Energy Policy 107 (2017) 225–233

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Persistence of the effects of providing feedback alongside smart metering devices on household electricity demand

Joachim Schleich[a,b,c,*], Corinne Faure[a], Marian Klobasa[b]

[a] Grenoble Ecole de Management, 12, rue Pierre Sémard, 38000 Grenoble, France
[b] Fraunhofer Institute for Systems and Innovation Research, Breslauer Straße 48, 76139 Karlsruhe, Germany
[c] Virginia Polytechnic Institute and State University, Hutcheson Hall, Blacksburg, VA 24061, USA

ARTICLE INFO

Keywords:
Smart metering
Feedback
Persistence
Household electricity consumption

ABSTRACT

Using large-sample high temporal resolution data from a smart metering field trial, we econometrically estimate the effects of providing feedback in addition to smart metering devices. We compare consumption levels and patterns between a pilot group that received feedback in addition to smart metering devices and a control group with only smart metering devices. We investigate, in particular, the persistence of the effects and whether the effects differ between periods of high and low household occupancy, i.e. between morning and evening periods, and between weekdays and weekend days. The findings show that feedback is effective, leading to about 5% electricity consumption reduction that is persistent over an eleven month period. Furthermore, our results show that this reduction affects both low and high occupancy periods, suggesting that feedback is associated with rather permanent changes in habitual behavior and/or investments in energy-efficient technologies.

1. Introduction

The roll-out of electricity smart metering devices is well under way in the European Union (EU), with a recent official report indicating that most EU member states are on track to achieve the target of 80% penetration by 2020 (European Commission, 2016). In recent years, many field studies have been conducted to assess the impact of introducing in-house displays on electricity consumption; most of these pilot studies have compared electricity consumption of households with or without in-house displays (or before-after the introduction of in-house displays). Providing households with information on their electricity consumption has mostly been found to be effective in reducing electricity demand (e.g. Wilhite and Ling, 1995; Matsukawa, 2004; Darby, 2006; EPRI 2009; Faruqui et al., 2010, Ehrhardt-Martinez et al., 2010; Gans et al., 2013; Gleerup et al., 2010; McKerracher and Torriti, 2013; Schleich et al., 2013; Houde et al., 2013). However, recent papers stress that providing in-house displays alone may not be sufficient. Tedenvall and Mundaca (2016), for instance, report a less than 2% reduction in electricity consumption over a long-term field study in Sweden; similarly, results from a meta-analysis (Delmas et al., 2013) indicate savings of less than 2% for "robust" studies (those including control groups or accounting for control factors). Such results lead authors to doubt the effectiveness of in-house displays per se and to recommend associating in-house displays with other mechanisms: Buchanan et al. (2015), for instance, recommend adding functions that increase user engagement with in-house displays; Tedenvall and Mundaca (2016) also recommend adding additional measures (especially awareness measures) with in-house displays. Such recommendations are consistent with Abrahamse et al.'s (2005) finding that feedback is particularly effective when it is provided together with information on energy-efficiency measures. These papers (and the fact that smart metering deployment is already well advanced) point to the need to investigate the impact of the presence of feedback *along* in-house displays. The present paper therefore examines households equipped with in-house displays and compares those that receive feedback to those that do not.

Two issues are of interest when focusing on the effects of feedback on in-house display users. First, if feedback affects electricity consumption, do the effects persist or disappear over time? Second, does feedback lead to changes in usage profile (for instance, reduction of the base load)? Households may respond to feedback on their electricity use in two manners: by changing habitual behaviors (such as turning off lights, reducing device usage, or switching off electronic devices rather than putting them in stand-by mode), or by investing in energy efficient technologies (such as purchasing electricity-saving appliances or power strips with on/off switches). While behavioral changes may only have a transitory effect on electricity use if households return to their long-practiced habits after a certain time (e.g. Allcott, 2011), the

* Corresponding author.
E-mail address: joachim.schleich@grenoble-em.com (J. Schleich).

http://dx.doi.org/10.1016/j.enpol.2017.05.002
Received 14 December 2016; Received in revised form 25 April 2017; Accepted 1 May 2017

effects of investments should be more persistent. To test persistence, it is necessary to follow consumption over a long period of time.

Usage profile is also susceptible to change based on feedback. Changes in behaviors or investments in air conditioning or electronic media devices are expected to primarily shape electricity consumption during periods of peak household activity (e.g. Torriti, 2012), that is, in the mornings and evenings on weekdays, and on weekends. In contrast, investments in energy efficient refrigerators or freezers should reduce the base load, and hence affect the entire electricity load profile of a household, and be particularly visible in off-activity times (at night and during the day on weekdays). Investigating usage profile changes therefore requires detailed consumption information at the household level and a systematic distinction between different hours and different days of the week.

So far, due to data availability limitations, few studies have explored whether feedback on electricity use resulted in persistent electricity savings or in changes in the usage profile. Relying on data from a field experiment with employees from Google in California, Houde et al. (2013) conclude that real-time feedback delivered via information and information and communication (ICT) technologies had only transitory effects; initial electricity savings disappeared after four weeks. They also find larger reductions during the morning and evening time intervals, i.e. during periods of high household occupancy. Thus, the findings by Houde et al. (2013) suggest that feedback on electricity use mainly leads to transitory changes in household habitual behaviors.

