NUDGE

Nudging consumers towards energy efficiency through behavioural science

Deliverable D1.1

Profiling of energy consumers: psychological and contextual factors of energy behavior

Authors: S. Van Hove, M. Karaliopoulos, L. Tsolas, P. Conradie, M. Amadori, I. Koutsopoulos, K. Ponnet

Project Coordinator: Filippos Anagnostopoulos

NUDGE has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 957012.

Project information

Project Title	Nudging consumers towards energy efficiency through behavioural science
Project Acronym	NUDGE
Project Number	927012
Project dates	September 2020 – August 2023

Rev.	Written by	Date	Checked by	Date
0.1	Merkouris Karaliopoulos	7/7/2021		
0.2	Stephanie Van Hove	14/7/2021		
0.3	Stephanie Van Hove, Peter Conradie, Koen Ponnet	28/7/2021	Filippos Anagnostopoulos	28/07/2021
0.4	Leonidas Tsolas, Merkouris Karaliopoulos	31/7/2021		
0.5	Stephanie Van Hove	6/8/2021		
0.6	Leonidas Tsolas, Merkouris Karaliopoulos	13/8/2021	Katharina Wohlfarth	14/8/2021
0.7	Stephanie Van Hove, Peter Conradie, Leonidas Tsolas, Merkouris Karaliopoulos	27/8/2021		
0.8	Merkouris Karaliopoulos, Leonidas Tsolas	3/09/2021	Stratos Keranidis, Heike Brugger	06/09/2021
0.9	Merkouris Karaliopoulos, Leonidas Tsolas, Stephanie Van Hove	13/9/2021		
0.95	Merkouris Karaliopoulos, Leonidas Tsolas, Peter Conradie, Iordanis Koutsopoulos	20/9/2021	Filippos Anagnostopoulos, Heike Brugger	21/9/2021 26/9/2021
1.0	Merkouris Karaliopoulos, Leonidas Tsolas, Peter Conradie, Manuela Amadori	28/9/2021		

Legal Notice

The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the European Union. Neither the EASME nor the European Commission is responsible for any use that may be made of the information contained therein.

NUDGE

Nudging consumers towards energy efficiency through behavioural science

About

Efforts to induce energy-friendly behaviour from end-users through behavioural interventions are characterized by a lack of customer personalization ("one-size-fits-all interventions"), a partial understanding about how different interventions interact with each other and contrasting evidence about their effectiveness, as a result of poor testing under real world conditions.

NUDGE has been conceived to unleash the potential of behavioural interventions for long-lasting energy efficiency behaviour changes, paving the way to the generalized use of such interventions as a worthy addition to the policy-making toolbox. We take a mixed approach to the consumer analysis and intervention design with tasks combining surveys and field trials. Firmly rooted in behavioural science methods, we will study individual psychological and contextual variables underlying consumers' behaviour to tailor the design of behavioural interventions for them, with a clear bias towards interventions of the nudging type.

The designed interventions are compared against traditional ones in field trials (pilots) in five different EU states, exhibiting striking diversity in terms of innovative energy usage scenarios (e.g., PV production for EV charging, DR for natural gas), demographic and socio-economic variables of the involved populations, mediation platforms for operationalizing the intervention (smart mobile apps, dashboards, web portals, educational material and intergenerational learning practices).

The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957012.

Project partners

IEECP – Institute for European Energy and Climate Policy
domX
Spring-Stof
imec
ZEZ – Zelena Energetska Zadruga
beegy – the energy manager
Active Citizenship Network – Cittadinanzattiva
Fraunhofer ISI
inegi – driving science & innovation
Athens University of Economics and Business

Table of Contents

List of Figures

List of Tables

Executive Summary

Within this report, the primary aim has been to generate a better understanding of energy consumers' behaviour in relation to energy efficiency and to further explore factors that either serve as barriers or facilitators to reduce energy consumption. To this end, in this report we invested effort in two main research activities. The first one consists in analyzing which factors can effectively predict the energy consumers' intent to reduce energy consumption. In the second one, we sought to identify distinct groups of energy consumers according to a range of psychosocial factors that modulate energy-saving behaviour in positive or negative manner. These two activities complement each other in unravelling the determinants of energy consumers' behaviour and yield critical insights for one of the milestone tasks of the overall project, i.e., identifying targeted interventions, primarily of the nudging type, that could leverage those determinants and achieve behavioural change.

Input for both activities were data collected through a Europe-wide online survey (n=3129, after pre-processing checks and filters were applied). The survey was made available in 15 languages and was completed by persons in 29 countries, as a result of a detailed dissemination strategy that involved several stakeholders of the public, private and civic sector. It broke fresh ground in the study of energy-related behaviour by operationalizing three theoretical models of human behaviour: the Theory of Planned Behaviour (TPB), the Value-Belief-Norm theory (VBN) and the Prototype Willingness model (PWM). Measuring the fifteen (15) psychosocial constructs that resulted from the synthesis of these 3 models was the subject of one of the five survey modules, the core one. Four more modules logged respondents' sociodemographic indicators, their residence properties, their current energy-saving behaviour in a number of activities, and their attitude against energy monitoring and control platforms, respectively.

The first level of data analysis concerned the derivation of descriptive statistics. With regards to energy-saving behaviour, we found that closing the windows when heating occurs frequently, with 87.2% state often/always. Water conservation is often related to the duration of hot water use, with people stating that they do not leave hot water running, or take shorter showers (respectively 70.1% and 45.6% often/always). Also notable is that turning down the lights is the most common saving behaviour, with 91.5% indicating that they often/always perform this action.

Our questions about data sharing preferences of energy consumers also revealed some interesting results, with the majority of our sample reporting unwillingness to share detailed energy data, i.e., on, at least, daily and up to real-time basis. Whereas 42.5% stated that monthly energy data is not shareable data, this proportion increases to 44% for daily and real-time data. Furthermore, people in our sample are more prone to share all types of energy data with their family members (67% real-time data and 90.3% monthly data).

Proceeding with the main analysis, in a first step, we explored the predictability of users' intentions to reduce energy consumption. We built a theoretical model, where we differentiate between specific intentions to reduce heating-related consumption and more general intentions towards energy-saving. Our analysis showed that the degree to which people perceive that they have the ability to conserve energy is an important consideration, above and beyond any considerations

such as environmental concern. We further found strong statistical evidence for the importance of subjective norms as antecedent for the intent to reduce consumption, pointing to the importance people attach to their neighbours and peers' attitudes as regulators of their own energy consumption. Overall, we found support for providing consumers with practical ways to reduce energy consumption.

In a second step, the goal was to identify groups (interchangeably: classes or clusters) of energy consumers with distinct characteristics that facilitate the selection of (nudging type) behavioural interventions for them. To this end, we sought to leverage the rich information about the respondents' psychosocial characteristics in the survey data, experimenting with two different approaches. The first one was based on clustering analysis, a common technique for statistical data analysis that systematically seeks to organize a set of objects into a number of groups (clusters) so that objects in the same cluster are more similar to each other, according to some criteria, than to objects in other clusters. Despite our experimentation with a range of clustering algorithms and trying many different parameterizations, including feature selection and transformation techniques, the obtained clustering structures routinely shared the same pattern: there would always be one cluster that scored top in all 15 constructs, one that would score second best in all of them, one that would score third best in all of them and so on. Namely, no groups with profound differentiation across the 15 features could be identified this way.

The second approach involved the a priori specification of energy consumer classes as combinations of conditions that the aforementioned constructs should satisfy. The acceptable value ranges for each energy consumer class involved several threshold values, which were determined as the solution of a non-linear optimization problem. We ended up with six distinct energy consumer classes: *Environmentally conscious and well-informed energy consumers*, *Concerned but comfort-oriented energy consumers*, *Concerned but lacking awareness energy consumers*, *Materialistic energy consumers escaping personal responsibility*, *Prone to social influence energy consumers*, and *Indifferent energy consumers.* The first group represents ideal energy-savers, whereas energy consumers in each of the other five groups share one or two distinct features that either serve as barriers towards their energy-saving intentions or prescribe specific type of intervention for strengthening these intentions. The within-group variation of the socio-demographic indicators (age, gender, education level) rather resembled the respective distribution across the full dataset, with a few rather mild exceptions.

This report is organized into four main chapters and a number of annexes. Chapter 1 provides a brief introduction. In chapter 2, we present the human behaviour theories underpinning the survey, the underlying research hypotheses and the structure of the survey into five distinct modules collecting different information about the respondent (residence properties and energy efficiency, current energy-saving activities, psychosocial constructs, attitude against energy monitoring and control platforms, and socio-demographic information). We also present statistical information about the survey respondents as well as the responses that were filtered out of the final dataset according to different criteria. Chapter 3 provides descriptive statistics out of the survey responses such as housing characteristics and the users' current energy-saving activities, looking also into regional differences across different European countries/regions. Chapter 4 presents findings relevant to the one of two main goals of this deliverable, i.e., the discovery of attitudinal and behavioural predictors of intent to reduce energy consumption, whereas Chapter 5 presents the outcomes of the experimentation with the two approaches towards the segmentation of energy consumers, the clustering and the classification approach.

These four main chapters are complemented by six annexes. Two of them (Annex I and II) provide additional results for Chapter 4 and 5, respectively; Annex III presents a brief review of other energy consumer segmentation studies in literature; Annex IV provides a table with statistical terms used in the document; Annex V reports material used in the context of raising awareness about the online survey; and, finally, Annex VI presents the full survey (question items as posed to the survey participants.

1. Introduction

Within the NUDGE project, our overall goal is to explore which (technical) interventions can be used to change domestic energy consumption behaviour, without the use of financial incentives, i.e.: through so-called nudges. Defined as any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any option or significantly changing their economic incentives (Thaler & Sunstein, 2008), nudges build on top of existing theories of behaviour, behavioural change and behavioural intent.

To this end, within NUDGE, a large-scale survey, available in 15 languages, was developed, with the overarching aim to find factors that can predict intent to reduce the use of energy. Centrally, the survey draws on three related but distinct models of human behaviour namely Theory of Planned Behaviour (TPB), the Value-Belief-Norm theory (VBN) and the Prototype Willingness model (PWM). More generally, our survey also looks at energy saving behaviours and willingness to share data about consumption with others.

Beyond exploring the predictors of energy efficient behaviour, we additionally use these three behavioural models as base to classify consumers into different classes with regard to their attitudes towards energy saving behaviour, which subsequently can be translated into nudges of interest.

While different theoretical models such as TPB, PWM and VBN have previously been successfully used in a variety of studies to assess sustainable behaviour (i.e.: TPB for recycling (Tonglet et al., 2004), PWM for environmentally friendly behaviour in general (Ratliff et al., 2017) and VBN for sustainable consumption (Thøgersen & Ölander, 2002)), our survey is the first – to our knowledge – to combine these models in one instrument. Of additional interest in our survey is the inclusion of other variables that further explore the rationale for energy saving such as environmental concern, awareness of energy consumption and the fear of losing comfort.

Moreover, our survey focusses specifically on heating, given its large share of energy related consumption in the EU[1], which at 63,6% compromises a majority of energy consumed in EU households, suggesting that improvements in this domain can have a significant impact on energy saving.

1 https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households

2. The NUDGE survey

As noted, our survey has two central objectives. Our first aim is to be able to measure intent to reduce heating related energy consumption in the winter. Our basis for this analysis is derived from existing models of behaviour, acting as a subsequent basis for our cluster analysis. We discuss these variables in more detail below, along with all our research hypothesis.

2.1. Theoretical framework and research hypotheses

Our survey operationalises three more general models of human behaviour: the Theory of Planned Behaviour (TPB), the Value-Belief-Norm theory (VBN) and the Prototype Willingness model (PWM). Moreover, we account for several covariates for the attitudinal component of the Theory of Planned Behaviour.

2.1.1. Theory of Planned Behaviour

Ajzen's TPB (Ajzen, 1991), is a theoretical framework aiming to explain people's behaviour and choices. TPB suggests that certain motivational factors, including attitudes toward a behaviour, subjective norm and perceived behavioural control impact people's intention to adopt a behaviour. Intent, in turn, can be associated with actual behaviour (Ajzen, 2002).

The attitudinal component refers to someone's evaluation or appraisal of an activity. Subjective norm refers to the perceived social pressure to engage in the activity, while perceived behavioural control is the perceived difficulty to perform the activity. TPB has been widely used in studies of sustainable behaviour. This includes recycling (Tonglet et al., 2004), food consumption (Olsen et al., 2008), energy consumption (Gadenne et al., 2011), household energy self-sufficiency (Engelken et al., 2018), or investment in community owned renewable energy sources (Proudlove et al., 2020).

In the following paragraphs we will briefly motivate our hypotheses based on previous research. According to TPB (Ajzen, 1991) (Abrahamse & Steg, 2011), a person's attitude towards a specific behaviour predicts their intent to engage in that behaviour. As a result, we propose that:

H1: Positive attitudes towards lowering the temperature in winter will be associated with a higher intent to reduce energy consumption by lowering the temperature.

Within the TPB modelling framework, subjective norm denotes the effect social pressure has on behavioural intention. In the case of energy reduction, we thus hypothesise that:

H2: There is a positive association between subjective norm and intention to reduce energy consumption by lowering the temperature in winter.

The third component of the TPB framework is the perceived behavioural control. According to Ajzen (Ajzen, 1991), someone's perceived ability to engage in a behaviour explains her intent to do so.

H3: There is a positive association between perceived behavioural control and intention to reduce energy consumption by lowering the temperature in winter

Finally, the intent to reduce energy consumption will be positively associated with actual energy consumption reduction behaviour:

H4: There is a positive association between intent to reduce energy consumption in winter by lowering the temperature and lowering the temperature in winter to conserve energy.

2.1.2. Antecedents of Attitude

While TPB by itself provides valuable insights into intent (and actual behaviour), extending the model with additional domain-related predictors offers a more in-depth view of how intent is influenced within the context of energy conservation (Perugini & Bagozzi, 2001). To this end, we introduce the following predictors in our model, specifically related to attitude.

Financial concern

In general, persons with high financial concern are more likely to have positive attitudes towards energy saving (Karlin et al., 2014; Long, 1993; Van Raaij & Verhallen, 1983). Financial concern is regarded as the extent to which people's decisions are motivated by economic considerations (Chen, Xu, Day, 2017).

H5: Financial concern is positively associated with the attitude towards reducing the temperature to conserve energy in winter.

Loss of comfort

The perceived risk of losing comfort due to energy-saving activities is negatively associated with the attitude towards reducing energy consumption (Wang et al., 2011, 2014).

H6: Perceived loss of comfort is negatively associated with the attitude towards reducing the temperature to conserve energy in winter.

Energy awareness

People with a high degree of knowledge about energy and its use will be more likely to have a positive attitude towards reducing their energy consumption (Van Raaij & Verhallen, 1983; Wang et al., 2014; Yoo et al., 2020)(Wang et al., 2011, 2014).

H7: Energy knowledge is positively associated with the attitude towards reducing energy consumption by lowering the temperature in winter.

2.1.3. Value-Belief-Norm Theory

A further component in our model originates from the Value Belief Norm theory (VBN-theory) (P. C. Stern et al., 1999). Conceptualized by Stern (P. Stern, 2000), the VBN-theory specifically considers sustainable behaviour. In summary, choosing sustainable behaviour can be the result of a perceived obligation to act environmentally consciously. It operationalises awareness of consequences, ascription of responsibility and personal moral norms, the latter being associated with intent. Given its conception as a model to explain sustainable behaviour, the VBN-theory has seen wide adoption in a variety of sustainability domains. This includes sustainable modes of travel (Lind et al., 2015) or sustainable consumption (Thøgersen & Ölander, 2002).

Household energy conservation is framed in terms of a social dilemma (Samuelson, 1990). The dilemma arises from the conflict that exists between individual and collective outcomes of energy conservation behaviour. While it would seem that energy use has many individual benefits (e.g., increased comfort and well-being), the negative side of the equation is, however, that it can lead to negative environmental consequences (e.g., depletion of energy sources, environmental degradation, carbon emissions) (Abrahamse, 2007). A more integrated approach, which considers both the individual and environmental beliefs, overcomes the weaknesses of models that only take into account self-interest variables (e.g., TPB) or only pro-social motives (e.g., Norm Activation Model, Value-Belief-Norm) (Gao et al., 2017).

