University of Neuchatel

Institute of Economic Research

IRENE Working paper 16-03

Move It! How an Electric Contest Motivates Households to Shift their Load Profile

Sylvain Weber

Stefano Puddu

Diana Pacheco

UNIVERSITÉ DE NEUCHÂTEL

Institut de recherches économiques

Move it! How an electric contest motivates households to shift their load profile*

Sylvain Weber† Stefano Puddu‡ Diana Pacheco§

July 10, 2017

Abstract

Photovoltaic systems generate electricity around noon, when many homes are empty. Conversely, residential electricity demand peaks in the evening, when production from solar sources is impossible. Based on a randomized control trial, we assess the effectiveness of alternative demand response measures aimed at mitigating these imbalances. More precisely, through information feedback and financial rewards, we encourage households to shift electricity consumption toward the middle of the day. Using a difference-in-differences approach, we find that financial incentives induce a significant increase of the relative consumption during the period of the day when most solar radiation takes place. Information feedback, however, pushes households to decrease overall consumption, but induces no load shifting.

JEL Classification: C93, D12, L94, Q41.

Keywords: household electricity usage; smart metering; demand response; randomized control trial; difference-in-differences.

*This research was financially supported by the Swiss Federal Office of Energy (SFOE), and is part of the activities of SCCER CREST, which is financially supported by the Swiss Commission for Technology and Innovation (CTI). Previous versions of this paper have been presented at the following events: *Enerday*, Dresden (Germany), April 17, 2015; *Workshop in Economic Evaluation of Energy Policy Measures in Switzerland*, Zurich (Switzerland), November 25, 2015; *SAEE/SCCER CREST Conference*, Lausanne (Switzerland), February 26, 2016; *SSES Annual Congress*, Lugano (Switzerland), June 10, 2016; *EAERE Annual Conference*, Zurich (Switzerland), June 25, 2016; *EMEE International Workshop*, Oviedo (Spain), July 7, 2016. We thank the participants for valuable comments that led to improvements. Any remaining error is our own.

†University of Neuchâtel, sylvain.weber@unine.ch.

‡University of Neuchâtel, stefano.puddu@unine.ch.

§University of Neuchâtel, diana.pacheco@unine.ch.

1 Introduction

Renewable energy sources are increasingly used in order to curb CO_2 emissions and mitigate climate change. In particular, solar photovoltaic systems are being deployed at an exponential rate (see e.g., IEA, 2016). However, this technology has the drawback of generating electricity when residential demand is relatively low. In fact, solar radiation is maximal around noon, when many homes are empty. Conversely, households' electricity usage peaks in the evening, when solar radiation is nil. This mismatch constitutes a potential obstacle to the diffusion of electricity produced from solar sources, even if it can be mitigated in different ways: Grid expansions would help in transporting electricity for remote consumption; Storage capacities would allow delayed consumption; Demand-response measures may be implemented so that electricity is used when and where it is produced.

While the first two solutions are supply-side measures requiring high installation costs (Beaudin et al., 2010) and involving losses (Denholm and Hand, 2011), demand-side management might prove relatively inexpensive and easy to implement. An additional advantage of demand-side interventions is that they can be implemented rapidly, while deployment of new technologies on a large scale can only be achieved in the long run (Dietz et al., 2009). It has also been argued that different solutions like storage and demand management should be considered as complementing each other for grid balancing (Elliott, 2016).

Mitigating imbalances between electricity production and usage is a particularly serious issue in remote areas. In fact, when large amounts of solar energy are produced in isolated regions at times of low demand, providers may find the technical installations necessary to accommodate excess supply too expensive and could prefer to literally destroy the electricity produced.[1] This dismal anecdote illustrates how supply-side measures might fail when it comes to integrating electricity from renewable sources in the grid. In such situations, demand-side management becomes crucial.

[1]We thank a representative of the electricity provider involved in this project for bringing this point to our attention.

Even if this paper focuses on the misalignment of solar energy production and electricity consumption, its scope extends to the general issue of the increasing importance of renewable technologies. Because the new renewable energy sources (sun and wind) are intermittent by nature, relying more on these will imply managing demand more actively in order to minimize the cost of using the electricity produced.

In this paper, using a randomized control trial, we investigate whether and by how much households shift intraday electricity usage when provided with incentives to do so. More precisely, the objective of our intervention is to encourage households to increase their share of electricity used between 11am and 3pm (when most solar radiation takes place), while discouraging any voluntary increase in overall consumption.[2] The first treatment consists in information feedback sent monthly via paper mail. Households included in this group obtain details about their own consumption, its pattern within the day, and similar information about households of comparable size. The second treatment is a contest in which the best performers (i.e., those who achieve the strongest shift in electricity usage while keeping total usage at reasonable levels) receive cash prizes determined by their position in the ranking.

Using a difference-in-differences approach, our results suggest that both treatments induce a reaction, but the type of reaction depends on the treatment. Households exposed to information feedback globally decrease electricity usage regardless of the hours of the day, but they do not shift consumption toward the indicated hours. Conversely, households exposed to financial incentives display a significant shift toward the solar energy production hours. They achieve an increase in their proportion of electricity used between 11am and 3pm by about 1 percentage point, which is not negligible considering this proportion is around 20% on average.

[2]For technical reasons and for facilitating participants' understanding, the "solar energy production hours" have been arbitrarily set from 11am to 3pm for the entire duration of the intervention. Appendix A displays radiation (a measure of the total energy delivered by sunlight over a given period) measured by *MeteoSwiss* in a weather station located in the area where our intervention takes place. About 50% of yearly radiation occurs during the 4 hours of interest, even though there are obvious seasonal deviations.

The remainder of the paper unfolds as follows. Section 2 discusses how our contribution fits into the recent literature. Section 3 presents our experimental design. Section 4 describes our dataset and provides a descriptive analysis of the impacts of our intervention. Section 5 explains our econometric strategy and reports our empirical results. Section 6 discusses the policy implications and concludes.

2 Literature review

Our paper is located at the intersection of two distinct bodies of literature. The first is composed by technical contributions that examine the feasibility of integration of large amounts of variable renewable energy sources (sun and wind) into the energy system (see the review by Kondziella and Bruckner, 2016). This branch of the literature focuses on engineering solutions and essentially analyzes supply-side measures, whereby the issue is to render electricity distribution amenable through grid expansion and/or storage, and it makes assumptions to simulate consumers' behavior. For instance, Denholm and Margolis (2007) evaluate scenarios where up to an arbitrary 10% of demand can be shifted to absorb excess photovoltaic generation, and Steinke et al. (2013) simply keep demand constant. In our study, we tackle a similar problem but look at it from an opposite perspective, that is, focusing on demand management. We obtain exact consumption data and observe actual households' reaction, but we make simplifying assumptions concerning the supply of renewable energy.

Among the technical studies, Pina et al. (2012) stand out as they analyze the impact of demand-side management strategies in an electricity mix characterized by high shares of renewable energy. Their results are nevertheless based on simulations, both for the installation of new generation capacity and for electricity usage. For example, they assume that because washing machines, dryers, and dishwashers can be easily programmed, these devices *will* be scheduled to operate when renewable electricity generation is high. In some of their scenarios, around 40% of these activities are thus shifted. The assumption that households (and other consumers) will actually change their habits is crucial, but far from obvious. In the present paper, we are testing this assumption by investigating if and how real households do actually shift load when faced with incentives to do so.

The second relevant branch of the literature is composed by economic studies dealing with the effects of feedback on household electricity demand. As documented in several literature reviews (Buchanan et al., 2015; Faruqui et al., 2010; Vine et al., 2013) and meta-analyzes (Delmas et al., 2013; Ehrhardt-Martinez et al., 2010), this research field is growing fast.[3] In these contributions, the usual objective is to induce electricity conservation. Instead, our goal here is to induce load shifting across hours of the day without requiring any decrease (but still prevent increases) in total electricity usage. As such, the present study is among the first economic contributions targeting load shifting, and the main novelty of our intervention is that we intend to attract electricity usage *toward* some specific (sunlight) hours instead of trying to push it *away* from some specific (peak) hours. In that sense, our intervention is opposed to the usual paradigm of peak shaving and in fact seeks to create a consumption peak that coincides with the peak of solar energy production.

Other remarkable features of our contribution are the duration of the experiment and the response rate of the pre-experiment survey. Our experiment was conducted over an exceptionally long period of 18 months: three months of pre-treatment observation, twelve months of treatment, and three further months of observation after the treatment was terminated. By comparison, 60% of the studies considered in the meta-analysis by Delmas et al. (2013) lasted three months or less. Most studies moreover discontinue observation once the treatment is terminated, so that the availability of a post-treatment observation period is also an exceptional feature. Furthermore, our pre-experiment survey reached a response rate of about 16%, a considerable

[3]Recent studies having targeted household electricity conservation include Allcott and Rogers (2014), Bernstein and Collins (2014), Chen et al. (2015), Degen et al. (2013), Di Cosmo et al. (2014), Ida et al. (2015), Ito et al. (2015), Jessoe and Rapson (2014), Lynham et al. (2016), McCoy and Lyons (2017), and Pellerano et al. (2017).

figure when compared to other studies.[4] In Ito et al. (2015) and Degen et al. (2013), the response rates are about 1.7% and 8.7%, respectively. Jessoe and Rapson (2014) sent 60,000 e-mails but use only 437 households in their analysis (0.7%). A high response rate tends to minimize the self-selection problems that might plague this type of experiment. Finally, our experiment fulfills all conditions to qualify as a "high-quality study" according to Delmas et al. (2013): dedicated control group, weather controls, demographic and household controls, and randomization.