Our paper adds to sparse empirical evidence on the long term effects of feedback on household electricity use and on a user's consumption profile. We employ large-sample high temporal resolution data from a 2010 smart metering field trial in the Austrian city of Linz to econometrically estimate the effects of providing feedback with in-house displays for each hour of the day (distinguishing between weekdays and weekend days). Following household consumption patterns over an eleven-month period, we analyze whether the effects are transitory or persistent, and whether the effects differ between hours of the day and especially between periods of high and low household occupancy, and between weekdays and weekend days. The findings allow us to explore whether the observed effects may be ascribed to changes in habitual behavior or rather to investments in energy efficient technologies.

The paper is organized as follows. The methodology Section 2 describes the field trial, econometric methods, data, and variables. Section 3 presents and discusses the results of the econometric analyses. The concluding Section 4 summarizes the main findings and derives policy implications.

2. Methodology

2.1. Field trial

The field trial in the city of Linz, Austria, originally involved a sample of more than 2000 households for whom the old electricity meters had been replaced by smart meters in 2009. These households were randomly assigned to two groups: the pilot group, in addition to the smart meters, received feedback on electricity consumption, whereas the control group had only the smart meter (no feedback). After correcting for households that either relocated during the field phase or encountered insoluble technical problems, data was available for 1525 households, 775 pilot group households and 750 control group households.

Pilot group households chose how they preferred to receive feedback on their electricity use: either via access to a web-portal or via written information by post. By accessing the web portal, households could see their electricity consumption patterns and electricity costs. Several types of charts and tables allowed for comparison of energy consumption and costs on a yearly (month-by-month comparison), twice-yearly (week-by-week comparison), monthly (day-by-day comparison), or daily basis (hour-by-hour comparison). The web portal also provided information on intermittent loads and (estimated) base loads (i.e. refrigerators and freezers) as shares of the total household electricity consumption. All data was available to the web portal users with a delay of, at most, one day. In comparison, the written feedback was sent to households once a month and consisted of two pages including color-printed information on daily, weekly, and monthly household electricity consumption. Both web portal and written feedback also provided practical information on how to save electricity.

The electricity consumption of households in both pilot control groups was recorded between December 2009 and November 2010. Since the written feedback could only be sent out after the first month of the trial, possible impacts of that feedback could only be expected from the second month onwards, i.e. for the period of January to November 2010. The smart metering systems provided hourly consumption data, which was read at the end of each day by a remote system. In addition to this detailed information on electricity use, information about household appliance stock and socio-demographic characteristics was available for both groups from computer-assisted telephone interviews. For more details on the design of the field trial and on the types of feedback provided, see Schleich et al. (2013). Unlike Torriti et al. (2015), for example, data on actual time use was not available. Finally, upon completion of the field phase, an additional survey asked participants to evaluate the quality of the feedback provided and whether they had implemented any energy-efficiency measures since the beginning of the field trial.

2.2. Statistical models

We employ several econometric models to (i) explore the average effects of providing feedback on electricity use for the entire duration of the field study, (ii) test for persistence of effects over the eleven-month period, and (iii) to test for differences in feedback effects across the 24 hours of the day on weekdays and weekend days.

To analyze the average effect of feedback on household electricity demand for the duration of the field study, we first estimate the following reduced form electricity demand equation

$$electricity_{it} = c + \delta\, feedback_i + \beta\ Z_i + \sum_{m=1}^{11} M_m + \sum_{h=1}^{24} H_h + \varepsilon_{it}, \qquad (1)$$

where $electricity_{it}$ is the (log of) electricity use by household i at hour t of a day (t=1–24) and c is a constant term. Electricity is calculated as the average electricity consumption at hour t per month. We thereby distinguish between weekdays (Monday to Friday) and weekend days (Saturday and Sunday), by estimating Eq. (1) separately for weekdays and for weekend days. *Feedback* is a dummy variable indicating that household i received feedback on electricity consumption.[1] Since we use the logarithm of electricity consumption as the dependent variable, δ measures the average percentage difference in hourly electricity consumption between households that received feedback and those that did not. Z_i is a vector of household socio-economic and appliance stock characteristics (which do not vary over time).

Variables reflecting household characteristics include income, level of education, and number of household members. The dummy variable *income* takes on the value of 1 if the household has a household disposable monthly income (including transfer payments) above 2500 €. Similarly, the dummy variable *education* equals 1 if the survey

[1] Note that we do not distinguish whether households received feedback via access to the web portal or via postal mail. We tested for such differences and did not find any statistical differences; as a consequence, we report both feedback types together. As an aside, in Darby's (2006) classification of direct versus indirect feedback, web portal information is somewhat ambiguous because it entails characteristics of direct feedback (immediate and interactive) but also of indirect feedback (it contains information that is processed by the utility company). Our results seem to suggest that web portal information has the same effects as postal mail (indirect feedback).

Table 1
Descriptive statistics.

		Pilot (N =599)				Control (N =283)			
Variable	Unit	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Income	dummy	0.38	0.49	0	1	0.36	0.48	0	1
Education	dummy	0.51	0.50	0	1	0.55	0.50	0	1
Working	dummy	0.68	0.47	0	1	0.74	0.44	0	1
Hhzize	number	2.49	1.17	0	6	2.52	1.16	0	6
Floorsize	m^2	107.47	49.63	25	538	100.25	37.75	36	380
Fridge	number	1.25	0.51	0	4	1.20	0.44	1	3
Freezer	number	0.75	0.55	0	3	0.75	0.58	0	2
Dryer	Number	0.40	0.49	0	1	0.34	0.47	0	1
Boiler	number	0.38	0.57	0	2	0.44	0.56	0	2
TV	number	0.87	0.83	0	5	0.81	0.74	0	3
Computer	number	1.29	0.90	0	5	1.24	0.88	0	5
Appliances	number	8.86	3.09	3	29	8.46	2.84	3	25

respondent has received at least 10 years of education. The dummy variable *working* reflects whether the respondent was working. *Hhsize* stands for the number of household members. *Floorsize* assesses the dwelling size (in square meters). Count variables are used to indicate the number of the following appliances: refrigerator, freezer, dryer, boiler, TV, and computer. For parsimony, we use only one count variable for the number of other *appliances* in the household such as dishwashers, air conditioners, espresso machines, microwaves or gaming consoles.

