How smart are electricity users with ‘Smart Metering’? A Behavioural Economics experiment

Bager, Simon; Mundaca, Luis

Published in:
38th International Association for Energy Economics (IAEE) Conference

2015

Citation for published version (APA):
How Smart are Electricity Users with ‘Smart Metering’? A Behavioural Economics Experiment

Simon Bager* & Luis Mundaca

International Institute for Industrial Environmental Economics at Lund University, Tegnérsplatsen 4, 22100 Lund, Sweden

Abstract
The purpose of this paper is to examine how behavioural biases affect consumers’ response to energy-use information provided through smart meters (SM). We take insights from behavioural economics and carry out two real-life experiments with SMs and electricity users. The experiments were conducted in Copenhagen (Denmark) to identify and assess the potential of two economic behavioural biases, salience and loss aversion. The results of the first experiment (i.e. installation of SM without further intervention) generally aligned with electricity use reductions found in previous research, and indicate that it may be reasonable to expect a reduction in electricity use in the medium-term (weeks/months) of ~5-7% approximately. Results of the second experiment (i.e. introduction of SM with and without intervention) show that subjecting participants to loss aversion and salience seems to affect their behaviour toward electricity use, as the intervention group reduced their consumption roughly twice as much as the reference group. With due limitations, the results suggest that the delivery of information to energy users needs to take into account not only its pure provision, but how it is designed, framed and presented. At all events, the results and reviewed studies strongly suggest that increased energy efficiency and energy conservation need to be addressed with a mix of policies – not only information schemes or the provision of feedback alone.

Keywords: Smart Meter, Energy, Households, Behavioural Economics, EU energy policy

* Corresponding author. Tel.: +45 27217414. E-mail: simonbager@gmail.com
1. Introduction

The amount of carbon dioxide (CO₂) and other greenhouse gases (GHG) in the atmosphere is increasing, which is leading to global climate change (IPCC, 2013). As the world’s energy system still largely relies on the consumption of fossil fuels, a significant part of global anthropogenic GHG emissions stem from energy use, in which the buildings sector is playing an increasing role. Emissions from this sector have more than doubled since 1970, growing from 3.8 gigatonne CO₂ equivalent (GtCO₂e) in 1970 to 6.3 GtCO₂e in 1990 and reached 9.18 GtCO₂e in 2010. This is equal to 19% of all global emissions, more than a third of which (3.5 GtCO₂e in 2010) can be directly attributed to electricity use in residential buildings (Lucon et al., 2014). Indirect CO₂ emissions from electricity use in buildings have grown considerably, quintupling since 1970, in contrast to direct emissions, which have roughly stagnated during this period (Lucon et al., 2014, p. 678). This makes reducing electricity consumption especially important for reducing GHG emissions from the sector.

Within this context, large cost-effective saving potentials through increased energy efficiency (EE) have been estimated for the EU residential sector (DG Energy, 2007; EC, 2010; EC, 2014b). Reducing GHG emissions in the residential sector through energy efficiency improvements can be done in two principle ways: 1) by reducing energy use through energy conservation measures, i.e. curtailing the use of products and services, such as switching off appliances, and 2) by increasing energy efficiency via technology means, i.e. decreasing the energy use needed to meet energy service demands and/or deliver goods and services consumed by households (Abrahamse et al., 2005). The full energy savings potential towards 2020 is estimated to be around 27% of total energy use by the sector. If this can be achieved, the EU’s total energy consumption would be reduced by ~11% compared to 2005 (DG Energy, 2007). The high energy use and large potentials for energy efficiency gains and reduction of consumption make the household sector an important target for policies aiming to reduce GHG emissions (e.g. Benders et al., 2006; Dietz et al., 2009; Attari et al., 2011). However, in spite of multiple economic, social and environmental benefits embedded in increased energy efficiency, a number of market failures and barriers have traditionally prevented efficiency improvements due to, for example: information asymmetries; negative externalities not reflected in energy prices; uncertainties about technical performance; etc. (see e.g. Gates, 1983; Howarth & Sanstad, 1995; Jaffe & Stavins, 1994a; 1994b; Metcalf, 1994; Ruderman et al., 1987; Sutherland, 1991; Train, 1985).

To correct or reduce many of these market failures and barriers, and thus tap EE potentials in the most cost-effective way, the EU and its Member States have deployed a great variety of specific policy instruments, including regulatory instruments, market-based instruments, and voluntary standards or informative policy instruments (Carter, 2007; Stavins, 2001), with the objective of addressing the unsustainable consumption of energy in the residential sector. Regulatory policies include building codes, energy performance standards and energy labelling of buildings, while market-based instruments include policies such as taxes, subsidies, rebate programmes and on-bill schemes, such as the UK Green Deal (Mundaca, 2008; EC, 2014). Informative policy instruments applicable to EE include communication campaigns and labels (Mundaca, 2008). Information generally has public-good attributes and tends to be underprovided by the market (Jaffe & Stavins, 1994a), meaning that individuals bear the cost (temporal, cognitive, and/or monetary) of obtaining the information. This means that consumers lack information or face asymmetric information, which, in the context of energy use, leads to lower uptake of EE measures and less energy reduction than predicted by economic theory; a market failure (Jaffe & Stavins, 1994a; 1994b). The rationale behind informative policy instruments is to correct this specific market failure by providing information, effectively reducing uncertainties inherent to many energy-related decisions (Mundaca et al., 2013, Somanathan et al., 2014). The theoretical assumption is that this information asymmetry is preventing consumers from exercising rational choice and maximising their personal utility (Micklitz et al., 2011).

