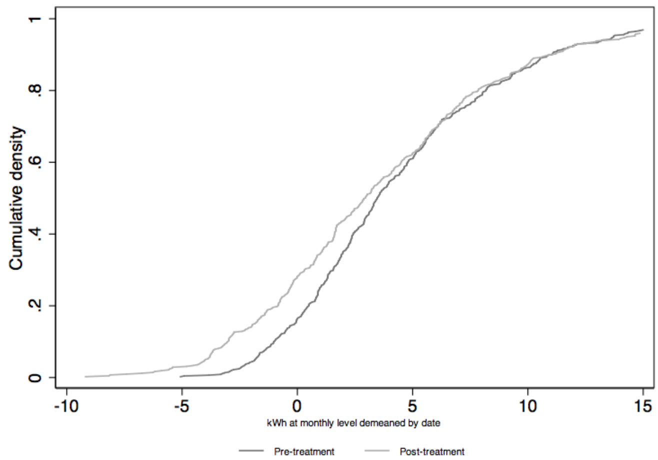


Figure 3: Cumulative Density Functions of Household Monthly Energy Use

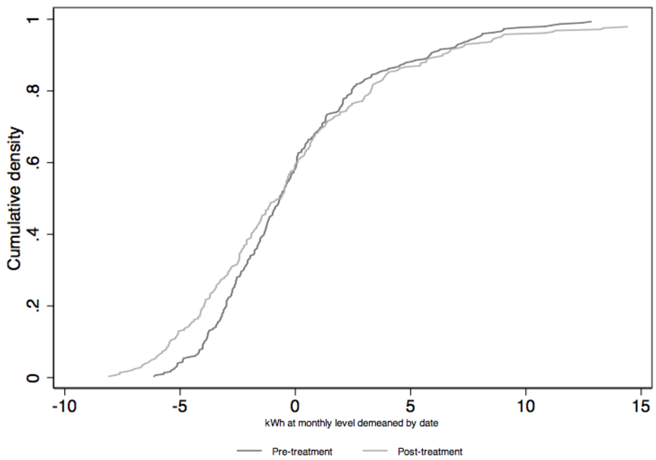
Panel A: Overestimate relative energy use