3 Experimental design

Our field experiment was conducted on households living in Cernier, a village of around 2,000 inhabitants located in the canton of Neuchâtel (Switzerland). In 2012, in the frame of a pilot study, the electricity provider *Groupe e* started to equip households with smart meters recording electricity usage in 15-minute intervals. In April 2013, we contacted 387 households already equipped or eligible to be equipped with smart meters, and invited them to fill an online survey.[5] The invitation letter remained intentionally vague about the objective of the study, in order not to influence the households that would later be assigned to the control group. It was simply stated that a "scientific study about electricity consumption" was undertaken to "investigate the flexibility of electricity consumption, in order to adapt it to the production of electricity from renewable sources (in particular photovoltaic production)." To foster participation, a lottery awarding CHF 200 in cash was organized among respondents, and a reminder letter was sent one month after the initial invitation letter.[6]

For households of the control group, these two letters constitute the only direct interaction with the project. A time lag of more than six months there-

[4]This figure represents the "clean" response rate, i.e., the ratio of usable households compared to the number of households initially contacted. The "gross" response rate, i.e., the ratio of the number of respondents to the number of households contacted, was around 27%.

[5]The survey (in French) is available from the authors on request. It collected information about household composition, respondent characteristics, and dwelling attributes.

[6]The invitation and reminder letters are available on request.

fore took place between the information received by these households and the outset of the intervention in January 2014, and it appears very unlikely that their behavior could be affected by our letters. Also, while we cannot formally exclude spillovers from households in the treatment groups to those in the control group, we argue that such effects are very unlikely to be strong, considering that the objective of the intervention is not straightforward. If anything, potential spillovers would moreover tend to minimize the observed impact of our intervention, so that the treatment effects we obtain are possibly downward biased and can be considered as a lower bound of the true effects.

By mid-December 2013, 131 households had completed the online survey. Among them, households facing a time-of-use tariff or already involved in other experiments were discarded, which left 65 households eligible to participate in our experiment.[7,8] The observation period lasted from October 2013 to March 2015, and treatments were administered from January to December 2014. October-December 2013 thus constitute the pre-treatment period, while January-March 2015 constitute the post-treatment observation period.

In order to guarantee the internal validity of our study, the 65 eligible households have been randomly assigned to one control and two treatment groups of equal size using a stratification procedure. The goal of the procedure was to create groups as similar as possible with respect to the following selected characteristics: pre-experiment electricity usage per household member, dwelling size, and highest education level achieved by a member of the household.

Two households filled the online survey after the stratification had been conducted but before the intervention started, and they have been included in the control group: The data (survey and electricity usage) are complete

[7]Households with a time-of-use tariff pay less for one kWh during the night than during the day. Such households can obviously not be included in an intervention where the objective is to shift consumption toward the daylight hours, as this would inflate their electricity bill.

[8]At the beginning of our study, different projects were ongoing in Cernier (see http://www.solution-concerto.org).

for these late respondents, but it was not possible to inform them in due time of the starting date of the treatment. During the observation period, four households moved out (two from the control and two from the treatment 1), and we discard these attrition households from the analysis because their observation period is truncated. Two out of these four households even moved out before the treatment started. Finally, one household (from the treatment 2) appears as a clear outlier and is also excluded from the analysis.[9] Our estimations are thus conducted on a sample comprising 62 households (with 22/19/21 in the control/treatment 1/treatment 2 groups).[10]

To assess the outcome of the stratification procedure, we test the average differences of the variables of interest between each treatment group and the control group. As reported in Table B.1 in Appendix B, the treatment and control groups appear largely comparable. Except for two categories of age, all tests do not reject the equivalence of the characteristics between groups.

3.1 Treatment 1: Information feedback

Treatment 1 consists in information feedback designed to foster shifts of electricity usage toward the period from 11am to 3pm. Through this intervention, our goal is to assess whether information alone is a sufficiently strong incentive for households to engage in load shifting. The feedback is delivered via monthly letters and intended to enhance households' knowledge

[9] This household's electricity usage corresponds to less than 10% of the average usage of all households, while its proportion of electricity used between 11am and 3pm is more than twice the average proportion of all households. The behavior of this household during the treatment even looks suspicious: During some months, 70% of all its electricity usage occurs between 11am and 3pm. More formally, several statistical tests nominate this household as a severe outlier. For example, a procedure for detecting outliers in multivariate data (Weber, 2010) shows it is the most distant from all others. Robust regressions also show the influence of this household and assign it the lowest weight.

[10] As robustness checks, we have conducted our estimations excluding the late respondents, including the attrition households, and including the outlier. The results are available on request. In all estimations based on alternative samples, the treatment effects obtained are stronger than the ones we display in the main text. In particular, when the outlier is included in the estimation sample, we obtain a twofold increase of the treatment effect. We therefore emphasize that our final sample was selected in the most defensive way.

and awareness of their electricity usage and load profile.[11] The feedback concerns the household's own electricity usage (what Fischer, 2008, defines as "historic" comparison) and that of households of similar size from the control group ("normative" comparison). More precisely, as the sample letter reproduced in Figure C.1 in Appendix C shows, the feedback refers to:

(i) The household's electricity usage in the last three months, along with average usage of similar households (top right graph).

(ii) The household's (average) daily load profile, highlighting usage between 11am and 3pm (bottom left graph).

(iii) The proportion of electricity used between 11am and 3pm and its evolution over the last three months (bottom right graph). For the last month, the proportion for similar households is also provided as a benchmark.

Moreover, the text of the letter emphasizes the importance of shifting load toward the sunlight hours, as this would reduce imbalances between solar energy production and electricity consumption. Every month, a different tip on how to shift electricity usage is also provided.[12]

In order to facilitate participants' understanding and mitigate potential boomerang effects (Schultz et al., 2007), we complement the letters with injunctive messages under the form of happy, neutral, or sad faces. The first faces are based on the monthly evolution of the proportion of electricity used between 11am and 3pm by the household. The second faces compare the proportion of the household against that of similar households. Green happy faces signal desired outcomes, that is, a relatively high proportion of electricity consumed between 11am and 3pm, yellow neutral faces signal not

[11] During the first month of the intervention (January 2014), letters were sent every week in order to draw households' attention and to ensure that the letters would not go unnoticed. Also, we note that households became aware of their participation in treatment 1 only when they received their first feedback letter, around January 7. No pre-information letter was sent to this group.

[12] For instance, "program the washing machine so that it starts at 11am", "start the dishwasher right after the lunch", "switch off electronic devices completely at night (no standby)", or "replace old light bulbs by more efficient ones".

good nor bad outcomes, and red sad faces signal undesired outcomes. We use a scale going from three happy faces to three sad faces.[13]

Similar interventions, based on information feedback containing descriptive and injunctive norms related to past electricity usage, have been implemented in previous contributions (e.g., Allcott, 2011; Allcott and Rogers, 2014). Again, their objective was to induce electricity conservation, while we here target load shifting.

3.2 Treatment 2: Financial incentives

Treatment 2 is implemented as a contest, in which cash prizes are awarded to the 15 top-ranked households on a monthly basis: Ranks 1-5, 6-10, 11-15 ascribe respectively CHF 50, 30, and 10.[14,15] Due to practical, political and legal aspects, the provider involved in the project decided not to implement any differentiated tariffs. This is the main reason why this treatment is designed as a contest. The objective of this intervention is to investigate whether monetary rewards without any precise information related to electricity usage can induce households to shift their load profile.

In December 2013, a letter was sent to the selected households to inform them about the starting date of the contest (January 1, 2014). The letter also explained intuitively what should be done to be well-ranked in the contest: maximize the proportion of electricity consumed during the period from 11am

[13]Appendix D provides detailed information about the thresholds defining the type and number of faces.

[14]For administrative reasons related to project funding, prizes were set before knowing the number of participants. Finally, 22 households participated in the contest, so that the probability of winning was extremely high. The results for treatment 2 are based on 21 households, one of the households being identified as an outlier and excluded from the analysis (see above).

[15]According to data from the Swiss Earnings Structure Survey 2014 (conducted by the Swiss Federal Statistical Office), monthly gross median wages in the region where the experiment took place are around CHF 6,000. Compared to wages, the monthly prizes distributed in our intervention are thus relatively small as they roughly correspond to the wage received for one hour of work. Nevertheless, we note that CHF 50 is sufficient to have dinner in a decent restaurant. Moreover, it turns out the average monthly electricity bill is around CHF 50 per household. Hence, on average, for households ranked 1-5, it is like if electricity had been free during one month. Overall, the cash prizes distributed in the contest are thus small in absolute terms, but certainly not negligible.

to 3pm, while keeping total consumption unaltered. The exact criteria used to establish the ranking (Appendix E) were provided in a webpage, but it was deemed unnecessary or even counterproductive to include all technical details in the information letter. In fact, knowing the precise rules would not confer any specific advantage, and the only reasons for making them available to the participants were transparency and fairness concerns.

From January to December 2014, at the end of every month, households were sent personalized mails indicating their position in the ranking (a sample letter is provided in Figure C.2 in Appendix C). The monthly letters also recalled the contest goal and provided tips on how to shift load. However, nothing concerning electricity usage was indicated. For those ranked in the top 15, prizes were included in the form of banknotes in the envelopes. Participants could check the detailed and complete ranking on `www.unine.ch/flexirank`, using a personalized password.[16]

We have to acknowledge that the design of our experiment makes rewards uncertain, unrelated to amounts of electricity used, and only indirectly linked to household's actions. In fact, a household increasing its proportion of electricity used between 11am and 3pm could remain unrewarded if others fared better. Conversely, a household not having changed anything could end up rewarded if others did worse. Our results thus cannot be used to infer about the price-elasticity of electricity demand, and the external validity of this treatment is difficult to assess.

4 Data and descriptive statistics

4.1 Data

Households' electricity usage comes from the electricity provider *Groupe e*. It was recorded in 15-minute intervals by smart meters, but consumers had no in-home display or any other device showing their real-time electricity

[16]This webpage now contains the monthly rankings for all the intervention period and can be accessed using the generic password "MoveItCernier". It turns out that only a handful of participants actually accessed this webpage; all of them visited only once and for most this visit took place shortly after receiving the first monthly letter.

use.[17] All feedback provided to households was thus indirect (Bernstein and Collins, 2014; Darby, 2006) and came exclusively from the letters we sent during the experiment.

The observation period ran from October 2013 to March 2015. The period from October to December 2013 has been used to monitor households' "standard" electricity usage and we define it as the before period. The treatment period expands from January to December 2014. The period between January and March 2015, defined as the after period, is used to investigate households' behavior once the treatment was over.