We also include monthly dummies M_m to capture variations in electricity demand across months, e.g. due to weather conditions. M_1 corresponds to January and M_{11} to November of 2010. Similarly, we include hourly dummies H_h to capture variation in electricity demand across the hours of a day, e.g. due to variations in household occupancy and activity. Finally, ε_{it} stands for the idiosyncratic error term.

Table 1 provides descriptive statistics of the variables used for the final sample in our econometric analyses. The number of observations in the final sample is somewhat lower than in the original sample. We lost observations because of missing values (e.g. 16% of households failed to report income), because we trimmed annual electricity consumption to the range of 700 kWh to 8000 kWh to exclude "unreasonable" consumption levels (resulting in a loss of about 3% of observations), and because of technical problems to record electricity use for some households (leading to a loss of about 12% of observations). The numbers in Table 1 suggest that for most variables, the pilot and control group are well matched. A noticeable difference exists for *floorsize*. We further address this issue in Section 3.4, where we report results from employing matching estimators.

Our second model allows for the feedback effect to change over the duration of the pilot phase. The second model therefore includes interactions of the feedback and monthly dummies, $feedbackM_m$:

$$electricity_{it} = c + \sum_{m=1}^{11} \rho_m feedbackM_{mi} + \beta\ Z_i + \sum_{m=1}^{11} M_m + \sum_{h=1}^{24} H_h + \varepsilon_{it}, \tag{2}$$

where ρ_m measures the percentage difference in hourly electricity consumption between households that received feedback and those that did not in month m.

The third model no longer assumes that the feedback effect is the same for each hour of the day, and therefore includes interactions of the smart and hourly dummies, $smartH_h$:

$$electricity_{it} = c + \sum_{h=1}^{24} \gamma_h feedbackH_{hi} + \beta\ Z_i + \sum_{m=1}^{11} M_m + \sum_{h=1}^{24} H_h + \varepsilon_{it}, \tag{3}$$

where γ_h measures the percentage difference in hourly electricity consumption between households that received feedback and those that did not in hour h.

Since data on electricity consumption prior to the pilot phase is not available, a before-after estimator cannot be employed. Instead, identification of the feedback effects rests on the (untestable) assumption that our regression analyses sufficiently control for differences in characteristics between the pilot and the control group such that the outcome which would result without feedback is the same. Invoking this "conditional independence" or "unconfoundedness" (see Imbens, 2004; Angrist and Pischke, 2009) assumption allows any difference between the pilot and the control group to be attributed to the feedback.

3. Results

All models were estimated via the GLS panel random-effects estimator. To account for serial correlation, the standard errors were clustered at the household level.

3.1. Average feedback effect

The results of estimating the average feedback effect for weekdays and weekend days appear in Table 2. Households receiving feedback used 5.5% less electricity on weekdays (5.1% less on weekend days) than households that did not receive feedback.[2] These effects are statistically significant at the 5% (10%) level, but the difference in feedback effects between weekdays and weekend days is not statistically significant.

Our point estimate for the average feedback effects are in the range of those found by Houde et al. (2013) (5.7% for participants in the US, mostly California), or by Schleich et al. (2013) (4.5% for almost the same participants as in this study).[3] Slightly lower effects were estimated by Matsukawa (2004) (1.5% for Japan), Tendenvall and Mindaca (2016) (ca. 2% for Sweden) and Gleerup et al. (2010) (3% for Denmark). In comparison, Gans et al. (2013) found significantly higher feedback effects (11 – 17% for Northern Ireland). Literature surveys of older studies (mostly in the US) report feedback effects of up to 20% (Faruqui et al., 2010; Ehrhardt-Martinez et al., 2010). Based on more recent studies, McKerracher and Torriti (2012) consider feedback effects to be in the range of 3–5%. Comparing findings across studies is challenging, however, since the pilot programs differ in methodology (e.g. participant selection, feedback duration, evaluation technique)

[2] To test for possible differences by feedback type, we split the sample and estimated Eq. (1) separately for pilot group households receiving feedback via postal mail and web access, respectively. For weekdays, the point estimates for feedback by postal mail is 5.0% and for feedback by web access 5.9%. For weekend days, the figures are 5.2% and 4.9%. However, these slight differences are not significant in a statistical sense. Hence, similar to Schleich et al. (2013), we found no indication that the effects differ between both types of feedback.

[3] Because of missing observations on hourly data, the sample in this study involves 184 fewer households than in Schleich et al. (2013).

Table 2
Average feedback effect (*p*-values in parentheses).