In order to address the lack of information (and strengthen the mix of the informative policy instruments) that prevent diffusion of profitable efficient technologies, the rollout of ‘Smart Meters’
is essential (Giordano et al., 2011; Christensen et al., 2013). Smart Meter (SM) functionalities vary from model to model, but generally include measuring and displaying electricity consumption in quantity (kWh), over time, and the cost of this, communication capabilities, storage and transfer of data, supporting dynamic tariffing and payment systems, communication with and remote disablement and enablement of electricity and individual devices within the home, and information transfer to a display or other equipment, such as a smartphone or tablet (after Owen & Ward, 2006; Hoenkamp et al., 2014). Historically, European consumers have inferred knowledge about their electricity demand through billing estimates and infrequent meter readings, which display the consumption in kilowatt-hours (kWh) as a cumulative total. This means that consumers have not had information on the impact of their energy-related actions, as the total figure obscures the impact of various daily energy use habits (van Elburg, 2009). In the EU, the development accelerated when the Energy Services Directive (ESD) (Directive 2006/32/EC) required energy suppliers to provide consumers with competitively priced, accurate, individual meters that provide information on time-of-use (art. 13, sub1) and accurate billing based on actual consumption (art. 13, sub 2), prompting the first installation of SMs in some EU countries. EU Member States were free to decide on their own implementation strategies (DG IPOL, 2012), which, consequently, have led to Member States taking different routes in terms of timing and technology regulations. A 50% rollout target is set for 2015 (Covrig et al., 2014), and by 2020 SMs should be installed in 80% of the households in the EU, where this is deemed a net benefit as assessed through cost-benefit analysis (CBA) (Directive 2009/72/EC); however, despite numerous policy expectations and arguments, at present, only Sweden, Finland and Italy have implemented rollout of SMs on a larger scale.

The EU expects that the introduction of SMs will result in a 10% reduction of energy use in the residential sector (EC, 2011); a policy approach that is based on an ‘information-deficit model’ that assumes a more rational behaviour by consumers if information asymmetries are reduced. A central assumption in this approach is that provision of real-time information via SMs enables end-users to make more rational decisions about their energy service demands (e.g. lighting). Much of European consumer policy, including the energy and environmental fields, is based on the information-deficit model and the assumption of the rational consumer (Micklitz et al., 2011; Gowdy, 2008). The implicit hypothesis of said policy model is that it is only asymmetries or lack of information that prevents people from making rational choices, and that if the right combination of regulations and economic policies are supported by informative policies delivering precise information to consumers, their behaviour will change accordingly.

Until recently, little was known about the effects of providing feedback on energy behaviour using SMs, but trials in the UK and Netherlands (Kinzig, 2013; Raw et al., 2011), as well as newer academic studies (Gans et al., 2013; Grønhøj & Thøgersen, 2011; Pyrko, 2011; Schleich et al., 2013) are making up for this lack of research. However, the amount of large-scale trials testing the effect on electricity consumption of installing a SM is somewhat limited. With this in mind, even if there is some evidence that real-time feedback to consumers drives a more efficient use of energy (Raw et al., 2011), the magnitude of this reduction has been debated, raising doubts about the effectiveness and economic efficiency of introducing SMs in the residential sector.

This paper aims to contribute to this growing body of knowledge by examining the effect of behavioural aspects on electricity use. We examine how behavioural biases affects consumers’ response to energy-use information provided through Smart Meters. Whether loss aversion can affect electricity use has, to the knowledge of the authors, not been previously investigated. Taking insights from behavioural economics (BE) as a starting point, we explore to what extent the way information is presented or framed to households has an impact on how the information is used and acted on. Research on BE argues that the (social) decision context in which information is presented, 1A total of 16 Member States will proceed with large-scale rollout of SMs by 2020 or earlier, three Member States have opted for selective rollouts to some customers, and four Member States had negative or inconclusive Cost-Benefit Assessments leading to no plan for a rollout (EC, 2014). Based on these assessments, Member States have committed to rolling out close to 200 million SMs, and it is expected that almost 72% of European consumers will have a SM for electricity by 2020 at a total investment cost of €35 billion (Covrig et al., 2014).
understood as the choice architecture, along with heuristics (rules-of-thumb), biases and salience, influence how humans respond to information and make decisions (Sunstein, 2013; Kahneman, 2003; 2011). Research on BE relies on empirical studies to infer the actual behaviour of individuals, rather than derive axiomatic assumptions from theory. To address the case at hand, our exploratory research used real-life experiments with SMs and electricity users conducted in Copenhagen, Denmark (details in next section) to identify and assess the potential of the above-mentioned economic behavioural biases. The central tenant of BE is that cognitive, emotional and social factors influences how information is understood and limits the possibility to display purely rational behaviour, in turn affecting human (economic) decision-making (Kolstad et al., 2014). Whereas much attention has been given to technological aspects and the pure provision of information (Fischer, 2008; Faruqui, 2010), less is known about the role of behavioural biases and cognitive issues (e.g. loss aversion) associated with SMs and energy use.

The rationale behind the SM experiment was thus two-fold: Firstly, to assess whether framing the total daily electricity use as an ‘economic loss’ can trigger loss aversion, meaning that any reduction in consumption will be seen as reducing a loss, which theoretically should induce a more significant change in behaviour than if electricity consumption reduction is seen as a gain (c.f. the value function in cumulative prospect theory) (Kahneman & Tversky, 1979, Tversky & Kahnemann, 1992). Secondly, to test whether the salience—understood as the ease with which data can be understood and processed by humans (Kahnemann, 2003)—of the information presented drives a change in final electricity use. This was tested by presenting the nightly electricity use (from 23 P.M. until 06 A.M.) as a cumulative total per day and per year in DKK\(^3\), rather than it not being visible, as is usually the case. It was assumed that the cumulative total would be more salient to the consumer than the daily figure. If salience does have an impact, this means that not only whether information is provided, but also how it is provided, can impact consumption. The hypothesis is that simply correcting the information asymmetry is not enough; but it should be carefully considered how the information is framed.

As a result, this paper sets out to answer what the expected energy efficiency improvements on electricity use as a result of SM deployment may be, particularly in the field of controlled customer feedback, with and without behavioural interventions? Furthermore, a critical literature review was carried out to compare our results with previous studies on the effects of providing ‘information feedback.’ We do this also taking into account sample size and intervention period.

The structure of the paper is as follows. Section 2 elaborates on the methodology, in which the experiments and data collection are described in more detail. Section 3 outlines the key findings of our research. Results are grouped according to the experiments carried out, including the comparison of our results to previous studies on information feedback. In light of our results, we draw some conclusions in Section 4; including methodological aspects.