Awareness of consequences and ascription of responsibility
Both the awareness of the consequences of high energy consumption for society and ascription of responsibility will have a positive impact on people's moral norm (Abrahamse, 2007; Abrahamse & Steg, 2011; Fornara et al., 2016; Guo et al., 2018)

H8a: Awareness of consequences is positively associated with personal moral norms.
H8b: Ascription of responsibility is positively associated with personal moral norms.
(Abrahamse et al., 2009; Abrahamse & Steg, 2011)

Persons who feel a moral obligation to act in pro-environmental ways will have stronger pro-environmental intentions and behaviours (Abrahamse, 2007; Abrahamse & Steg, 2011; Fornara et al., 2016; Song et al., 2019; Steg et al., 2005; Wittenberg et al., 2018)

H9: Personal moral norms are positively associated with intent to reduce energy consumption by lowering the temperature.

2.1.4. Prototype Willingness Model

Alongside TPB and VBN-theory, we take into account a third theoretical model, the Prototype Willingness Model (PWM) (Gerrard et al., 2008). Originally developed as a model to assess health-related risk behaviour in adolescents (Gerrard et al., 2008), PWM argues that behaviour isn't necessarily planned, as proposed by the TPB model. Rather, PWM suggests that behaviour is often a result of *risk conducive social situations* (Gibbons et al., 2020). PWM consists of two pathways, one based on reasoning (reasoned pathway) and another based on social reaction (social reactive). Prototype, in the context of PWM, also refers to images of people who engage in a particular (risky) behaviour (Gerrard et al., 2008). TPB also emphasises behavioural intent, while PWM focuses on behavioural intent and behavioural willingness (i.e., what one is willing to do versus what one is planning to do).

Although the PWM model is originally used to assess risky adolescent behaviour, it has also been used to assess positive health-related behaviour such as exercise (Rivis et al., 2006). It has also been used in tandem with TPB (Rivis et al., 2006)(Van Gool et al., 2015). Finally, and of special interest for this study, PWM has been used to assess sustainable behaviour such as cycling (Frater et al., 2017) or general environmentally friendly behaviour (Ratliff et al., 2017).

H10: Prototype similarity is positively associated with the willingness to reduce energy consumption behaviour by lowering the temperature in winter
H11: Prototype favourability is positively associated with the willingness to reduce energy consumption behaviour by lowering the temperature in winter.

Finally, following Gerrard et al. (Gerrard et al., 2008) , willingness can be associated with intention to reduce energy consumption.
H12: Willingness is positively associated with the intention to reduce energy consumption behaviour by lowering the temperature in winter.

2.1.5. Control variables

Finally, our model contains the following control variables: Gender, Country of Residence, Income, Age and Level of Education. Pending significance of these variables, they will be excluded from our structural equation model, shown as a whole in Figure 1.

2.1.6. Conceptual relationship to nudging

Our survey as a whole does not specifically ask questions about particular nudges or nudge techniques. This is a deliberate decision since nudges are typically evaluated experimentally, as opposed to through a survey. However, despite the lack of explicit nudging assessment, the different theoretical behaviour models presented in our survey do capture and align with several types of nudges.

Specifically, loss aversion has practical links with 'financial concern' and 'awareness of consequences'. Individuals scoring high on financial concern also tend to focus on the possible costs instead of the pleasure of gains. By means of fear and confronting nudges, namely emphasising energy losses or the nearby and immediate impact of excessive energy consumption, can be responded to the loss aversion bias.

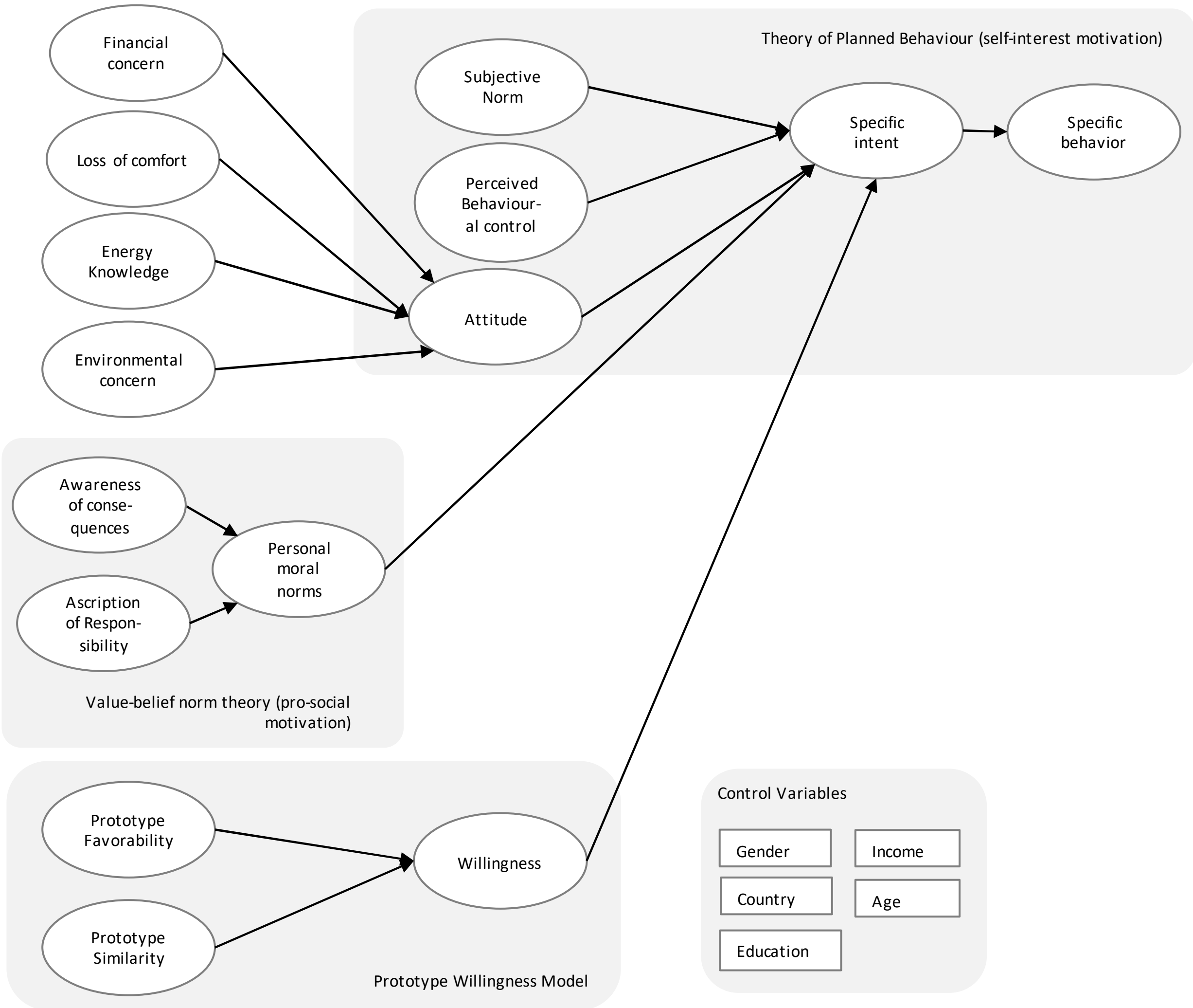


Figure 1: Complete research model with elements from three theoretical frameworks, the TPB model, the VBN-theory and the Prototype Willingness model and the control variables under consideration.

Further, the availability bias draws on readily available information in memory, which is related to one's energy knowledge (Caraban et al., 2019). Home audits, where households receive personalized information about energy-saving measures, capitalize on the availability bias and result in increased energy conservation knowledge (Winett et al., 1985).

Also, the confirmation bias or the tendency to be attentive to belief-conforming information (Yoo et al., 2020) has been incorporated in our survey using the concept of environmental concern. Being conscious of the natural environment might also influence how information is searched and selected that fits one's beliefs.

Further, the herd instinct bias refers to people's tendency to copy others' actions (Caraban et al., 2019). By means of the subjective norm concept we measured people's susceptibility to others' preferences and support for behaviour, which might be indicative of people's tendency to social conformity (Abrahamse et al., 2007; Staats et al., 2004). Furthermore, the status quo bias shows linkages to the concept of 'perceived behavioural control'. This is the extent to which individuals feel capable of performing a particular behaviour, such as conserving energy (Abrahamse, 2007). Facilitating nudges, which provide practical information about energy conservation, trigger an automatic response or make alternatives more salient (Caraban et al., 2019) cater to the status quo bias.

2.2. Survey structure

In this section, we provide an overview of the survey structure and its according modules. The complete survey can be retrieved in Complete Survey.

2.2.1. Residence properties and energy efficiency

The goal of the first module was to collect general information on the physical characteristics of people's main residence, its energy efficiency, and possible energy production and consumption facilities. The country of residence variable originally consists of 15 countries but the open textbox, letting people state other options, has resulted in an additional 14 countries. After data cleaning and having coded answers in the 'other' textbox, participants from 29 different European countries remain. Given the small sample sizes in some countries (i.e., Austria, Finland), we recategorized our sample to larger regional categories according to the United Nations Geographical Scheme2. The only exception is Malta, which has been added to the Southern Europe region.

The question addressing house surface was informed by the H2020 Penny project and addresses four categories, I.e., less than 20, ... living area (m²), 400 or more, and I don't know. It applies to the total living area of the dwelling and not the per person space. The answers have been recoded into 10 categories with a fixed range of 50 except for the first two and last categories, i.e. (1) less than 20, (2) 21-50, (3) 51-100, ..., (9) 351-400, (10) more than 400. To some questions (such as heating source, renewable energy system, energy monitoring system, and thermostat system) the 'other' category has been added, which provided respondents with the opportunity to enter a unique answer if it was not part of the predefined categories. These descriptive answers have been re-coded, and if they were deemed substantial, additional categories have been added to the original variable. Respondents estimated the energy performance of their residence on a colour scale ranging from green to red, which resulted in a percentage estimate of the energy-efficiency of the respondent's residence.

2 https://unstats.un.org/unsd/methodology/m49/

2.2.2. Current state of users' energy-saving activities

The second module assessed the "actual" energy-saving behaviour of respondents. By evaluating the current state of energy-saving behaviour, we can find out where there is still room for improvement and which behaviours can be capitalized by nudging mechanisms.

The module is further structured into four distinct blocks grouping statements about energy-saving activities. The first block asks questions related to saving energy related to heating, i.e.: *Wearing more clothes instead of turning the heating up.* The second block asks questions about savings that take place in the bathroom, relating most significantly to warm water use: *Preferring a shower over bathing*. Our third block focusses on saving behaviour in the kitchen *Only using dishwasher when fully loaded.* Finally, our fourth block addresses miscellaneous other saving behaviours un and around the home: *Switch of the TV when no-one is watching*. Frequencies had to be indicated on a 5-point Likert scale from 'Never (before)' to 'Always'. If an activity was not possible in one's household (e.g., because an air-conditioning system, a dishwasher, a tumble drier, or an electric vehicle were not available), respondents had to indicate 'not applicable'. Additionally, respondents were asked to indicate their *perceived impact* on a ladder with 9 steps ranging from '1 = not at all energy conscious and relative high energy bills' to '9 = very energy conscious and relative low energy bills'.

2.2.3. Measuring constructs

The third module consisted of a series of attitudinal, motivational and behavioural constructs measuring the underlying theoretical model. In essence, it entails the above introduced Theory of Planned Behaviour (Ajzen, 1991), the Value-Belief-Norm Theory (P. C. Stern et al., 1999) and the Protype Willingness Model (Gerrard et al., 2008). In this module, participants were asked to imagine a concrete energy-saving action, i.e., '*saving energy by lowering the temperature setting in winter*' with accompanying questions about their attitudes, subjective norms and perceived behavioural control about this action.

This energy saving activity has been determined based on its prevalence among Europe and its substantive impact on energy conservation. Moreover, the more tangible the situation, the better respondents can assess their according behaviour in that particular situation. The TPB model prescribes that attitude, subjective norm, and perceived behavioural control determine an individual's degree of determination and willingness to undertake a particular activity (Ajzen, 1991). Given the specificity of our construct for intent (i.e.: savings related to heating), it isn't clear that our results can be generalised. However, as discussed below, we also include a construct that measures general intent to save energy at home, allowing estimations of how specific and general intent relate.

All TPB-constructs, except attitude, have been measured on a 5-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree.

- **Attitude (abbr.: ATT)** is the degree to which a person evaluates (un)favorably a particular behaviour. Exceptionally, it was measured with five 7-point semantic differential scales, which have been informed by general and topic-related research: *useless – useful, foolish –*

wise, disadvantageous – advantageous, ineffective – effective, dull – interesting (Webb et al., 2013).

- **Subjective norm (abbr.: SN)** is an individual's assessment of others' preferences and support (Ajzen, 1991). It consisted of four items, e.g., '*Most people who are important in my life would approve that I save energy by lowering the temperature setting in Winter*'.
- **Perceived behavioural control (abbr.: PBC)** is the degree to which a person feels capable of performing a particular activity (Ajzen, 1991) and was measured by three items. An indicative item is '*I have the capabilities to save energy by lowering the temperature setting in Winter*'.
- **Behavioural intent (abbr.: INT)** is an individual's degree of determination and willingness to perform an activity, here saving energy at home (Ajzen, 1991). This variable has been measured at both general (abbr.: INT_GEN) and specific (abbr.: INT_SPEC) level. Both intent constructs consisted of three items with '*I intend to save energy at home/by lowering the temperature setting in winter*' as an exemplary item.

The value-belief-norm theory (P. C. Stern et al., 1999) states that people engage in a given pro-environmental behaviour, as they feel the moral obligation (i.e., moral norm) to behave properly if they feel responsible (i.e., ascription of responsibility) for the impact of their actions on the environment (i.e., awareness of consequences). All VBN-constructs have been measured on a 5-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree.

- **Moral norm (abbr.: PERS_NORM)** was covered by three items (Abrahamse, 2007) with '*I feel morally obliged to reduce my energy use, regardless of what other people do*' as one of the items.
- **Ascription of responsibility (abbr.: ASCR_RESP)** was addressed by three items (Abrahamse, 2007). One of the items was: '*I take joint responsibility for the depletion of energy resources*'.
- **Awareness of consequences (abbr.: CONSEQ_AWARE)** was measured by three items: '*Energy conservation contributes to a reduction of global warming*' (Abrahamse, 2007); '*The increasing energy demand is a serious problem for our society*'; '*The increasing shortage of energy sources is a serious problem for our society*' ().

The Prototype-Willingness model assumes that individuals have clear social images of a person their age who engages in a given activity. The degree of liking (i.e., prototype favourability) and the degree of similarity to oneself (i.e., prototype similarity) determines one's willingness to engage in that particular activity (Gerrard et al., 2008). Respondents were asked 'to think about someone who saves energy by lowering the temperature setting in winter'.

- **Prototype favourability (abbr.: PROT_FAV)**. Respondents were asked to rate the favourability of the energy-saver persona on a 5-point scale (1 = not at all to 5 = totally) using five adjectives: *conscious, progressive, smart, green, responsible* (Van Gool et al., 2015).

- **Prototype similarity (abbr.: PROT_SIM)** was assessed with four items (Elliott et al., 2017) on a five-point scale. An example of this construct's items is '*Do you resemble the typical person who saves energy by lowering the temperature setting in Winter?*' (1 = no to 5 = yes).
- **Willingness (abbr.: WILL)** was measured by asking responders to specify how frequently they perform four specific actions (Frater et al., 2017) on a 5-point Likert scale from 1 = Extremely unlikely to 5 = Extremely likely. Situations: '*You lower the temperature setting in all unused rooms when you are at home all day*'; '*You lower the temperature setting when you leave home*'; '*You keep the doors closed to prevent heat loss*'; '*You go to sleep and you lower the temperature setting*'.

Finally, since previous research has established the relation between attitude and energy-saving on the one hand, and financial concern (Karlin et al., 2014), loss of comfort (Wang et al., 2014), energy knowledge (Han & Cudjoe, 2020; Wang et al., 2014) and environmental concern (Karlin et al., 2014; Taufique & Vaithianathan, 2018) on the other, the following variables were also included.

- **Financial concern (abbr.: FIN_CONCERN)** (Chen et al., 2017) e.g., '*I pay attention to energy-saving tips to reduce my electricity bills'.*
- **Loss of comfort (abbr.: LOSS_COMFORT)** (Abrahamse, 2007; Sütterlin, B., Brunner, T. A., & Siegrist, M, 2011), e.g., '*Energy conservation means I have to live less comfortably'.*
- **Energy knowledge (abbr.: ENERGY_AWARE)** (Dianshu et al., 2010; Wang et al., 2014), e.g., '*I know energy-saving methods well'.*
- **Environmental concern (abbr.: ENV_CONCERN)** (Kilbourne & Pickett, 2008), e.g., '*I am very concerned about the environment*'.

Each of these variables consisted of three items and was measured on a 5-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree.