Panel B: Underestimate relative energy use



Panel C: Correct about relative energy use



Panel D: Non-respondents to energy use survey

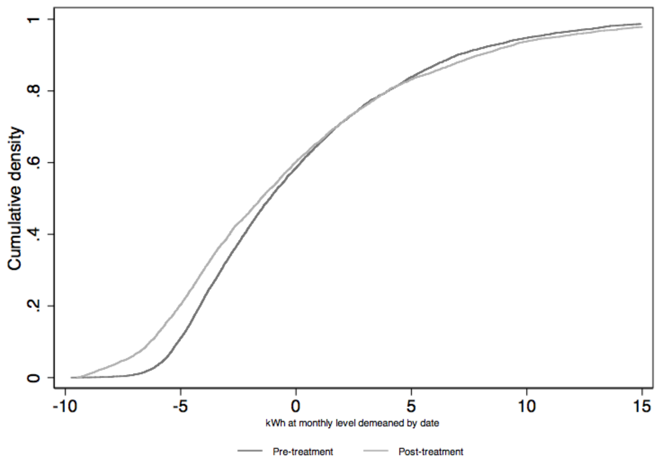


Figure 4: TIME-VARYING TREATMENT EFFECTS

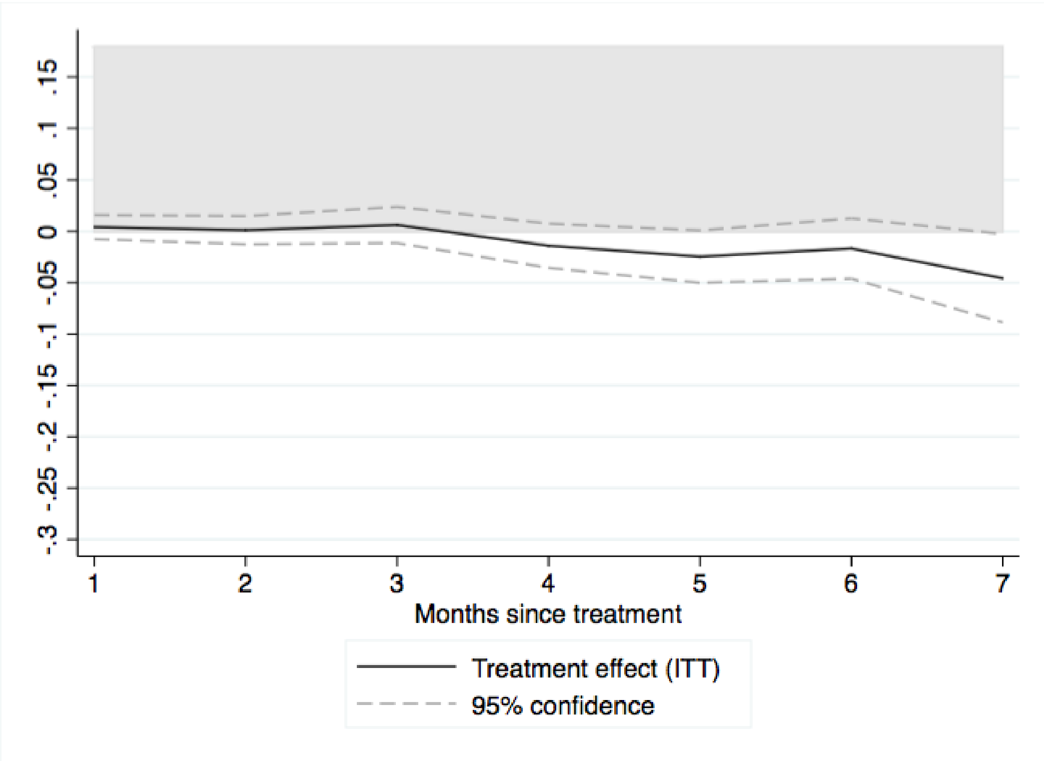
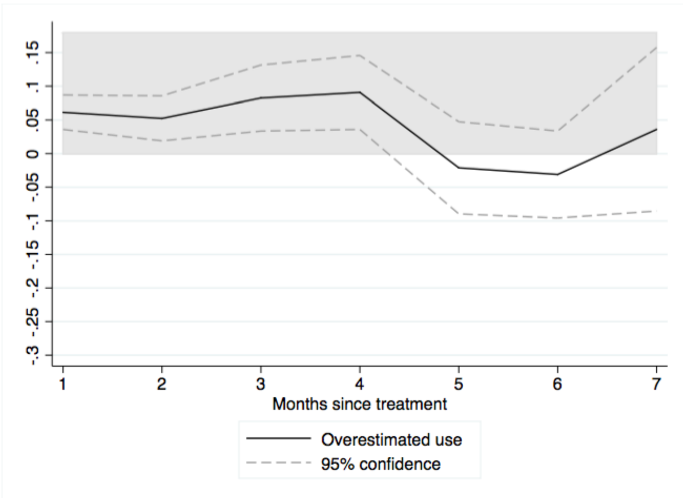
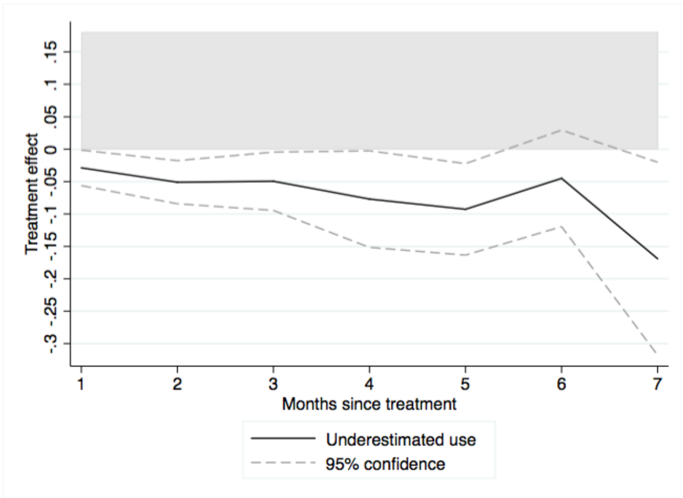


Figure 5: TIME-VARYING TREATMENT EFFECTS BY INFORMEDNESS

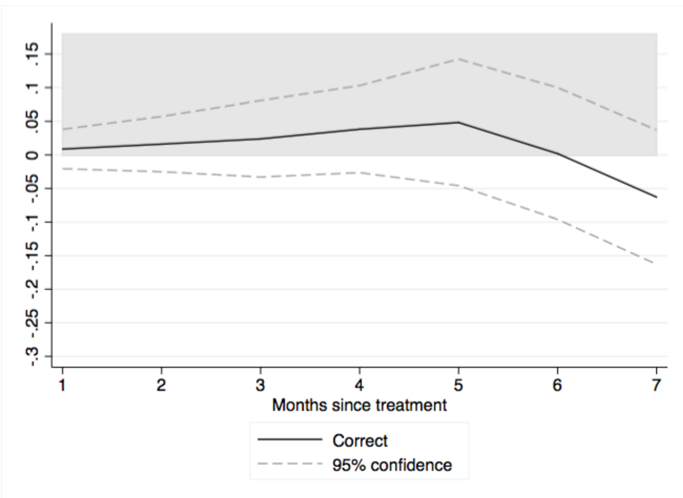
Panel A: Overestimate relative energy use