	Weekday	Weekend
Feedback	−0.055** (0.031)	−0.051* (0.058)
Income	0.083*** (0.003)	0.097*** (0.001)
Education	−0.071*** (0.006)	−0.082*** (0.002)
Working	−0.099*** (0.002)	−0.015 (0.640)
Hhsize	0.105*** (0.000)	0.089*** (0.000)
Floorsize	0.002*** (0.000)	0.002*** (0.000)
Fridge	0.067** (0.028)	0.064** (0.040)
Freezer	0.119*** (0.000)	0.114*** (0.000)
Dryer	0.124*** (0.000)	0.120*** (0.000)
Boiler	0.080*** (0.000)	0.078*** (0.001)
TV	0.039** (0.015)	0.047*** (0.004)
Computer	0.035** (0.046)	0.034* (0.053)
Appliances	0.021*** (0.000)	0.025*** (0.000)
Month dummies	YES	YES
Hour dummies	YES	YES
R^2 (overall)	0.4172	0.4375
Observations	229,794	229,370
Households	886	886

Note:
*** Significant at $p < 0.01$.
** Significant at $p < 0.05$.
* Significant at $p < 0.1$.

and in the type/technology of feedback provision (e.g. monthly feedback via postal mail versus real time feedback via modern ICT; presence or not of additional information on energy savings measures).

The results presented in Table 2 further suggest that household electricity use is positively related with *income*. High income households are associated with about 8% higher electricity use on weekdays (10% on weekend days) compared to low income households. For weekdays and weekend days, electricity use is also positively related to the number of household members (*hhsize),* the size of the dwelling (*floorsize*), and to all appliances, but negatively related to *education*. Unsurprisingly, *working* is negatively related to electricity use for weekdays, but not for weekend days. Overall the models explain a fairly large share of the overall variation in hourly electricity use (as indicated by the values for R^2); the coefficients exhibit the expected signs, are statistically significant, and take on reasonable values.

3.2. Persistence

Fig. 1a and b provide visual results of estimating Eq. (2) for weekdays and weekend days. Detailed regression results appear in Annex Table A1. The results are qualitatively very similar for weekdays and weekend days and suggest that hourly electricity use is lower for the pilot group households than for the control group households for the entire duration of the pilot program, but it takes about a month before these effects are statistically significant. Most notably, and in contrast to the findings by Houde et al. (2013), Fig. 1 does not suggest that feedback effects fade over time. In fact, feedback effects appear to increase somewhat over the duration of the pilot study. Our findings therefore imply that providing feedback with smart meters is associated with rather permanent changes in habitual behavior and/or investments in energy efficient technologies. This interpretation is also consistent with the findings for month 8 (August 2010). For example, if pilot group households invested in energy efficient refrigerators or freezers, or re-programmed the hot water boiler to be in synch with actual needs (or switched off appliances during vacation time), the feedback effects (in percentage terms) would be particularly large in a month where base consumption is low, i.e. during vacation time. We can speculate on the differences between Houde et al. (2013) and our findings. Higher electricity prices, and the higher income expenditure share for electricity of participants in our study compared to the Google employees in Houde et al. (2013) may provide stronger financial incentives to adjust habitual behavior or investment, and may therefore contribute to explaining the different findings. In addition, most participants in the study by Houde et al. (2013) were recruited from California, where air-conditioning accounts for a substantial share of electricity consumption. Furthermore, all participants in the study by Houde et al. (2013) were volunteers who worked for the same high-tech company; this might have led to a smaller difference between pilot and control groups, as well as to atypical relationships with technological devices. In contrast, the sample used in the present study consisted of representative households that had received the smart meter as part of a normal replacement.

It appears that receiving feedback on electricity use is not sufficient to alter household electricity usage profiles. Without additional measures, receiving real-time feedback as in Houde et al. (2013) may primarily prompt transitory habitual behavioral change related to immediate activities only. In the field trial presented in this paper, pilot group households also received information and advice on energy-efficiency alongside electricity consumption feedback. We speculate that this helped consumers overcome information-related barriers to energy efficiency technology adoption (Abrahamse et al., 2005; Ehrhardt-Martinez et al., 2010) such as lack of information about technologies and cost-savings potential (e.g. Palmer et al., 2012), thus explaining our finding on the persistence of feedback effects. In fact, as in previous studies, our econometric analysis does not allow us to separate the effects of providing consumption feedback from the effects of providing information on electricity-saving measures.

3.3. Feedback effects by time of day

To gain further insights into the differential effects of feedback on energy consumption, we estimated Eq. (3) which allows feedback effects to vary by the hour of the day. Fig. 2a display the findings for weekdays and Fig. 2b for weekend days. Results are shown for the hour seen by the electricity user. As such, changes from daylight saving and standard time are taken into account. Detailed regression results appear in Annex Table A2. Accordingly, the feedback effects appear to be rather constant in percentage terms through the hours of the day for weekdays and (to a slightly lesser extent) for weekend days. In absolute terms, the feedback effects are higher during periods of high electricity use. A noticeable exception seems to be during hour 6, i.e. the time between five and six o′clock in the morning. This "phenomenon" may be explained by that fact that during this time of the day, the hot water generation (including circulation pump etc.) is typically scheduled to start, thus overriding any other effects on electricity use. Overall, the findings on the feedback effects by time of day presented in

Fig. 2a and b corroborate the previous results on the persistence of feedback effects. We observe feedback effects during night and early morning hours; these effects are likely to stem from adjustments in behavior or from investments in energy-efficiency measures which lower electricity use during these hours. Similar to the findings for all appliances discussed above, however, we do not find evidence that pilot group household propensity to invest in new refrigerators or freezers during the pilot phase was higher than for control group households. In addition, since percentage changes during times of high occupancy correspond to higher absolute savings, providing feedback appears to also be associated with changes in habitual behavior and /or with investments in energy efficient technologies, which are used during these times only.