2.2.4. Energy monitoring and control platforms

A fourth module explored the potential of energy platforms that provide real-time energy monitoring but also control and automate energy flows. This module consisted of the following two question blocks:

- The first question assessed the interest in seven types of energy data, depending on data accuracy. The energy data types ranged from 'average monthly energy usage' to 'real-time usage of many appliances'. Interest had to be indicated on a 5-point Likert scale ranging from 1 =None at all to 5 = A great deal.
- The second question evaluated whether people were willing to share these energy data types with six particular parties: family members, neighbourhood, energy provider, energy distributor, third parties and the government.

2.2.5. Socio-demographic indicators

A fifth and last module included socio-demographic indicators such as gender, age, household type, household composition, educational attainment, career status, and income. More specifically:

Gender: To indicate gender, participants were asked to not state what their gender is on their national ID or passport.

Income: Income was presented as 15 categorical options, ranging from below €501 per month, €501 - €1000, €1001 - €1500 etc., to above €7000. Two more options, 'No answer' and 'I don't know' was also provided. It was explicated that *income* denotes the total net household income in 2020.

Education: The module distinguished between seven ordinal categorical levels: None, Primary education, Lower secondary education, Upper secondary education, Bachelor's or equivalent level, Master's or equivalent level, Doctoral or equivalent level.

A complete overview of all demographic factors can be found in our survey, in Annex VI

2.3. Sample of participants and data pre-processing

2.3.1. Recruitment of participants

To maximixe the citizens' response to the onlne survey, the project came up quite early with a concrete survey dissemination strategy that involved different external stakeholders, both private and public. Cittadinanzattiva were the primarily responsible to plan and carry out this strategy.

In particular, contacts were made with a number of civic and consumer organizations across Europe that have shown interest in the issues of sustainability and energy transition and the overall purpose of the Nudge project such as the Croatian Association for Consumer Protection (HUZP - Croatia), the Center for education and informing consumers (CEIP - Croatia), the association InfoCons (Romania), the Union of Working Consumers of Greece (EEKE - Greece), the Association of Consumers Organizations in Slovakia (Slovakia), the association Talented Borders (Latvia), the Confederation of Consumers and Users (Spain), IFOK (Germany), the National Association Saugok Save (Lithuania), Lithuanian Consumer Association (Lithuania), Indecosa (France), the Consumers Association of Malta (Malta), Association for Consumer Rights (Malta), Consumur (Spain); Social-Mentes Canarias (Spain); the Slovene Consumers' Association (ZPS – Slovenia), UNWE ECO Club from the University of Sofia (Bulgaria), and the European Consumer Union through its member organizations. These associations were actively involved in raising awareness about the survey and disseminating the link to the online location of the survey through newsletters, their Web and social media pages, and emails to their members.

These associations, in collaboration with the Consortium partners, had the primary role in translating the survey in their native language. As a result, the survey was available online in the following languages: English, Dutch, French, Italian, Portuguese, Croatian, Greek, German, Lithuanian, Latvian, Romanian, Slovenian, Slovak, Spanish, and Bulgarian.

Furthermore, national consumer bodies accredited by the EU were requested to disseminate the survey through their website and constituencies. The European Consumer Centers from the Netherlands, Bulgaria, Poland, and Spain have positively embraced our initiative and posted the link to the survey in their social media and newsletters.

The project partners got in direct contact with different consortia of European projects in which they have been involved in the past or are still actively participating. Some of the partners of these consortia have supported the NUDGE project by filling in the survey and/or disseminating the survey via their social media platforms. Likewise, the 50 associations that support the Inter-Institutional Group "SDGs for well-being and consumers' protection" were contacted to this end.

Moreover, the project has established different media partnerships at international, regional, and national level. The media partners include The Innovation Platform, SyncSci Publishing, and the Italian online magazine Canale Energia. Similarly, the international not-for-profit ecolabel for energy EKOenergy provided its support via its social media channels.

In Annex V of this deliverable we list some of the promotional material we used for disseminating the survey.

2.3.2. Final sample of participants and preliminary socio-demographic analysis

Originally, 7089 people opened the webpage of the survey, but throughout the process, a portion of them dropped out: 954 (13.46%) respondents dropped out after reading the introduction and 536 (8.74%) after reading the privacy statement. Another 689 dropped out after the third module consisting of attitudinal, motivational and behavioural constructs because they falsely answered one or both control questions. Additionally, the Flemish survey applied recruitment quota of age, gender and educational degree in order to strive for a representative sample of the Flemish population, and contained two control questions to warrant data quality. 1298 Flemish people started the survey but matched with quota that were already fulfilled. Hence, they were not able to proceed to the next 'housing characteristics' block of questions. In total, 3173 individuals completed the survey.

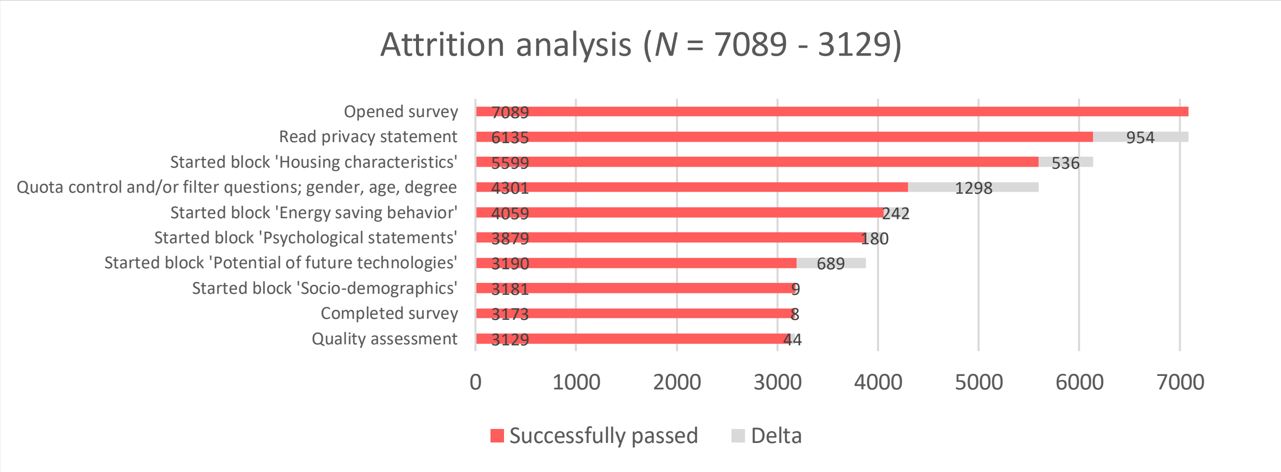


Figure 2: Decrease of respondent participation throughout survey

Our data cleaning resulted in a final sample of n=3129; we subsequently refer to it as sample. Below we discuss these removals in more detail. Only people between 18 and 100 years (6 people were removed) and Europeans (15 respondents from outside Europe, i.e., India, South Africa, Kuwait, Turkey, etc. were removed) were retained. To warrant data quality, respondents who have carelessly filled out multiple questions throughout the survey have been flagged (i.e., living surface outside range of [15;1000], zero standard deviation for 3 attitudinal constructs and all saving behaviours, survey completion time of less than 300 sec, test responses). Consequently, 44 respondents with 4 or more flags have not been considered in data analyses. Out of those 3129 respondents, 1521 (48.6%) participants are female, 1592 (50.9%) male, with 16 (0.5%) people stating that their gender was neither female or male. An independent samples t-test revealed that men ($M = 53.35$, $SD = 16.19$) are significantly older than women ($M = 47.34$, $SD = 15.25$) in our sample ($t(3140) = -10.727$, $p = .000$). This is also visible in the figure below.

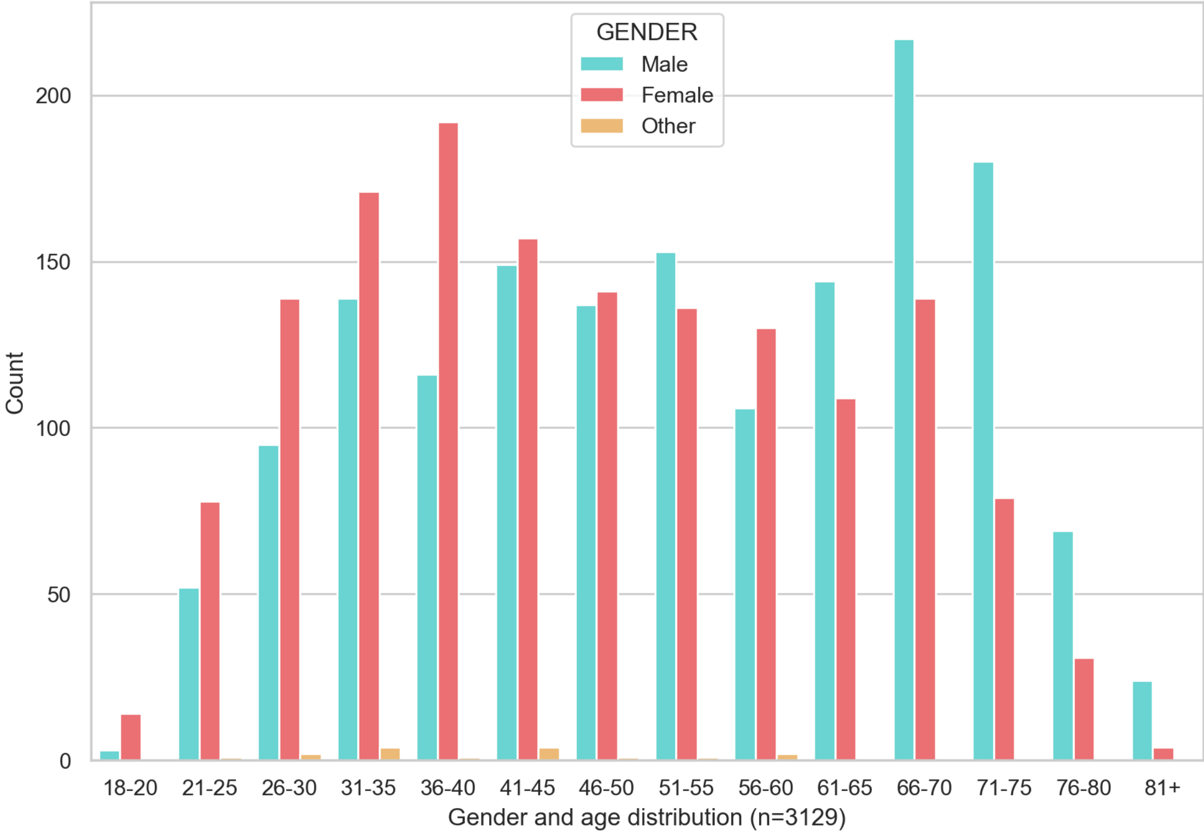


Figure 3: Gender per age group

Some groups in relation to age, gender, and educational attainment are slightly over or underrepresented. However, we decided to not use weighted data since the purpose of this report is to present results from inferential statistics (and far less descriptive statistics) and no consensus exists about this topic in academic literature (Gelman, 2007; Kott, 2007; Winship & Radbill, 1994).

We use odds ratio thresholds of 1.5 and 0.5, respectively, to declare a group over or underrepresented in a sample from a specific country. With respect to gender representation, we see an overrepresentation of females in the Bulgarian, Spanish, Croatian, and Latvian samples, whereas males are overrepresented in the Dutch sample. With regard to age, respondents in the age interval 20-39 are overrepresented in the Bulgarian, German, Greek, French, Croatian, Latvian, and Portuguese samples. Respondents of the Bulgarian and the Italian samples have higher education (from upper secondary education to doctoral level, ISCED3-8) compared to their national populations, whereas people who have attained lower levels of education (up to lower secondary education, ISCED0-2) are overrepresented in the German, Greek, Spanish, Latvian, Lithuanian, and Portuguese samples.

After data cleaning, participants from 29 different European countries remain. Given the small sample sizes in some countries (i.e.: Austria, Finland), we re-categorized our sample according to the United Nations Geographical Scheme[3], with the exception of Malta, which was added to Southern Europe. We discuss in more detail the differences between geographic regions in section 3.4.

Table 1: Assignment of countries to geographic region

Country Code	Eastern Europe, N = 235[1]
SK	79 (34%)
BG	31 (13%)
fCZ	3 (1.3%)
PL	3 (1.3%)
BA	2 (0.9%)
RO	117 (50%)
	Northern Europe, N = 148[1]
UK	4 (2.7%)
LT	36 (24%)
NO	3 (2.0%)
DK	2 (1.4%)
FI	2 (1.4%)
LV	100 (68%)
IE	1 (0.7%)
	Southern Europe, N = 1,114[1]
ES	46 (4.1%)
SI	272 (24%)
GR	234 (21%)
IT	208 (19%)
HR	180 (16%)
MT	16 (1.4%)
PT	156 (14%)
CY	1 (<0.1%)
XK	1 (<0.1%)
	Western Europe, N = 1,665[1]
FR	74 (4.4%)
AT	6 (0.4%)
NL	299 (18%)
DE	148 (8.9%)
BE	1,136 (68%)
CH	1 (<0.1%)
LU	1 (<0.1%)

3 https://unstats.un.org/unsd/methodology/m49/

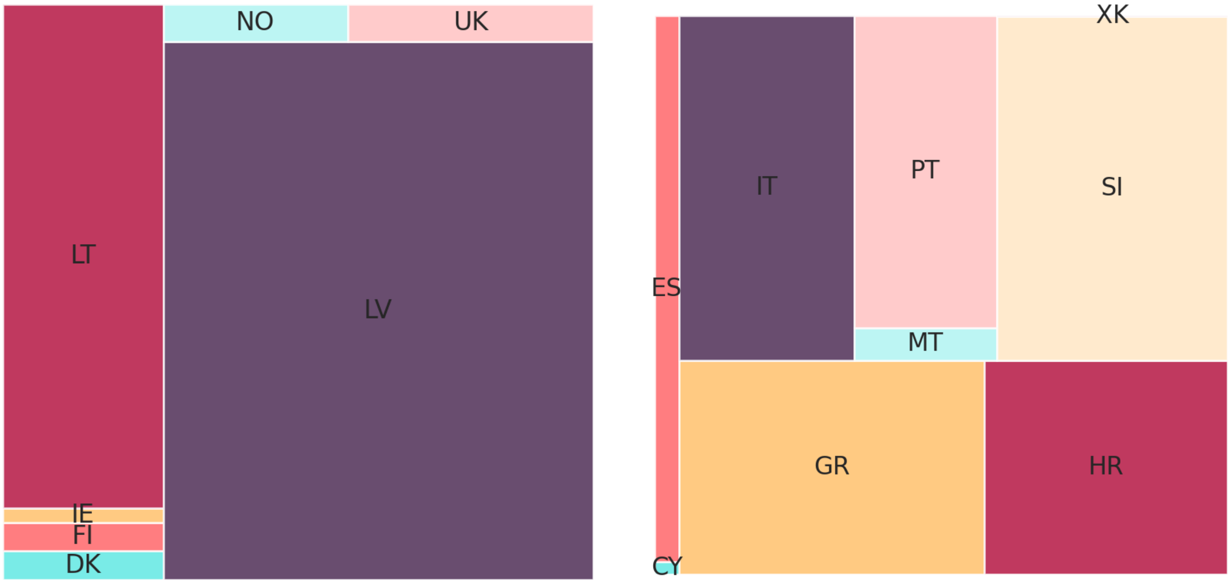


Figure 4: Proportion of respondents from countries in Northern Europe (left) and Southern Europe (right) in our sample

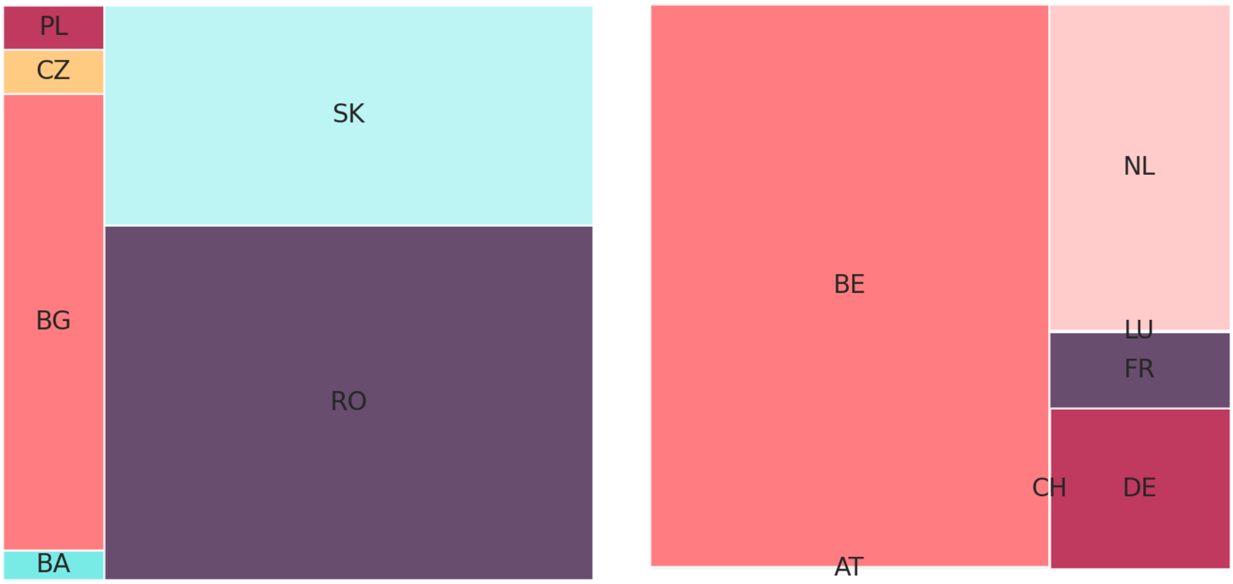


Figure 5: Proportion of respondents from Eastern Europe (left) and Western Europe (right) countries in our sample