2. Method

2.1. Experiments

To meet the objective, experimental research was conducted through two experiments: First, a simple experiment took place: whether the installation of SMs yielded (or not) reductions in electricity use. To that end, the electricity use data of 92 households in Copenhagen, Denmark with SMs installed was collected, with no other intervention being made. Customers with such meters were provided with near-real time feedback, as their consumption was reported and made available

---

2 Modern research on human behaviour in response to energy use information is generally grounded in two main shifts away from the information-deficit model and the understanding of the consumer as a fully rational agent: a practice theoretical approach focusing on the power of social norms and practices (e.g. Hargreaves, 2011; Gram-Hanssen, 2013) and a behavioural economics approach proposing the limited rationality of humans and the influence of automatic judgments as explanations of why humans do not make energy-saving decisions to the degree predicted by neoclassical economic theory (e.g. Baddeley, 2011; Newell & Siikamäki, 2013; Gilbert & Zivin, 2014; Kallbekken et al., 2013).

3 DKK = Danish Krone. 1 DKK = 0.13141 Euro, 1 DKK = 0.18099 USD (as of 23 July 2014) (www.oanda.com).
in 15-minute intervals. The data was analysed to assess whether any effects greater than those expected from natural yearly fluctuations could be discerned. The customers had access to a portal where they could check their consumption in kWh and DKK on an hourly, daily, weekly, monthly or yearly basis. Thus, unless they accessed this portal, they received no other information than what a conventional electricity meter can provide (see Figure 1). The rationale behind this experiment was to test whether the EC prediction of a reduction in electricity use of 10% from installation of SMs in Europe is in any way reflected in electricity use profiles of customers with SMs installed.

Second, a residential SM experiment was conducted in collaboration with NorthQ, a producer of SMs. In the experiment, electricity use in the participating households was monitored, and BE-inspired feedback was provided. The experiment tested the effect of two behavioural biases, namely salience and loss aversion, on consumer behaviour with regards to electricity use and energy-related decisions. Salience is understood as the ease with which data can be understood and processed by humans, which is influenced by the degree or extent to which a given piece of information stands out relative to other information. In other words, the mental capacity needed to understand, process and act upon an event, such as information on ones electricity consumption (Kahneman, 2003). Theoretically, it is easier to process a single, aggregate number (e.g. total cost) rather than numerous, disaggregated values, a feature we included in the design of the feedback. Loss aversion is defined as the tendency for people to put substantially greater weight on relative losses than on gains of the same magnitude, or the idea that the value function employed when making a decision is non-linear (Kahneman et al., 1990), an insight formulated in cumulative prospect theory (CPT) (Kahneman & Tversky, 1979, Tversky & Kahnemann, 1992). Given that humans apply a non-linear value function, it is hypothesised that framing a cost as avoiding a loss, rather than obtaining a gain, should have an effect on behaviour; as such, the cost of electricity was presented to participants as an economic loss, which could be reduced by curtailing electricity use (see e.g. Kahneman, 2011; Weber, 2013).

The participating households were divided in two groups: a reference group, and an intervention group, who were subjected to salience and loss aversion. The reference group received information on household electricity use in kilowatt-hours (kWh), and information on how much their consumption aligned with a pre-determined budget (in DKK per year), which had been set by the household – usually this was equal to the amount of money they had spent on electricity in the previous calendar year. The test group received the same information, along with information on the running costs of electricity use and the estimated weekly cost (framed as a loss), and the cost of

Figure 1 – Snapshot of consumption information available to SM customers online. ©NorthQ (used with permission) (www.northq.com).
passive and standby electricity use per day and per year under the assumption that nightly electricity consumption could be taken as a proxy for standby electricity consumption (framed as a loss and made salient). The widget in the software interface read: “Money lost from electricity consumption” and then stated the value. The widget displayed the amount spent per day as a running total, which was updated every few seconds and reset every day, meaning that it looked like the money was “flowing” out of their (figurative) pockets. The estimated weekly cost was updated every 15 minutes. The standby electricity consumption data was updated once a day to reflect the standby consumption of the previous night.

In order to be able to compare the effect found in the two experiments against the ‘normal’ effect of feedback, a literature-based analysis was conducted. Though some literature studies exist, they were, with the notable exception of Ehrhardt-Martinez et al. (2010), deemed to contain an insufficient amount of information to conclude on the magnitude of the effect of feedback, as basic statistical tests, as well as comparisons against intervention time or sample size, were generally lacking.

Note that the experiments conducted and analysed for this paper can be defined as field experiments as opposed to lab experiments. The experiment took place in a real-life setting where consumers actually used and paid for their electricity and can be seen as a real-life situation, where the behaviour takes place in the actual setting. This means that external validity is high while internal validity is rather low (Roe & Just, 2009). However, the very specific context makes replicating the experiments difficult, which means that the estimated effect cannot be assumed to be of the same magnitude once scaled to a population. This has implications for the conclusions that can be drawn based on the experiment and the data collected. In general, as pointed out by Levitt & List (2007), there is a lack of knowledge on how to scale economic behaviour as found in a lab or artificial settings to the market.

2.2. Data collection and processing

The historical consumption data collected from the SMs (experiment 1) was sampled from 92 meters (nSM=92). For privacy considerations, participants did not provide data in relation to their income status, education, and other contextual factors. The observation from each meter was of varying time span and covered a period from the fall of 2012 to early August 2014. Note that the number of observations was much higher in the final part of the period, due to a higher number of SMs being installed at that point. Before analysis, data from weeks with less than 7 observations in one week, i.e. weeks with data missing for any days, were discarded, as was weeks that returned a total consumption of “0” (zero), as this indicated something had gone wrong when transmitting the data. Meters with observations spanning less than 3 weeks in total were excluded. This yielded at total sample of 51 meters. Furthermore, two meters with extremely high readings were also excluded based on an assumption that this data was incorrect. This yielded a final sample of 49 meters (nSMfinal =49). Taking into account that there are 2.6 million households in Denmark (Statistics Denmark, 2014), and using a confidence level of 95%, results are representative when the margin of error is 14%.6 To approximate the potential effect, the deviation in weekly consumption from expected consumption was compared. Electricity consumption data from 3,000 Danish households was used to create a yearly demand profile for an average Danish household, against which the consumption data from the SMs could be compared (NorthQ, 2013). The data was based on daily readings, but was

---

4 The criteria for including a study in the review was: (a) the presence of a control/no effect group, (b) a quantitative estimate of the effect of the intervention provided, (c) a test group larger than 10 individuals/households, (d) the feedback could not be normative (social), i.e. compared to the consumption of other households. Where available, the intervention length was also included.