It should be noted that the evidence for feedback effects at different times of the day is rather weak. Fig. 2a, Fig. 2b, and Table A1 imply that the P-values of most hourly feedback effects are just below or above the 10% significance level, arguably due to a lack of statistical power. To allow for a more comprehensive picture, we aggregate the hours of the day into three categories based on intensity of usage and allow feedback effects to vary across those periods. The first category, *feedback_night* (hour 1 to hour 5 and hour 24) focuses on nighttime hours, during which little activity is expected The second (*feedback_day* (hour 10 to hour 18)) and third (*feedback_presence* (hour 6 to hour 9 and hour 19 to hour 23)) categories both focus on consumption

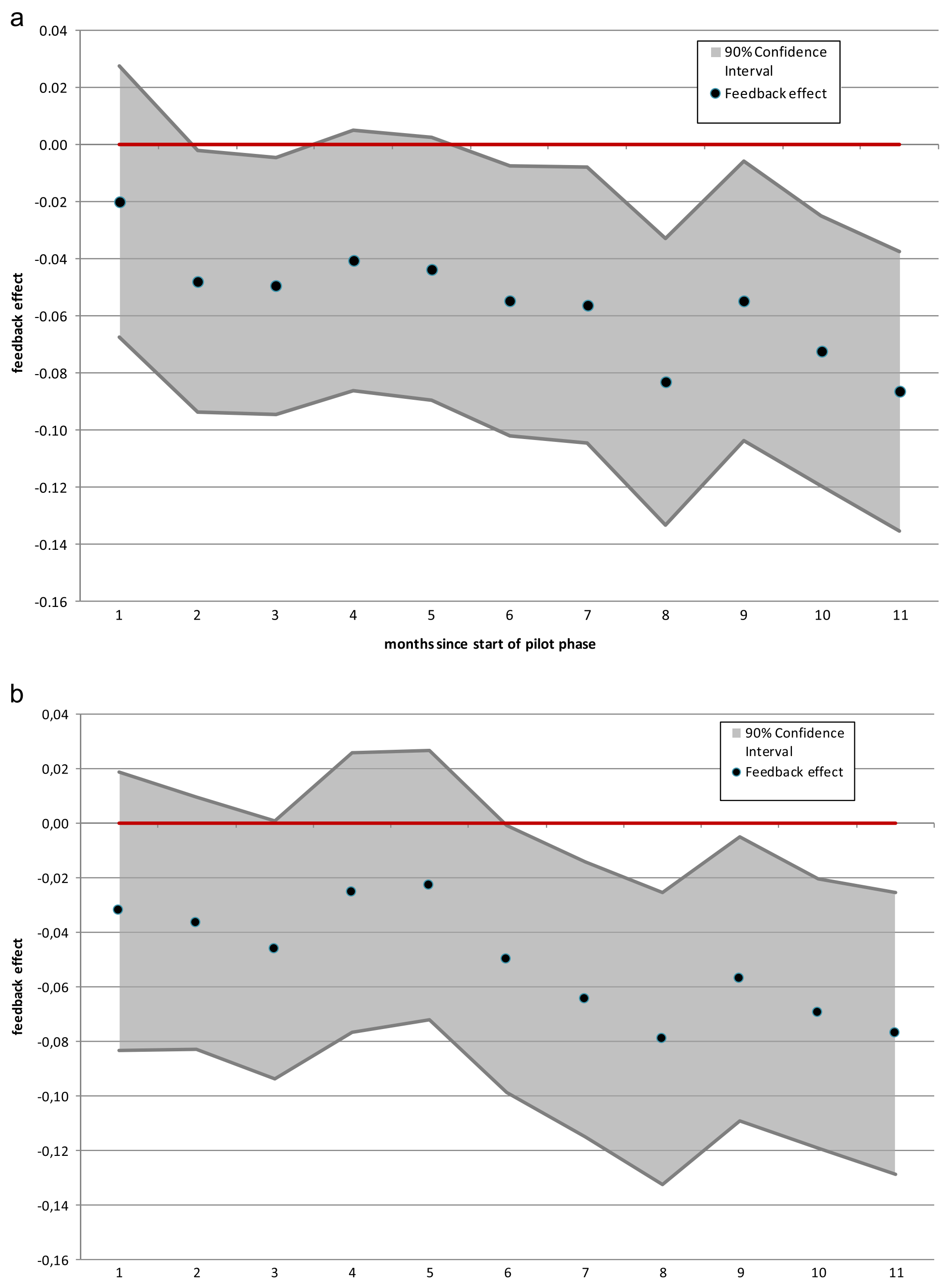

Fig. 1. a: Average feedback effect for different months (weekday). b: Average feedback effect for different months (weekend).

during the day, with a distinction between times of intense electricity use (feedback_presence) when most working participants are at home and times of lower intensity use (feedback_day) when working participants are at work. In particular, *feedback_presence* is expected to capture habitual behavior and technology use during periods of high household occupancy. Clearly, the distinction between times of high and low occupancy may not apply to some respondents (such as retirees or unemployed persons). However, these periods are based on aggregate household consumption patterns and for simplicity are here considered at the aggregate level. We also tested whether households where the respondent was unemployed responded stronger to feedback on electricity consumption than households where the respondent was employed. To do so, we estimated Eq. (2) separately for the sub-samples of employed and unemployed respondents. While we find some evidence that the feedback effects for households with unemployed respondents are stronger than for employed respondents during low occupancy times (between hour 10 and hour 18), these differences were not statistically significant.

