Effective energy saving policy requires causal evidence

Kees Vringer
PBL Netherlands Environmental Assessment Agency
Bezuidenhoutseweg 30
2594 AV The Hague
the Netherlands
kees.vringer@pbl.nl

Daan van Soest
Tilburg University
Warandelaan 2
5037 AB Tilburg
The Netherlands
d.p.vansoest@tilburguniversity.edu

Mirthe Boomsma
Tilburg University
Warandelaan 2
5037 AB Tilburg
The Netherlands
mirthe.boomsma@gmail.com

Keywords

evaluation methods, smart metering, energy policy, impact evaluation, causal evaluation practises, residential energy conservation

Abstract

Residential energy use is the source of 15–20 % of CO_2-emissions in the Netherlands. Reaching the Dutch and European climate goals is impossible without a substantial decrease in household energy consumption. Previous research documented that feedback on one's household's energy use, based on smart meter data, can induce energy savings. However, effective energy savings due to the wide enrolment of the smart meter were in the Netherlands much smaller than expected: about 0.9 % for only natural gas (the main source of residential heating in the Netherlands) and no savings for electricity, compared to a predicted 3.5 % reduction for both. To close this gap of about 2.7 % savings, the Dutch government and energy companies decided to improve the already widely used Home Energy Reports (HERs), a form of feedback, delivered bimonthly via mail or email. Unfortunately, the revised monthly HER did not lead to higher energy savings, as proven by a large Randomized Controlled Trial (RCT). To illustrate the importance of methods to get robust insights into energy saving effects like RCTs for evidence based policy making, we present in this paper impact estimates of three different feedback mechanisms – an app, email, and an In Home Display (IHD). Impacts were measured using RCTs, implemented in the Netherlands. Our results confirm earlier findings that feedback is effective if it is real time and continuously visible. For a simple IHD applied in the Netherlands we found savings of more than 2 % for electricity and nearly 7 % for natural gas. Our research also illustrates that impact estimates from observational studies and field-experimental evidence are sensitive to the country-specific environment, as impacts measured abroad can differ substantially from those that are materialized in the Netherlands. The use of causal ex-ante impact analyses, like RCTs, implemented in the country context and among the target population, is indispensable for evidence based policy making.

Introduction

Residential energy use is the source of 15–20 % of CO_2-emissions in the Netherlands (PBL, 2020). Households are thus an important contributor to climate change, while the potential for energy savings (and hence CO_2 emission reductions) is substantial. Residential energy use is determined by the energy efficiency of the home as well as by its occupants' behaviour. While more stringent building codes are effective in reducing energy consumption of new homes (Vringer et al., 2016), the Netherlands' greenhouse gas emission goals cannot be achieved without substantial reductions in energy consumption by existing houses. And with respect to changing behaviour, occupants can save energy by means of a large number of actions, such as systematically closing doors and switching of lights, turning off appliances when they are not in use, and purchasing more energy efficient appliances. The energy-saving potential of residential energy use is estimated to be substantial, because occupants do not even take all actions that are deemed cost-effective (Allcott and Greenstone 2012, Gillingham and Palmery 2014). This may be the case because, among other reasons, households may not be aware of how expensive energy actually is (and hence how costly it is to waste energy; Brounen et al. 2013) or because they underestimate how effective the various measures are that can be taken to save energy (Attari et al. 2010).

Feedback about energy use can be an effective and efficient way to stimulate energy users to take better decisions and to save energy (see e.g. Abrahamse et al. 2005; Allcott 2011; Darby 2006; Dromacque and Grigoriou 2018; Ehrhardt-Martinez et al. 2010; Fischer 2008; Jessoe and Rapson 2014). The notion that providing individual feedback can substantially reduce households' energy consumption is one of the key reasons why the European Union has decided to mandate the deployment of smart energy meters in all member states. More specifically, the European Commission decided that by the end of 2020 at least 80 % of EU households must be equipped with intelligent metering systems, such as smart meters (EU Directive 2009/72). The large-scale roll-out of smart meters in the Netherlands started in 2015, and in 2019 more than 80 % of the homes have been equipped with a smart meter (RVO, 2020). The installation of smart meters will, however, not reduce energy consumption unless the meter's information is relayed to the household. One way to transmit this information is via so-called Energy Consumption Managers (ECMs, like an app or an In Home Display) so that the household has access to its past and present energy consumption data. Another way is that the smart meter's information is collected by the energy supplier and then relayed back to the household at regular intervals (monthly or bimonthly), in the form of so-called Home Energy Reports (HERs).

To date, the smart meters have not fully lived up to their promise. While van Gerwen et al. (2010) predicted savings of about 3.5 % for both natural gas and electricity (based on a combination of field-experimental evidence from abroad and the outcomes of observational studies in the Netherlands), Uitzinger and Uitdenbogerd (2014) estimate actual savings of about 0.9 % for natural gas (the main source of residential heating in the Netherlands) and no savings for electricity. According to Vringer and Dassen (2016) the main reason for these disappointing savings is the low usage of above-mentioned ECMs. While the smart meter roll-out was accompanied by energy suppliers sending out bimonthly HERs to all households with a smart energy meter, a mere 30 % of them installed an ECM.