Panel B: Underestimate relative energy use



Panel C: Correct about relative energy use



Panel D: Non-respondents to energy use survey

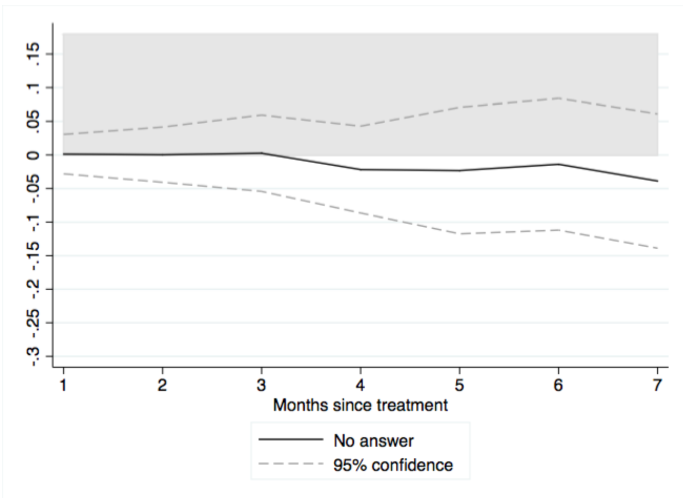


Table 1: BALANCE OF PRETREATMENT CONSUMPTION BY WAVE AND MONTH OF USE

	Wave 1	Wave 2 & 3	Wave 1 -	Wave 2	Wave 3	Wave 2 -
	Mean	Mean	(Wave 2 + 3)	Mean	Mean	Wave 3
July 2012	13.80	13.80	-0.00			
			(0.394)			
August 2012	13.12	12.88	-0.23			
			(0.36)			
September 2012	11.16	10.87	-0.29			
			(0.296)			
October 2012	10.34	9.80	-0.54			
			(0.297)*			
November 2012				9.17	8.57	0.60
						(0.381)
December 2012				9.08	8.53	0.56
						(0.352)
January 2013				9.40	8.86	0.54
						(0.332)
February 2013				10.05	10.17	-0.13
						(0.306)

Notes: Electricity consumption data is reported at the monthly level (average of daily kWh) for any month before a household receives treatment. Difference in means between Wave 1 and its control group (households in Waves 2 and 3) reported in Column (3) and between Wave 2 and its control group (households in Wave 3) in column (6). Standard errors of differences reported in brackets are clustered at postcode level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. See the text for wave definitions.

Table 2: BALANCE ON OBSERVABLE CHARACTERISTICS BY WAVE

	Wave 1	Wave 2	Wave 3			
	Mean	Mean	Mean	Wave 1-Wave 2	Wave 2-Wave 3	Wave 1 - Wave 3
Survey data						
Has air conditioning	0.70	0.65	0.70	0.050	-0.050	-0.000
				(0.060)	(0.042)	(0.053)
Has gas appliances	0.71	0.65	0.69	0.062	-0.040	0.022
				(0.074)	(0.053)	(0.054)
Has swimming pool	0.02	0.02	0.03	-0.005	-0.012	-0.017
				(0.023)	(0.017)	(0.018)
Number of residents	2.27	2.27	2.43	-0.001	-0.162	-0.163
				(0.156)	(0.119)	(0.161)
Number of bedrooms	2.38	2.33	2.47	0.051	-0.141	-0.091
				(0.191)	(0.148)	(0.179)
Is a freestanding house	0.50	0.38	0.47	0.115	-0.090	0.025
				(0.102)	(0.077)	(0.096)
Census data						
Average weekly income (AUD)	843.94	815.44	799.94	28.501	15.498	43.999
				(68.111)	(48.618)	(55.681)
Average age	37.09	36.83	36.66	0.266	0.164	0.431
				(1.297)	(1.024)	(1.327)
Proportion renters	0.39	0.40	0.40	-0.011	0.003	-0.008
				(0.051)	(0.040)	(0.050)
Proportion of labor force employed full time	0.42	0.42	0.41	0.008	0.010	0.019
				(0.030)	(0.022)	(0.026)
Has above median vote for Green Party	0.58	0.58	0.60	0.002	-0.016	-0.014
				(0.116)	(0.077)	(0.098)
Number of households	761	4,545	3,264			

Notes: Means and difference in means between waves reported. Standard errors of differences reported in brackets clustered at postcode level. Census and voting data reported at Statistical Area Level 1 (SA1). * $p<0.10$, ** $p<0.05$, *** $p<0.01$ See the text for wave definitions.