5 Validity is the question of whether a particular conclusion represents a good approximation of the true conclusion, i.e. whether the methods of research reflects the truth. Generally, one distinguishes between internal and external validity. Internal validity can be defined as “the ability of a researcher to argue that observed correlations are causal,” whereas external validity can be defined as “the ability to generalize the relationships found in a study to other persons, times, and settings” (Roe & Just, 2009). Finally, ecological validity is understood to be the “extent that the context in which subjects cast decisions is similar to the context of interest,” i.e. whether the experimental settings reflect the settings where decisions would be made in real life.

6 With standard margin of error (5%), the sample should have been 385 to be representative.
converted into weekly values and normalized, displaying the weekly consumption as a percentage of the yearly consumption. The obtained profile can be seen in Figure 2.

![Figure 2 – Seasonal fluctuations in consumption (Data from NorthQ, 2013).](image)

To approximate and thus assess the potential effect of installing a SM, the consumption from the first week after installation was taken as a starting point, assumed to be the ‘normal’ consumption from which an effect should be discerned. Based on the yearly consumption profile, the expected weekly consumption was estimated, while taking into account the time of the year the meter was installed (week = 0) (e.g. in week 27 consumption was expected to be 1.644% of total yearly consumption because of the seasonal fluctuations in electricity consumption). This means that the starting point varies according to the installation time of the meter (1 ≥ X ≤ 52). The actual consumption in week 0+X was then compared to the expected consumption in week 0+X, and the relative change in consumption from the expected consumption was found.

$$E_{\text{SmartMeter}} = \frac{C_{\text{Actual}} - C_{\text{Expected}}}{C_{\text{Expected}}} \times 100\%$$

Data for the second experiment on loss aversion and salience was collected over a 5-week period between July and August 2014. 85 households located in Copenhagen, none of which took part in experiment 1, were invited to participate in the experiment, but only 63 chose to do so by providing their written consent. Even though the experiment was postponed two weeks to increase the sample size, out of the 63 households that accepted the invitation, only 16 installed the meter in time to be part of the experiment, meaning that the two groups were reduced to a test group of 11 households (n_{test} = 11) and a reference group of 5 households (n_{ref} = 5). The sample data contained information on daily consumption (kWh/day), weekly consumption (kWh/week), nightly consumption (passive and stand-by) (kWh/night [23-06]), as well as average daily consumption (kWh) and average weekly consumption (kWh). The households had the same type of SM installed as those meters from which the long-term data was collected, along with software to monitor the use of electricity. The change in electricity use as a result of the intervention was analysed using two different methods: i) one tracking the absolute change from first to last week; and ii) the other calculating the relative change over time using the average consumption in the first week as baseline. In the first analysis, the effect was assessed by comparing average use in the beginning of the period with average use at the end of the intervention period. In the second analysis, the potential effect was approached by comparing the average deviation from the baseline consumption in a manner similar to that applied for experiment 1. Where applicable, statistical tests of significance were conducted. Unless anything else is stated, _significant_ means significant at a 95%-confidence level (p-value <0.05).

The households participating in the second experiment have similar building characteristics, as all houses are semi-detached houses, constructed around the same time, all connected to the district-heating grid, none of them using electric heating, with sizes varying from 103-130 m², and are all located in the same neighbourhood in Copenhagen, Denmark (see Figure 3). The house owners were
assumed to be middle class or richer in order to afford these houses, meaning that the findings from this study are not necessarily representative of residents in low-income areas\textsuperscript{7}. Similarly, it is plausible that people with an interest in electricity conservation are overrepresented in the sample, as all the participants in the loss aversion experiment indicated that they considered themselves to be “environmentally aware.” If a specific finding is inferred on a narrow sub-set of the population (e.g. high-income, urban dwellers), the relationship may not exist in more diverse samples (Henrich et al., 2010)\textsuperscript{8}.

Figure 3 – The participating households (the black-roofed buildings in the centre) are all located in a residential area in the north-western part of central Copenhagen. Images courtesy of Google©

3. Results and discussion
The following section presents the findings from the various analyses and experiments conducted for this paper. First, the results of the analysis of the data from the first SM experiment are presented. Then the findings from the second experiment on SMs with loss aversion and salience framing are offered, in order to answer the research question. This is followed by the results from the literature-based analysis on the effect of feedback provision, conducted in order to have something to compare the results of the SM experiments against.

3.1. Experiment 1: Installation of Smart Meter without further intervention
Based on the yearly demand profile, the relative effectiveness of the meters, i.e. the deviation in percentage from the expected consumption, was found for the 49 SMs included in the sample. Data was quite scant for the weeks prior to week 43 in 2013 (less than 10 observations), due to a lack of

\textsuperscript{7}Romanach et al. (2013) found that “financial savings” was the main reason low-income household participated in energy-use programmes. As this would not necessarily be the case for high-income households (e.g. it might be social norms or other non-monetary reasons), there is a need to be cautious against assuming that these findings are applicable in any social settings in high-income countries, as low-income households are usually underrepresented in energy use studies (Romanach et al., 2013).

\textsuperscript{8}Note that (economic) experiments have historically relied on a distinct set of people (university undergraduate students are highly overrepresented), so there is a need to test these interventions on a broader group of people having “participants that are representative of whole countries or cultures” (Fehr & Fischbacher, 2003, p. 790).
installed meters in the early period. For this reason, the early results are not included herein. Based on the applied method (i.e. using week 1 as baseline and correcting the measured data for climate factors), large deviations (±50%) from the expected consumption were found. The data plot can be seen in Figure 4. The black line displays the average (the arithmetic mean) of all the data points, while each of the coloured lines represent an individual SM.