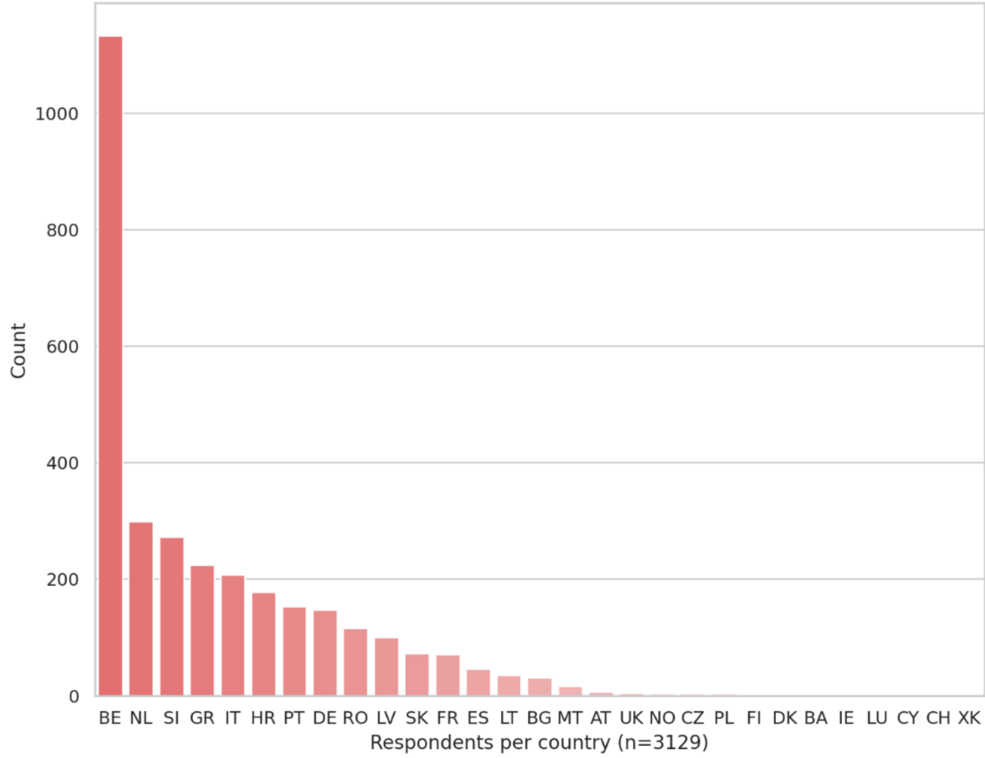


Figure 6: Participants per country of residence in our sample

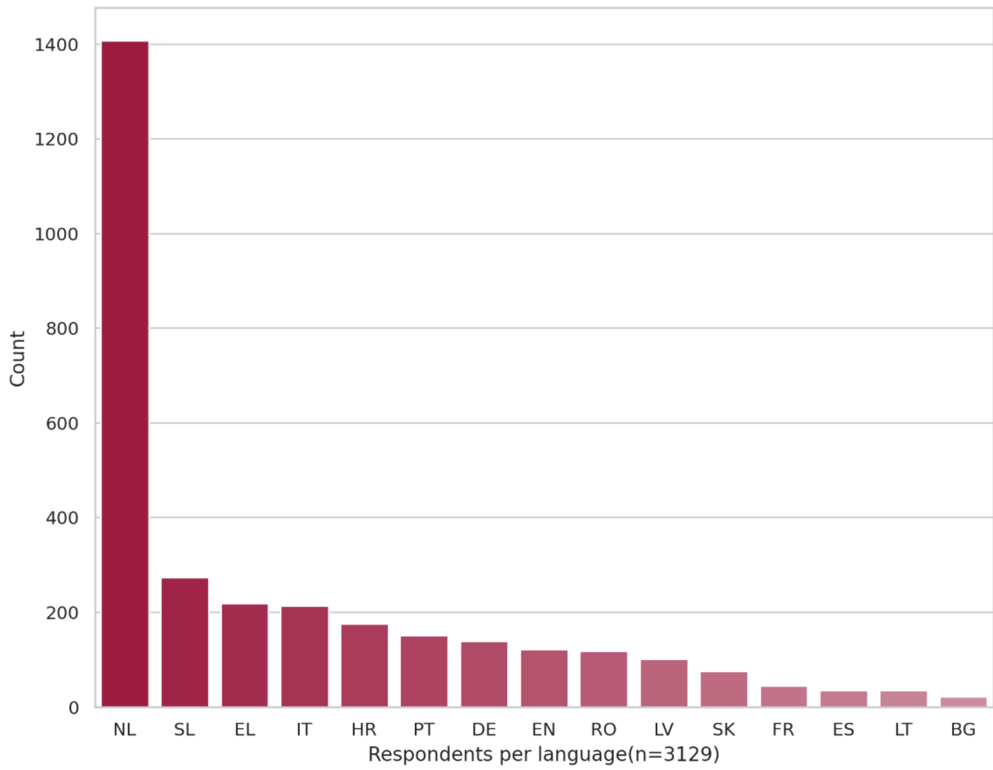


Figure 7: Native language distribution in our sample

3. Descriptive statistics

3.1. Housing characteristics

Three out of four people in our sample own their house or apartment (76.5%), one out of five are renters and the remaining 3% lives in a free residence. Two out of five people in our sample live in an apartment (41.5%). We also considered maisonettes, student dormitories, porched houses, quadrant houses, panel buildings, and boarding houses, as mentioned in the open text-box answers, to be part of the 'apartment' category. Almost one third (31.1%) lives in a detached house (also bungalows, cottages, and farm houses). The remaining families live in terraced houses (also townhouses) (14.3%) and semi-detached houses (also duplex) (12.5%). More than one third (36%) indicates that their house has not been renovated (as far as they are aware of).

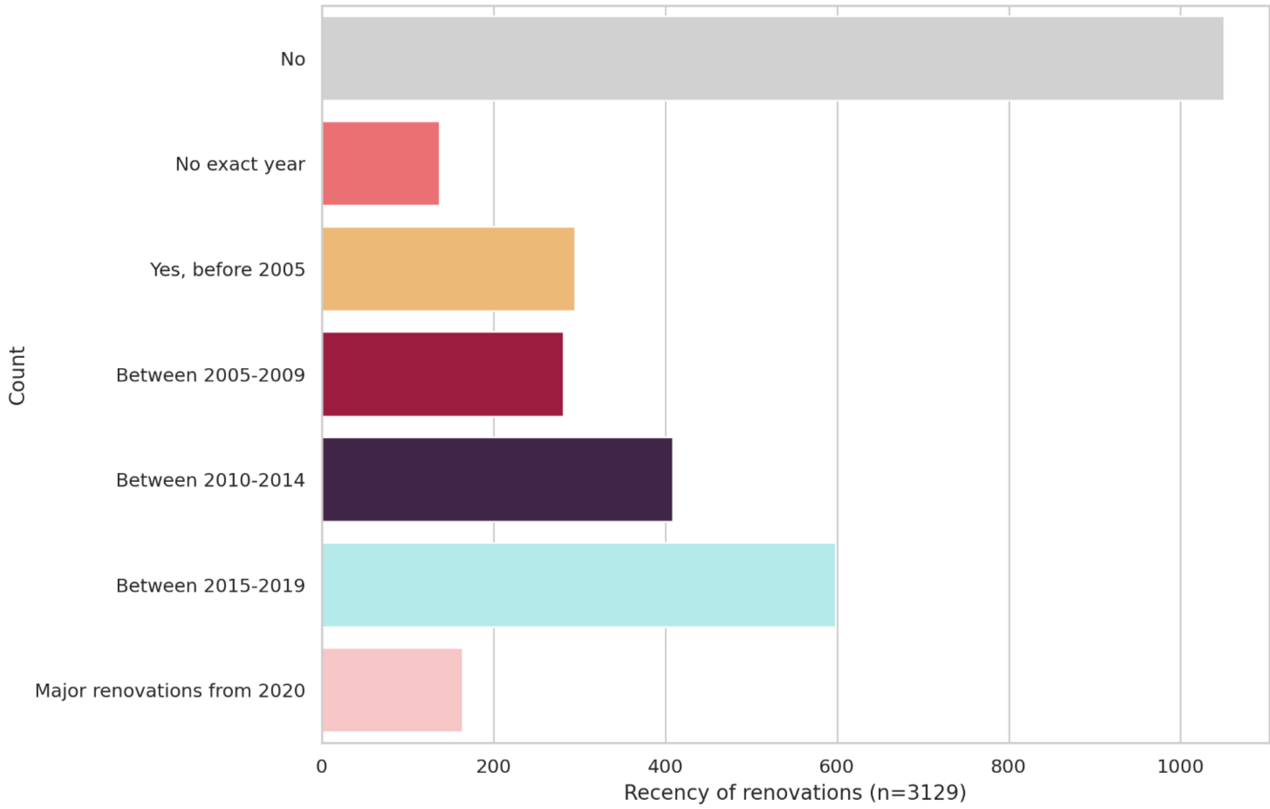


Figure 8: Recency of housing renovations among participants in our sample

3.2. Energy-saving behaviour types

3.2.1. Heating and cooling

The first set of energy-saving behaviours is related to **heating and cooling**. The most common energy saving behaviour is *closing the windows when heating* (87.2% state 'often'/'always'). Since many families do not have air-conditioning installed, a percentage of 65.3% mentioned that turning the heating off while air-conditioning is on, is not applicable. Based on the share of participants who answered *Not applicable* per question, wearing more clothes is the most accessible behaviour in this category, nonetheless, it is the second least undertaken behaviour with only 47.4% mentioning that they 'often'/'always' wear more clothes if it feels cold inside.

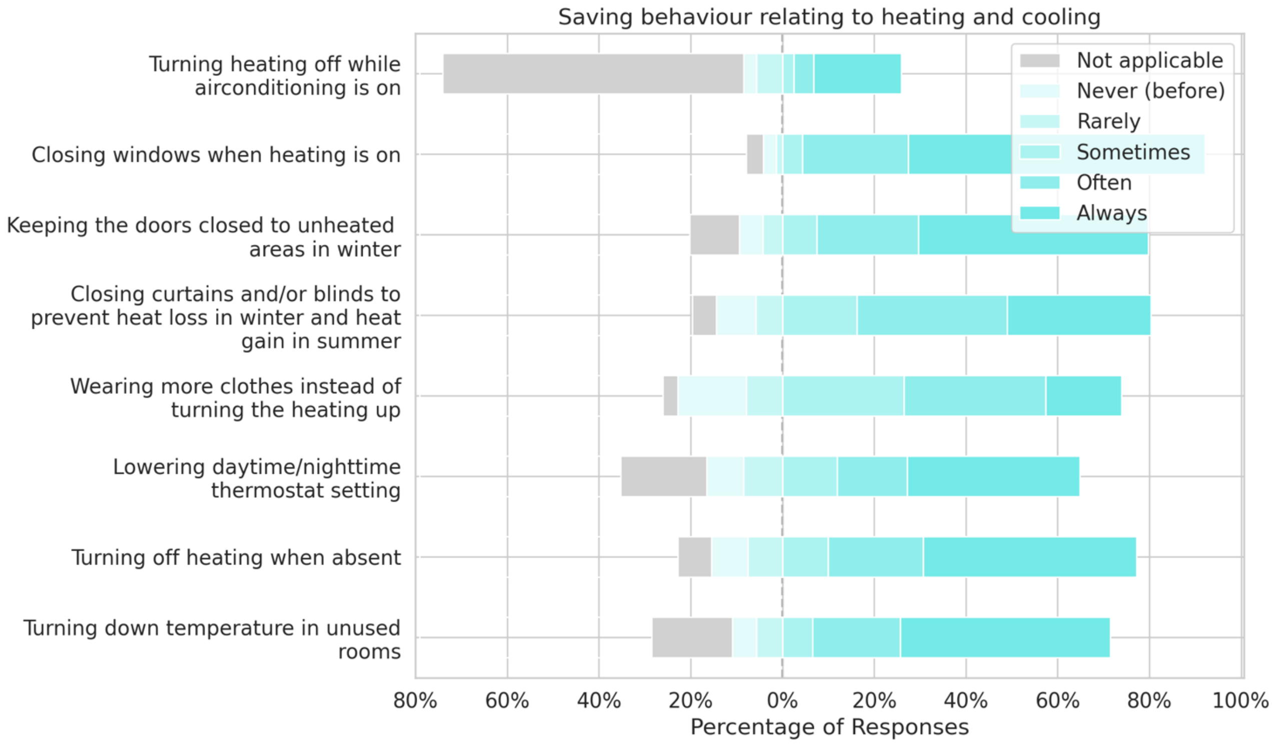


Figure 9: Energy-saving behaviour with respect to heating and cooking

3.2.2. Hot water usage

The second set of energy-saving behaviours relates to **hot water usage in the bathroom**. In comparison to heating and cooling saving behaviours, more variation can be noticed among these hot water saving behaviours. It appears that it is not the frequency, but rather the duration of hot water saving behaviours that is adjusted, i.e., not leaving the water running, taking shorter showers (respectively 70.1% and 45.6% often/always). Across all categories, people are least inclined to reduce their number of showers (55% never/rarely), which indicates that the majority is not willing to give up comfort and personal hygiene to save energy. A percentage of 80.3% prefers indicated preferring showering *often* or *always* to bathing, however we cannot deduce from this question if people only have a shower or both a shower and a bath. Moreover, taking a bath vs showering could also be related to personal preference, above and beyond any conservation motives.

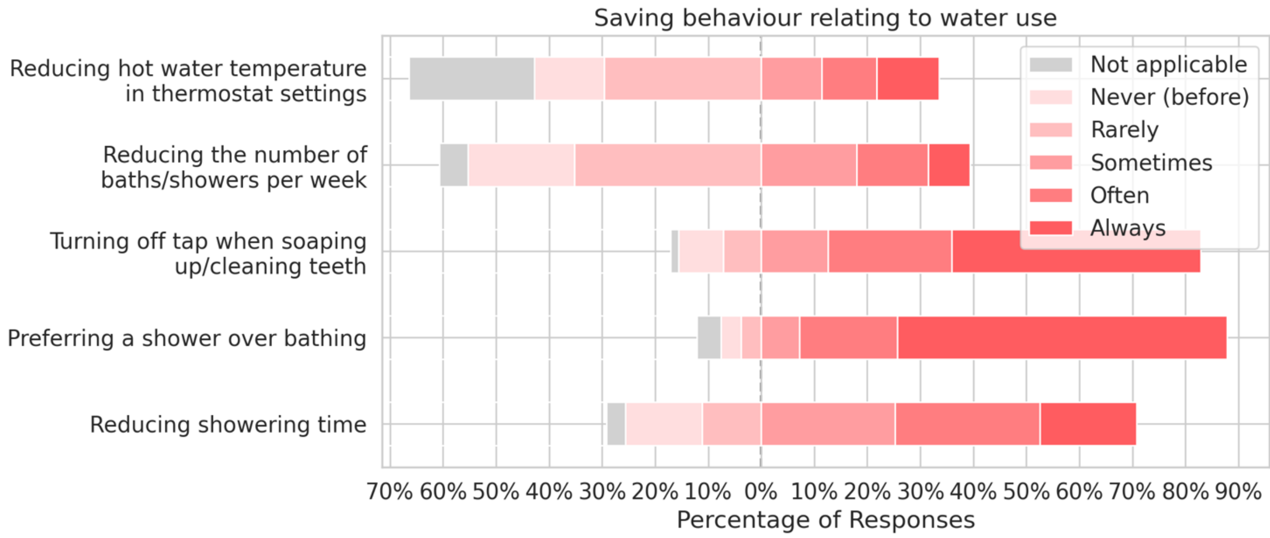


Figure 10: Energy-saving behaviour related to the use of hot water

3.2.3. Kitchen activities

Approximately 28% of respondents indicate not having a dishwasher. The ones who have a dishwasher more often fully load the dishwasher (Median = 4) than using the energy-saving mode (Median = 3) (Wilcoxon signed-rank test, Z = -17,004, p = .000). The most frequent energy-saving action performed in the kitchen is covering cooking pots (81.6% often/always), whereas using the energy saving program of the dishwasher occurs least often (38.4% often/always). Nonetheless, this result should be nuanced, given the large share of persons who indicated *not applicable* for this option.