Table 3: Comparison of beliefs and actual energy use in quantiles

Panel A: Number of households

Survey data (priors)	Actual pre-treatment use 1-20%	20-40%	40-60%	60-80%	80-100%	Total
1-20%	**22**	8	10	8	4	52
20-40%	84	**58**	36	42	28	248
40-60%	118	143	**153**	136	135	685
60-80%	12	23	34	**46**	45	160
80-100%	2	3	9	11	**18**	43
Total	238	235	242	243	230	1188

Panel B: % of households

Survey data (priors)	Actual pre-treatment use 1-20%	20-40%	40-60%	60-80%	80-100%	Total
1-20%	**1.9**	0.7	0.8	0.7	0.3	4.4
20-40%	7.1	**4.9**	3.0	3.5	2.4	20.9
40-60%	9.9	12.0	**12.9**	11.4	11.4	57.7
60-80%	1.0	1.9	2.9	**3.9**	3.8	13.5
80-100%	0.2	0.3	0.8	0.9	**1.5**	3.6
Total	20.0	19.8	20.4	20.5	19.4	100.0

Notes: Table shows numbers (%) of households in each cell. Survey data (priors) = prior beliefs of use quantiles from pre-treatment survey. Actual use = actual quantile of energy use conditional on same number of bedrooms.

Table 4: Average treatment effects

	Daily data (1)	Monthly data (2)
Received information	0.005	0.003
	(0.006)	(0.005)
Household fixed effects	Yes	Yes
Date fixed effects	Yes	No
Month fixed effects	No	Yes
Number of observations	1093516	53709
Number of households	8578	8389
R Squared	0.143	0.204

Notes: Dependent variable for columns (1) and (2) is log(kWh/day), and for columns (3) and (4) is monthly average log(kWh/day). Standard errors clustered at household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: TREATMENT EFFECTS BY LEVEL OF INFORMEDNESS

	Same No. Bedrooms		Same No. Residents	Same Retailer	Drop "Guesses"	Mean Reversion by quantile	Mean Reversion by group	Only Survey Respondents	Monthly Data	Attrition Correction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Overestimated X received information	0.070***	0.072***	0.071***	0.073***	0.042**	0.047***	0.047*	0.048***	0.045***	0.032**
	(0.012)	(0.013)	(0.012)	(0.013)	(0.020)	(0.012)	(0.026)	(0.018)	(0.012)	(0.013)
Correct X received information	0.005	0.007	0.013	0.009	-0.015	0.003	-0.011	0.004	0.011	0.008
	(0.014)	(0.015)	(0.016)	(0.015)	(0.025)	(0.015)	(0.023)	(0.021)	(0.014)	(0.016)
Underestimated X received information	-0.047***	-0.045***	-0.052***	-0.039***	-0.027	-0.029**	-0.007	-0.013	-0.035***	-0.032**
	(0.012)	(0.012)	(0.011)	(0.011)	(0.017)	(0.012)	(0.023)	(0.018)	(0.012)	(0.014)
No response X received information		0.004								
		(0.006)								
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Month fixed effects	No	No	No	No	No	No	No	No	Yes	Yes
Group X date fixed effects	No	No	No	No	No	No	Yes	No	No	No
Quintile X date fixed effects	No	No	No	No	Yes	Yes	No	Yes	No	No
Quintile X month fixed effects	No	No	No	No	No	No	No	No	Yes	Yes
Number of observations	1,093,516	1,093,516	1,093,516	1,093,516	980,142	1,093,516	1,093,516	198,133	53,709	42,140
Number of households	8,578	8,578	8,578	8,578	7,888	8,578	8,578	1,251	8,389	7,462
R Squared	0.143	0.143	0.143	0.143	0.151	0.150	0.145	0.156	0.220	0.209

Notes: Dependent variable for every model is log(kWh/day) except columns (8) and (9) for which is average log(kWh/day) at the monthly level. Standard errors are clustered at household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Over (under) estimated implies that a household's prior belief about their use was an over (under) estimate of their household's actual consumption. In all columns except for Columns (3) and (4) actual consumption is assessed relative to households with the same number of bedrooms. Column (3) compares consumption for same number of residents and Column (4) compares within the entire sample. Estimates in Columns (5), (6) and (8) through (10) include corrections for mean reversion by use quintile. In Column (5) the sample is restricted to households that stated that they believed they were strictly above or strictly below average. In Column (8) the sample is restricted to households who answered the survey, i.e. with known number of bedrooms. Estimates in column (10) are computed using Inverse Probability Weights (IPW) to account for sample attrition. The first stage attrition model for column (10) includes one period lags of the treatment indicator and electricity consumption as well as indicators for distributor and price plan.