The low levels of realized energy savings are thus caused by both a low level of adoption of ECMs as well as by the HERs being less effective than predicted. The Dutch government's response to the low realized savings was to negotiate a covenant with the Dutch energy companies to obtain additional savings of 2.7 % on both gas and electricity (or 10 PJ; see Covenant, 2017). As changing the HERs' design is easier (and less costly) than implementing policies to raise ECM usage, the focus was on improving the effectiveness of the HERs. Indeed, evidence from the United States suggests HERs should be able to provide energy savings of about 2 % (see Allcott and Rogers 2014; but also see Andor et al. 2020 who document HER-induced energy savings of just 0.7 % in the case of Germany). Based on the recommendations of Menkveld et al. (2017), the effectiveness of HERs was to be improved by, amongst other changes, using more figures and graphs and less numbers, and by a change in the frequency with which HERs are sent out (from bimonthly to monthly).[1] As robust impact estimates are not available for the Netherlands (Menkveld et al., 2017; Vringer et al., 2021), the 2017 covenant also posited that the impact of the changed HER design was to be measured by means of a Randomized Controlled Trial (RCT). By randomly assigning a sufficient number of households to a treatment and control group (here, being sent either the revised or the standard version of the HER), differences in outcome between the groups can then only be the result of the intervention (the changes in HER design). RCTs are the most robust to uncover the causal impact of an intervention and, as such, have been labelled the gold standard of impact evaluation in the natural sciences (including the medical sciences), but also in the social sciences.

The RCT, implemented by Paradies et al. (2020), uncovered that the revised HER design did not result in any additional savings than the standard version. The potential of HERs for energy savings in the Netherlands is thus limited at best, and hence it is important to assess whether ECMs may prove to be more effective.

In this paper, we aim to illustrate the importance of evidence-based policy making. We do this by presenting impact estimates for three different feedback mechanisms to save household energy use in the Dutch context. We tried to get robust impact estimates by implementing three RCTs to measure the causal effects of three different types of ECMs that are available in the Netherlands: a simple In Home Display (IHD), an app and an email with information.

Method / Approach

To get a better insight into the effectiveness of the various types of ECMs in the Dutch context, we decided to test the effectiveness of an IHD, an app with historic feedback, and an email with feedback information. As mentioned above, an ECM's effectiveness crucially depends on its design. The effectiveness is higher the more detailed the information is, and also the higher the frequency with which the information is provided (ideally real-time); see Darby (2006) and Fischer (2008). Furthermore, the duration is important too, as impact tends to strengthen if the feedback is provided over a longer period of time (Ehrhardt-Martinez et al. (2010). Also, whether the information is relayed automatically (or whether the user needs to actively request it, for example by logging on to a web site) influences the energy saving effects (Menkveld et al. 2017). IHDs are likely to be especially effective because they can give detailed, real-time information, continuously visible on the display, and present in the home over a longer period of time. However, IHDs are quite expensive – in terms of purchase and sometimes there is also a monthly subscription fee for the use of a server for the energy data. Websites and apps are much cheaper (and often costless for the energy user), but have the disadvantage that they typically are unable to provide real-time detailed feedback[2] (often

1. Indeed, the impact of feedback strongly depends on the context the feedback is given and its design. Trials show large differences in how much energy can be saved with feedback (see e.g. Darby, 2006; Darby et al. 2015; McKerracher and Torriti 2013). Based on a literature review, Menkveld et al. (2017) conclude that the savings percentages for HERs and ECMs, fed by data from the smart meter, can vary from 0 to 15 percent.

2. A few apps give real-time feedback, but then they also require an additional piece of hardware, and they typically also charge a subscription fee.

not more up to date than yesterday's usage or earlier), and they also require the user to actively seek access to the information. Information provision via email is also relative cheap, and it also has the disadvantage of just information about the (distant) past. In terms of the issue of the user needing to actively seek the information, e-mail messaging takes on an intermediate position; it requires more action than consulting the IHD (as one has to open an email), but less so than with for example an app (as people are likely to check their email regularly anyway).

To develop a proven effective energy policy, we need to estimate the effectiveness of each of the three ECMs with a robust method. That means that the impact evaluation needs to be implemented such that it is valid for the current policy context and that the observed effects can only be attributed to the ECM. Estimating an impact is always challenging because the counterfactual is typically unobserved; how much gas and electricity would a household have used had it not been provided with feedback on its energy use.

As already stated before, the most robust method to measure the causal impact of feedback on energy consumption is to randomly assign households to a treatment group (with receiving feedback as treatment) and a control group (not receiving feedback), and then compare the energy consumption of households between the two groups. With large enough sample sizes, random assignment ensures that the two groups are identical – in terms of their observable characteristics (like year of construction of the dwelling, and the number of household members), but also in terms of their unobservable characteristics (like the insulation status of the home, or the occupants' stance with respect to climate change). If the energy consumption of the treatment group differs from that from the control group, this difference can only be caused by the feedback given to the treatment group. However, implementing RCTs is challenging in the social sciences. As mentioned earlier, to get robust results the RCTs have to be implemented in it current policy context and also the size of the control and treatment groups are preferably large. McKerracher and Torriti (2013) show that new Trials with a better methodological design, higher number of participants and a better recruitment and selection of these participants, find smaller effects than older Trials. In their meta study Ehrhardt-Martinez et al. (2010) concluded that energy savings are lower if the intervention period lasts longer and the control and treatment groups are larger. They report savings by IHDs of between 3 and 5 % in pilots with larger sample sizes, while 6 to 10 % is found in smaller studies. However, because Ehrhardt-Martinez et al. (2010) and McKerracher and Torriti (2013) analyze the outcomes of both RCTs and observational studies, causality is not beyond dispute and hence that other factors may be at play too.