Data on weather (temperature and radiation) were gathered from *MeteoSwiss*, the Swiss Meteorological Institute. The data are available at the hour level and collected in a meteorological station located in Chaumont, which is 6 km away from Cernier and exactly at the same elevation. Local holiday periods were retrieved from the official calendars of the Canton of Neuchâtel.[18]

Our final dataset contains a total of 33,914 household-day observations: 62 households followed over 18 months (i.e., 547 days).[19] When using hourly data, the dataset contains 813,812 household-hour observations.[20]

[17] In fact, some households had access to an online platform showing their (almost) real-time consumption, a service offered by the provider. These households are excluded from our analysis. Some results obtained for these households are available in Perret et al. (2015).

[18] See `www.ne.ch/themes/travail/Pages/jours-feries.aspx` for the official holidays and `www.ne.ch/themes/enseignement-formation/Pages/calendrier-scolaire.aspx` for the school calendar.

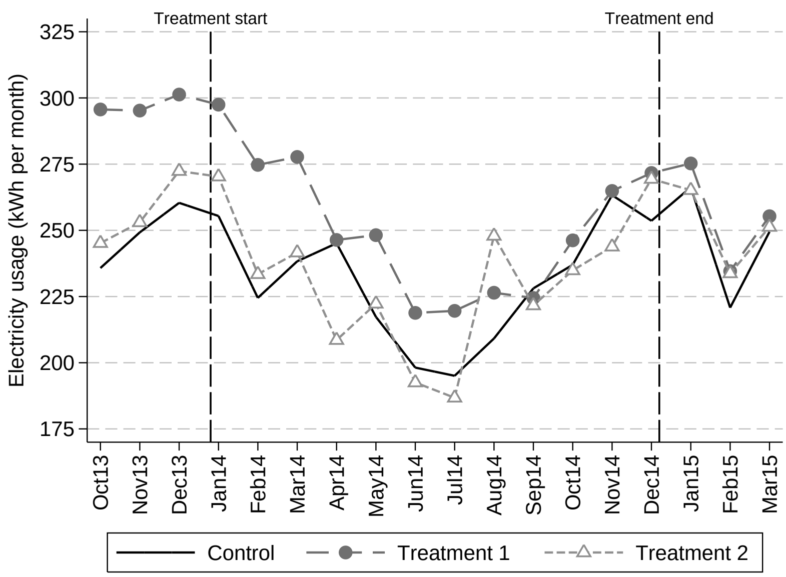
[19] There were in fact 8 daily observations (concerning two households), for which electricity usage was missing because of technical issues encountered with the smart meters. In order to avoid an unbalanced panel because of so few missing values, which would prevent us from using some estimation techniques, we imputed these observations by using the predicted values of a regression of the daily proportion of electricity used between 11am and 3pm on fixed effects for households, months, and days of the week. Alternative ways of imputing missing values, such as simply copying the values observed for the same household and the corresponding day one week earlier or later, do not affect the results. As robustness checks, the estimations were also conducted without imputing the missing values (and using estimation techniques compatible with unbalanced panels). The results (available on request) are similar.

[20] We imputed 1,001 hourly missing values using a similar strategy as for daily observations: We predicted them from a regression of hourly electricity usage on fixed effects for households, months, days of the week, and hours of the day. The findings are unaltered if we exclude the observations with missing values.

Figure 1 shows the monthly electricity usage for the control and treatment groups. Average annual consumption is around 2,850 kWh per household in our sample. This value is slightly larger than the figures reported by Degen et al. (2013) for their sample (2,300 kWh) and the city of Zurich (2,600 kWh). Nevertheless, it only represents one fourth of the electricity used by an average US household, which is larger than 11,000 kWh (see Allcott and Rogers, 2014). It is also twice smaller than the values reported by Ito et al. (2015) concerning Japan. By international standards, electricity used by Swiss households is thus low.

Before the beginning of the intervention, average household consumption in treatment 1 was above the corresponding series of the other two groups.[21] Since the beginning of the treatment, we observe a relative decrease for treatment 1, such that the series come closer during the experiment. Electricity used by the control and treatment 2 groups evolves quite similarly. At the end of the observation period, the spread between the three groups is very narrow.

Figure 1: Monthly electricity usage per household

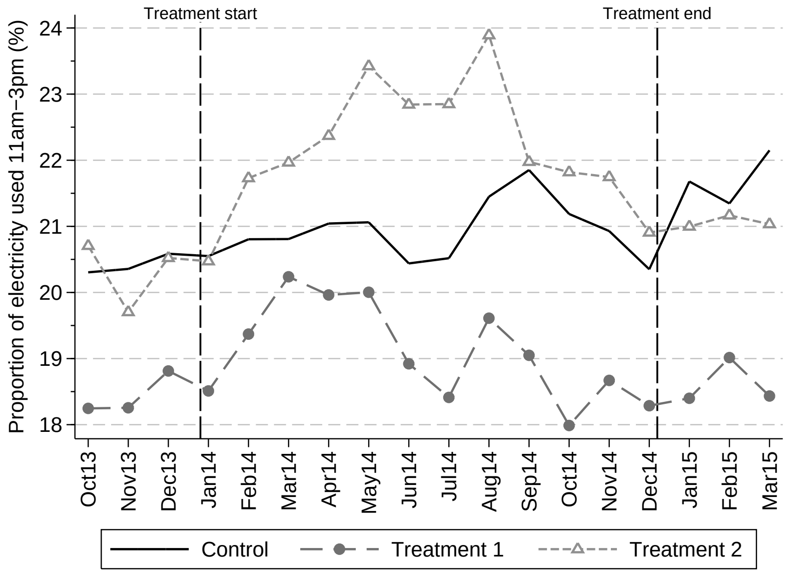


[21]Even though average consumption may seem largely different between groups, the difference is not significant as documented in Table B.1.

Figure 1 moreover highlights the seasonal pattern of electricity usage. When the weather is warm and sunny, people naturally tend to stay outside longer, eat lighter and colder, use less artificial lighting, and wear lighter clothes. These behaviors lead to lower consumption in the period of the year when weather conditions are good. Note also that air conditioning, which could involve significant electricity usage during the warm season, is virtually non-existent in Swiss households because of tight regulation (see Winkler et al., 2014). In 2014, space cooling and ventilation only accounted for 2.6% of final energy consumption whereas heating accounted for 28.9% (SFOE, 2015).

Figure 2 displays the proportion of electricity used between 11am and 3pm by group. While the control group displays relatively little variation over time, there is a clear increase at the beginning of the intervention period for both treatment groups. Their proportion reaches a maximum after some months of treatment and it thereafter decreases, gradually returning toward its pre-experiment level.

Figure 2: Proportion of electricity used between 11am and 3pm



We observe that households in treatment 2 performed particularly well (i.e., increased their proportion of electricity used between 11am and 3pm) from April to August. Two alternative explanations (not mutually exclusive) might support this outcome. First, the duration of the intervention might influence its effectiveness: Households gradually changed their habits at the beginning of the treatment, but they stopped their efforts after some months. The reversal observed after August would then be explained by some loss of enthusiasm or weariness caused by the intervention duration. Second, because of the seasonal effects described above, it might be easier to shift load when the weather cooperates.

Because our intervention lasted 12 months and not more, it is complicated to assert which of these two explanations is correct. One should either conduct a similar experiment over a longer period, or repeat the same experiment launching cohorts at different periods of the year to identify whether the pattern is aligned with calendar year or with intervention duration.

In their field experiment on a sample of Swedish households, Bartusch et al. (2011) make very similar observations: After the introduction of peak (i.e., weekdays between 7am and 7pm) and off-peak pricing, they obtain substantial shifts in electricity consumption from peak to off-peak periods during the summer seasons but almost no change in the winter seasons. Because their intervention spans two complete years, their findings support the explanation according to which the ability to shift load is related to the period of the year rather than to the duration of the intervention.

In Figure 2, we finally observe that the average proportion 11am-3pm of the control group becomes larger than that of the two treatment groups in the after period. This observation suggests that the intervention did not form long-run habits in the households.

4.2 Possible strategies

In order to achieve increases in the proportion of electricity used between 11am and 3pm, households could follow different strategies:

(i) Increase usage during the period 11am-3pm.

(ii) Decrease usage outside the period 11am-3pm.

(iii) (Ideally) do both (i) and (ii), i.e., actually shift load toward the solar energy production hours.

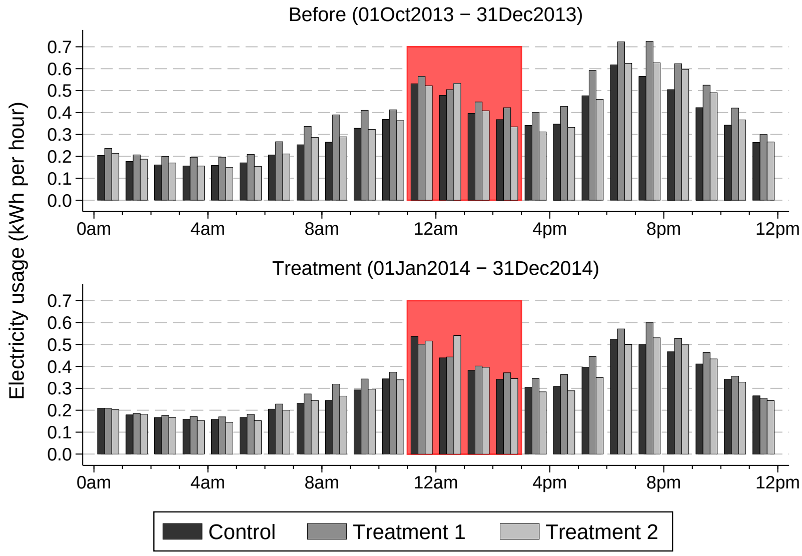
In the letters sent to treated households, we tried to prevent strategy (i) by highlighting that increases in overall consumption were to avoid as much as possible and providing tips on how to shift consumption. Yet, we had no means to ascertain that the households would correctly interpret the letters.[22] Figure 3 shows the daily average load profile for each group, separating before and treatment periods. Panel A displays electricity usage for every hour, while panel B collapses the hours into ranges that are relevant to our analysis. Panel B is less detailed but allows to spot more easily how electricity usage evolved between before and treatment periods for each group. If our intervention was effective, we should observe relative increases between 11am and 3pm and/or relative decreases outside this period for the treatment groups compared to the control group.