Table 6: TREATMENT EFFECTS BY QUINTILE OF USE AND INFORMEDNESS

	Overestimated X received information	Correct X received information	Underestimated X received information
Lowest use	0.094***	0.062	
	(0.018)	(0.038)	
Lower than average use	0.055***	0.032	0.060
	(0.017)	(0.020)	(0.057)
Average use	0.038	0.013	-0.036
	(0.035)	(0.016)	(0.024)
Higher than average use	0.119	0.013	-0.012
	(0.133)	(0.016)	(0.020)
Highest use		-0.031	-0.079***
		(0.034)	(0.017)
Household fixed effects		Yes	
Date fixed effects		Yes	
Number of observations		1,093,516	
Number of households		8,578	
R squared		0.150	
Number of households by category	439	452	297

Notes: Dependent variable for model is log(kWh/day). Standard errors are clustered at household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Over (under) estimated implies that a household's prior belief about their use was an over (under) estimate of their household's actual consumption. Actual consumption is assessed relative to households with the same number of bedrooms. Estimate includes day by sample quintile fixed effects to account for potential mean reversion amongst quintiles.

Table 7: IMPACT OF DISTRIBUTION OF INFORMEDNESS PRIORS ON ESTIMATED EFFECT SIZE

	Many Underestimators (1)	(2)	(3)	(4)	(5)	(6)	(7)	Few Underestimators (8)	(9)
Panel A: Pooled estimates									
Received information (ITT)	-0.002	-0.001	0.001	0.002	0.004	0.006**	0.007***	0.009***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Panel B: Time varying estimates									
One month since X	-0.006*	-0.005	-0.004	-0.003	-0.002	-0.001	-0.000	0.001	0.001
received information (ITT)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Two months since X	-0.005	-0.003	-0.002	0.000	0.002	0.004	0.006	0.007*	0.009**
received information (ITT)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Three months since X	0.003	0.004	0.006	0.007	0.009*	0.010**	0.011**	0.013**	0.014***
received information (ITT)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Four months since X	-0.015*	-0.011	-0.008	-0.005	-0.002	0.002	0.005	0.008	0.011
received information (ITT)	(0.008)	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
Five months since X	-0.022**	-0.021**	-0.020**	-0.020**	-0.019**	-0.018**	-0.017**	-0.016**	-0.016**
received information (ITT)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
Six months since X	-0.012	-0.013	-0.016*	-0.018**	-0.020**	-0.022**	-0.025***	-0.026***	-0.029***
received information (ITT)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Seven months since X	-0.044***	-0.043***	-0.043***	-0.043***	-0.041***	-0.040***	-0.039***	-0.038***	-0.037***
received information (ITT)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Over-estimators/Total known	7%	13%	20%	27%	33%	40%	47%	53%	60%
Number of observations	67,129	67,101	67,074	67,042	67,094	66,992	66,949	66,922	66,895
Number of replications	500	500	500	500	500	500	500	500	500

Notes: Outcome variable is average log(kwH/day) at the monthly level. Mean estimates from Monte Carlo replications reported. Standard errors in brackets (standard deviation of treatment effect estimates over replications). Number of observations is the average sample size over all 500 replications. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: PREDICTING INFORMEDNESS USING CENSUS DISTRICT CHARACTERISTICS