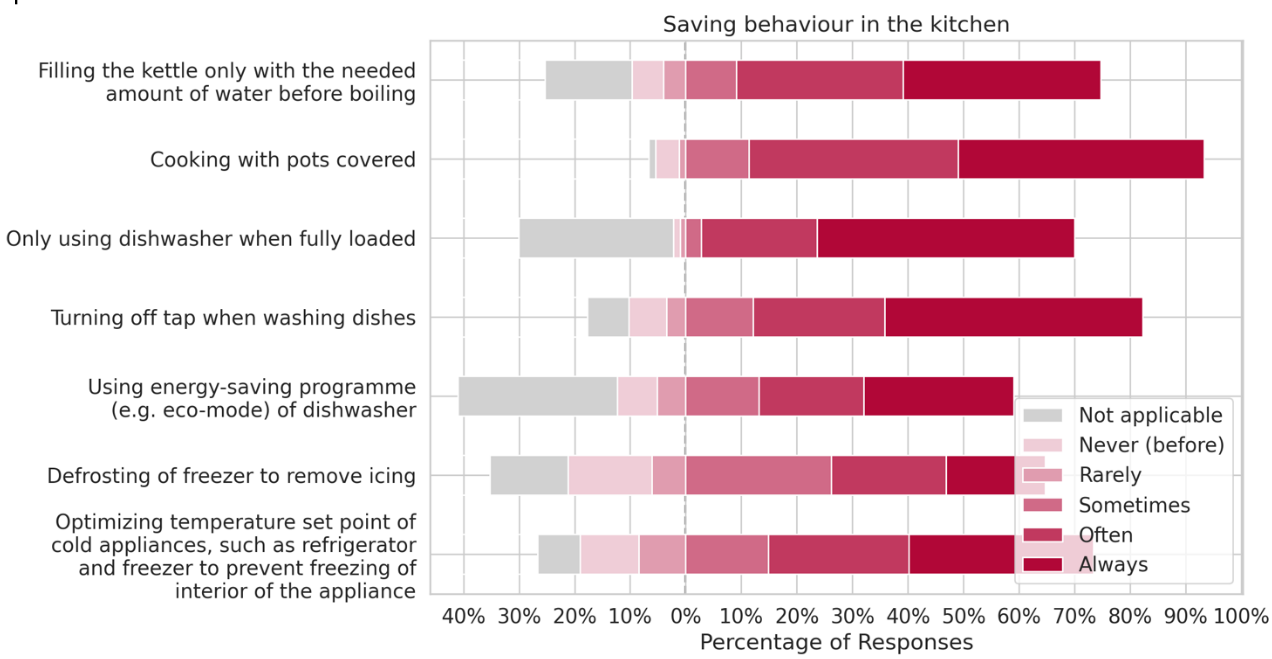


Figure 11: Energy-saving behaviour in the kitchen

3.2.4. Use of appliances

The last category examines the **energy saving use of appliances**, which represents the category of extremes: the most frequently performed but also 'not applicable' actions are part of this category. 66% of respondents indicate not having PV panels and 90.9% of respondents state not having either PV panels or an electric vehicle. Also, but to a lesser extent, almost half of the sample does not possess a tumble dryer (48.1% not applicable). Conversely, turning off the lights is the most common behaviour across all categories with 91.5% indicating that they '*often* up to *always*' perform this action. Turning off unnecessary appliances, such as unplugging set-top boxes and not overcharging devices, are not commonly occurring energy-saving behaviours. Almost a third of respondents rarely or never perform this action (31.5%). Except for the previous actions, there is a higher degree of similarity among the remaining energy-saving actions: switching off TV, and loading and "eco" mode of washing machine. These actions are prevalent among the majority of our sample with frequencies of '*often/always*' statements varying from 68% up to 82%.

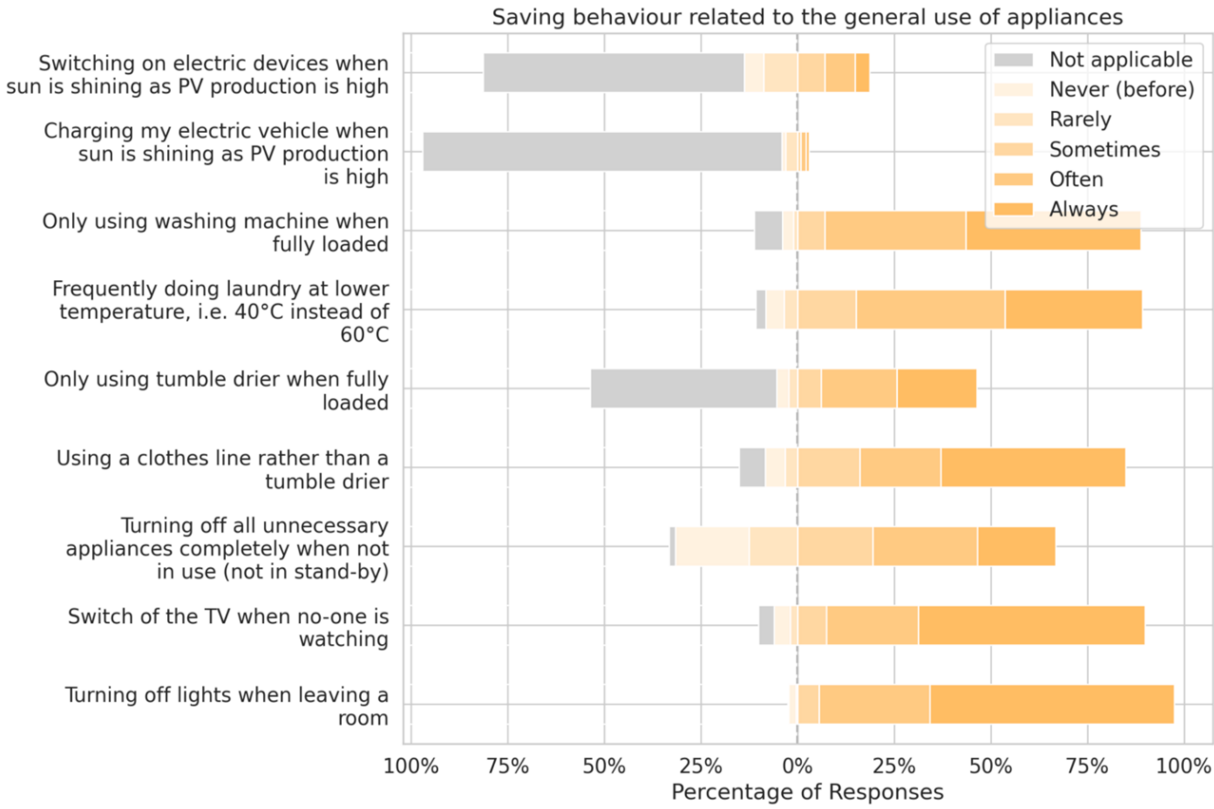


Figure 12: Energy-saving behaviour related to the general use of appliances

3.3. Interest in sharing energy data

People were asked to what extent they were interested in getting more insights into the way they consume energy. In general, people are mostly interested in real-time energy usage with 6 out of 10 people indicating a lot or even great interest. In contrast, people tend to be neutral towards data revealing their daily and occupancy of property and its activity (the 2nd and 3rd largest categories).

With regard to energy data sharing, three notable tendencies are apparent. The majority seems unwilling to share the detailed energy data, i.e., on a daily to real-time basis. Whereas 42.5% feels that monthly energy data is not shareable data, this proportion increases to 44% for daily and real-time data. Furthermore, people in our sample are more prone to share all types of energy data with their family members (67% real-time data and for 90.3% monthly data). However, people are less willing to share detailed energy data with all other entities. **On average, 88.1% of respondents are willing to share monthly energy data with their neighbours, energy providers and distributors, third parties, and the government.** Additionally, people seem to have the most diverse willingness to share their energy data with their neighbours. The willingness to share real-time data with their neighbours decreases with almost 48% from monthly to real-time data.

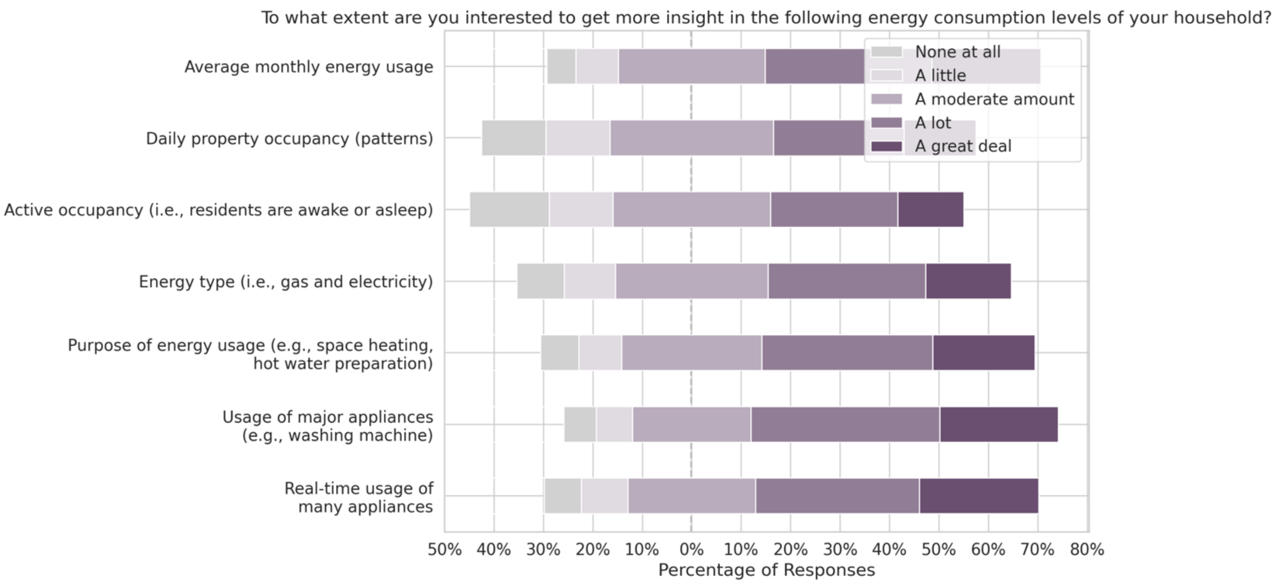


Figure 13: Interest in more detailed information about household energy consumption

3.4. Regional differences across Europe

We report some regional differences between the four geographic regions of Europe. However, these results are produced as a matter of form, rather than drawing any absolute conclusions, given the extremely limited sample sizes, notably for Northern Europe, containing only n=148 participants, the vast majority of whom are Latvian (n=100) and Lithuanian (n=36). We thus strongly caution against over-interpretation of these results.

The Organisation for Economic Co-operation and Development (OECD) income of households in Northern Europe in our sample is significantly lower than those in Southern and Western Europe ($F(3,1207) = 58.06$, $p < .000$). However, this data is based almost exclusively on Latvian ($n = 34$, $M = 522.06$ EUR) and Lithuanian ($n = 21$, $M = 767.22$ EUR) households that shared their average monthly income and as a result is not at all reflective of actual Northern European levels of income. For reference, the OECD income of Western European citizens in our sample is 1535.5 EUR ($n = 451$).

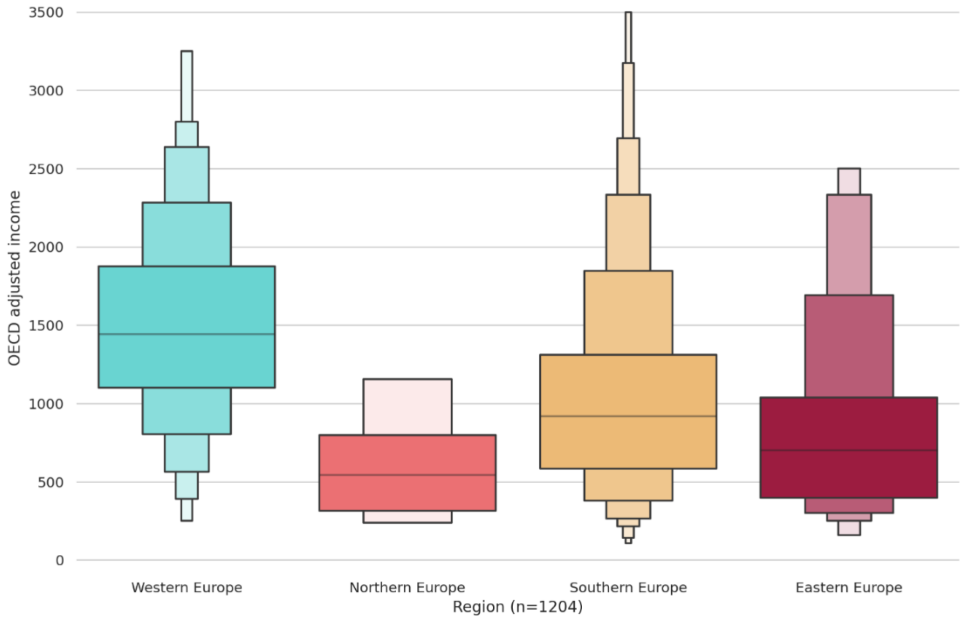


Figure 14: OECD adjusted income per region (outliers removed)

Respondents were asked where they perceive themselves on a 'conservation ladder' with 9 steps, ranging from (1) *not energy conscious at all* (and having relatively high energy bills) to (9) *very energy conscious* (and having relatively low energy bills).

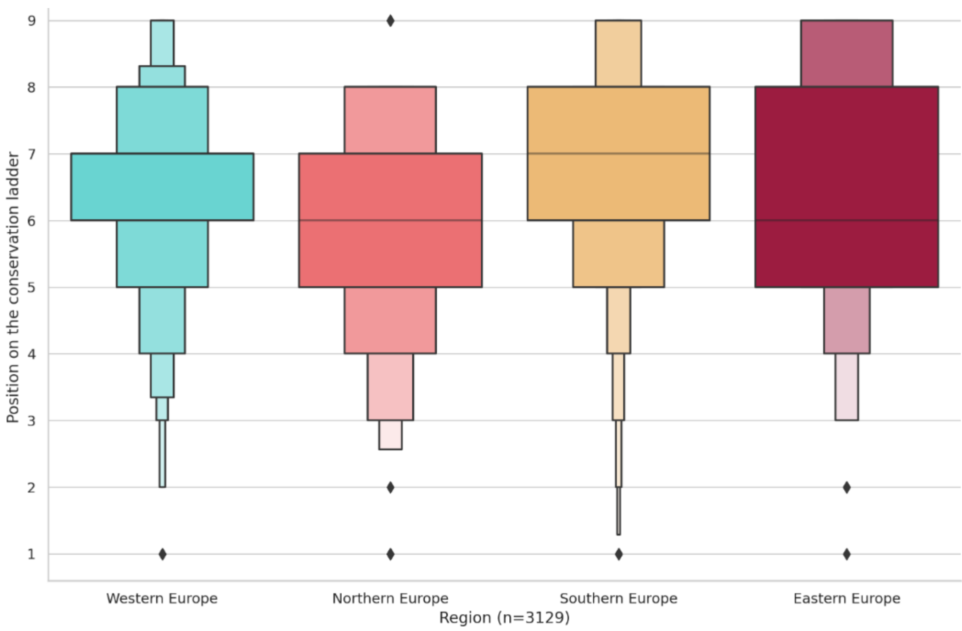


Figure 15: Self-positioning on the energy-conservation ladder per region

Respondents from Northern Europe perceive themselves significantly less energy conscious compared to the other European regions ($F(3,3112) = 19.087$, $p = .000$, see Figure 15). This may partly be explained by particular uses and behaviours that are common across regions. This finding is also apparent in individual's intent to lower the temperature setting in winter. The lowest intent is measured in Northern Europe ($M = 2.89$, $SD = 1.10$), which significantly differs from other European regions ($F(3,3125) = 39.095$, $p = .000$). Nonetheless, we again emphasise that our total northern European sample is modest, a mere 147 participants, and as such, we caution against overinterpretation of these results.