Figure 3 first shows that electricity usage is lower in the treatment period than in the before period. This is primarily explained by seasonality: The before period runs from October to December (i.e., high-consumption months) while the treatment period spans an entire year. Second, relative to the control group, treatment 1 households decrease usage in every hour of the day. While their electricity usage was larger than that of the control group in the before period, the gap is much smaller in the treatment period. This finding suggests that the content of the letters was probably not interpreted as intended. What households might have understood from

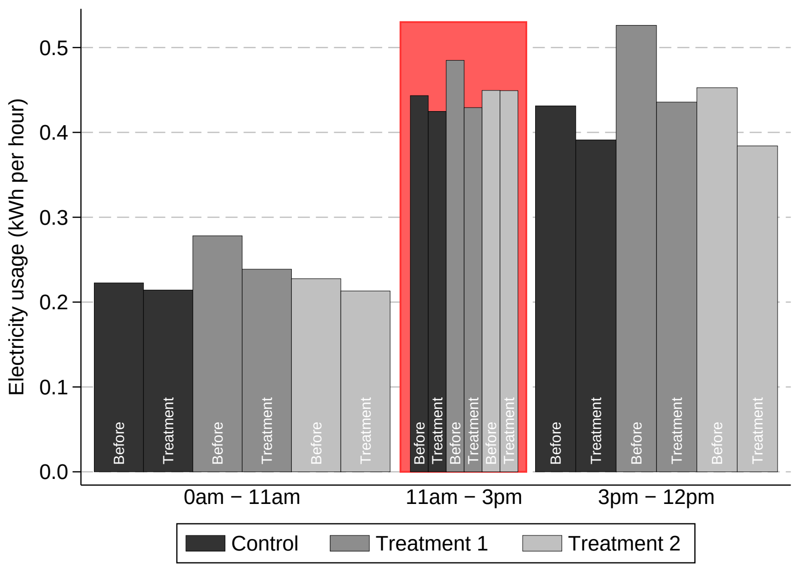
[22]In July 2015, we conducted a follow-up survey of participating households in order to get feedback about their feelings about the project. Only a handful of households responded, but the answers collected offer some interesting anecdotal evidence. In most cases, respondents claim that the letters were easy to understand. However, regardless of that previous answer, several fail in correctly stating the goal of the experiment, mentioning electricity conservation instead of load shifting. Furthermore, feedback received from participants over the course of the experiment also points in that direction: Several of them asked to be excluded from the experiment because they considered our letters as too general and not useful for "decreasing electricity usage".

Figure 3: Daily load profiles

A. Hours



B. Hour ranges



Note: Red shaded areas indicate the “solar energy production hours” (11am-3pm). In panel A, electricity usage is displayed for every hour, separating before (top) and treatment (bottom) periods. In panel B, the width of the bins is proportional to the number of hours in each range. The area of the bins thus represents electricity usage in each hour range.

receiving frequent feedback is that they should conserve electricity, which is the message traditionally transmitted by actors of the energy sector. For treatment 2, however, we observe a different pattern: Relative to the control group, electricity usage decreased substantially in the evening and to a lesser extent in the morning, while it slightly increased between 11am and 3pm. Treatment 2 households thus seem to have succeeded in shifting electricity usage toward the solar energy production hours.

5 Econometric strategy and results

5.1 Treatment effect on proportion 11am-3pm

To assess the impact of information feedback (treatment 1) and financial incentives (treatment 2), we use the following difference-in-differences specification:

$$\text{prop}_{igt} = \sum_{g \in \{1,2\}} \beta_g T_{gt} + \delta_t + \alpha_g + \varepsilon_{igt} \tag{1}$$

where prop_{igt} is the proportion of electricity used between 11am and 3pm by household i of group g in day t, and ε_{igt} is an error term.[23]

The variables of interest are the treatment group indicators, T_{gt}, which take the value 1 if household i is in treatment group g (= 1, 2) and day t is in the treatment period, and 0 otherwise. The time fixed effect δ_t is set equal to 1 during the treatment period, that is, for every day of the year 2014, and 0 otherwise. A group fixed effects α_g is included for each of the two treatment groups. In alternative specifications, we include household fixed effects and additional controls (or day fixed effects).

Since the dependent variable is a proportion, the coefficients in (1) must be interpreted as changes in percentage points associated with a change in

[23] Our dependent variable being a proportion, it is bounded between 0 and 100%. In practice, however, censoring does not appear to be an issue: In our sample, from October 2013 to March 2015, a proportion of 0% is observed for 3 observations out of 33,914. On the upper side of the distribution, 94% is the highest value observed, the second highest value being 75%. The 1[st] quartile is at 13% while the 3[rd] quartile is at 26%. Therefore, censoring does not appear to be a serious practical issue and we find it useless to consider more sophisticated econometric models for so few observations.

the independent variables. Also, the time unit is the day, since the proportion of electricity used over 11am-3pm can obviously not be computed at a higher frequency.

Note that our dataset is a panel with few individuals (62 households), very few clusters (3 groups), and a very large number of time periods (547 days). This setting renders serial correlation and heteroskedasticity very likely.[24] We therefore need to carefully choose the estimation technique in order to appropriately quantify the uncertainty surrounding our estimates. More generally, there are some serious warnings in the literature about potential failures in difference-in-differences estimations. In an influential paper, Bertrand et al. (2004) show that because of serial correlation, conventional standard errors may be severely underestimated. They recommend the use of generalized least squares (GLS) for the estimation of difference-in-differences models. Brewer et al. (2013) and Cameron and Miller (2015) argue in the same direction and show that feasible generalized least squares (FGLS) offer substantial efficiency gains compared to OLS, that is, FGLS estimations combined with robust inference technique can increase statistical power considerably while maintaining correct test size. Beck and Katz (1995) nevertheless demonstrate that FGLS may underestimate the standard errors in the context of longitudinal data, and they advocate using panel corrected standard errors (PCSE) instead. For all these reasons, we will report results obtained using both FGLS and PCSE techniques. In all cases, we allow for cross-sectional correlation and serial correlation with an AR(1) coefficient common to all households.

Table 1 reports the results obtained using various specifications of equation (1). Using FGLS, column (1) reports the basic difference-in-differences regression, without either fixed effects or any control variables. Column (2) contains household fixed effects and additional controls for outdoor temper-

[24]The presence of serial correlation has been assessed using Wooldridge's (2002) test, which is implemented in Stata by Drukker (2003). The presence of heteroskedasticity was tested by estimating our models by GLS once using a homoskedastic error structure and once using a heteroskedastic one. The former model being nested in the latter, a likelihood ratio test then constitutes a test for heteroskedasticity at the panel-level. Homoskedasticity is clearly rejected at any conventional significance level in our data.

ature (in and out of the period 11am-3pm), holiday periods, and weekend days. Finally, column (3) includes household fixed effects and day fixed effects. Columns (4) to (6) repeat similar estimations using PCSE instead of FGLS.

In all estimations, the two treatment effect coefficients display positive signs, implying that treated households tended to react as expected. However, the coefficients for the two treatments are of considerably different magnitude. For the information feedback, the impact is small (0.1-0.5 percentage point) and at best weakly significant. In contrast, the effect of financial incentives is larger (0.9-1.4 percentage point) and always highly significant. Households who were administered this treatment have thus increased their share of electricity used between 11am and 3pm from 20% to around 21%, which corresponds to a 5% increase of this share.

As already noted, the cash prizes distributed to households in the contest were generous: Ranging from CHF 10 to 50, they were in fact as large as the electricity bills for some households in some months. In the end, the cost of this intervention appears high compared to the amount of load that was shifted. Yet, one might wonder if the reaction of households would have been different with lower cash prizes and/or more participants in the

Table 1: Treatment effects: proportion of electricity used between 11am and 3pm

	FGLS AR1			PCSE AR1		
	(1)	(2)	(3)	(4)	(5)	(6)
Information feedback	0.0041	0.0048^{*}	0.0051^{*}	0.0014	0.0014	0.0014
	(0.0032)	(0.0025)	(0.0026)	(0.0043)	(0.0034)	(0.0034)
Financial incentives	0.0126^{***}	0.0085^{***}	0.0085^{***}	0.0133^{***}	0.0135^{***}	0.0135^{***}
	(0.0033)	(0.0026)	(0.0026)	(0.0045)	(0.0035)	(0.0035)
Household FE	NO	YES	YES	NO	YES	YES
Additional controls	NO	YES	NO	NO	YES	NO
Day FE	NO	NO	YES	NO	NO	YES
# Observations	28,334	28,334	28,334	28,334	28,334	28,334
# Households	62	62	62	62	62	62
R^2				0.0086	0.2246	0.2428

Notes:
- $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.
- Standard errors clustered at the household level in parentheses.
- Before period = 01Oct2013-31Dec2013, Treatment period = 01Jan2014-31Dec2014.
- Additional controls: outdoor temperature (in and out of the period 11am-3pm), holiday periods, weekend days.

contest. Considering that the cash prizes received by one household were uncertain and only indirectly linked to this household's actions, one possible conjecture is indeed that lower expected gains might not necessarily trigger weaker reactions by the participants, thereby reducing the cost of shifting one kWh. Our results must therefore simply be interpreted as evidence that financial incentives have the potential to induce behavioral changes and load shifting, but we cannot provide any estimation of the households' sensitivity to a given subsidy or a given price variation.