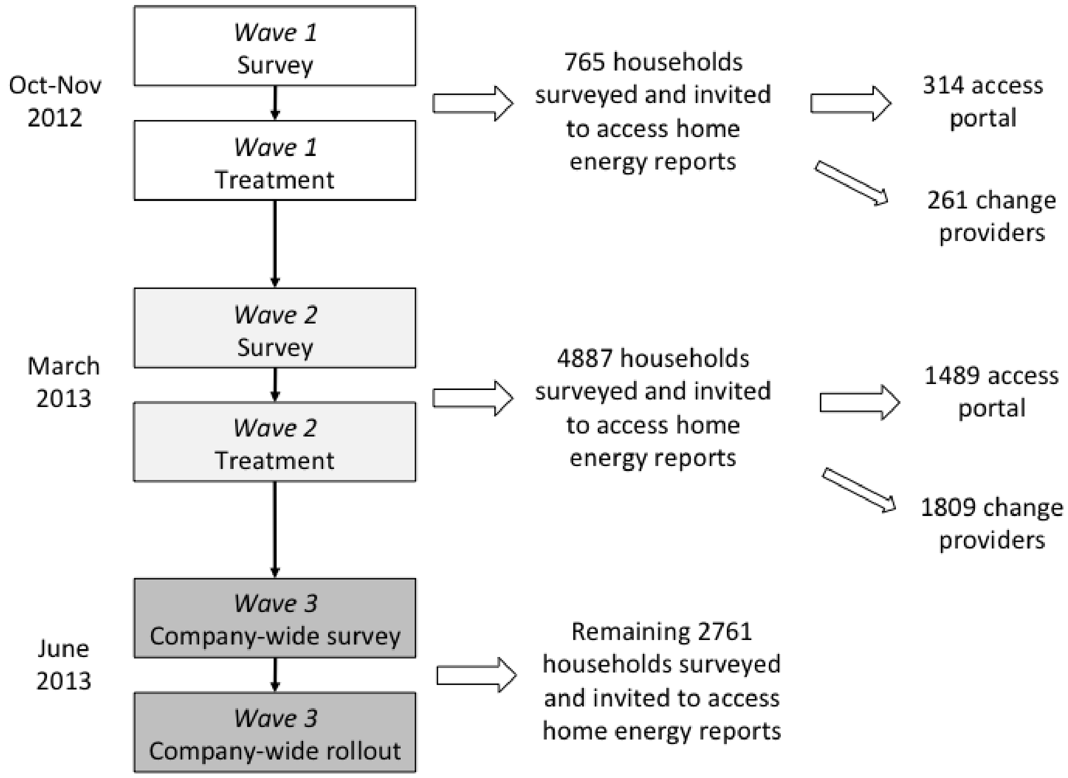
	Household informedness					
	(1)	(2)	(3)	(4)	(5)	(6)
Overestimated						
Average pretreatment kWh/day	-1.239***	-1.140***	-1.207***	-1.134***	-1.155***	-1.187***
	(0.159)	(0.148)	(0.160)	(0.149)	(0.167)	(0.169)
Has above median primary vote for Green Party	-0.403***		-0.396			-0.700
	(0.142)		(0.563)			(0.597)
Average weekly income (2011 AUD)		-0.791***	-0.550	-0.856***	-0.783***	-0.805
		(0.282)	(0.524)	(0.290)	(0.293)	(0.532)
Above median vote Greens X Income			0.129			0.461
			(0.668)			(0.698)
Average age				-0.020*		-0.023**
				(0.012)		(0.012)
Proportion rental properties					-0.159	0.127
					(0.414)	(0.478)
Underestimated						
Average pretreatment kWh/day	1.272***	1.130***	1.293***	1.133***	1.367***	1.397***
	(0.180)	(0.167)	(0.181)	(0.168)	(0.192)	(0.192)
Has above median primary vote for Green Party	0.518***		-0.868			-1.007
	(0.147)		(0.634)			(0.642)
Average weekly income (2011 AUD)		0.605**	-1.098*	0.616**	0.389	-0.913
		(0.300)	(0.593)	(0.301)	(0.311)	(0.609)
Above median vote Greens X Income			1.719**			1.604**
			(0.746)			(0.775)
Average age				0.001		-0.003
				(0.011)		(0.011)
Proportion rental properties					1.410***	1.023**
					(0.431)	(0.467)
Number of households	1560	1560	1560	1560	1560	1560
Pseudo R Squared	0.102	0.095	0.104	0.097	0.102	0.108

Notes: Coefficients of multinomial logit explaining informedness relative to houses with the same number of bedrooms. The omitted category is households who are correct regarding their relative usage. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Census district is the SA1 level from the 2011 Australian Census of Population and Housing

A Appendix

A.1 Supplemental figures

Figure A.1: TIMELINE FOR TREATMENT WAVES



A.2 Supplemental tables

Table A.1: BALANCE ON OBSERVABLE CHARACTERISTICS BY SURVEY RESPONSE STATUS

	Non Respondent	Respondent	Non respondent -
	Mean	Mean	Respondent
Average pretreatment kWh/day	10.68	11.04	-0.361
			(0.255)
Census data			
Average weekly income (AUD)	809.89	826.01	-16.130
			(19.399)
Average age	36.70	37.28	-0.575
			(0.504)
Proportion renters	0.40	0.37	0.032
			(0.016)*
Proportion of labor force	0.41	0.42	-0.012
employed full time			(0.010)
Proportion primary vote for	0.33	0.33	-0.004
Australian Labor Party			(0.008)
Has above median vote for	0.51	0.47	0.038
Australian Greens Party			(0.039)
Number of households	5,995	1,070	

Notes: Electricity consumption data is reported for the period before a household receives treatment. Survey administered prior to treatment. Census and voting data reported at Statistical Area Level 1 (SA1). Difference in means reported by survey response status. Standard errors of differences in brackets clustered at postcode level.