Except for respondents from Northern Europe, gas is the most common heating source across Europe (39.1% in Southern Europe, up to 66.6% in Western Europe). Latvian and Lithuanian respondents, who are almost exclusively populating the Northern Europe subsample, use more frequently district heating (43.2%).

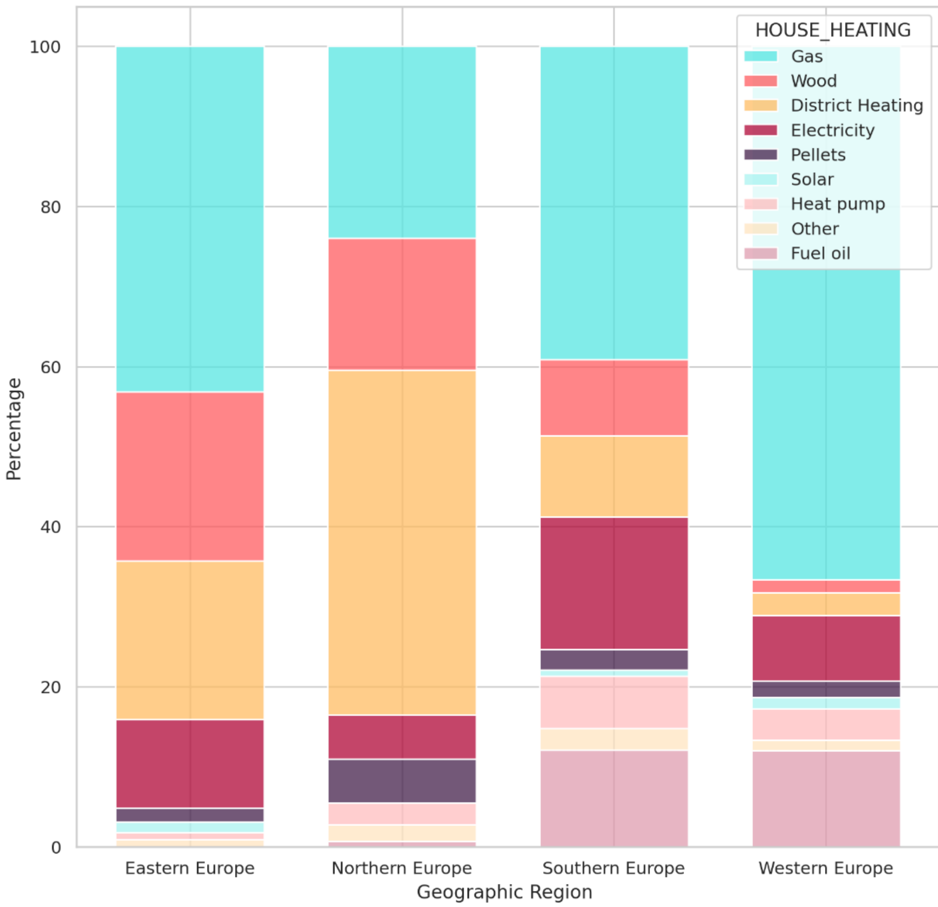


Figure 16: Use of heating technologies/means across the four European regions

We furthermore note significant differences between the self-reported estimates of the Energy Performance Certificate (EPC). Participants were asked to indicate on a colour gradient where they felt their household scored. Scores were normalised to range between 0 and 1 and subsequently reversed so a higher score indicated a better EPC rating (i.e., better performance). As noted, this is a purely subjective rating and deviates from the formal EPC, given the slight national variations in the implementation of the certificate. Overall analysis of variance showed significant results ($F(3,2704) = 31.39$, $p < .001$).

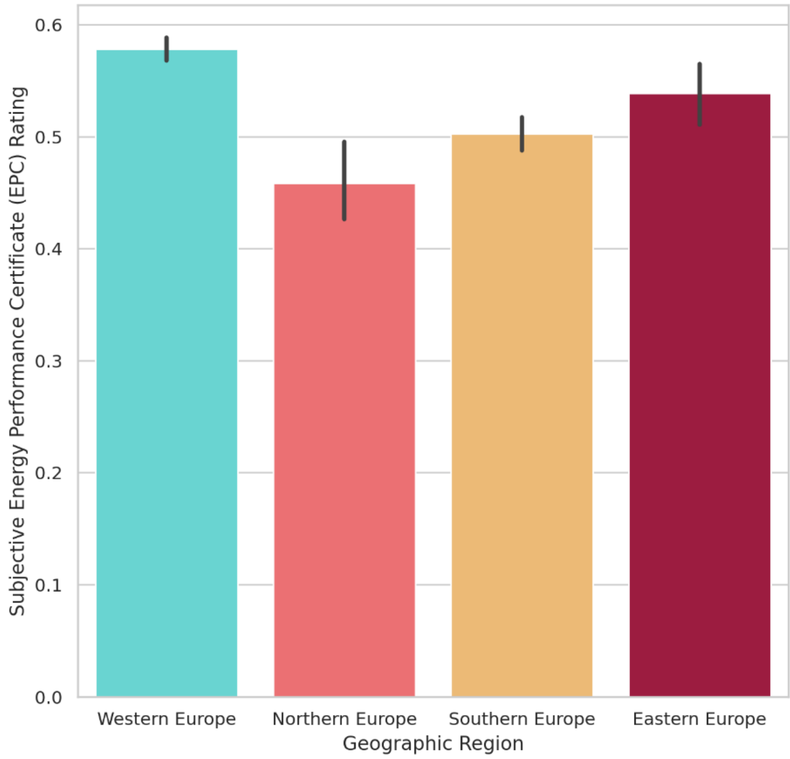


Figure 17: Regional differences in Energy Performance Certificate (EPC) self-ratings

4. Explaining intention to improve energy efficiency

A primary goal of our survey is to find attitudinal and behavioural predictors of intent to reduce energy consumption, focussing both on the very specific aim to reduce heating related consumption (*I intend to save energy by lowering the temperature setting in winter)* and more general reductions of energy consumption (*I intend to save energy at home*). As discussed earlier, our attitudinal and behavioural predictors were derived from three theoretical frameworks (TPB, VBN and PWM) (see section 0 for a complete overview of hypothesis and rationale for the use of these framework).

Below, we will first review the reliability of our instrument, followed by descriptive statistics. We conclude with a summary of findings related to both intent to reduce heating related energy consumption in general.

4.1. Reliability of all constructs

When performing survey research, it is important to consider how consistent our participants answered related questions of a particular construct (i.e.: Willingness). To do so, we performed Cronbach's α analysis of all constructs in our model. The Cronbach's α value for any particular construct should be higher than 0.7. Cronbach's α analysis of all constructs from our three theoretical models and all remaining variables were satisfactory, with the lowest value of .77 for Willingness, well above the customary 0.7 threshold. In table 2 the results of the reliability analysis are presented per construct

Table 2: Cronbach α values of all constructs used in our regression model

Model construct	Cronbach value
Specific intent (INT_SPEC)	0,90
Subjective norm (SN)	0,83
Attitude (ATT)	0,91
Perceived behavioural bontrol (PBC)	0,82
Prototype favourability (PROT_FAV)	0,92
Prototype similarity (PROT_SIM)	0,95
Willingness (WILL)	0,77
Financial concern (FIN_CONCERN)	0,80
Loss of comfort (LOSS_COMF)	0,90
Energy knowledge (ENERGY_AWARE)	0,94
Environmental concern (ENV_CONCERN)	0,82
Awareness of consequences (CONSEQ_AWARE)	0,78
Ascribing responsibility (ASCR_RESP)	0,93
Moral norm (PN)	0,80
General intent (INT_GEN)	0,84

4.2. Correlations

Our correlation matrix shows significant positive correlations between all measured constructs, with the exception of the *Loss of comfort* variable, which is negatively correlated with all other variables.

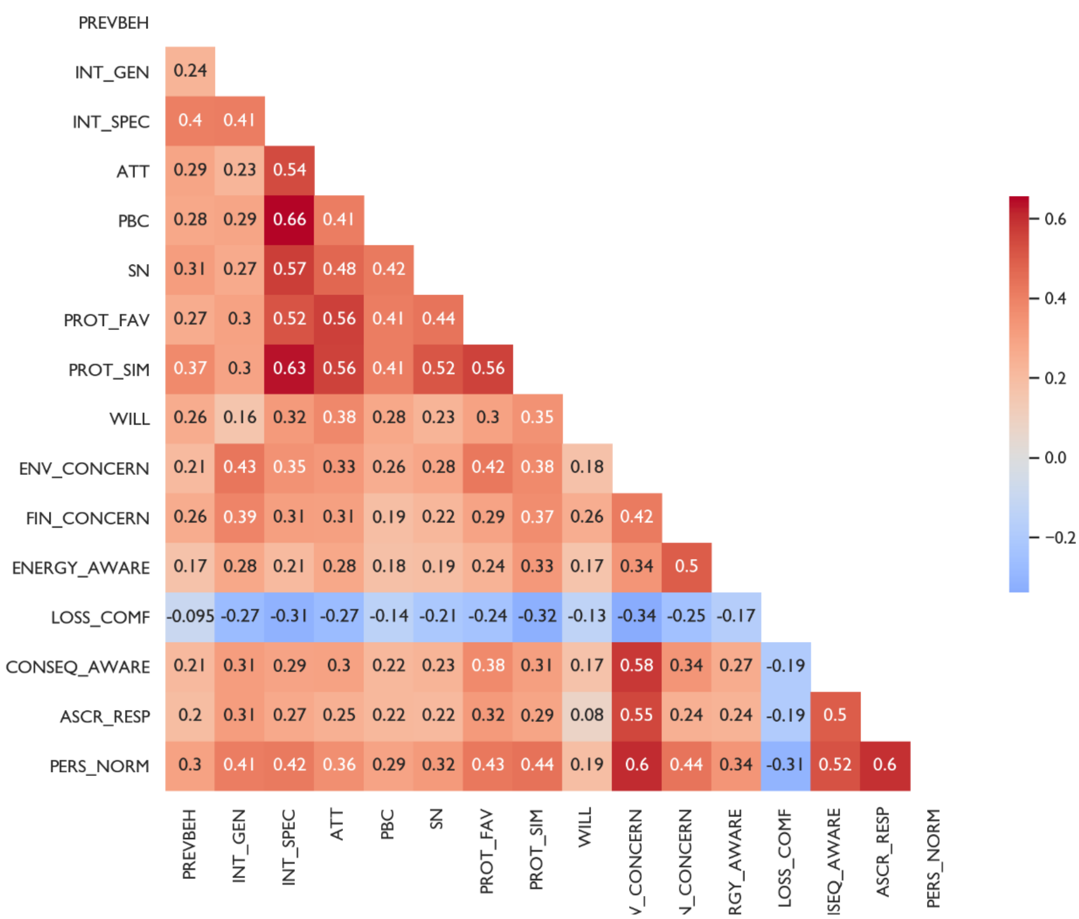


Figure 18: Visualised correlation matrix

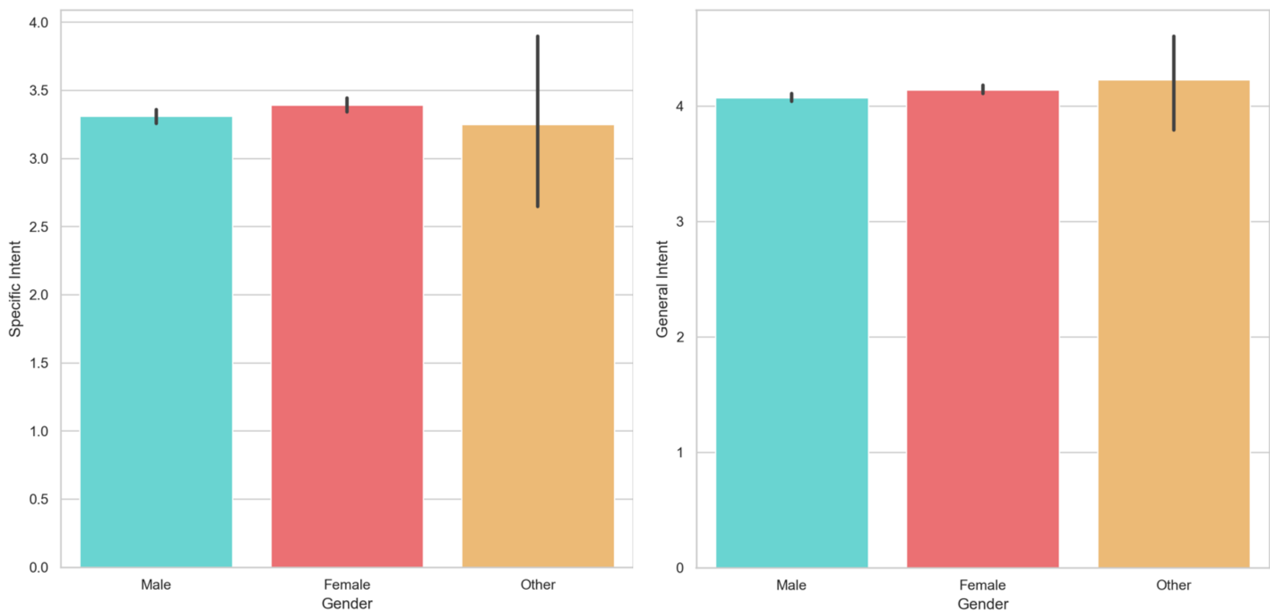
4.3. Predicting intent to reduce energy consumption

We present two series of linear regression analysis, both related to the consumption of energy. As noted earlier, we first examine the specific intent to reduce energy use related to heating, while in our second series of models we examine more general intent. Our complete regression models for both our outcome variables can be found on page 86. Table 12 details specific intent, while Table 13 looks at general intent. In both cases, we first present our socio-demographic variables (model 1a and 2a). This is followed by our three theory of planned behaviour variables (model 1b and 2b). The third model in the series examines our remaining attitudinal variables (model1c and model 2c). Finally, we present the complete analysis, containing all variables (model 1d and 2d). We limit our discussion here to the results in our final models (model 1d and 2d).

More specifically, we report the statistical significance for all relationships (p) taking .05 as threshold of significant. Furthermore, we report B, which denotes the strength of the relationship between our measured variable (i.e.: *attitude*) and our outcome (i.e.: *specific intent*) while controlling for all the others variables in our model. A complete overview of statistical abbreviations can be found on page 99.

4.3.1. Gender inclusivity

As discussed earlier, we asked our participants to indicate their gender using the following question: *What is your gender, as indicated on your national ID or passport?* Our sample contained n=1521 females, n=1592 males, while 16 persons indicated '*other'.* The very small number of people indicating '*other'* at first precludes their inclusion in our regression analysis. At the same time, we are aiming for an inclusive approach so that these respondents can further me included in the analysis. We therefore tested for differences in the responses based on gender. Using a one way analysis of variance test, we find that gender is not associated with specific intent (p=0.076), but a significant difference between men and women for general intent (p=.02). To further assess the impact of gender, we included it as a categorical predictor (and with removal of our 16 participants who indicated '*other')* in both our models for specific and general intent. We find no statistically significant result once other factors are taken into consideration. Given this, we have elected to include our 16 participants who indicated '*other'* while subsequently not using gender as a predictor of intent to ensure gender inclusivity.



a) Specific Intent b) General Intent

Figure 19: Gender's association with specific and general intent

4.3.2. Specific Intent

Our first analysis concerns predicting specific intent, which relates to reductions of energy use related to heating during the winter. We find strong and significant associations between all three main variables of Theory of Planned Behaviour (Attitude, Perceived Behavioural Control and Subjective Norms) and our main outcome variable, specific intent. We report the coefficient (B) and the statistical significance (p). Attitude (B=0.14, p<0.001), Perceived Behavioural Control (B=0.47, p<0.001) and Social Norms (B=0.28, p<0.001) were all significantly associated with intent. Of our remaining variables, financial concern is positively associated with specific intent (B=0.10, p<0.001), whereas both loss of comfort (B=-0.10, p<0.001) and energy related knowledge (B=-0.05, p<0.001) are negatively associated.

These results can be seen in Figure 20 and Figure 21. However, these results only show the direct correlations between, for example, Perceived Behavioural Control and Intent, without consideration of any other variables (as is the case for our complete regression). As a result, the strength of the association (or even its direction) can change once all other variables are taken into account. This can be seen for energy related knowledge, with a slight negative association in our overall model, while having a positive association when seen in isolation with intent.

egional effects, while present, are muted, only Western Europe expressing significantly lower impact when compared to our reference category, Eastern Europe (B=-0.12, p<0.05). Similarly, age appears to have a very modest impact (B=0.01, p<0.001). In total, region, age and degree explain a mere 4% of variance.