To measure the causal energy saving effect of the three ECMs we executed three RCTs, taking into account the Dutch context and focusing on the impacts on both natural gas and electricity.[3] Most Dutch households use natural gas for space heating and warming water, which is about twice the direct primary energy required to generate the used electricity. Below we briefly describe the three Trials.

Figure 1. The Geo-trio II, the In Home Display used for Trial 1.

TRIAL 1: IN HOME DISPLAY

First we discuss the measured effects of a relatively simple IHD, the Geo Trio II, which was new on the Dutch market at the time this study was implemented; see Figure 1. The display is a small colour-screen monitor which can be placed wherever it is convenient – e.g., in the living room or kitchen. The device displays the household's real-time energy consumption (both natural gas and electricity) in energy units or financial costs. A light on top signals the current level of energy use (green – orange – red). Historic consumption is also available via a graphical interface. The IHD is expected to help induce the household to save energy because high consumption (overall, or in peak periods) can be viewed as an indication of whether energy can be saved and also when, and hence also possibly how.

For the RCT we recruited more than 900 households in seven regions in the Netherlands. We advertised the possibility to receive an IHD, but also that not all applications could be granted (to be determined by a lottery). Applicant households had to be living in gas-heated houses with smart meters, and should give permission to collect their smart meter data. Identification of the treatment effect is by means of a random rejection design, via which half of the applicants receive a device, and the other half does not. The IHD was installed by an energy coach who briefly explained what information the IHD can provide and also how to operate it. The information was also provided in the form of a leaflet, to be consulted later if further questions would arise. We measured daily energy consumption using remote smart meter readings and estimated the impact on energy use by comparing the daily smart meter electricity and gas consumption data between the treatment and control group. For 802[4] households the energy usage data was collected between December 2017 and June 2020 for a period of between 7 and 18 months after the display had been installed. After the monitoring period we probed the impact of the device on behavioural factors using a survey for a better understanding of how the display affected behaviour. Households were invited by an email and the survey participation was incentivized with a lottery: twenty gift cards of 50 Euros were raffled off.

3. In this paper we do not give a complete description of the three Trials, due to space limitations. Extended descriptions of the Trials can be found in Boomsma (2021) and Van Soest and Vringer (2021)

4. For 124 households that participated in the experiment it was not possible to collect the energy data.

This led to a 55 % response rate (N=505)[5], with the response rate being slightly higher among households that did receive the IHD (61 %) than among those whose application was randomly rejected (48 %). The survey questions aimed to measure the household's knowledge of the largest contributors to their energy bill and household's knowledge of their monthly energy charges. For more detailed information see Boomsma (2021).

TRIAL 2: APP, GIVING HISTORIC FEEDBACK

For the second Trial we partnered with a medium-sized Dutch energy company to measure the impact of an app aimed at reducing household energy consumption. The app was free for the customers of the energy company and could be used on a smartphone or tablet. The app provided historic feedback (with a 24 hour delay) of the household's electricity and natural gas consumption. The information was provided both in physical units (kWh and m^3) as well as in monetary values – in amounts, but also in figures. The app estimated how much energy was consumed by which appliance, based on some household characteristics provided by the household itself, and its pattern of energy use. The app was thus able to provide detailed energy saving suggestions that were tailored to the household. We expected that the households with the app would open it at their device and receive the feedback shown and subsequently act on the suggestions to save energy.

The energy company used the app as a marketing tool, to differentiate itself from the other energy companies supplying to the Dutch market. Identification is by means of a so-called encouragement design. The app was available to all customers who had not installed it yet; the company typically advertised it to all of its clients. Per our request, the energy company sent out an additional email to a random subset of their customers to encourage them to install the app. If this email is able to increase the share of households installing the app above and beyond the share of adopters in the control group, any difference in outcomes between the two groups can only have been caused by a larger share of households using the app in the encouraged group. The difference then reflects the change in energy consumption by those households that adopt the app in the encouraged group, whereas they would not have done so had they not been sent the additional email.

We measured the impact on the energy consumption due to offering the app by using smart meter data of the daily gas and electricity consumption of 139,176 households over an 18-month period, from July 2018 to December 2019. Half of those households received an email at the end of June 2019 encouraging to download the app.[6] So we have information on energy consumption in the 12 months before and in 6 months after the invitation email was sent. In this way we can measure the average impact of sending an email to encourage to install the app (by comparing the means of energy consumption in the encouraged group to that of the control group). However, we do not have information on which or how much households installed the app. And we were not able to send a survey to probe the impact of the app on behavioural factors. An extended description of this Trial can be found in van Soest and Vringer (2021).

TRIAL 3: EMAIL WITH INFORMATION

For the third Trial we again partnered with a Dutch energy company. In this Trial we aimed to test whether financial information on a household's energy consumption can help induce energy savings. Almost all households in the Netherlands pay their energy costs in monthly instalments. The household's energy company makes a forecast on how much a customer is expected to pay for a whole year– based on the historic household's natural gas and electricity consumption, expected changes in energy prices and energy taxes, etc. The company then divides this amount by 12, which is the monthly instalment the customer needs to pay. Each year, actual expenses are compared to the sum of the instalments paid, and any difference is settled with the customer. Because it is easier for a company to pay a refund than to make the customer pay extra, the monthly instalment is typically set such that the sum of the instalments is slightly higher than the expected annual bill.