Table A.2: BALANCE OF PRETREATMENT CONSUMPTION BY ATTRITION STATUS WITHIN WAVE

	Wave 1			Wave 2			Wave 3		
	Retained Mean	Attrited Mean	Retained - Attrited	Retained Mean	Attrited Mean	Retained - Attrited	Retained Mean	Attrited Mean	Retained - Attrited
July 2012	13.59	14.23	-0.64 (0.626)	13.43	14.53	-1.10 (0.496)**			
August 2012	12.93	13.49	-0.56 (0.575)	12.54	13.57	-1.04 (0.427)**			
September 2012	10.96	11.57	-0.61 (0.454)	10.63	11.48	-0.86 (0.340)**			
October 2012	10.33	10.38	-0.05 (0.490)	9.69	10.29	-0.60 (0.284)**			
November 2012				9.05	9.38	-0.32 (0.246)			
December 2012				8.95	9.36	-0.41 (0.229)*			
January 2013				9.26	9.76	-0.50 (0.242)**			
February 2013				9.92	10.36	-0.45 (0.297)			
March 2013							10.45	12.75	-2.30 (0.664)***
April 2013							10.27	12.67	-2.39 (0.716)***
May 2013							12.23	14.80	-2.57 (1.204)**

Notes: Electricity consumption data is reported at the monthly level (average of daily kWh) for any month before a household receives treatment. Means reported for households retained and households who attrited before June 2013. Means reported only for months with observations for at least 100 households per wave. Differences and standard error of difference (in brackets, clustered at postcode level) reported. * $pvalue < 0.10$, ** $pvalue < 0.05$, *** $pvalue < 0.01$ Wave 1 households were assigned to the treatment group prior to March 2013. Wave 2 households were assigned to the treatment group between March 2013 and June 2013. Wave 3 households are control households and assigned to treatment at the end of our sample period (June 2013).

Table A.3: BALANCE ON OBSERVABLE CHARACTERISTICS BY ATTRITION STATUS

	Wave 1			Wave 2			Wave 3		
	Retained Mean	Attrited Mean	Retained - Attrited	Retained Mean	Attrited Mean	Retained - Attrited	Retained Mean	Attrited Mean	Retained - Attrited
Survey data									
Has air conditioning	0.70	0.68	0.024	0.64	0.75	-0.106	0.70	0.80	-0.102
			(0.126)			(0.079)			(0.209)
Has gas appliances	0.71	0.72	-0.012	0.66	0.54	0.117	0.69	0.60	0.090
			(0.102)			(0.097)			(0.212)
Has swimming pool	0.02	0.00	0.019	0.02	0.00	0.022	0.03	0.20	-0.171
			(0.027)			(0.036)			(0.079)**
Number of residents	2.27	2.28	-0.011	2.26	2.37	-0.102	2.43	3.00	-0.575
			(0.285)			(0.196)			0.564
Number of bedrooms	2.38	2.40	-0.020	2.34	2.27	0.069	2.48	2.40	0.075
			(0.277)			(0.243)			(0.530)
Is a freestanding house	0.48	0.56	-0.079	0.38	0.41	-0.035	0.47	0.40	0.072
			(0.153)			(0.130)			(0.267)
Census data									
Average weekly income (AUD)	859.41	812.72	46.687	839.37	777.44	61.927	802.54	739.92	62.623
			(54.266)			(51.416)			(90.284)
Average age	37.38	36.52	0.861	37.03	36.50	0.538	36.64	37.22	-0.583
			(1.130)			(0.997)			(2.279)
Proportion renters	0.39	0.39	-0.000	0.39	0.41	-0.016	0.40	0.32	0.082
			(0.039)			(0.039)			(0.084)
Proportion of labor force employed full time	0.43	0.41	0.019	0.43	0.40	0.027	0.41	0.39	0.012
			(0.024)			(0.023)			(0.043)
Proportion primary vote for Australian Labor Party	0.34	0.32	0.011	0.34	0.33	0.007	0.33	0.33	-0.002
			(0.025)			(0.024)			(0.043)
Has above median vote for Australian Greens Party	0.59	0.56	0.032	0.59	0.57	0.013	0.60	0.53	0.067
			(0.106)			(0.082)			(0.131)
Number of households	503	258		2,738	1,807		3,099	165	

Notes: Wave 1 households were assigned to the treatment group prior to March 2013. Wave 2 households were assigned to treatment group between March 2013 and June 2013. Wave 3 households are control households and assigned to treatment at the end of our sample period (June 2013). Census and voting data reported at Statistical Area Level 1 (SA1). Differences are reported for tests of differences in means between households retained throughout the sample period, and households who attrited within a Wave. Standard errors of differences (clustered at postcode level) reported in brackets.