The results presented in Table 3 are consistent with the notion that on weekdays and weekend days, the feedback effect (in percentage

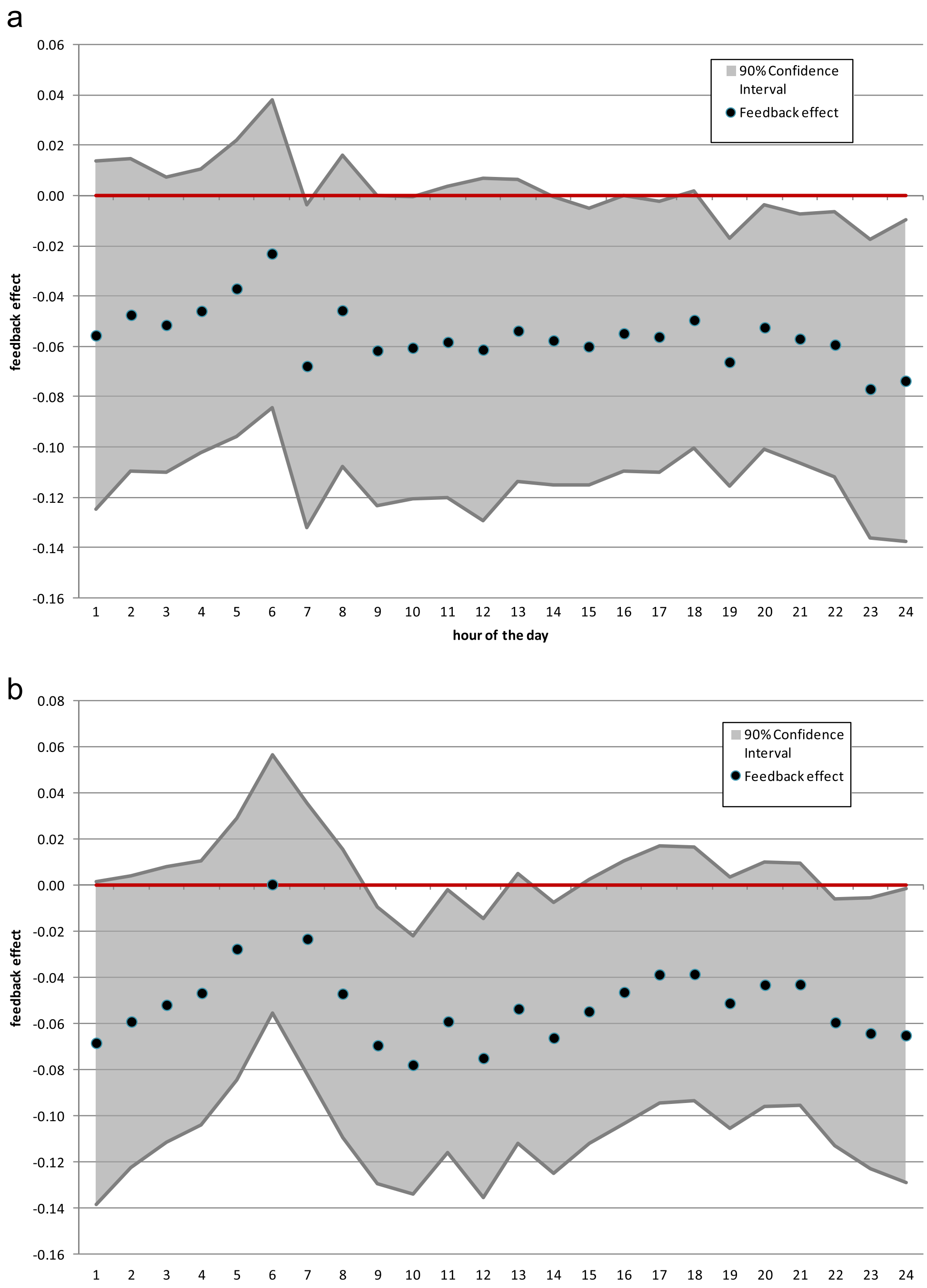

Fig. 2. a: Feedback effect at different time of the day (weekday). b: Feedback effect at different time of the day (weekend).

Table 3
Feedback effect by time period of day (*p*-values in parentheses).

	Weekday	Weekend
Feedback_night	−0.050	−0.043
	(0.102)	(0.186)
Feedback_day	−0.057*	−0.057*
	(0.066)	(0.065)
Feedback_presence	−0.060**	−0.054*
	(0.025)	(0.053)
Household characteristics	YES	YES
Month dummies	YES	YES
Hour dummies	YES	YES
R^2 (overall)	0.4172	0.4375
Observations	229,794	229,370
Households	886	886

Note:
***Significant at p < 0.01.
** Significant at p < 0.05.
* Significant at p < 0.1.

terms) is rather constant across different periods of the day. For weekdays, *feedback_night* is just shy of significance at the 10% level, but becomes statistically significant if hour 6 is dropped from that period.