The energy company we partnered with for this Trial sends emails to its customers (with smart meters) with information how their current energy consumption compares to their monthly instalment. The company views this as a service to its customers, because this information can help to prevent them from having to pay extra at the end of the year – if their actual energy consumption is higher than expected. This information is provided by means of regular emails (typically bimonthly), and the email may also include a suggestion to adjust the instalment – to increase or to decrease it, depending on whether the company expects the household's bill at the end of the year to be higher or lower than the sum of the monthly instalments. There could be four different messages in the email, ranging from "[your] instalment is currently no longer in line with your energy consumption. It is possible that you will have to pay additionally when you receive your annual bill", to "[you] can lower your monthly instalment if you want to. With your current energy consumption you will probably get a refund when you get your annual bill". In the cases of smaller deficits the word "small" was added. The email also contained a recommended instalment; "[Your] instalment amount is now €XX. Increase/decrease your instalment to €YY, and it will match your energy consumption." Based on this information a household can decide whether it is necessary to take action. Finally the email included a link to the energy company's web page where customers can adjust their instalment and that web page also contain links to more detailed insight into the individual energy consumption, and a link to (general) energy saving tips.

We expected that the recipient household would open the email, take in the information and, if the signal is that their consumption is higher than expected, reduce their energy consumption to avoid paying extra when receiving the annual bill. This is not the only possible response, however; the household could also simply increase its monthly instalment. The customers who received the message that their energy consumption was below the instalments, may also respond in various ways

5. The survey included the full experimental population of 926 households, including the households of which there was no energy data.

6. With over 135,000 households in this Trial, a random attribution to the control and treatment group makes the probability very high that the composition of the two groups is nearly identical. To make the probability that this is indeed the case we applied a stratified randomization to make sure that the type of dwelling is comparable for both control- and treatment groups.

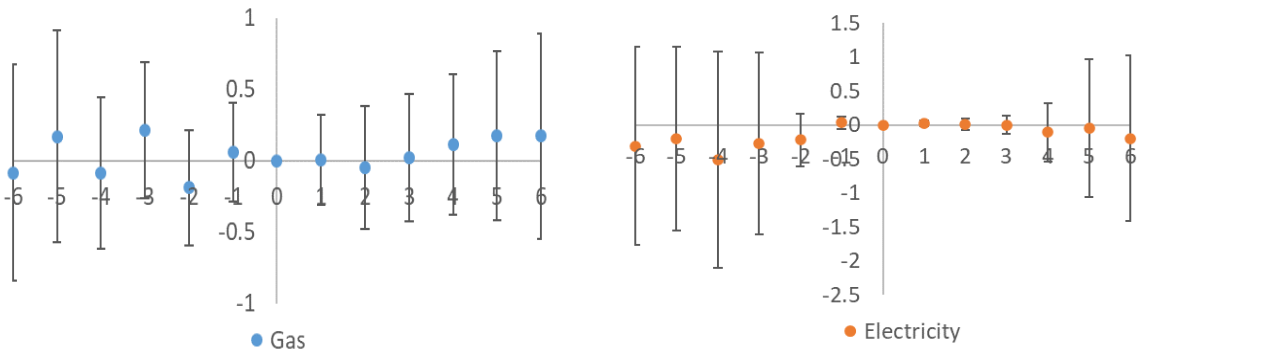


Figure 2. The difference in monthly gas and electricity consumption (measured in m3 and kWh, respectively) between households in the treatment group and those in the control group, in the six months before and in the six months after sending out the email inviting them to install the app.

– do nothing, reduce their instalment, or actually *increase* their consumption, if they would interpret the message that they were using relatively little energy as a 'license to consume more'. The energy company agreed to send a subset of their customers the email with information every two months (resulting in sending one to five emails); and no emails to a control group. Because we did not know which households actively opened the emails, the analysis is an intention-to-treat estimate (ITT) – the comparison of the average energy consumption between the treatment and control group. This comparison reflects the overall impact of sending the email, and thus the expected effect for a similar population to our research group, of sending out one or more emails.

We received energy consumption data (gas and electricity) from 138,596 households of which 70,032 were assigned to the intervention group, and 67,314 to the control group[7,8]. We followed their energy consumption for 15 months, partly before sending the emails[9]. After the monitoring period the energy company invited a part of the intervention and control groups to fill in our survey to probe the impact of the app on behavioural factors, comparable with the survey send to Trial 1 (IHD Trial). 24,819 households were invited by an email and the survey participation was incentivized with a lottery: twenty gift cards of 50 Euros were raffled off. This led to a 14 % response rate (N=3383), with the response rate being slightly higher among treated households (51 %) than among control households (49 %). For more information see van Soest and Vringer (2021).

7. We excluded 877 households from the analysis due to the fact that, at some point during the period they were followed, they had both negative consumption of electricity and gas. We also do not use the information on the 0.5 % of households with the very highest gas consumption, nor that of the 0.5 % of households with the very lowest gas consumption; we do the same for the consumption of electricity. This means that we analysed the gas consumption of 136,110 households, and the electricity consumption of 135,963 households.