Although the potential effect on electricity use reduction is difficult to quantify due to large fluctuations, an attempt was made nonetheless. The deviations are negative 75% of the time (31/40 weeks) (i.e. consumption is below the expected value). Overall, the average change from the expected consumption is -6.7% (±41%, n=47). However, the standard deviation is 0.43, or 43% (n=47), meaning that there is a large variation in the effect found. The weekly deviations are negative in 309 instances and positive in 341 instances, indicating that the average negative deviation is slightly larger than the average positive deviation.

![Figure 4 – Deviations in electricity consumption from expected values for the Smart Meters (SM) analysed. The numbers refer to individual SM, while the black line is the average (arithmetic mean) of all the SM.](image)

The results were roughly consistent with the reductions found in the literature, as an effect of -6.7% (±41%) falls within most of the studies reviewed. In our case, a total of 19 separate Smart Meters was installed. Based on the applied method (i.e. using week 1 as baseline and correcting the measured data for climate factors), large deviations (±50%) from the expected consumption were found. The data plot can be seen in Figure 4. The black line displays the average (the arithmetic mean) of all the data points, while each of the coloured lines represent an individual SM.

Although the potential effect on electricity use reduction is difficult to quantify due to large fluctuations, an attempt was made nonetheless. The deviations are negative 75% of the time (31/40 weeks) (i.e. consumption is below the expected value). Overall, the average change from the expected consumption is -6.7% (±41%, n=47). However, the standard deviation is 0.43, or 43% (n=47), meaning that there is a large variation in the effect found. The weekly deviations are negative in 309 instances and positive in 341 instances, indicating that the average negative deviation is slightly larger than the average positive deviation.

The results were roughly consistent with the reductions found in the literature, as an effect of -6.7% (±41%) falls within most of the studies reviewed. In our case, a total of 19 separate Smart Meters was installed. Based on the applied method (i.e. using week 1 as baseline and correcting the measured data for climate factors), large deviations (±50%) from the expected consumption were found. The data plot can be seen in Figure 4. The black line displays the average (the arithmetic mean) of all the data points, while each of the coloured lines represent an individual SM.

Although the potential effect on electricity use reduction is difficult to quantify due to large fluctuations, an attempt was made nonetheless. The deviations are negative 75% of the time (31/40 weeks) (i.e. consumption is below the expected value). Overall, the average change from the expected consumption is -6.7% (±41%, n=47). However, the standard deviation is 0.43, or 43% (n=47), meaning that there is a large variation in the effect found. The weekly deviations are negative in 309 instances and positive in 341 instances, indicating that the average negative deviation is slightly larger than the average positive deviation.

The results were roughly consistent with the reductions found in the literature, as an effect of -6.7% (±41%) falls within most of the studies reviewed. In our case, a total of 19 separate Smart Meters was installed. Based on the applied method (i.e. using week 1 as baseline and correcting the measured data for climate factors), large deviations (±50%) from the expected consumption were found. The data plot can be seen in Figure 4. The black line displays the average (the arithmetic mean) of all the data points, while each of the coloured lines represent an individual SM.

Although the potential effect on electricity use reduction is difficult to quantify due to large fluctuations, an attempt was made nonetheless. The deviations are negative 75% of the time (31/40 weeks) (i.e. consumption is below the expected value). Overall, the average change from the expected consumption is -6.7% (±41%, n=47). However, the standard deviation is 0.43, or 43% (n=47), meaning that there is a large variation in the effect found. The weekly deviations are negative in 309 instances and positive in 341 instances, indicating that the average negative deviation is slightly larger than the average positive deviation.
Meter intervention studies testing the effect of feedback were found (see e.g. Gans et al., 2013; van Dam et al., 2010; Ersson & Pyrko, 2009a, 2009b; Raw et al., 2011; Schleich et al., 2011; 2013; and Grønhøj & Thøgersen, 2011). As a whole, the average change to electricity use in the studies reviewed was found to be -1.6%±9.7%, 1st-3rd quartile: -6.15% to 0.05%, M=-2.9%, n=19. If meters with displays are excluded, the effect changes to -0.7%±11.2%, M=-2.4%, n=14. Given the different nature of the studies, the lack of longitudinal studies, and the general context in which these studies were conducted (different countries, different setups, different time of the year), it is very difficult to generalise from these results. However, that the median value is lower (more negative) than the mean indicates that a larger number of studies found low values (i.e. reductions), and perhaps the presence of a single (or a few) studies (e.g. Ersson and Pyrko, 2009a; 2009b) finding high values (i.e. large increases in consumption) distorts the picture. Finally, it must be mentioned that Schleich et al. (2011; 2013) and Gans et al. (2013) both use regression analysis, but find fairly different effects, e.g. Gans et al. (2013) find that education and job status impacts consumption, while Schleich et al. (2013) find that previous consumption levels matter most. This indicates that other contextual factors also matter, and should be taken into account when analysing electricity use data.

3.2. Experiment 2: Installation of Smart Meter with and without further intervention (loss aversion and salience)

Although data was somewhat scattered and fluctuated significantly within and between households, the experiment indicated that end-users are prone to behavioural biases when faced with decisions relating to electricity use. Note that the analysis ignores the fact that there is a difference in the composition of the households (e.g. number of householders, income, educational level, etc.).

Using the first analytical approach (absolute change), the group not subjected to loss aversion and salience (i.e. reference group) reduced their daily electricity consumption by 7% on average, while those subjected to loss aversion and salience (intervention group) reduced their consumption by 18% (see Table 1). Incidentally, the same reduction was found for the reference group as for the SM group as a whole, but this is assumed to be a coincidence and further research is needed to address causality10. As the experiment took place over just 5 weeks in the summer months, no climate correction was done11.