5.2 Treatment effect on electricity usage

The analysis offered in previous section about the proportion of electricity used between 11am and 3pm investigates whether households did alter their load profile or not. It does however not unravel the strategy followed to reach this target. Therefore, it is of interest to evaluate how electricity used in different periods of the day has been affected by the intervention. To this end, we separate the day into three periods (before 11am, from 11am to 3pm, and after 3pm) on which the treatments are expected to exert different effects:[25]

$$\ln(kwh_{ight}) = \sum_{g\in\{1,2\}} \left(\sum_{\substack{h^\star\in\{0\text{am}-11\text{am},\\ 11\text{am}-3\text{pm},\\ 3\text{pm}-12\text{pm}\}}} \beta_{gh^\star} T_{gh^\star t} \right) + \delta_t + \alpha_g + \lambda_{h^\star} + \varepsilon_{ight} \quad (2)$$

where kwh_{ight} is electricity usage by household i of group g in hour h of day t. We transform this dependent variable using logarithms, so that the

[25] We consider these three time periods because they coincide with the indications provided to the treated households, who were simply requested to shift load toward the 11am-3pm period. Estimating distinct treatment effects for each hour of the day yields similar but less robust results (available on request).

coefficient of a dummy variable approximately indicates its percentage effect on electricity usage.[26] For this analysis, the time unit is the hour.[27]

The variables $T_{gh^{\star}t}$ take the value 1 if household i is in treatment group g (= 1, 2), hour h falls in hour range $h^{\star}$ and day t is in the treatment period, and 0 otherwise. This equation also encompasses an hour-range fixed effect $\lambda_{h^{\star}}$ for each period of the day (0am-11am, 11am-3pm, 3pm-12pm). If households did actually shift load toward the solar energy production hours in response to our intervention, we should observe for treatment groups a decrease in electricity usage in the morning (0am-11am) and/or the evening (3pm-12pm) combined with an increase in the period from 11am to 3pm.

Results are reported in Table 2. For the information feedback, in all periods of the day and all estimations, we obtain negative and (most often) significant coefficients. We thus obtain electricity conservation regardless of the hour of the day. This is not the expected outcome, but it might be explained if households interpreted the feedback "superficially". As admitted above, information contained in the letters was somewhat abundant and sophisticated. Households might therefore have understood that something related to electricity was going on, and they adopted electricity-saving behaviors without distinction between periods of the day. Such findings can also be related to the Hawthorne effect (see Schwartz et al., 2013). In the context of energy, attention alone might change behavior because people are already aware of the issue and they (think they) know how to respond. It is thus possible that our feedback simply increased salience of electricity usage, which triggered the observed reaction.

We have moreover conducted analyzes to investigate potential heterogeneous effects across households by interacting the treatment dummies with selected characteristics (results available on request). Interestingly, one ro-

[26]The precise percentage effect on electricity usage can be computed as $100 \cdot \{e^{\beta} - 1\}$ (see Halvorsen and Palmquist, 1980).

[27]In principle, the best strategy would be to exploit the most disaggregated data available (i.e., 15-minute intervals) like in some other studies (e.g., Jessoe and Rapson, 2014, or Ito et al., 2015). However, as stated above, electricity usage is comparatively low in our sample of Swiss households, and zeros recorded in high-frequency data are very common. Electricity usage in 15-minute intervals is thus a very asymmetric variable and is hardly transformable using logarithms, so that we refrain from analyzing it as such.

bust finding we obtain is that households with a high level of education have achieved a larger shift in their consumption. These households have increased their proportion of electricity used between 11am and 3pm by around 3 to 4 percentage points more than others, a reaction three times stronger than that of the average household. Such a finding supports the hypothesis that the information provided was complicated: Households with a higher level of education interpreted the letters more correctly and reacted more as requested.

It is worth mentioning that the reduction of consumption induced by information feedback appears substantial in some hour ranges (up to 7% for 11am-3pm according to FGLS), but not implausible and in line with the findings of recent studies. In their literature review, Buchanan et al. (2014)

Table 2: Treatment effects: electricity usage

	FGLS AR1			PCSE AR1		
	(1)	(2)	(3)	(4)	(5)	(6)
Information feedback						
0am - 11am	-0.0352^{***}	-0.0357^{***}	-0.0337^{***}	-0.0063	-0.0081	-0.0072
	(0.0122)	(0.0078)	(0.0077)	(0.0143)	(0.0095)	(0.0093)
11am - 3pm	-0.0773^{***}	-0.0717^{***}	-0.0703^{***}	-0.0398^{***}	-0.0475^{***}	-0.0485^{***}
	(0.0127)	(0.0086)	(0.0085)	(0.0153)	(0.0110)	(0.0108)
3pm - 12pm	-0.0239^{*}	-0.0199^{**}	-0.0187^{**}	-0.0195	-0.0148	-0.0157
	(0.0123)	(0.0080)	(0.0079)	(0.0145)	(0.0098)	(0.0096)
Financial incentives						
0am - 11am	0.0222	0.0207^{**}	0.0256^{***}	0.0159	0.0130	0.0167
	(0.0138)	(0.0090)	(0.0088)	(0.0173)	(0.0112)	(0.0110)
11am - 3pm	0.0585^{***}	0.0584^{***}	0.0626^{***}	0.0469^{**}	0.0476^{***}	0.0496^{***}
	(0.0144)	(0.0099)	(0.0097)	(0.0183)	(0.0128)	(0.0126)
3pm - 12pm	-0.0032	-0.0103	-0.0065	-0.0285	-0.0332^{***}	-0.0312^{***}
	(0.0140)	(0.0091)	(0.0090)	(0.0175)	(0.0115)	(0.0113)
Household FE	NO	YES	YES	NO	YES	YES
Additional controls	NO	YES	NO	NO	YES	NO
Day FE	NO	NO	YES	NO	NO	YES
# Observations	679,954	673,816	679,954	679,954	673,816	679,954
# Households	62	62	62	62	62	62
R^2				0.0194	0.2328	0.2353

Notes:
- $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Robust standard errors clustered at household level in parentheses.
- Before period = 01Oct2013-31Dec2013, Treatment period = 01Jan2014-31Dec2014.
- Additional controls: outdoor temperature, holiday periods, weekend days.

indeed observe that energy savings caused by information feedback on energy consumption typically fall in the region of 5-20%.[28]

Note also that the results we obtain here are not in contradiction with our previous findings about the proportion of electricity used between 11am to 3pm, which (if anything) slightly increased. Even though we observe a stronger decrease for the hours between 11am and 3pm than for other periods, the former is only four hours long. The hourly decrease is more modest during the rest of the day, but the greater number of hours leads to a large absolute decrease of consumption, henceforth yielding an increase in the proportion 11am-3pm.

Financial incentives induced different reactions. Even though coefficients vary somewhat from one specification to another, we consistently obtain negative coefficients for the evening period (3pm-12pm). Moreover, coefficients for the 11am-3pm period are positive and highly significant. Hence, it turns out households exposed to financial incentives did actually shift electricity use from the evening period (3pm-12pm) toward the solar energy production hours (11am-3pm). The two effects combined led to the increase of the proportion of electricity used between 11am and 3pm already discussed. These findings are also consistent with the load profiles displayed in Figure 3. More detailed estimations with distinct treatment effects for each hour of the day (available on request) reveal that households shifted load mainly from the period 3pm-7pm toward 12am-1pm.

5.3 Evolution of the treatment effect

As already documented in Figures 1 and 2, electricity consumption is characterized by seasonality: This feature may have driven households' reactions during the intervention. The literature on energy conservation (e.g., Allcott and Rogers, 2014; Lynham et al., 2016) moreover shows that treatment effects tend to decline over time. To shed more light on this issue, we

[28]For instance, Bartusch et al. (2011) find a decline from 11 to 14% in total electricity consumption in the first two years following a switch to time-of-use pricing. Also, Ito et al. (2015) observe that economic incentives have an impact on electricity consumption reduction between 14 and 17%, depending on the level of the peak price.

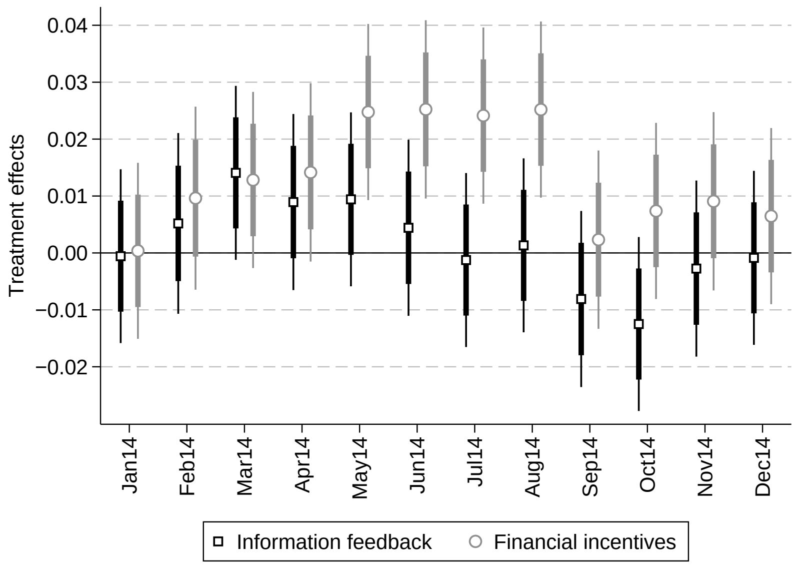
analyze how the treatment effects unfold during our intervention by disaggregating the treatment period into months:

$$\text{prop}_{igt} = \sum_{g\in\{1,2\}} \left(\sum_{m=\text{Jan14}}^{\text{Dec14}} \beta_{gm} T_{gm} \right) + \delta_m + \alpha_g + \varepsilon_{igt} \tag{3}$$

where T_{gm} takes the value 1 if household i is in treatment group g (= 1, 2) and the month is m, and 0 otherwise. Month fixed effects δ_m and group fixed effects α_g are also included. This equation allows to estimate monthly treatment effects, by comparing each month of the treatment period against the before period (October-December 2013). The treatment coefficients (β_{gm}) are presented graphically in Figure 4.

We observe it takes some months for the treatment effects to unfold, the coefficients being close to zero in January and increasing during the first three

Figure 4: Evolution of the treatment effect



Notes: This Figure shows the β_{gm} coefficients obtained using (3). Thick (thin) whiskers indicate 90% (99%) confidence intervals. Estimation technique is panel corrected standard errors. Household and month fixed effects are included.

months of intervention. For both treatments, the effect becomes statistically significant in March.