8. With over 135,000 households in this Trial, a random attribution to the control and treatment group makes the probability very high that the composition of the two groups is nearly identical. To make the probability that this is indeed the case we applied a stratified randomization to make sure that the type of dwelling is comparable for both control- and treatment groups.

9. The monitoring period before and after the intervention emails differs per household. Also the moments of receiving the emails differ.

Results

RESULTS TRIAL 1: THE IMPACT OF THE IHD

Using a fixed-effects regression model with standard errors clustered at the household level,, we find that households that had randomly been endowed with an IHD in Trial 1 used 2.2 %less electricity and 6.9 % less natural gas than the control group. These impacts are significant at the five percent, or better (as the coefficient for electricity consumption had a p-value of 0.035, and that for gas consumption had a p-value of 0.014). That means that owning an IHD yields an average reduction of energy consumption of about 5 percent and financial savings of about 100 Euros per year. We did not find any differences between the groups with and without display concerning their attitude towards the energy transition or self-reported energy behaviour. The group with display did have some less general knowledge about the large energy use for heating for Dutch households, while their savings for gas could mainly be attributed to cold days. So, the established saving seems not to be driven by more general knowledge about energy use of households. But the group with display was better able to estimate their energy bill for natural gas and they also appreciated the display. The majority wanted to keep the display after the Trial. Also most of the displays were placed in the living or kitchen (85 %) and most of the respondents said they still watched regularly the display at the end of the monitoring period (85 %). This result can be interpreted as that the display makes the household energy use visible and keeps it visible.

RESULTS TRIAL 2: APP, GIVING HISTORIC FEEDBACK

With the app Trial we did not find evidence that sending the email with the invitation to install the app resulted in a change in the energy use. Neither on the short term, nor the longer term up to six months after sending the invitation. The dots in Figure 2 represent the average difference in energy consumption between the treatment and control groups, in the relevant month before or after the first email was sent. The vertical line segments in this figure represent the 95 % confidence interval – the minimum and maximum difference in energy consumption between which the true difference lies with 95 % confidence.

RESULTS TRIAL 3: EMAIL WITH INFORMATION

With the email Trial we did not find evidence that sending an email with information about the correctness of the monthly instalment leads to energy savings. On the contrary, we found that households who received one or more emails were going to use more natural gas and electricity. The dots in Figure 3 represent the average difference in energy consumption between the treatment and control groups, in the relevant month before or after the first email was sent. The vertical line segments in this figure represent the 95 % confidence interval – the minimum and maximum difference in energy consumption between which the true difference lies with 95 % confidence.

Most of the household in the treatment group did receive the message that they probably would get a refund when making up their annual bill[10], and their instalment could be lowered. Some households did actually lower their instalment, but in a small degree, on average about 1 %. It is assumable that most households concluded that with a small increase in their energy consumption they would not have to pay extra when the annual bill would be made up. And so, they did not have to be very restrictive by consuming energy.

The survey results show that the rise of the energy use can be attributed to a lack in energy saving measures for the households who did get the emails. However, they plan in a higher degree to take energy saving measures compared to the households who did not receive the emails. It is possible that sending the mails will result in energy saving on a longer term.

Further, the survey results did not give any indication that the emails led to changes in the general attitude towards energy saving or the energy transition. We also did not find evidence that the given information has lead to better informed customers or their ability to take energy saving measures. However, about ¾ of the households is satisfied with receiving the emails and want to keep receiving them.

COMPARISON OF OUR RESULTS TO THOSE OF OTHER STUDIES

As mentioned earlier there are nearly no robust effect studies for ECMs in the Dutch context. Below we first compare the results of our RCTs with earlier estimations for the Dutch context according to Menkveld et al. (2017), who did an extensive literature study.

The savings in natural gas consumption that we found for the IHD Trial (Trial 1) are in line with the savings predicted by Menkveld et al. (2017), but for electricity we document an effect which is just half of the predicted savings. Menkveld et al. (2017) based their estimates on two studies for Powerplayer and TOON, of which the TOON – an extended In Home Display – is on the Dutch market. The study for the effects of the Powerplayer was implemented among a relatively small sample of very motivated households, and the impacts were based on a before-after comparison (STEDIN, 2013). The TOON study (Ramondt, 2015) did have large samples, but the impact was assessed by comparing the energy consumption of households that did purchase the IHD to that of households that did not. As treatment status was not randomized, the impact estimate can be confounded by selection effects. So, neither STEDIN (2013) nor Ramondt (2015) are able to provide unbiased, causal evidence on the impact of the IHD that they studied. Having said that, it is interesting to note that the impact estimates of Ramondt (2015) – 2.9 % savings for electricity and 5.5 % for natural gas – are close to ours.