Table A.4: MONTHLY ATTRITION RATES BY WAVE

	Wave 1 Attrition rate	Wave 2 Attrition rate	Wave 3 Attrition rate	Wave 1 - Wave 2	Wave 2 - Wave 3	Wave 3 - Wave 3
Monthly attrition	0.04	0.06	0.01	-0.028	0.050	0.022
				(0.008)***	(0.004)***	(0.007)***
Number of households	765	4,548	3,265			

Notes: Monthly attrition rates and differences by wave reported. Standard errors of differences reported in brackets. * $pvalue < 0.10$, ** $pvalue < 0.05$, *** $pvalue < 0.01$. Wave 1 households were assigned to the treatment group prior to March 2013. Wave 2 households were assigned to the treatment group between March 2013 and June 2013. Wave 3 households are control households and assigned to treatment at the end of our sample period (June 2013).

Table A.5: PREDICTING INFORMEDNESS USING HOUSEHOLD CHARACTERISTICS

	Household informedness		
	(1)	(2)	(3)
	Same No. Bedrooms		Drop "Guesses"
Overestimated			
Has air conditioning	-0.167	-0.374*	-0.339
	(0.158)	(0.217)	(0.212)
Has gas appliances	0.128	0.070	-0.015
	(0.189)	(0.252)	(0.246)
Has swimming pool	-0.028	0.192	-0.356
	(0.486)	(0.559)	(0.562)
Number of residents	0.012	0.011	-0.001
	(0.081)	(0.108)	(0.113)
Number of bedrooms	0.764***	0.753***	0.445***
	(0.126)	(0.173)	(0.162)
Is a free-standing house	-0.044	-0.032	-0.252
	(0.208)	(0.289)	(0.293)
Proportion consumption 9am-5pm	-2.303*	-1.124	-0.346
	(1.183)	(1.475)	(1.561)
Is a rental property		0.277	
		(0.228)	
Underestimated			
Has air conditioning	-0.166	-0.220	-0.286
	(0.172)	(0.230)	(0.229)
Has gas appliances	-0.326*	-0.345	-0.368
	(0.187)	(0.246)	(0.249)
Has swimming pool	-0.208	-0.315	-0.887*
	(0.343)	(0.434)	(0.522)
Number of residents	0.006	-0.068	0.015
	(0.079)	(0.104)	(0.111)
Number of bedrooms	-1.104***	-1.107***	-0.680***
	(0.133)	(0.175)	(0.167)
Is a free-standing house	0.292	0.357	0.336
	(0.224)	(0.297)	(0.313)
Proportion consumption 9am-5pm	-1.559	-1.122	-0.938
	(1.165)	(1.466)	(1.599)
Is a rental property		-0.555**	
		(0.235)	
Pre-treatment quintile use FE	yes	yes	yes
Number of households	1673	940	742
Pseudo R Squared	0.316	0.311	0.181

Notes: Coefficients of multinomial logit explaining informedness relative to houses with the same number of bedrooms. The omitted category is households who are correct regarding their relative usage. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (2) includes the variable 'Is a rental property' which was only asked during later waves of survey data collection. Column (3) Households who did not consider themselves average.

A.3 Testing for common trends in pre-treatment consumption

Our identification of the treatment effect of the information program relies on the assumption that treatment and control groups share a common linear time trend. We are not able to test this assumption over the experimental window. However, we can provide some evidence for common trends in the pre-treatment period. To do so, we divide our sample into the three groups corresponding to the three waves of treatment. We focus on the period prior to the first treatment date in our sample (August 2012). We specify the following model for consumption of households in each wave:

$$Y_{it} = \tau_1 \times w_{it1} \times t + \tau_2 \times w_{it2} \times t + \tau_3 \times w_{it3} \times t + \mu_{i0} + \eta_{t0} + \epsilon_{it0} \tag{3}$$

where Y_{it} is the logarithm of daily energy use (in kWh) for household i on date t, and w_{itj} is a dummy that equals one if household i is a member of wave j. The composite error includes a household-specific effect μ_i, a date-specific effect η_t, and an idiosyncratic shock ϵ_{it}. The parameter τ_j captures the linear time trend of households in wave j, conditional on household and date fixed effects. We then test whether $\tau_1 = \tau_2 = \tau_3$. The P-value for this test is 0.69. This provides some reassurance that at least in the period just before the initial wave of treatment that the common trends assumption cannot be rejected.