For the information feedback, the maximum effect is observed after three months already, and the coefficients thereafter decrease and return to zero in July. Near the end of the treatment period, the point estimates even become negative (and slightly significant in October), which is in contradiction with what was requested in the intervention.

The effect of the financial incentives, on the other hand, rises steadily during the first five months, stabilizes during the summer, and declines afterward. The coefficients estimated from May to August are large, all of them being around 0.025 while all others are below 0.015. This evolution support the idea that the treatment effect is related to seasonal effects. Increasing the proportion of consumption between 11am and 3pm appears to be easier during the summer months thanks to the different factors mentioned earlier. During this period of the year, most activities take place outside, meals are lighter so that less cooking is done, and the light requirements are lower.

5.4 Habit formation

In previous section, we find that the treatment effects evolve non-linearly during the intervention. An additional interesting question is then whether our intervention induced habit formation, that is, changes in households' behavior that would continue even when treatments are withdrawn. In order to explore potential habit formation, we collected electricity usage during three months after the treatments were terminated, from January to December 2015. Households did not receive any direction or information at all during this period.

We assess lasting effects by estimating equations (1) and (2) using data from October-December 2013 (before period) and January-March 2015 (after period). If households' habits were permanently shaped by our intervention, the after coefficients should be similar to those obtained for the treatment period (Tables 1 and 2). As shown in Table 3, this is however not the case for the proportion of electricity used between 11am and 3pm. The coefficients obtained for the after period are negative and sometimes even significant, which implies the proportion 11am-3pm was not impacted in the long run by our intervention. While households reacted during the treatment (at least those exposed to financial incentives), their proportion returned to its pre-experiment level after the end of the treatment or even became smaller than before the experiment.

However, when conducting a similar exercise on electricity usage (Table 4), we obtain largely different results. Households included in the financial incentives treatment display no difference in usage before and after the intervention. Hence, these households appear to have returned to their pre-experiment behavior in every respect once the treatment was withdrawn. On the contrary, households assigned to the information feedback intervention show a strong decrease in their electricity usage, not only during, but also after the intervention ended.

In the literature, only few studies investigate a period after withdrawing the treatment, and results are mixed. Ito et al. (2015) find evidence of habit formation (i.e., consumption decrease in the long run) for their economic incentives group, but not for their moral suasion group. This is the opposite to what we obtain here. Allcott and Rogers (2014) show that the treatment effects decay faster if the intervention is short (two to ten weeks), but the

Table 3: Habit formation: proportion of electricity used between 11am and 3pm

	FGLS AR1			PCSE AR1		
	(1)	(2)	(3)	(4)	(5)	(6)
Information feedback (after)	-0.0099^{***}	-0.0102^{***}	-0.0108^{***}	-0.0114^{**}	-0.0116^{***}	-0.0116^{***}
	(0.0030)	(0.0025)	(0.0027)	(0.0054)	(0.0043)	(0.0043)
Financial incentives (after)	-0.0035	-0.0054^{**}	-0.0055^{**}	-0.0058	-0.0058	-0.0058
	(0.0034)	(0.0026)	(0.0026)	(0.0052)	(0.0041)	(0.0041)
Household FE	NO	YES	YES	NO	YES	YES
Additional controls	NO	YES	NO	NO	YES	NO
Day FE	NO	NO	YES	NO	NO	YES
# Observations	11,284	11,284	11,284	11,284	11,284	11,284
# Households	62	62	62	62	62	62
R^2				0.0105	0.2133	0.2236

Notes:
- $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at the household level in parentheses.
- Before period = 01Oct2013-31Dec2013, After period = 01Jan2015-31Mar2015.
- Additional controls: outdoor temperature (in and out of the period 11am-3pm), holiday periods, weekend days.

Table 4: Habit formation: electricity usage

	FGLS AR1			PCSE AR1		
	(1)	(2)	(3)	(4)	(5)	(6)
Information feedback						
0am - 11am	-0.0970^{***}	-0.1045^{***}	-0.0972^{***}	-0.1271^{***}	-0.1555^{***}	-0.1284^{***}
	(0.0159)	(0.0097)	(0.0102)	(0.0196)	(0.0167)	(0.0136)
11am - 3pm	-0.1304^{***}	-0.1418^{***}	-0.1336^{***}	-0.1371^{***}	-0.1810^{***}	-0.1540^{***}
	(0.0174)	(0.0119)	(0.0123)	(0.0226)	(0.0205)	(0.0181)
3pm - 12pm	-0.0626^{***}	-0.0728^{***}	-0.0645^{***}	-0.1237^{***}	-0.1417^{***}	-0.1143^{***}
	(0.0163)	(0.0102)	(0.0106)	(0.0204)	(0.0175)	(0.0146)
Financial incentives						
0am - 11am	0.0456^{**}	0.0308^{***}	0.0422^{***}	0.0288	-0.0062	0.0239
	(0.0181)	(0.0117)	(0.0119)	(0.0232)	(0.0190)	(0.0157)
11am - 3pm	0.0010	-0.0097	0.0017	0.0255	-0.0017	0.0283
	(0.0199)	(0.0142)	(0.0144)	(0.0265)	(0.0233)	(0.0206)
3pm - 12pm	0.0105	-0.0028	0.0086	-0.0255	-0.0508^{**}	-0.0205
	(0.0186)	(0.0122)	(0.0124)	(0.0240)	(0.0199)	(0.0168)
Household FE	NO	YES	YES	NO	YES	YES
Additional controls	NO	YES	NO	NO	YES	NO
Day FE	NO	NO	YES	NO	NO	YES
# Observations	270,754	266,414	270,754	270,754	266,414	270,754
# Households	62	62	62	62	62	62
R^2				0.0213	0.2425	0.2433

Notes:
- $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at the household level in parentheses.
- Before period = 01Oct2013-31Dec2013, After period = 01Jan2015-31Mar2015.
- Additional controls: outdoor temperature, holiday periods, weekend days.

effects become more persistent as the intervention continues. Lynham et al. (2016) obtain substantial electricity conservation during one month even in households where in-home displays have been removed.

Overall, it appears that shaping households' behavior in a sustainable way is difficult, and habit formation might depend on the objective and the treatment. In order to affect electricity usage in the long run, permanent measures (or at least longer than one year) should probably be implemented. In (relatively) short interventions like ours, even though longer than most of the other experiments presented in the literature, it seems unlikely that households undertake "structural changes" such as replacing old devices by more recent ones. Nevertheless, such changes appear as one way of achieving lasting effects, in particular when the objective is to shift electricity usage. If new devices can be programmed to start at a given time, load shifting would become easier. Indeed, several participants in our experiment complained they could not do much about load shifting because they were out all day for work and thus had no grasp on their electricity usage between 11am and 3pm.

6 Policy implications and conclusions

Demand-side management is gaining momentum in the electricity sector. On the one hand, decentralized generation of electricity is growing thanks to photovoltaic systems that are becoming widespread. On the other hand, advanced technologies such as smart meters and in-home displays improved in the past few years and their cost has plummeted. As a consequence, households are increasingly being equipped with such devices that provide real-time information on electricity usage. These innovations transform households from mere electricity consumers to electricity producers, and they contribute in making them more aware of their electricity usage.

One particular issue with solar energy is that its production takes place at times when demand for electricity is not necessarily high. Households' electricity usage indeed peaks in the evening, when solar production energy is at best low. Furthermore, the mismatch is not limited to hours of the day, but it also extends to periods of the year and to geographical regions. A system relying on large amounts of solar (and wind) energy therefore needs storage capacities, backup power, and strong grid interconnections. While such measures are necessary to reduce the gap between supply and demand, they are extremely costly. Moreover, because of the time required for their implementation, they can only be considered as long-run measures. In the meantime, complementary measures targeting the demand for electricity could be devised with the aim of reducing the gap in the short run.

In this paper, using a randomized control trial, we assess the impact of an intervention that seeks to influence residential electricity usage with information feedback and financial incentives. While policies and research experiments generally push for electricity conservation, we here target load shifting: We stimulate households to shift consumption toward the period of the day when the production of solar energy is expected to be the largest.

The final objective is to investigate whether imbalances between solar energy production and residential electricity usage can be reduced via load shifting, that is, changes in the patterns of consumption. The relevance of such a study is enhanced in the current context, where many countries engage in their energy transition and plan to increase the share of renewable energy in they energy mix.[29]

Our first treatment consisted in information feedback. Every month, treated households have been provided with detailed information about their electricity usage and their typical daily load profile. As a benchmark and in order to induce social competition, they have also been provided with comparable information from households of similar size. The feedback letters moreover explained the importance of aligning electricity consumption with solar energy production and included tips to achieve this outcome. No other motivation has been given.

The households of this treatment group did not succeed in shifting their load profile. However, they achieved significant electricity conservation, which tends to indicate that participants interpreted our letters superficially and reacted as "traditionally expected" in this field, disregarding more sophisticated information. Interestingly, this finding points to imperfections of the billing system currently used in Switzerland (and many other countries). In fact, the only feedback received by most households on their electricity consumption is a bill every second month that does not even reflect true consumption because it is based on usage forecasts. Actual electricity usage is measured only once a year through traditional meter readings and the balance is then billed or paid back to the consumer. Kempton and Layne's (1994) humorous comparison of the electricity market with a supermarket where prices are not displayed and the shopper receives a bill where only the total amount is indicated hence applies literally.

Two easy and inexpensive changes in the billing system could thus be devised to improve households' knowledge of their electricity usage. First, an increase in the frequency of billing would make electricity usage more salient.