For apps Menkveld et al. (2017) estimated the effect to be 4 % savings for natural gas and 2 % for electricity. These estimates were based on the results of three studies that had been implemented in the Netherlands. Two of the studies were executed for real-time apps (Energiekrijgers and Ectual), which are rare on the Dutch market (RVO, 2020). However, also these two studies are not RCTs. In the Ectual study (Geelen et al. (2019), after randomisation of the control and treatment group, the households attributed to the treatment group had to make a decision to participate or not, resulting in a large drop-out. The control groups were not approached at all, so there were no dropouts. This led to a comparison between a self-selected treatment group and a non-self-selected control group. Geelen et al. (2019) did not find any saving effects for the Ectual app. For the Energiekrijgers study (Liander, 2013), an effect was reported for electricity (-3 %) and natural gas (-4 %). But this study was based on before-after measurements. The third study on which Menkveld et al. (2017) based their estimation was done for the Anna app (Rigo, 2019), which gives just historic feedback, comparable with the app used in our Trial 2. Rigo (2019) estimated savings of 2 % for electricity and no savings for natural gas. However, also this study was not an RCT because the treatment group had to sign up themselves, while the control group consisted of app-users who did use the app in the past and stopped their use.

For the email Trial (Trial 3) there is no comparable study about its effectiveness. So we are not able to compare our results with other studies.

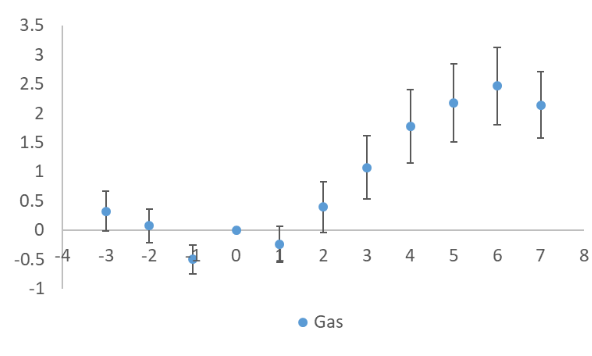
Menkveld et al. (2017) also provided an estimate for the effectiveness of web applications. They estimated an energy saving of 2 % for both electricity and natural gas. This estimate is based on one Dutch study for a web application that provided historic feedback (Sluis et al., 2011). However, Sluis et al. (2011) compared the energy use of a treatment group and a control group, but there was no randomised assignment of the participants to the groups. They compared current users of the web application with no web application users. Also, they did a survey before and after the monitoring period. So the measured difference in energy use between both groups cannot be solely attributed to the use of the web application.

In Table 1 we summarise our RCT results and the estimates made by Menkveld et al. (2017). We included the earlier mentioned RCT results for the Dutch Home Energy |report (HER) of Uitzinger and Uitdenbogerd (2014) and for the improved HER of Paradies et al. (2020). The estimates made by Menkveld et al. (2017) are based on non-randomised studies and where a comparison is possible, they deviate from the full RCT results in the field. As explained above, we consider the RCT results to be more reliable.

Discussion

As mentioned above, evidence-based energy policy requires prior testing using methods that can create a plausible and robust counterfactual. Our results show that energy savings of en-

10. In the monitoring period there was a relative warm winter in the Netherlands, while the monthly instalment is based on an average outside temperature.



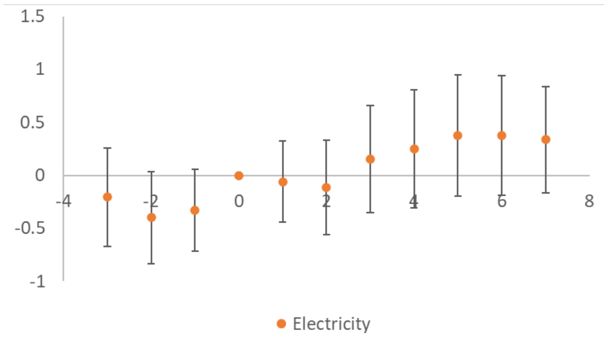


Figure 3. The difference in monthly natural gas and electricity consumption (measured in m³ and kWh, respectively) between households in the treatment group and those in the control group, in the three months before and in up to 7 months after receiving the first email.

Table 1. Final energy savings due to Energy Consumption Managers, estimation based on non-RCTs and based on RCTs, valid for the Dutch context.

ECM	Estimated savings[a] (in %) gas	electr	Number of ECMs in the Netherlands by end 2019[b]	Energy saving effects found with RCT Trials (in %) gas	electr.	Realised energy saving based on RCT trials (in PJ)
1. In Home Display	-5	-6	0.4 million	-7	-2	1.25
2. App – historic feedback	-2	-4	1.2 million	no saving evidence		-
3a. E-mail with information	n.a.		n.a.	no saving evidence		-
3b. Web applications	-2	-2	0,8 million	no study		-
4a. HER	n.a.		> 7 million	-0.9 [c]	0 [c]	2.65
4b. Improvement HER	-2.7	-2.8	> 7 million	0.0 [d]	0.0 [d]	0

[a] *Menkveld et al. (2017);* [b] *** RVO (2020);* [c] *Uitzinger and Uitdenbogerd (2014);* [d] *Paradies et al. (2020).*

ergy consumption managers based on available RCTs deviate from saving estimates based on non-randomised studies. This underlines the value for evidence-based policy of robust impact estimates by RCTs because we consider the RCT results to be more reliable. However, RCTs are in a lot of cases difficult (and also costly) to implement. It requires a lot of time to get a good control of the conditions and the necessary partners in the field should have a very high willingness to help implementing. Despite our effort to execute full RCTs of which impact results can solely be attributed to the interventions investigated, our Trials have also limitations. These are:

For the IHD RCT-Trial (Trial 1) we should note that the participating households were more than average interested in the IHD. This does not affect the study's internal validity (because this holds for both the control and treatment groups), but it may affect the external validity – the impact of the IHD may well be very different if offered to households that are less interested in technology or energy. It is possible that Trial 1 provides an overestimate because the participants were highly motivated to participate. But it is also possible that we underestimate the savings because our participants may have taken more actions to reduce energy consumption (like having installed insulation measures), and hence the potential for energy savings may have been lower. We have no conclusive indication which effect is larger, so we have no reason to believe that the relatively high motivation to participate affected the obtained energy savings. In fact, we think that our results are more reliable than those of Ramondt (2015), because in that study the changes in energy consumption of households who decided to purchase a TOON were compared to those of (probably less motivated) households that had not decided to purchase a TOON.