3.4. Robustness check employing matching estimators

The descriptive statistics in Table 1 suggest that although households were randomly assigned to the pilot and control groups, their covariates differ somewhat in the final sample. For example, the average dwelling in the pilot group is about 7 m^2 larger than in the control group. We therefore employ coarsened exact matching (CEM) (Iacus et al., 2011, 2012) to construct a new sample which balances the distributions of the relevant covariates in the pilot and the control groups via appropriate weighting. The CEM algorithm first transforms the continuous variables into strata, which enables exact matching of pilot and control group observations. Unmatched observations receive a weight of zero. Each matched pilot group household receives a weight of 1. The weight of a matched control group household reflects the relative frequency of pilot to control group households in the control group household's stratum, compared to the relative frequency of the total number of pilot group households to the control group households in the matched sample. Matching estimators do not rely on specific functional forms for the covariates and are therefore less prone to specification bias. In comparison to CEM, the popular propensity score matching estimators often fail to meet the balancing assumption, i.e. treatment and control group members in the post matching subsample exhibit systematic differences in the covariates. In addition, CEM has been also found to outperform propensity score matching and distance-based matching estimators in terms of bias and variance (Iacus et al., 2011, 2012). We use *income_high*, *hhsize* and *floorsize* as matching covariates, which lead to matches of more than 92% of the households in the final pilot group sample for weekdays and more than 96% for weekend days.[4] These weights are then applied to the covariates (and the constant) for the matched subsample using the original (and not the coarsened) set of covariates when estimating Eqs. (1), (2) and (3). Employing CEM leads to almost identical findings as those presented in section 3.1. to 3.3. For example, the point estimate for the average feedback effect is 4.9% for weekdays and 4.7% for weekend days.[5] Thus, our findings do not appear to be affected by the slight differences in the covariates between the pilot and the control group.

4. Conclusions and policy implications

We employ econometric analysis to analyze the effects of feedback by relying on high resolution household electricity consumption data of pilot and control group households equipped with in-house displays from a field trial. This field trial took place in the Austrian city of Linz over a period of eleven months. In particular, we analyze whether the effects are transitory or persistent, and whether they differ over the hours of the day between periods of high and low household occupancy. Thus, we seek evidence on whether feedback effects may be ascribed to changes in habitual behavior or rather to investments in energy efficient technologies. Our results also provide insights for policy making.

First, we found that average electricity consumption in the pilot group was about 5% lower than in the control group. This effect appears to be the same for weekdays and weekend days. Second, and most interestingly, our findings suggest that these feedback effects were persistent over the course of the field trial. Persistence of effects is crucial for policies promoting the diffusion of smart meters to be effective and cost efficient. In this sense, our findings indicate that the roll-out of smart meters to 80% of consumers in EU Member States by 2020, as foreseen by the EU Electricity Directive 2009/72/EC, could benefit from being accompanied by electricity consumption feedback. Third, our findings provide (weak) evidence that feedback not only lowers peak load but also base load, thus alleviating intermittency problems associated with the integration of fluctuating electricity sources such as solar or wind power. In sum, our findings suggest that feedback on energy consumption is likely to have prompted investments in more energy efficient technologies (such as refrigerators or freezers), or led to permanent habitual changes (such as switching off or re-programming appliances) in the pilot group, when compared to the control group. We speculate that these investments and changes may have been supported if not triggered by information on electricity saving measures provided to pilot group households. Future research could attempt to disentangle the effects of providing consumption feedback from the effects of providing information on electricity saving measures. Finally, since the city of Linz is not representative of Austria (nor of the EU), we can only speculate whether the feedback effects found would be lower or higher outside of this sample; future studies could explore to what extent our findings are generalizable to a wider geographical area.

Role of funding source

This research was partly funded by the German Federal Ministry of Education and Research in the socio-economic research funding program "From Knowledge to Action – New Paths towards Sustainable Consumption".

Acknowledgements

The authors gratefully acknowledge the contributions of Sebastian Gölz, Konrad Götz, and Georg Sunderer for carrying out the field study and of Mark Brunner and Yashar Bashirzadeh for their help in preparing the data.

[4] Using more variables as matching covariates reduces the number of matched pilot group households but does not alter the key findings of the paper.

[5] All results not shown to save space are available upon request from the authors.

Annex

See Tables A1 and A2.

Table A1
Average feedback effects for different months (*p*-values in parentheses).

	Weekday	Weekend
FeedbackM1	−0.020 (0.490)	−0.032 (0.299)
FeedbackM2	−0.048* (0.086)	−0.037 (0.191)
FeedbackM3	−0.049* (0.071)	−0.046 (0.108)
FeedbackM4	−0.041 (0.144)	−0.025 (0.417)
FeedbackM5	−0.044 (0.119)	−0.023 (0.450)
FeedbackM6	−0.055* (0.057)	−0.050* (0.095)
FeedbackM6	−0.056* (0.056)	−0.065** (0.035)
FeedbackM8	−0.083*** (0.007)	−0.079** (0.015)
FeedbackM9	−0.055* (0.067)	−0.057* (0.073)
FeedbackM10	−0.072** (0.012)	−0.070** (0.021)
FeedbackM11	−0.086*** (0.004)	−0.077** (0.014)
Household characteristics	YES	YES
Month dummies	YES	YES
Hour dummies	YES	YES
R^2 (overall)	0.4173	0.4376
Observations	229,794	229,370
Households	886	886

Note:
*** Significant at $p < 0.01$.
** Significant at $p < 0.05$.
* Significant at $p < 0.1$.

Table A2
Feedback effect at different time of the day (*p*-values in parentheses).