Using the analytical approach of relative change, the change in electricity consumption was 2% for the reference group and -5% for the intervention group (see Table 2). However, and with due limitations, the findings were consistent with the results found in the first analytical approach: the behavioural intervention has a larger effect than when no framing was applied. Compared to related research, findings revealed that reductions in electricity use were also larger than the average electricity reduction found in other studies of feedback on electricity use.

Table 1 – Daily average electricity use for households in experiment 2 using analytical approach 1 (absolute change)

<table>
<thead>
<tr>
<th>Analytical approach 1: daily consumption</th>
<th>First week Average (kWh)</th>
<th>Mid-period week Average (kWh)</th>
<th>Last week Average (kWh)</th>
<th>Change in consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test group (Loss Aversion) (n=11)</td>
<td>5.72</td>
<td>4.88</td>
<td>4.68</td>
<td>-18%</td>
</tr>
<tr>
<td>Reference group (No intervention) (n=5)</td>
<td>5.10</td>
<td>--</td>
<td>4.72</td>
<td>-7%</td>
</tr>
</tbody>
</table>

that participants equipped with a SM achieved average savings after four months of 7.8% (SD = 13.8%, n = 54), but report that the effect dropped to an average of 1.9% (SD = 11.8%, n = 54) after 15 months.

10 Due to the limited sample size available for regression analysis, a multiple regression analysis of consumption was finally dismissed after statistically insignificant results were computed.

11 Based on the yearly consumption data used to correct the Smart Meter data, consumption is expected to be about 5% higher at the end of the test period than at the beginning, so if anything, not correcting for climate impacts yield answers that might be too low. As so many other variables were also not accounted for, it was assumed that not correcting for climate variability would have little influence.
As is the case with the daily data series, the nightly data displayed fluctuations within and between households, but the internal variation was smaller, suggesting that standby consumption was comparatively stable. Using the analytical approach of absolute change, the reduction in standby consumption was 3% for the reference group, but 28% for the intervention group (see Table 3). Using the method of relative change, the change in standby consumption was -13% on average for the intervention group, and a 3% increase in consumption for the reference group (see Table 4).

However, the findings were consistent with the results found using the first analytical approach: the loss aversion and salience framing has a larger effect than when no loss framing is applied. For the nightly values, the difference between the two types of framing was markedly larger than for the daily values. As nightly values were labelled standby consumption and aggregated to a yearly figure, this suggests that the salience of the cost of standby electricity consumption increased, which would further contribute to the reduction in consumption. This would be in line with other research (e.g. Gilbert & Zivin, 2014) and indicate that highlighting the aggregate figure can increase the effect of feedback. The effect of the intervention analysed using the two analytical approaches can be seen in Figure 5.

Table 2 – Effect of framing on daily electricity use using analytical approach 2 (relative change)

<table>
<thead>
<tr>
<th>Analytical approach 2: daily consumption</th>
<th>Test group (loss aversion, salience) (n=11)</th>
<th>Reference group (No intervention) (n=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change, all days included (also first week)</td>
<td>-4.3%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Change, all days after day 7 (not first week)</td>
<td>-5.2%</td>
<td>2.2%</td>
</tr>
</tbody>
</table>

Table 3 – Standby electricity use for households in experiment 2 using analytical approach 1 (absolute change)

<table>
<thead>
<tr>
<th>Analytical approach 1: standby electricity use</th>
<th>First week Average (kWh)</th>
<th>Mid-period week Average (kWh)</th>
<th>Last week Average (kWh)</th>
<th>Change in consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test group (Loss Aversion) (n=11)</td>
<td>1.04</td>
<td>0.94</td>
<td>0.75</td>
<td>-28%</td>
</tr>
<tr>
<td>Reference group (No intervention) (n=5)</td>
<td>0.89</td>
<td>--</td>
<td>0.86</td>
<td>-3%</td>
</tr>
</tbody>
</table>

Table 4 – Effect of framing on standby electricity use using analytical approach 2 (relative change)

<table>
<thead>
<tr>
<th>Analytical approach 2: standby electricity use</th>
<th>Test group (loss aversion, salience) (n=11)</th>
<th>Reference group (No intervention) (n=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change, all days included (also first week)</td>
<td>-10.5%</td>
<td>2.4%</td>
</tr>
<tr>
<td>Change, all days after day 7 (not first week)</td>
<td>-12.8%</td>
<td>3.2%</td>
</tr>
</tbody>
</table>
Figure 5 – Changes in electricity consumption as a result of information feedback

The effect approximated using the second analytical approach was slightly smaller than with the first one, possibly because the first seven days (where the meter is already installed) are taken as a starting point, which means that any reduction made in that period would lead to a lower starting point from which change should then be discerned. However, both analyses found that the framing led to larger reductions in electricity use and that this is especially pronounced for standby electricity consumption.

The larger reduction for the participants subjected to loss aversion suggested that seeing a reduction in electricity as ‘avoiding a loss’ (as done by the intervention group) provided a stronger incentive to reduce this loss than seeing a reduction in electricity use as gaining a benefit (as done by the reference group). This is in line with the value function of CPT and previous research on loss aversion in other areas (e.g. Genesove & Mayer, 2001; Tom et al., 2007). Although the approximated reductions in electricity use do have a high level of uncertainty, they roughly correspond with the general notion that losses are felt twice as much as gains (c.f. Kahneman, 2011).