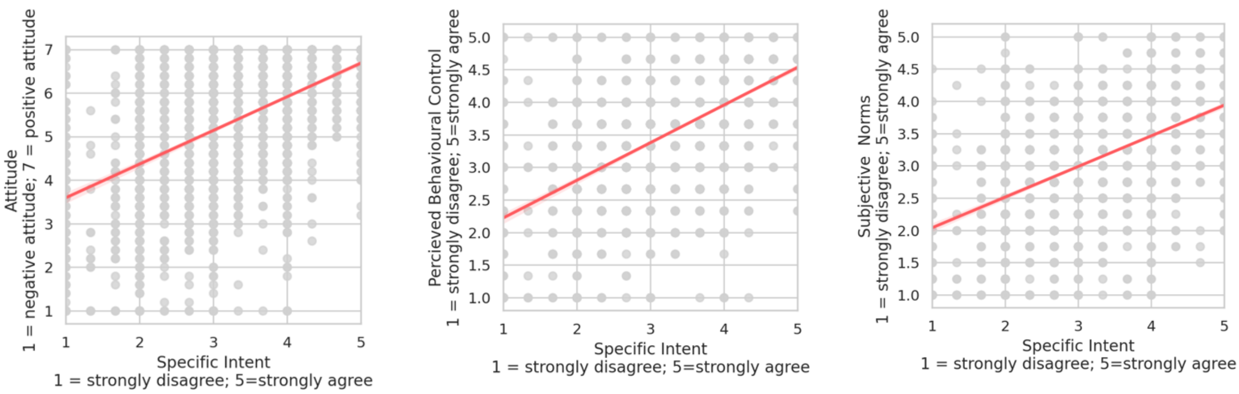


Figure 20: Attitude, Perceived Behavioural Control and Subjective Norm and their relationship with Specific Intent to reduce energy consumption related to heating

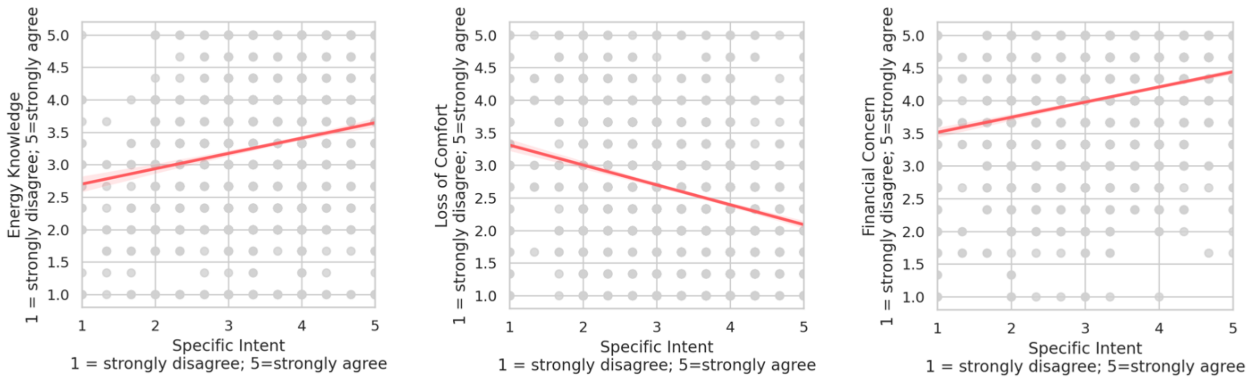
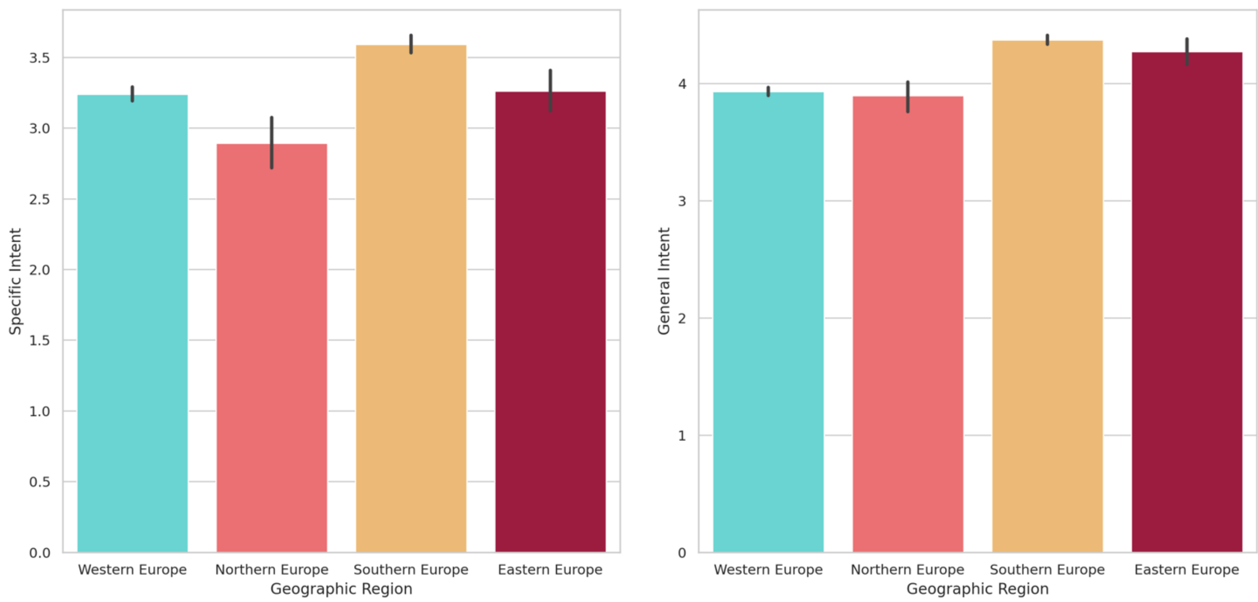


Figure 21: Impact of Energy Knowledge, Loss of Comfort and Financial Concern on the Specific Intent to reduce energy consumption related to heating



(a) Specific intent

(b) Generic intent

Figure 22: Regional differences in intent to reduce energy consumption

Of our remaining variables, Financial Concern is positively associated with specific intent, while both Loss of Comfort and Energy-related Knowledge are negatively associated. The overall variance explained by our model is 62%.

4.3.3. General Intent

Our second series of analysis examined which factors predict general intent to reduce energy consumption. As expected, the more general phrasing of our outcome variable results in a large reduction of explained variance, from 62% for specific intent to 34% for general intent.

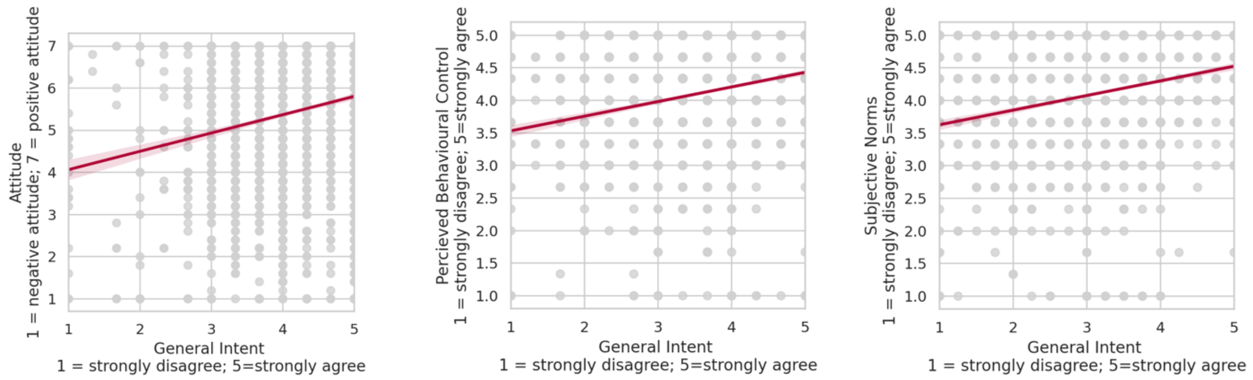


Figure 23: Impact of Attitude, Perceived behavioural bontrol and Subjective borm constructs on General intent to reduce energy consumption

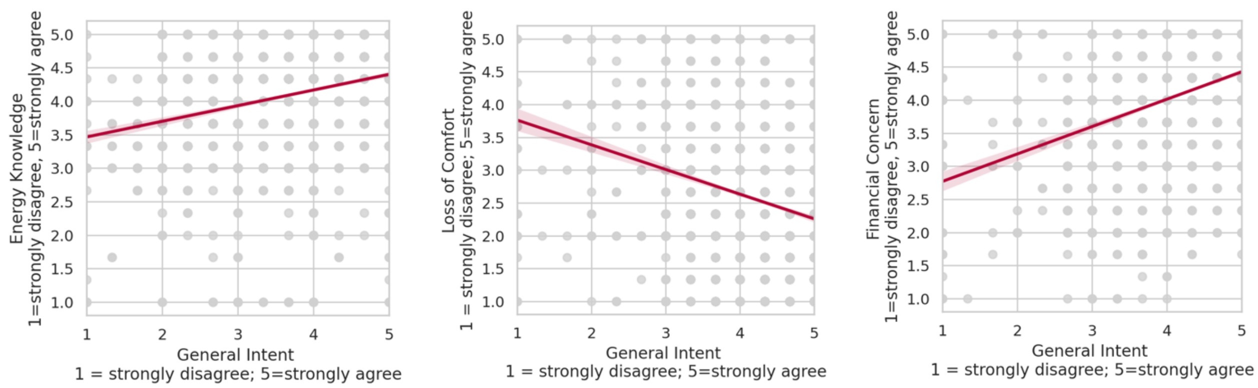


Figure 24: Impact of Energy knowledge, Loss of comfort and Financial concern constructs on General intent to reduce energy consumption

On their own, our demographic variables explain 9% of variance. The impact of age is modest (B=-0.01, p<0.001) but significant, while Northern Europe (B=-0.11, p<0.001) and Western Europe (B=-0.22, p<0.001) both have lower general intent when compared to Eastern Europe, our reference category. Attitude has a very modest but statistically significant negative impact (B=-0.02, p<0.05), while Perceived Behavioural Control (B=0.11, p<0.001) and Social Norms (B=0.06, p<0.001) are both positively associated with general intent.

Beyond this, Financial Concern (B=0.20, p<0.001), Energy knowledge (B=0.05, p<0.001, Environmental concern (B=0.14, p<0.001), and Personal Norms (B=0.06, p<0.001) are all positively associated with General Intent, while Loss of Comfort (B=-0.07, p= <0.001) has a negative association. As before, Figure 23 and Figure 24 display the simple correlations between, for example, Energy Awareness and General Intent, not taking into account other variables.

4.3.4. Conclusions

To conclude, we reflect on some of our results of the survey, most notably our efforts to assess the intent reduce consumption. As suggested by regression, we find strong support for perceived behavioural control, social norms and attitudes towards energy-saving as predictors of intent to reduce heating related energy consumption.

For perceived behavioural control in particular, this points towards providing consumers with practical ways through which consumption might be reduced. The importance of subjective norms additionally suggests that emphasising what others think of lowering the temperature is an important lever to nudge consumers to reduce less. Existing attitudes, by contrast, also has an important role in shaping intent, but is somewhat lower in importance. Nonetheless, it remains an important tool, suggesting that educational campaigns on energy consumption can have a positive outcome.

Financial savings, while significant also has a role to play, but as seen in our regression, has a muted impact when compared to perceived behavioural control and subjective norms. In addition, the fear of losing comfort has a negative association with intent, but should similarly be viewed within the context of our remaining variables.

As expected, our explained variance for general intent to save energy is far lower than our more specific questions(62% vs 34%). Additionally, our central variables within TPB do not accurately capture the intended behaviour (i.e.: heating vs generally consumption). Given this, over-interpretation of our results should be avoided. However, we still find support for perceived behavioural control and social norms, strengthening the suggestion that what close family and friends think about consumption is important, while practical ways through which energy might be saved should also play a role in any efforts to reduce consumption.

The impact of loss of comfort remains broadly similar, while do find a stronger effect for energy knowledge, which aligns with the earlier findings that perceived behavioural control is an important predictor.

5. Segmentation of energy consumers

5.1. Objective and high-level methodology

The main objective of the data analysis presented in this section is to identify groups of energy consumers with distinct characteristics that facilitate the selection of tailored interventions. To this end, we have experimented with two different approaches.

The first approach is based on clustering. Clustering (or cluster analysis) is a common technique for statistical data analysis that aims to organize a set of objects into a number of groups (clusters). Each object (here: energy consumer) is described by a number of features (here: variables/constructs such as those described in sections 2 and 3) and the goal of clustering is to group them so that objects in the same cluster are more similar to each other, according to some criteria, than to objects in other clusters. Cluster analysis provides a rich analytical framework highly differentiated by the techniques used to select the features describing the objects, the criteria used to assess the similarity of objects and the quality (fitness) of a particular clustering structure as well as the algorithms that carry out the actual clustering task.

The second approach consists in a priori specification of classes of users as logical conjunctions of conditions that the energy consumer variables should satisfy. Each energy consumer can then be separately classified into one or more of these energy-consumer classes.

The main advantage of the first approach is that it provides a solid analytical framework and that we can leverage a rich arsenal of techniques to come up with clustering structures. On the other hand, these structures need to be subsequently analysed and it cannot be taken for granted that they will be highly informative regarding interventions that are appropriate for each cluster. On the contrary, with the a priori specification of classes, we can fully take into account the set of interventions at hand and specify classes in ways that the mapping of interventions is straightforward. Nevertheless, the overall process is not automated as clustering is and we cannot take for granted that all energy consumers will be covered by one of those classes.

These points will become more obvious in what follows, as we describe the two approaches in more detail and look into their outcomes. For the time being, here are some methodological choices that are common across the two user segmentation approaches:

- The reference set of features consists of the 15 energy-related psycho-social variables (constructs) measured in the survey (sections 2.3.1-2.3.4). Socio-demographic variables are used in a second step, to describe the identified clusters. This choice is in line with reported experience in literature (Rossiter and Persy, 1987), (Suetterlin et al, 2011).
- The score of each object in a given feature/construct is computed as the average of her scores to the set of items measuring the construct. All 15 features are ordinal, with individual question items' scoring on a 5-point Likert scale so that the average scores over all items pertaining to a specific construct take a finite set of not-necessarily integer values in [1,5] -

the single exception is the ATT construct with items measuring on the 7-point Likert scale so that the average score over all its items ranges in [1,7].

- The sample of survey responses we consider for our analysis consists of the 3129 objects (i.e., energy consumers) that survived the initial filtering of survey responses (ref. section 2.5).

Figure 25 shows the histograms of the normalized scores achieved by all 3129 objects in all 15 features. Normalized score of 0 (zero) corresponds to 1 in the Likert scale, while normalized score of 1 corresponds to 5 in the Likert scale (7 in the case of the ATT variable). These plots provide a first clear indication that there is significant differentiation across the responses of the survey participants. The intensity of differentiation is not identical for all constructs: for example, the responses demonstrate high concentration at the high scores (> 3) for the Environmental Concern, Awareness of Consequences and Financial Concern variables, but they are more uniformly spread for the Prototype Similarity and the Loss of Comfort variables.

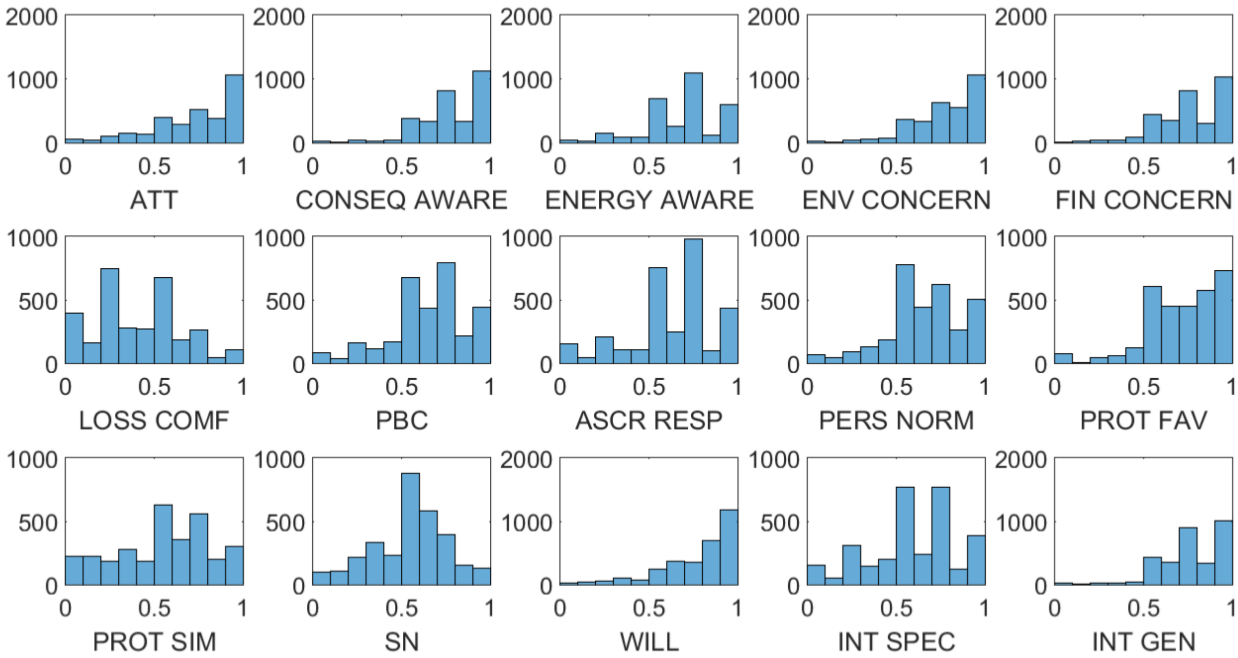


Figure 25 Distribution of average scores in the 15 features subsequently used for the segmentation analysis of energy consumers. The x-axis values mark normalized scores in [0,1], while the y-axis counts number of objects (i.e., energy consumers)

It is the task of the analysis that follows to figure out whether there is some useful structure in the way these features vary across users, or whether this variance is completely random.