Although statistically the app Trial (Trial 2) had a good chance to measure an effect on the energy consumption, this Trial has some problems. The invitation by email was sent only once to the treatment group. And we did not receive any information how many customers installed and used the app. If there were only a few customers who did, it is possible that the effect per customer is considerable while we were not able to detect any energy saving on group level. This means that based on this Trial we cannot conclude that the app had no impact on energy savings. We can only state that sending an invitation to install the app did not lead to energy savings on group level.

Finally, Trial 3, with the email with information about the actual consumption versus the instalment, did not show energy saving effects. As mentioned above this intervention was not specifically targeted to induce energy savings; its main objective was to prevent households having to pay much extra if their current consumption was much higher than their monthly instalment.

Conclusions

Having insight into one's own energy consumption is seen as an important – though not strictly necessary – condition for energy savings. Feedback about energy use can be an effective and efficient way to stimulate energy users to take better decisions and to save energy. The aim of this study was to illustrate the importance of evidence based policy making, by getting robust insights into energy saving effects from Energy Consumption Managers (ECMs) in the Dutch context. We did this by measuring the causal effects with RCTs for three ECMs with a different design, available in the Netherlands; a simple In Home Display (IHD), an app and an email with information. Next we compared our results with other non-randomised studies.

The simple IHD we tested for Dutch Households saves 2.2 % electricity and 6.9 % natural gas. We found evidence that households reduced gas consumption due to being better informed of the (for Dutch Households) relatively high gas expenditures. Households did appreciate having the IHD, they wanted to keep the device after the monitoring period. We also found indications that households frequently consulted the device, which suggests that sustained feedback may have played a role in reducing energy consumption. These findings contribute to the recent literature that stresses the importance of concrete and relevant information to optimize the effect of energy consumption feedback. The relative high costs for the hardware will be earned back in less than a year on average.

In neither of the two Trials for the app and the email with information, we found evidence that offering energy consumption information leads to reduction in energy consumption. We found no evidence that broadcasting the invitation to install the app leads to a statistically significant change in energy consumption – neither in the short term, nor in the longer term (up to 6 months after the invitation to install the app). As for sending out emails with information about the correctness of the monthly instalment the household is paying, we even find that consumption of both natural gas and electricity increases as a result of receiving such information. Based on survey results for the email intervention, we found that offering the information prompted households to adopt energy-saving daily behaviours somewhat less strictly. Most households received information that their current energy consumption was relatively low to their monthly instalment; the households in question did slightly respond by reducing their monthly advance (by about 1 % of the monthly energy bill) as by, for example, turning up the thermostat a little higher, or defrosting the freezer just a little less often. On the other hand, we also found that receiving the emails slightly increased the willingness to take energy saving measures. So on the long term the email can still lead to energy savings. However, the overall conclusion is that informing households about their energy consumption does not necessarily lead households to reduce their energy consumption.

We conclude that for policy making the use of causal ex-ante impact analyses, like RCTs, implemented among the target population, is indispensable for effective, evidence-based policy making. A lot of (field)studies do not use a randomized attribution to control and treatment groups, which leads to energy saving results which cannot be solely attributed to the intervention. Non RCT Trials or Trials executed abroad can give results which deviate largely from real achieved energy savings locally. Based on the available RCT Trials, executed in the Dutch context, we were able to attribute an annual household energy saving of nearly 4 PJ to the use of energy consumption managers in 2019, which is slightly more than 1 % of the Dutch household consumption.

References

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. Journal of Environmental Psychology, 25 (3), 273–291.

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

Allcott, H. and M. Greenstone (2012), Is there an energy efficiency gap? Journal of Economic Perspectives, 26 (1), 3(28).

Allcott, H. and T. Rogers (2014), The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. Amersican Economic Review. Vol.104, No.10, October 2014. (pp. 3003–37).

Andor, Mark A., Andreas Gerster, Jörg Peters and Christoph M. Schmidt (2020). Social Norms and Energy Conservation Beyond the US. Journal of Environmental Economics and Management 103.

Attari, S.Z., M.L. DeKay, C.I. Davidson & W. Bruine de Bruin (2010), Public perceptions of energy consumption and savings. Proceedings of the National Academy of Sciences, 107 (37), 16054{16059. Retrieved from http://www.pnas.org/cgi/doi/10.1073/pnas.1001509107

Boomsma, M. (2021). On the transition to a sustainable economy: Field experimental evidence on behavioral interventions. CentER, Center for Economic Research. https://doi.org/10.26116/center-lis-2113

Brounen, D., N. Kok and J.M. Quigley (2013), Energy literacy, awareness, and conservation behavior of residential households. Energy Economics 38, pp.42–50.