3.3. The effect of feedback on electricity use

Compared to the number of studies employing SMs to provide feedback, there is a relative abundance of studies that have tested the specific effect of information provision to consumers. In fact, several review studies have assessed the overall effect of these interventions (Abrahamse et al., 2005; Darby, 2006; Fischer, 2008; Ehrhardt-Martinez et al., 2010; Vine et al., 2013. In general, the studies reviewed for this paper indicated that providing feedback results in electricity reductions of 1-13% (m=−6.9±5.8%, M=−6.5%, n=23). This range is in line with the ranges given in recent review studies (Abrahamse et al., 2005; Darby, 2006; Fischer, 2008), though in the lower end. The mean and median are almost similar, indicating that there are just as many studies finding no effects as studies finding large effects. The standard deviation is around ±5%, which shows that there is a relatively large spread in the estimated effect. However, some researchers (e.g. Gans et al., 2013) caution against taking the reductions reported too literal, as many of the studies and reviews do not contain any statistical tests of the significance of the results, mostly due to the relative low number of households/participants in the experimental group, as well as the large within-group variation (Osbaldeston & Schott, 2011; Abrahamse et al., 2005). Similarly, almost no studies employ multiple regression analysis to test the influence of other variables, and it is thus difficult to know whether there is really an effect or it is merely the influence of another variable, such as housing size that is found (Silver, 2012). This suggests that it may be common to have very low-level (statistical) sample data, meaning that the sample size and statistical power of the experiments conducted for this paper
are the norm rather than the exception. Obviously, this calls for further research to address this issue.

Ehrhardt-Martinez et al. (2010, p. 53) find that “average energy savings were higher for shorter studies (10.1%) than for longer studies (7.7%).” This is reflected by the finding that in the situations where a follow-up to a short-term study was conducted (e.g. van Dam et al., 2010; Raw et al., 2011), the effect had often diminished over time (c.f. Abrahamse et al., 2005; Owen & Ward, 2006; Ehrhardt-Martinez et al., 2010). The SM data collected for this experiment also covered relatively short time periods (less than 3 months) for most of the data points in the sample. Some studies also speculate that small-sample studies find higher levels of effect than large-scale studies, e.g. Ehrhardt-Martinez et al. (2010) find that “average energy savings across large-sample [>100] studies is roughly 6.6% compared to average savings of 11.6% across small-sample [<100] studies.” To assess whether this was the case, an f-test was conducted to test whether the length of the intervention period or the sample size had a significant influence on the variability in electricity reduction reported (see Table 5). Applying a 95%-confidence level, none of the two can be said to have a significant influence. However, changing the confidence level to 90% brings the feedback as function of sample size within range of the critical value, and changing the confidence level to 85% brings it below. However, applying such a low confidence interval is not the norm, and therefore our study does not draw any conclusions from this. However, due to the limited sample size (n=20), an effect cannot be ruled out, especially not for sample size, and further research is warranted.

Table 5 – Statistical test results for the f-test

<table>
<thead>
<tr>
<th>Statistical parameters</th>
<th>Feedback as a function of length of intervention period</th>
<th>Feedback as a function of sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>r, r²</td>
<td>-0.165, 0.027</td>
<td>-0.339, 0.115</td>
</tr>
<tr>
<td>f-test</td>
<td>0.557</td>
<td>2.597</td>
</tr>
<tr>
<td>f-critical (95%, 90%, 85%)</td>
<td>4.351, 2.975, 2.241</td>
<td>4.351, 2.975, 2.241</td>
</tr>
<tr>
<td>p-value</td>
<td>0.465</td>
<td>0.123</td>
</tr>
</tbody>
</table>

Ehrhardt-Martinez et al. (2010) further caution that only two household-level studies in recent times (post-1995) have found energy savings in excess of 10% as a result of feedback, a finding which indicate that the very large numbers reported in older literature (e.g. 18.8% reduction in Midden et al., 1983) may not be a reflection of the reductions that one can expect from SM interventions today. They further speculate that reviews including studies older than 20 years (as is done by e.g. the widely cited study by Darby, 2006) might result in “inflated expectations regarding potential energy savings today” (Ehrhardt-Martinez et al., 2010, p. 74). As more than half of the studies reviewed are short term and small sample, this could lead to inflated expectations about possible energy savings if these studies generally report larger electricity reductions.

Comparative feedback, i.e. having your electricity consumption compared to your neighbour’s consumption rather than your own previous consumption, is another way of providing feedback, which has not been explored in detail in this study, but is worth considering. Only a limited amount of academic studies explicitly testing the effect of comparative feedback could be found (Midden et al., 1983; Schultz et al., 2007; Abrahamse et al., 2007; Nolan et al. 2008), and these all find reductions in energy use. Excluding the oldest study (Midden et al., 1983) yields an effect slightly higher than that found for individual feedback (m=−7.2%±2.4%, n=3). It should be noted that these reductions are notably larger than those found in a large-scale study (n>3,000) conducted by the American utility OPOWER, which found reductions of 2.24%[95%-confidence interval: 1.91%; 2.56%] (Klos, 2009, p. 2), and a large-scale British trial (Raw et al., 2011) which found effects ranging from 1–2% reduction in energy use from comparative feedback. As these trials are larger and of newer date, they probably provide a more accurate view of the reductions, which can realistically be expected from feedback.
Using the large-scale trials of comparative feedback as a benchmark, the effects found in large-scale real-life trials are assumed to be about a quarter to a third of effect found in the smaller studies. The second SM experiment in our study is also short-term and small-scale, as most studies reviewed, meaning that if other feedback studies provide any indication, the effect will be smaller in the population than in the study sample and likely diminish over time (as was found in van Dam et al., 2010). Assuming previous large-scale trials using comparative feedback provide any indication of what can be expected with behavioural feedback, effects in a population of applying individual feedback with loss aversion and salience suggests that feedback should thus result in reductions of 4-6% in daily consumption and 6-8% reduction in standby consumption – reductions of this magnitude would also align with the reductions expected by, for example, Ehrhardt-Martinez et al. (2010) and large scale trials such as Klos (2009) and Raw et al. (2011). In any case, the indication of an effect in both instances (daily and standby electricity use), and the likelihood of replication in real-life situations, calls for large-scale trials to further test this hypothesis.

Whereas the analysis of our experiments did not consider socio-economic aspects (e.g. income, house size, education), it has to be acknowledged that the variability of the reviewed studies in this respect adds complexities to any comparison. The reviewed studies include participants with highly different housing types, household composition, income, education level, cultural backgrounds, and climatic conditions (van Elburg, 2009), which are highly likely to affect the results. Furthermore, many of the studies reviewed were “opt-in” trials, which, based on findings from behavioural economics (e.g. Kahneman, 2011), could have an effect on the outcome, as it must be assumed that those deciding to “opt-in” have a reason to do so, and thus are more willing to change behaviour than the general population, which could lead to higher reductions in the study samples than in the population (Raw et al., 2011).