[29]See in particular Switzerland's *Energy Strategy 2050*, Germany's *Energy Turnaround* (*Energiewende*), or United States' *Clean Power Plan.*

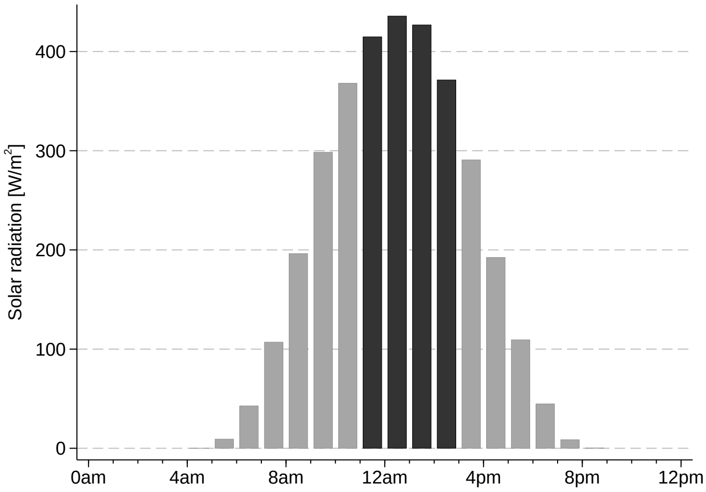
Second, charging the actual amount instead of a forecast would reinforce consumers' awareness of the consequences of their actions. When provided with more frequent information feedback about their actual electricity usage, households would indeed be better informed and will likely become more careful about consumption. This conclusion has already received substantial support in the literature (e.g., Bernstein and Collins, 2014; Vine et al., 2013), but it has only hardly been implemented in practice.

Our second treatment was designed as a contest. Every month, competing households have been ranked according to their proportion of electricity used during the "solar energy production hours" (i.e., from 11am to 3pm) and its evolution, controlling for changes in overall usage. The top-ranked households earned cash prizes ranging from CHF 10 to 50. Participants were informed about the rewards and about what to do to rank well, but they did not receive detailed information concerning their electricity usage. These households did shift some of their electricity usage, raising their proportion of electricity used between 11am and 3pm from 20% to 21% on average. This outcome was achieved by shifting electricity usage from the evening period (3pm-12pm) toward the middle of the day (11am-3pm).

Even though we acknowledge the external validity of our intervention is difficult to assess, both because of the small number of households involved and because of the design of the contest, such a finding suggests that financial incentives matter for residential electricity demand. In this context, time-of-use pricing with attractive tariffs when solar production is abundant could prove an effective tool to induce changes in behavior and load shifting. The cost of such a measure would obviously be much lower than the contest we organized. Further research in this direction is however required to establish whether incentives would then be sufficient from the households' perspective.

Appendix A Radiation

Figure A.1: Average solar radiation per hour, 2014



Data source: *MeteoSwiss*. Weather station: Chaumont (Switzerland).

Appendix B Descriptive statistics

Table B.1: Descriptive statistics, by group

	Control		Treatment 1			Treatment 2		
	Before	Treatment	Before	Treatment	Difference	Before	Treatment	Difference
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Bef. C − T1 [SE]	Mean (SD)	Mean (SD)	Bef. C − T2 [SE]
Electricity use (kWh/day)	8.102	7.576	9.736	8.265	−1.621	8.376	7.598	−0.274
	(4.808)	(5.276)	(9.020)	(7.356)	[1.982]	(5.739)	(5.655)	[1.352]
Electricity use 11am-3pm (kWh/day)	1.773	1.699	1.940	1.717	−0.165	1.798	1.797	−0.025
	(1.496)	(1.576)	(2.152)	(1.863)	[0.454]	(1.719)	(1.757)	[0.338]
Proportion 11am-3pm	0.204	0.209	0.184	0.191	0.020	0.203	0.222	0.001
	(0.098)	(0.106)	(0.093)	(0.092)	[0.014]	(0.102)	(0.121)	[0.015]
# Observations	2,024	8,029	1,741	6,935	41	1,932	7,665	43
High education	0.318		0.211		0.108	0.333		−0.015
	(0.477)		(0.419)		[0.141]	(0.483)		[0.146]
Age: < 40	0.318		0.421		−0.103	0.286		0.032
	(0.477)		(0.507)		[0.154]	(0.463)		[0.143]
Age: 40-49	0.364		0.105		0.258*	0.238		0.126
	(0.492)		(0.315)		[0.132]	(0.436)		[0.142]
Age: 50-64	0.091		0.368		−0.278**	0.238		−0.147
	(0.294)		(0.496)		[0.125]	(0.436)		[0.113]
Age: 65+	0.227		0.105		0.122	0.238		−0.011
	(0.429)		(0.315)		[0.119]	(0.436)		[0.132]
Household size	2.545		2.684		−0.139	2.429		0.117
	(1.471)		(1.529)		[0.469]	(1.287)		[0.422]
# Households	22		19		41	21		43

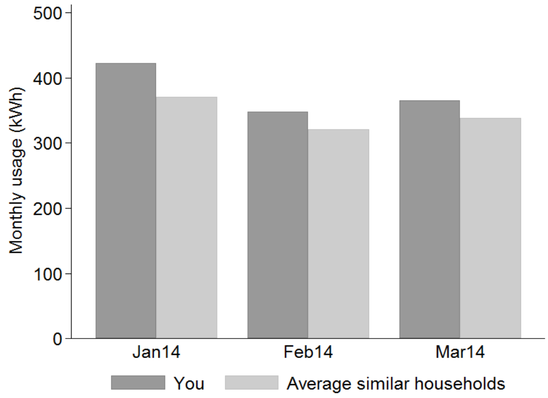
Notes:
- Before period: 01Oct2013-31Dec2013. Treatment period: 01Jan2014-31Dec2014.
- Standard deviations (SD) reported in rounded parentheses.
- Columns entitled "Difference" show the differences in the means (of before period) between the control group and treatment group 1/2 (Bef. C − T1/T2).
- Standard errors of the differences in means [SE] reported in squared brackets. $^{*}p < 0.10$, $^{**}p < 0.5$, $^{***}p < 0.01$.
- Minor discrepancies might arise because of rounding.

Appendix C Sample letters

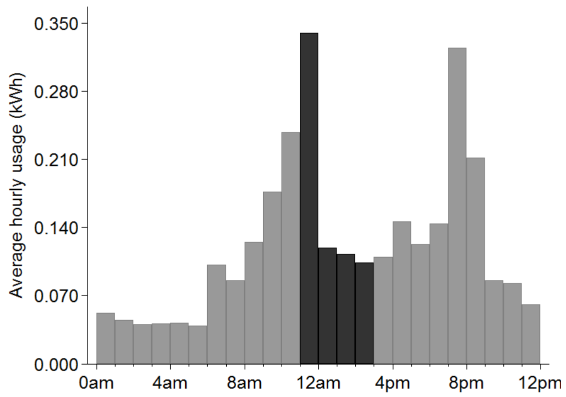
Figure C.1: Sample letter for treatment 1

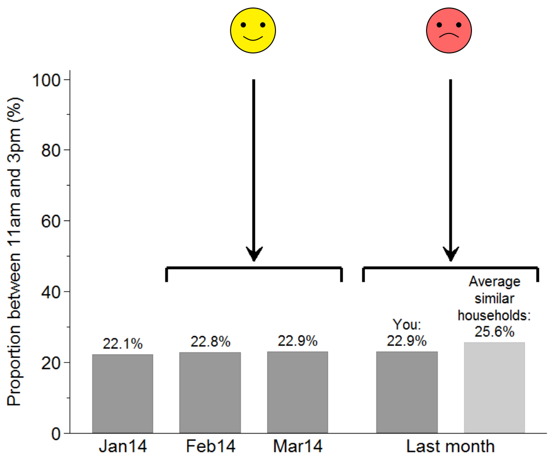
Dear Madam, Dear Sir,

This month, you consumed 365 kWh of electricity. The monthly cost of electricity thus amounts to CHF 72.54 (VAT and communal fees not included). These values are indicative. Legally, only the invoice you will receive from *Groupe E* will be valid.
Compared to last month, your consumption has decreased. Your consumption is currently lower than the average consumption of similar households.



Your hourly consumption profile for the past month is shown below (left graph). Your consumption between 11am and 3pm represents 23% of your total consumption (right graph). Compared to last month, you maintained this proportion. For the past month, your performance is **worse** than the average of similar households. *Shifting* your consumption toward this period of the day, without increasing your total consumption, makes it possible to limit the gap between solar production and electricity consumption. One way of achieving this goal would be, for example, to program your washing machine so that it starts at 11am.





Note: The letters were originally in French.

Figure C.2: Sample letter for treatment 2

Dear Madam, Dear Sir,

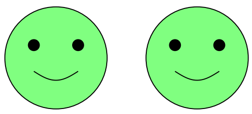
As part of the Flexi project, your household was selected to take part in an experiment which will take place between January and December 2014. Your participation does not involve any cost or commitment, but may allow you to earn cash rewards.

Every month, the 22 households participating in the experiment will be ranked according to the proportion of electricity consumed between 11am and 3pm. Your goal is to maximize the *proportion* of electricity you consume in that time slot. In other words, your goal is to *shift* your consumption toward that time slot, without increasing your total consumption. One way of achieving this goal would be, for example, to iron at the end of the morning or in the early afternoon. Households whose consumption increases artificially will be excluded from the classification.

Each month, the following prizes will be distributed (in cash) :

1$^{\text{st}}$-5$^{\text{th}}$: CHF 50 | 6$^{\text{th}}$-10$^{\text{th}}$: CHF 30 | 11$^{\text{th}}$-15$^{\text{th}}$: CHF 10

Last month, your household ranked **6$^{\text{th}}$** and you thus receive **CHF 30** :



Detailed information on how the ranking is established is available on the project website. The anonymised classification is available at `http://www.unine.ch/flexirank`. Your household is identified by the code `WQ36rb`.

Note: The letters were originally in French.

Appendix D Number of happy or sad faces

Table D.1: Faces for treatment 1

Household's own change	Household compared to others	Faces
$> +5$	$> +10$	🙂 🙂 🙂
$> +2.5$	$> +5$	🙂 🙂
$> +0.5$	$> +1$	🙂
$[-0.5; +0.5]$	$[-1; +1]$	😐
< -0.5	< -1	🙁
< -2.5	< -5	🙁 🙁
< -5	< -10	🙁 🙁 🙁

Table D.2: Faces for treatment 2

Rank	Prize (CHF)	Faces
1-5	50	🙂 🙂 🙂
6-10	30	🙂 🙂
11-15	10	🙂
16-20	—	🙁
21-22	—	🙁 🙁
out	—	🙁 🙁 🙁

Appendix E Contest rules (treatment 2)

Treatment 2 is implemented as a contest among households. This Appendix describes in details the rules that have been used for establishing the ranking every month during the intervention.