Convenant (2017), Convenant 10 PJ energiebesparing gebouwde omgeving (https://www.rijksoverheid.nl/documenten/convenanten/2017/05/23/convenant-energiebesparing-gebouwde-omgeving).

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.

Darby, S., C. Liddel, D. Hills and D. Drabble (2015), Smart Metering Early Learning Project. Research conducted for Department Of Energy and Climate Change (DECC), Environmental Change Institute, Oxford , the University of Ulster and the Tavistock Institute.

Dromacque, Ch. & R. Grigoriou (2018), The Role of Data for Consumer Centric Energy Markets and Solutions. VaasaETT, Helsinki 11-12-2018.

Ehrhardt-Martinez, K., K.A. Donelly & J.A. Laitner (2010), Advanced metering initiatives and residental feedback programs: A meta review for Household electricity-saving opportunities. Report no E105. ACEEE, Washington, June 2010.

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

Geelen, D. R. Mugge and S. Silvester & A. Bulters (2019), The use of apps to promote energy saving: a study of smart meter-related feedback in the Netherlands. Energy Efficiency 12, 1635–1660. https://doi.org/10.1007/s12053-019-09777-z

Gerwen, R. van, F. Koenis, M. Schrijner and G. Widdershoven (2010), Intelligente meters in Nederland. Herziene financiële analyse en adviezen voor beleid. KEMA, Arnhem.

Gillingham, K. and K. Palmery (2014), Bridging the energy effiency gap: Policy insights from economic theory and empirical evidence. Review of Environmental Economics and Policy 8 (1), 18–38.

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

Liander (2013), Resultaten Alternatieve Feed Back Pilot. Energiekrijgers.

McKerracher, C., and J. Torriti (2013), Energy consumption feedback in perspective: Integrating Australian data to meta-analyses on in home displays. Energy Efficiency 6(2), 387–405. https://doi.org/10.1007/s12053-012-9169-3

Menkveld, M., M. Rietkerk, J. Mastop, C. Tigchelaar and K. Straver (2017), Besparingseffecten van slimme meters met feedbacksystemen en slimme thermostaten', Amsterdam, 5 april 2017. ECN Notitie N-17-017.

Paradies, G., L. Dreijerink and M. Menkveld (2020), Effectmeting verbeterd verbruiks en kostenoverzicht. TNO Rapport 2020 P10380.

PBL (2020), Klimaat en Energie Verkenning (KEV) 2020. Planbureau voor de Leefomgeving, Den Haag, 2020.

Ramondt, D. (2015), Energy savings from smart thermostats with energy displays. Working Paper, Vrije Universiteit Amsterdam. November 2015.

Rigo (2016), Verbruiksmonitoring met de slimme meter Een experiment in Rotterdam.

RVO (2020), Monitoringrapportage 2019 convenant Gebouwde Omgeving. Rijksdienst voor ondernemend Nederland.

Sluis, M. v.d., A. Hesselink & R. Jonkers (2011), Energieadvies op maat via gebruik van slimme meters. Determinanten- en effectonderzoek. Onderdeel van Projectencatalogus Energie-innovatie. ResCon, Amsterdam.

STEDIN (2013), Publieksrapport onderzoek powerplayer. STEDIN.

Uitzinger, J., and D. Uitdenbogerd, (2014), Monitoring en evaluatie van de slimme meter en het tweemaandelijkse verbruiksoverzicht. Amsterdam: IVAM rapport O-1417.

van Soest, D. and K. Vringer (2021), De invloed van energieverbruiksinformatie op energiebesparing: Effectonderzoek naar twee diensten. Planbureau voor de Leefomgeving, April 2021.

Vringer, K., M. van Middelkoop and N. Hoogervorst (2016), Saving energy is not easy: An impact assessment of Dutch policy to reduce the energy requirements of buildings. Energy Policy 93, 23–32.

Vringer, K. and T. Dassen (2016), De slimme meter, uitgelezen energie(k)? Achtergrondstudie, Planbureau voor de Leefomgeving, Den Haag, 17 november 2016.

Vringer, K., M. Boomsma and D. van Soest (2021), Energieverbruiksmanagers in Nederland. Energie besparen met de slimme meter, Den Haag: PBL.

Acknowledgements

The implementation of the three RCTs that we discuss in this paper required dedicated cooperation of many people and organizations. We would like to thank Ernestine Elkenbracht (Quintens), Thomas Dirkmaat and Jentse Hoekstra (Ministry of Economic Affairs and the Climate), Henk van Elburg (RVO), Stephanie Rosenkranz (Utrecht University), Michel Handgraaf (Wageningen University), Walter Ruijgrok (Energie-Nederland), Marijke Menkveld (TNO), Bart Rietveld, Allard Warrink, Nico Hoogervorst, Jetske Bouma and Jarry Porsius (PBL), Vattenfall and the member institutes of the Energy Display Alliance (Stichting !Woon (Amsterdam), municipality of Utrecht, de Woonbond with tenant organisations Albanianae (Alphen a/d Rijn) and MEVM (Assen), Energieloket AGEM (8 municipalities in the Achterhoek), Woningcorporatie U-Woon (Harderwijk and surroundings), Energierijk Houten).