	Weekday	Weekend
FeedbackH1	−0.056 (0.187)	−0.068 (0.109)
FeedbackH2	−0.047 (0.209)	−0.059 (0.125)
FeedbackH3	−0.051 (0.150)	−0.052 (0.153)
FeedbackH4	−0.046 (0.182)	−0.047 (0.178)
FeedbackH5	−0.037 (0.302)	−0.028 (0.424)

Table A2 *(continued)*

	Weekday	Weekend
FeedbackH6	−0.023 (0.537)	0.000 (0.989)
FeedbackH6	−0.068* (0.082)	−0.023 (0.515)
FeedbackH8	−0.046 (0.226)	−0.047 (0.217)
FeedbackH9	−0.062 (0.100)	−0.069* (0.057)
FeedbackH10	−0.060* (0.097)	−0.078** (0.022)
FeedbackH11	−0.058 (0.122)	−0.059* (0.089)
FeedbackH12	−0.061 (0.139)	−0.075** (0.042)
FeedbackH13	−0.054 (0.142)	−0.054 (0.131)
FeedbackH14	−0.058* (0.098)	−0.066* (0.064)
FeedbackH15	−0.060* (0.072)	−0.055 (0.117)
FeedbackH16	−0.055 (0.100)	−0.046 (0.182)
FeedbackH17	−0.056* (0.088)	−0.039 (0.254)
FeedbackH18	−0.049 (0.112)	−0.039 (0.251)
FeedbackH19	−0.066** (0.027)	−0.051 (0.124)
FeedbackH20	−0.052* (0.077)	−0.043 (0.181)
FeedbackH21	−0.057* (0.060)	−0.043 (0.177)
FeedbackH22	−0.059* (0.065)	−0.059* (0.069)
FeedbackH23	−0.077** (0.033)	−0.064* (0.073)
FeedbackH24	−0.074* (0.059)	−0.065* (0.093)
Household characteristics	YES	YES
Month dummies	YES	YES
Hour dummies	YES	YES
R^2 (overall)	0.4172	0.4376
Observations	229,794	229,370
Households	886	886

Note:
*** Significant at $p < 0.01$.
** Significant at $p < 0.05$.
* Significant at $p < 0.1$.

References

Abrahamse, W., Steg, L., 2005. A review of intervention studies aimed at household energy conservation. J. Environ. Psychol. 25, 273–291.

Allcott, H., 2011. Social norms and energy conservation. J. Public Econ. 95, 1082–1095.

Angrist, J., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, New Jersey.

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

Darby, S., 2006. The Effectiveness of Feedback on Energy Consumption. A Review for DEFRA of the Literature on Metering, Billing and Direct Displays. Environmental Change Institute. University of Oxford, Oxford.

Delmas, M.A., Fischlein, M., 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.

Ehrhardt-Martinez, K., Donnelly, K.A., Laitner, J.P., 2010. Advanced Metering Initiatives and Residential Feedback Programs: a Meta-Review for Household Electricity-Saving Opportunities (Report No. E105). American Council for an Energy-Efficient Economy, Washington, D.C.

Electric Power Research Institute (EPRI), 2009. Residential electricity use feedback: a research synthesis and economic framework (Report 1016844). Electric Power Research Institute, Palo Alto, CA.

European Commission, 2016. Smart Metering Deployment in the European Union. ⟨http://ses.jrc.ec.europa.eu/smart-metering-deployment-european-union⟩.

Faruqui, A., Harris, D., Hledik, R., 2010. Unlocking the $53 billion savings from smart meters in the EU: how increasing the adoption of dynamic tariffs could make or break the EU's smart grid investment. Energy Policy 38, 6222–6231.

Gans, W., Alberini, A., Longo, A., 2013. Smart meter devices and the effect of feedback on residential electricity consumption: evidence from a natural experiment in Northern Ireland. Energy Econ. 36, 729–743.

Gleerup, M., Larsen, A., Leth-Petersen, S., Togeby, M., 2010. The effect of feedback by text message (SMS) and email on household electricity consumption: experimental evidence. Energy J. 31 (3), 111–130.

Houde, S., Todd, A., Sudarshan, A., Flora, J., Armel, K.C., 2013. Real-time feedback and electricity consumption: a field experiment assessing the potential for savings and persistence. Energy J. 34 (1). http://dx.doi.org/10.5547/01956574.34.1.4.

Iacus, S.M., King, G., Porro, G., 2011. Multivariate matching methods that are monotonic imbalance bounding. J. Am. Stat. Assoc. 106 (493), 345–361.

Iacus, S.M., King, G., Porro, G., 2012. Causal inference without balance checking: coarsened exact matching. Political Anal. 20, 1–24.

Imbens, G.W., 2004. Non parametric estimation of treatment effects under exogeneity: a review. Rev. Econ. Stat. 86, 4–29.

Matsukawa, I., 2004. The effects of information on residential demand for electricity. Energy J. 25 (1), 1–17.

McKerracher, C., Torriti, J., 2013. Energy consumption feedback in perspective: integrating Australian data to meta-analyses on in-home display. Energy Effic. 6 (2), 387–405.

Palmer, K., Walls, M., Gordon, H., Gerarden, T., 2012. Assessing the energy-efficiency information gap: results from a survey of home energy auditors. Energy Effic. 6, 271–292.

Schleich, J., Klobasa, M., Gölz, S., Brunner, M., 2013. Effects of feedback on residential electricity demand -Findings from a field trial in Austria. Energy Policy 61, 1097–1106.

Tedenvall, M., Mundaca, L., 2016. Behaviour, context and electricity use: Exploring the effects of real-time feedback in the Swedish residential sector. Paper presented at the 39th IAEE International Conference 'Energy: Expectations and Uncertainty' Bergen, Norway, 19-22 June 2016.

Torriti, J., 2012. Demand side management for the European Supergrid: occupancy variances of European single-person households. Energy Policy 44, 199–206.

Torriti, J., Hanna, R., Anderson, B., Yeboah, G., Druckman, A., 2015. Peak residential electricity demand and social practices: deriving flexibility and greenhouse gas intensities from time use and locational data. Indoor Built Environ. 24 (7), 891–912.

Wilhite, H., Ling, R., 1995. Measured energy savings from a more informative energy bill. Energy Build. 22, 145–155.