Inclusion: Household i is included in the ranking in month t only if the following condition is met:

$$\frac{\Delta kwh_{it}}{kwh_{i,t-1}} \leq \frac{\Delta kwh_{\bullet t}}{kwh_{\bullet,t-1}} + 0.1$$

where kwh_{it} indicates electricity usage in kWh by household i in month t, $kwh_{\bullet t}$ is average electricity usage by the households participating in the contest (in fact a trimmed mean obtained by removing two observations on each side of the electricity usage distribution), and Δ is the first-difference operator.

This criterion is intended to exclude participants who would increase their proportion of electricity used between 11am and 3pm simply by consuming (much) more electricity during these hours without decreasing usage in other periods of the day.

Note that a simpler criterion excluding all households increasing usage above some fixed threshold is not applicable because of the seasonal pattern of electricity usage documented in Figure 1. We therefore use a relative benchmark and tolerate a 10 percentage-point deviation to allow for relatively small variations due, for instance, to vacations.

Because the top 15 households are supposed to receive a reward, we bring back the "least worst" in case fewer households satisfy the inclusion condition in a given month.

Rank 1: For households satisfying the inclusion criterion, a first ranking is established according to the proportion of electricity used between

11am and 3pm:

$$\begin{aligned} R1_{it} &= 1 + \sum_{j=1}^{N} \mathbb{1} \left\{ \frac{kwh_{it}^{11am-3pm}}{kwh_{it}} < \frac{kwh_{jt}^{11am-3pm}}{kwh_{jt}} \right\} \\ &= 1 + \sum_{j=1}^{N} \mathbb{1} \left\{ \text{prop}_{it}^{11am-3pm} < \text{prop}_{jt}^{11am-3pm} \right\} \end{aligned}$$

where $kwh_{it}^{11am-3pm}$ is the electricity used by household i during month t between 11am and 3pm, N is the number of households included in the contest, and $\mathbb{1}\{\cdot\}$ is an indicator function taking the value 1 if the condition in braces is true and 0 otherwise. A larger proportion implies a better (i.e., smaller) rank.

Rank 2: A second ranking, independent from the first, is established based on the monthly changes in the proportion of electricity used between 11am and 3pm:

$$R2_{it} = 1 + \sum_{j=1}^{N} \mathbb{1} \left\{ \Delta\text{prop}_{it}^{11am-3pm} < \Delta\text{prop}_{jt}^{11am-3pm} \right\}$$

Once again, a larger change in the proportion implies a better rank.

Final rank: The final rank is given by the following weighted average:

$$\text{Rank}_{it} = 0.25 \cdot R1_{it} + 0.75 \cdot R2_{it}$$

In case of ties in the final rank, $R2$ takes precedence. The overweight given to $R2$ is intended to reward the change of behavior and give the opportunity to households with a low ranking to improve rapidly. We also expect some sort of asymptotic limit for the proportion of electricity used between 11am and 3pm (which might depend on household structure), so that rewarding the proportion more than its change could have conducted to a very stable ranking month after month.

References

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9–10):1082–1095.

Allcott, H. and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10):3003–3037.

Bartusch, C., Wallin, F., Odlare, M., Vassileva, I., and Wester, L. (2011). Introducing a demand-based electricity distribution tariff in the residential sector: Demand response and customer perception. *Energy Policy*, 39(9):5008–5025.

Beaudin, M., Zareipour, H., Schellenberglabe, A., and Rosehart, W. (2010). Energy storage for mitigating the variability of renewable electricity sources: An updated review. *Energy for Sustainable Development*, 14(4):302–314.

Beck, N. and Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *The American Political Science Review*, 89(3):634–647.

Bernstein, M. and Collins, M. (2014). Saving energy through better information: A new energy paradigm? *Contemporary Economic Policy*, 32(1):219–229.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Brewer, M., Crossley, T. F., and Joyce, R. (2013). Inference with difference-in-differences revisited. IZA Discussion Papers 7742, Institute for the Study of Labor (IZA).

Buchanan, K., Russo, R., and Anderson, B. (2014). Feeding back about eco-feedback: How do consumers use and respond to energy monitors? *Energy Policy*, 73:138–146.

Buchanan, K., Russo, R., and Anderson, B. (2015). The question of energy reduction: The problem(s) with feedback. *Energy Policy*, 77:89–96.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Chen, V. L., Delmas, M. A., Kaiser, W. J., and Locke, S. L. (2015). What can we learn from high-frequency appliance-level energy metering? results from a field experiment. *Energy Policy*, 77:164–175.

Darby, S. (2006). The effectiveness of feedback on energy consumption. a review for DEFRA of the literature on metering, billing and direct displays. Technical report, Environmental Change Institute, University of Oxford.

Degen, K., Efferson, C., Frei, F., Goette, L., and Lalive, R. (2013). Smart Metering, Beratung oder sozialer Vergleich: Was beeinflusst den Elektrizitätsverbrauch? SFOE publication 290850, Swiss Federal Office of Energy.

Delmas, M. A., Fischlein, M., and Asensio, O. I. (2013). Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy*, 61:729–739.

Denholm, P. and Hand, M. (2011). Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. *Energy Policy*, 39(3):1817–1830.

Denholm, P. and Margolis, R. M. (2007). Evaluating the limits of solar photovoltaics (PV) in electric power systems utilizing energy storage and other enabling technologies. *Energy Policy*, 35(9):4424–4433.

Di Cosmo, V., Lyons, S., and Nolan, A. (2014). Estimating the impact of time-of-use pricing on Irish electricity demand. *Energy Journal*, 35(2):117–136.

Dietz, T., Gardner, G. T., Gilligan, J., Stern, P. C., and Vandenbergh, M. P. (2009). Household actions can provide a behavioral wedge to rapidly reduce

US carbon emissions. *Proceedings of the National Academy of Sciences*, 106(44):18452–18456.

Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models. *Stata Journal*, 3(2):168–177.

Ehrhardt-Martinez, K., Donelly, K. A., and Laitner, J. A. (2010). Advanced metering initiatives and residential feedback programs: A meta-review for household electricity-savings opportunities. ACEEE Report E105, American Council for an Energy-Efficient Economy.

Elliott, D. (2016). A balancing act for renewables. *Nature Energy*, 1.

Faruqui, A., Sergici, S., and Sharif, A. (2010). The impact of informational feedback on energy consumption – A survey of the experimental evidence. *Energy*, 35(4):1598–1608.

Fischer, C. (2008). Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency*, 1(1):79–104.

Halvorsen, R. and Palmquist, R. (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review*, 70(3):474–475.

Ida, T., Murakami, K., and Tanaka, M. (2015). Electricity demand response in Japan: Experimental evidence from a residential photovoltaic generation system. Discussion papers e-15-006, Graduate School of Economics Project Center, Kyoto University.

IEA (2016). *2015 snapshot of global photovoltaic markets*. Report IEA PVPS T1-29:2016, International Energy Agency.

Ito, K., Ida, T., and Tanaka, M. (2015). The persistence of moral suasion and economic incentives: Field experimental evidence from energy demand. NBER Working Paper 20910, National Bureau of Economic Research.

Jessoe, K. and Rapson, D. (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review*, 104(4):1417–1438.

Kempton, W. and Layne, L. L. (1994). The consumer's energy analysis environment. *Energy Policy*, 22(10):857–866.

Kondziella, H. and Bruckner, T. (2016). Flexibility requirements of renewable energy based electricity systems – a review of research results and methodologies. *Renewable and Sustainable Energy Reviews*, 53:10–22.

Lynham, J., Nitta, K., Saijo, T., and Tarui, N. (2016). Why does real-time information reduce energy consumption? *Energy Economics*, 54:173–181.

McCoy, D. and Lyons, S. (2017). Unintended outcomes of electricity smart-metering: Trading-off consumption and investment behaviour. *Energy Efficiency*, 10(2):299–318.

Pellerano, J. A., Price, M. K., Puller, S. L., and Sánchez, G. E. (2017). Do extrinsic incentives undermine social norms? Evidence from a field experiment in energy conservation. *Environmental and Resource Economics*, 63(3):413–428.

Perret, L., Fahrni, J., Wyrsch, N., Riesen, Y., Puddu, S., Weber, S., and Pacheco, D. (2015). FLEXI: Determining the flexibilization potential of the electricity demand. SFOE publication 291032, Swiss Federal Office of Energy.

Pina, A., Silva, C., and Ferrão, P. (2012). The impact of demand side management strategies in the penetration of renewable electricity. *Energy*, 41(1):128–137.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5):429–434.

Schwartz, D., Fischhoff, B., Krishnamurti, T., and Sowell, F. (2013). The Hawthorne effect and energy awareness. *Proceedings of the National Academy of Sciences*, 110(38):15242–15246.

SFOE (2015). *Analyse des schweizerischen Energieverbrauchs 2000-2014 nach Verwendungszwecken*. Swiss Federal Office of Energy.

Steinke, F., Wolfrum, P., and Hoffmann, C. (2013). Grid vs. storage in a 100% renewable Europe. *Renewable Energy*, 50:826–832.

Vine, D., Buys, L., and Morris, P. (2013). The effectiveness of energy feedback for conservation and peak demand: A literature review. *Open Journal of Energy Efficiency*, 2(1):7–15.

Weber, S. (2010). bacon: An effective way to detect outliers in multivariate data using Stata (and Mata). *Stata Journal*, 10(3):331–338.

Winkler, R., Almer, C., Bader, C. C., Gonseth, C., Laurent-Lucchetti, J., Thalmann, P., and Vielle, M. (2014). Energy consumption of buildings - direct impacts of a warming climate and rebound effects. In CH2014-Impacts, editor, *Toward quantitative scenarios of climate change impacts in Switzerland*, pages 99–105. OCCR, FOEN, MeteoSwiss, C2SM, Agroscope, and ProClim, Bern, Switzerland.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT Press, Cambridge (MA).