4. Conclusions

The purpose of the paper was two-fold; firstly, to establish whether the provision of information through the use of a SM leads to reductions in electricity use, and secondly, whether the way in which this information is provided has an effect on the size of said reductions. At the risk of oversimplifying, results suggest that the way information on electricity use is presented to households has an impact on how it is perceived and acted on. From a policy perspective, and with due limitations, the research conducted for this paper has implications for policy-makers, because it highlights that information is not just about quantity (i.e. correcting the market failure), but that a policy prescribing the delivery of information to consumers needs to take into account that how the information is presented, framed, and designed affects the impact that this information will have on consumer behaviour. The research highlighted that insights from BE could explain why some information is more effective than others, and that salience and loss aversion seem to affect behaviour. This implies that when designing informative policies, there is a need to look beyond the information-deficit model, and view information as much in terms of quality as quantity.

The results of the first experiment (i.e. introduction of SM without any further intervention) showed reductions in electricity use as claimed in the related policy literature. Results were also generally aligned with electricity reductions found in previous research on the topic. Taken together, the experiment and subsequent review indicate that it may be reasonable to expect a reduction in electricity use in the medium-term (weeks/months) of ~5-7% approximately as a result of installing a SM. However, the findings also suggest that this effect may diminish over time. The large standard deviation indicates that one probably should not take the results too literal. In addition, the small sample size and the large fluctuations mean that one “odd” sample have a great influence on results, which means that the results are very sensitive to deviating observations. As no background information on the households included in the sample was available (household size, income,

---

12 Effect on daily consumption: 18%. 18*0.33 = 6%. 18*0.25 = 4.5%. Effect on standby consumption: 25%. 25*0.33 = 8.33%. 25*0.25 = 6.25%
housing type, etc.) it was not possible to test for the effects of these variables using econometric analysis. Obviously, this also limits the strength of the results. Using the first week as “background consumption” also introduces an error (and thus uncertainty), as an unusually high or low first week will lead to large deviations later on. In order to have a firm baseline from which to estimate a change in consumption, a much longer time series pre-installation of the meter is needed. As the climate lead to seasonal fluctuations, correcting the consumption for yearly fluctuations should increase the relative strength of the results.

The results of the second experiment (i.e. introduction of SMs with and without intervention) on loss aversion and salience found that subjecting participants to loss aversion and salience do seem to modify their behaviour toward electricity use. Testing the effect on daily electricity use, the intervention group had a larger reduction than the reference group in both analytical approaches. This was also the case for both analytical approaches taken when analysing the standby electricity use, except that the effect was much more pronounced here. As the standby consumption was framed as a loss and made more salient by aggregating costs for a year, this supports the hypotheses that increasing the salience and framing reductions as avoiding a loss rather than obtaining a gain, trigger behavioural responses, as theorised. However, it must be also noted that the results are neither statistically robust nor representative for the case area. The approximated reduction effect in electricity use found in our experiment was larger than most reviewed studies testing the effect of feedback, and especially the reduction in standby consumption was larger. Whether or not the intervention used in the second experiment can actually improve information provision hinges to large extent on the scalability of such an intervention. The indication of an effect in both instances (reductions in daily and standby electricity use), and the likelihood of replication in real-life situations, suggest that there is a strong need to conduct a study similar in design to what was originally intended for this study, i.e. a large-scale, longitudinal study where effects are determined using regression analysis, in order to establish whether these interventions can be said to work over time, and what the effect is likely to be.

In fact, the reviewed studies were also limited by low statistical explanatory power due to limited sample sizes and lack of background information. This suggests that feedback studies generally are very difficult to conduct and that the work conducted for this paper, which also was limited by this, might not be the exception. Contrary to other reviews of feedback (e.g. Darby, 2006; Fischer, 2008; Vine et al., 2013), we are very cautious against concluding that an effect of the magnitude found in a sample will arise in a population, as a number of factors potentially affecting the magnitude of the reduction needs to be further explored. The studies reviewed in our paper, as well as the results from the experiments conducted for this research, yield support to the concern voiced by Ehrhardt-Martinez et al. (2010) that the provision of feedback alone is unlikely to lead to reductions above 10%, meaning that the policy goal set by the EC (EC, 2011) is unlikely to be met without further (policy) measures.

There are two reasons why it is worth reflecting on how the design of the experiments impacted the results obtained. This has implications for the results and thus the conclusions that can be drawn. It also has implications for future research. There is some evidence that people that are aware of their participation in an experiment modify their behaviour because they are partaking in said experiment; an effect sometimes labelled ‘the Hawthorne effect’ (Draper, 2013; Olson et al., 1994). It is not confirmed whether this effect actually exists, but based on the importance of randomised-controlled trials in other research areas, especially medicinal and pharmaceutical sciences, it seems plausible that such an effect in fact exist (Haynes, 2012). For the loss aversion experiment, the possibility of a Hawthorne effect is thus possible. Both groups knew they were part of the experiment, since written consent had to be obtained before the intervention was set in motion. This entails that there is a risk that people wanted to perform well (reduce consumption) because they knew they were being observed; an effect that would not be found under normal circumstances.
As a whole, and with due uncertainties, the results suggest that feedback information can encourage efficient electricity use and thus contribute to meeting the goal of reducing household energy consumption through the use of SMs. However, the (expected) effects and their order of magnitude may heavily depend on how feedback is designed, framed and presented. Again, large-scale trials are needed for more conclusive results. In addition, the study of other policy instruments is needed as they may also lead to the sustained behavioural change required to meet EU policy goals, with correct energy and carbon pricing playing a critical role. Therefore, the deployment of SMs should not be conceived only about the provision of the ‘right’ information (i.e. reduction of information asymmetries), but ‘how’ information is actually provided within a mix of effective and efficient policy instruments. Thus, findings also suggest that the right combination of behavioural insights, policy instruments and SM technologies needs to be investigated further.

References