5.2. Segmentation of energy consumers with clustering techniques

The goal of clustering in this study is to identify groups of energy consumers who share similar psycho-social characteristics (e.g., concerns about environmental and/or financial aspects of energy-saving, vulnerability to social pressure, presence or absence of certain values and beliefs), as these are described in sections 2.1-2.4. The existence of groups with distinct characteristics would enable addressing them with targeted interventions.

Our clustering analysis has progressed along three main phases, as depicted in Figure 26:

- **Data pre-processing**, involving the imputation of missing values in the dataset, the normalization of features and a first check on their clusterability.
- **Clustering algorithm selection and parameterization**, including the selection of features that are input to the algorithm, the number of clusters when this is input to the algorithm and metrics for assessing the (dis)similarity between two objects or clusters.
- **Experimentation**, involving the actual derivation of clusters and their analysis.

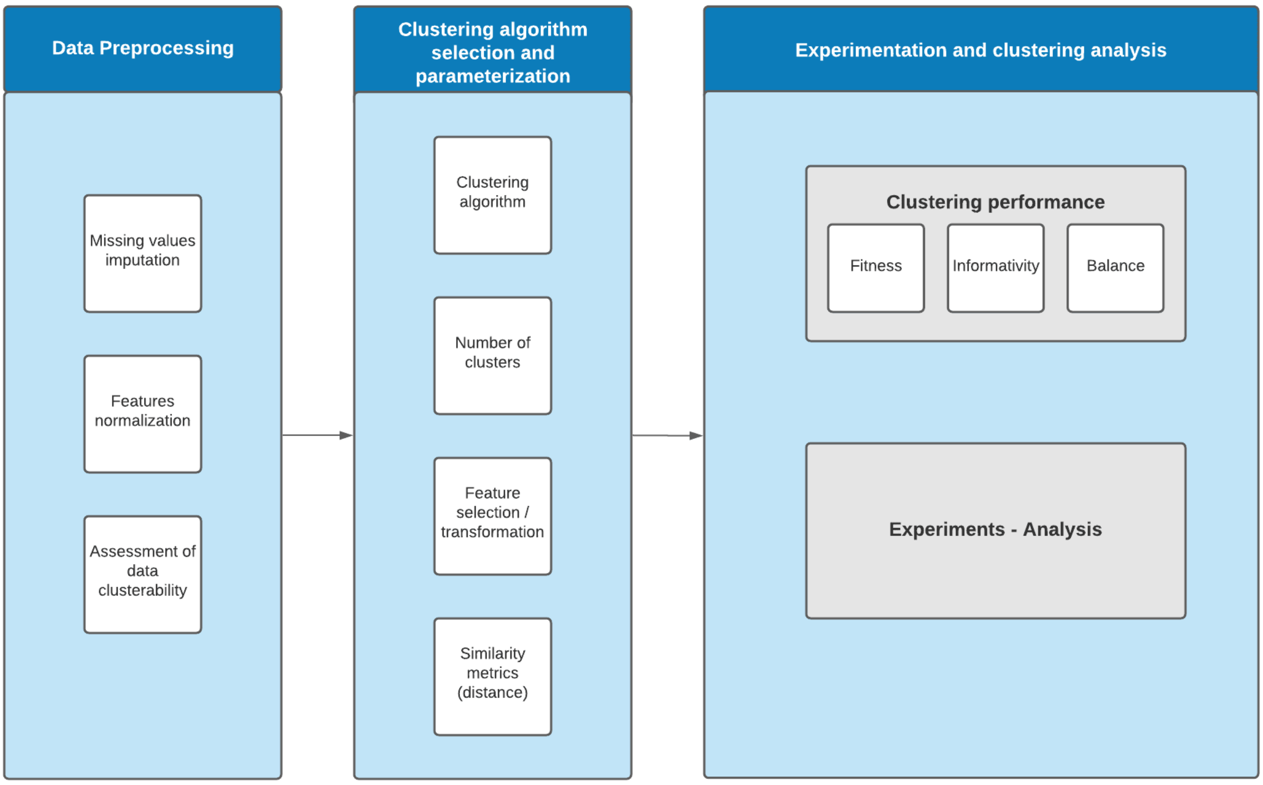


Figure 26: The three phases of our clustering analysis and the tasks involved in each of them.

5.2.1. Data pre-processing

The data pre-processing phase includes a number of steps applied to the initial dataset before the clustering algorithm is executed:

- *Imputation of missing values*: A few of the 3129 records of survey responses that survived the filtering steps in section 2.1, 0.99% or 31 records, contained unanswered questions in at least one of the 15 variables of interest for the clustering study, resulting in missing values. In order to proceed with the clustering approach, the imputation of these values was carried out. The k-nearest neighbours imputation algorithm is used to this end: for each missing value in a record, the 10 nearest neighbours with a non-missing value in the unanswered item/feature are identified based on the Euclidean distance and the missing value is set to the weighted average of those 10 values, where weights are assigned inversely proportionally to the distance between each neighbour and the record at hand.
- *Normalization*: The features of the dataset were transformed so that they are in the same scale. In particular, we applied the Min-Max normalization method to map each feature score on the [0,1] scale. If x_{uf} is the score of record *u* on feature *f*, the normalized score is given by

$$f(x_{uf}) = \frac{x_{uf} - \min_u x_{uf}}{\max_u x_{uf}}$$

5.2.1.1 Clustering tendency analysis

At this first phase, we have also carried out a clustering tendency analysis on the 3129 x 15 dataset (Cross & Jain, 1982). The goal of this analysis is to check whether the dataset indeed possesses a non-random structure that can yield meaningful clusters. Such analysis is deemed useful because clustering algorithms tend to produce clustering structures even when the dataset does not possess one and this typically results in random clusters without much value.

The Hopkins' statistic is an established and powerful statistic for measuring the clustering tendency of a dataset (Hopkins & Skellam, 1954). Its computation for our dataset proceeded as follows:

- A random sample of m data points was chosen, with m ranging, as prescribed, within [1/10,1/20] of the dataset size, and their distances to their nearest neighbours, $x_i, i \in \{1..m\}$ were computed (we discuss distance metrics later in section 5.2.2.4).
- A second set of m uniformly distributed points was generated and their distances, , $y_i, i \in \{1..m\}$ to their nearest neighbours in our dataset were computed.
- The statistic H was then computed as

$$H = \frac{\sum_{i=1}^{m} y_i}{\sum_{i=1}^{m} y_i + \sum_{i=1}^{m} x_i}$$

The Hopkins statistic lies in the range (0, 1). Uniformly distributed data have a Hopkins statistic value around 0.5 since x_i and y_i tend to assume similar values. On the contrary, when data is clustered, the values of x_i are typically much smaller than y_i, pushing the Hopkins statistic value closer to 1. Hence, the higher the value of the Hopkins statistic is, the stronger the indication we have for a clear clustering structure in the dataset.

We computed the statistic H multiple times, with different sets of dataset points and randomly generated sets of size m, to increase the confidence in its value. We also applied the statistic not only to the full set of 15 features but also to particular subsets of features; in particular, to all 455 subsets of 3 features, 1365 subsets of 4 features and 3003 subsets of 5 features. In each case, the test was repeated for 100 times with sample sizes ranging from 1/20 to 1/10 of the full dataset. Table 5.1 reports intervals of the average values the statistic assumes over the 100 runs in the case of 3, 4 and 5 features.

Table 3: Range of variation of Hopkins statistic values over all possible feature sets of size 3, 4, 5 and 15.

Feature set size	**15**	**5**	**4**	**3**
Hopkins statistic value range	0.678	0.65-0.79	0.71-0. 0.86	0.8-0.94

The main remarks out of Table 4, presenting the combinations with top Hopkins statistic scores, are:

- The value of the statistic increases as fewer features are retained to characterize the users, pointing to tighter clusters.
- For a fixed number of features, the value of the statistic varies significantly depending on which features are selected/retained for clustering. This is leveraged when using the Hopkins statistic for feature selection purposes (see section 5.2.3.2.1).
- The statistic assumes smaller values when more dimensions are considered but even for the full set of 15 features the value of 0.678 is significant.

Overall, the test suggests that there is potential for obtaining "good tight clusters", in particular if the feature space is reduced down to 3-5 features. This is pursued later in section 5.2.3.2.1 with the feature selection process.

5.2.2. Clustering algorithm selection and parameterization

In this second phase, we studied the following issues in the context of the clustering procedure:

5.2.2.1 Number of clusters

The optimal number of clusters is initially unknown and several executions of the algorithm may take place to determine the optimal choice. The number of clusters, when it constituted input to the clustering algorithm, ranged from 2 to 6 throughout our experimentation.

5.2.2.2. Set of features to consider in clustering

Feature selection and feature transformation are two methods that determine the actual subset of existing features (in case of feature selection) or new features (in case of feature transformation) that the clustering algorithm works with.

The feature selection process, in particular, was carried out in three different ways, which take into account the set of interventions to different extent:

- *Standard/intervention-unaware*: At one extreme, the feature selection did not account at all for the interventions at hand and proceeded all the way to select the most promising features through the use of standard techniques such as the Hopkins statistic that do not account for the interventions at hand (we described the Hopkins statistic earlier a measure of clustering tendency in 5.2.1.1, and later we will discuss its use for feature selection).
- *Intervention-driven:* At the other extreme, the feature selection algorithm was completely bypassed and the set of features for clustering was chosen according to how informative they can be for selecting interventions. For example, Subjective Norm (SN) is a feature that can identify energy consumers possibly responding to interventions leveraging social pressure; or, FIN_CONCERN may point to users who are more interested in monetary incentives or should be targeted with real-time feedback on monetary implications of their choices (e.g., every time they try to change their thermostat settings towards a higher value).
- *Intervention-assisted:* This was the intermediate case between the two extremes, the standard and the intervention-aware feature selection. In this case, we first used the standard feature selection process to come up with "good" enough feature sets and then we considered their relevance to interventions as a second-level criterion to choose the feature set to work with in clustering.

Feature transformation is an alternative to the feature selection process. The original features are subject to dimensionality reduction using the Principal Component Analysis (PCA), which yields a set of new features called Principal Components (PCs). The number of PCs that were input to the clustering analysis was either as a fixed input parameter or it was determined as the number of PCs that could explain a certain percentage varex of data variance. The total explained variance is the sum of the explained variances by each PC. When varex is used to determine the number of PCs, this is taken to be the minimum number needed such that the total explained variance by those PCs exceeds varex. (Lever et al, 2017).

5.2.2.3. Clustering algorithm

We primarily worked with two popular clustering algorithms:

- *k-means*: k-means is the most popular unsupervised machine learning algorithm. The parameter *k* indicates the number of clusters that will be formed and forms an input of the algorithm. Every observation is allocated to the nearest centroid, based on a distance metric, and the objective is to minimize the within-cluster sum of squares distance (variance). Given an initial allocation of cluster centroids, which can be either random or driven by some criteria, an iterative process takes place, at each step of which the cluster centroids are updated so that the objective function is minimized. The algorithm ends when the cluster centroids have been stabilized.

- *Agglomerative hierarchical clustering*: The objective in agglomerative hierarchical clustering algorithms is to create a hierarchy of clusters, i.e., the outcome of the algorithm is not just one clustering structure but rather *N* different clustering structures with cluster size ranging from 1 to *N*, where *N* is the number of observations. The number of clusters and the corresponding clustering structure can then be chosen *a posteriori* according to different criteria that evaluate the fitness of the structures. The algorithm proceeds in a number of sequential steps. In the beginning, each observation is viewed as one cluster (level-*N* clustering structure). Then, the two most similar clusters are merged into a larger one giving rise to the level-(*N*-1) clustering structure consisting of *N*-1 clusters. The algorithm proceeds with merging two existing clusters in each step, thus reducing the number of cluster by one, till all observations are merged into one big cluster (level-1 clustering structure).

5.2.2.4. Measures of distance/similarity between observations

Distance or similarity measures are core components used by clustering algorithms to group similar data points into the same clusters, while dissimilar data points are placed into different clusters. The distance measures we use are the Euclidean and Manhattan distances and the Cosine Similarity.

Both Euclidean and Manhattan distances belong to the Minkowski family. The Minkowski distance is defined as:

$$d_{min} = \left(\sum_{i=1}^{n} |x_i - y_i|^m\right)^{1/m}, m \geq 1,$$

where m is a positive real number and x and y are two vectors (observations) in the n-dimensional feature space and x_i, y_i their values in feature i. Then:

- The *Manhattan distance* is a special case of the Minkowski distance for m=1. This method is very sensitive to outliers. The fact that the survey data are on a 5-point Likert scale (7 for ATT), guarantees that there are no significant outliers, making it a proper choice.
- The *Euclidean distance* is a special case of the Minkowski distance for m=2. The Euclidean distance is very popular for clustering. This method is less sensitive to outliers than the Manhattan distance metric. One important characteristic is that the largest-scaled features dominate the others. This is the reason why data scaling (normalization) was deemed important during the data pre-processing phase.
- On the other hand, the *Cosine Similarity* measure is defined as:

$$CosineSim(x, y) = \frac{\sum_{i=1}^{n} x_i y_i}{\|x\|_2 \|y\|_2}$$

where $\|y\|_2$ is the Euclidean norm of vector $y = (y_1, y_2, \ldots, y_n)$ defined as:

$$\|y\|_2 = \sqrt{y_1^2 + y_2^2 + \cdots + y_n^2}$$

An obvious advantage of Cosine Similarity is that it does not depend on the number of features (length of vector y).

5.2.3. Experimentation and clustering analysis

5.2.3.1 Measures of clustering performance

Besides the extent to which the resulting clustering structure facilitates the identification of proper behavioural interventions, we are also interested in the following aspects that characterize it:

- *Clustering Fitness*: It assesses how "similar" are energy consumers assigned to the same cluster as opposed to those assigned to different clusters. We measured it using the Silhouette Score. The Silhouette score ranges from −1 to +1 and a high value indicates that an object is well matched with its own cluster and poorly matched with other clusters. The average Silhouette score over all users counts as a metric to assess the fitness of the overall clustering structure.
- *Cluster Balance*: It considers the distribution of cluster sizes, i.e., whether there are some extremely large or small clusters. At the one extreme, one large cluster gathering almost all consumers would not give us much information about targeted interventions. At the other extreme, many small clusters with tens of users would not be very reliable as valid targets of distinct behavioural trends. We have not used a particular metric to assess the balance of the clusters, but we used instead the rule of thumb that each cluster should not represent less than 5% of the observations.

5.2.3.2 Clustering analysis

We have carried out a number of experiments that span in different ways the parameterization space described in section 5.2.2. Since their outcome is similar (as it will be shown and explained in what follows) we report an indicative subset of those experiments, which can be categorized under:

- clustering with feature selection based on the Hopkins statistic; and
- clustering with feature transformation with Principal Components Analysis (PCA).

Additional experiments with different parameterization are presented in Annex II.

5.2.3.2.1 Clustering with feature selection

The parameterization of this experiment is as follows:

Feature selection: intervention-assisted. We retrieve combinations of n features, n in {3,4,5}, that score in the top-10 in the Hopkins statistic for each value of n, as shown in Table 4. We then select sets of features that are more informative regarding applicable interventions and execute clustering with 2-5 clusters as input.

Clustering algorithm: k-means. The Euclidean distance is employed as the distance metric and the number of clusters is given as an argument to the algorithm.

Clustering fitness assessment: Silhouette score.

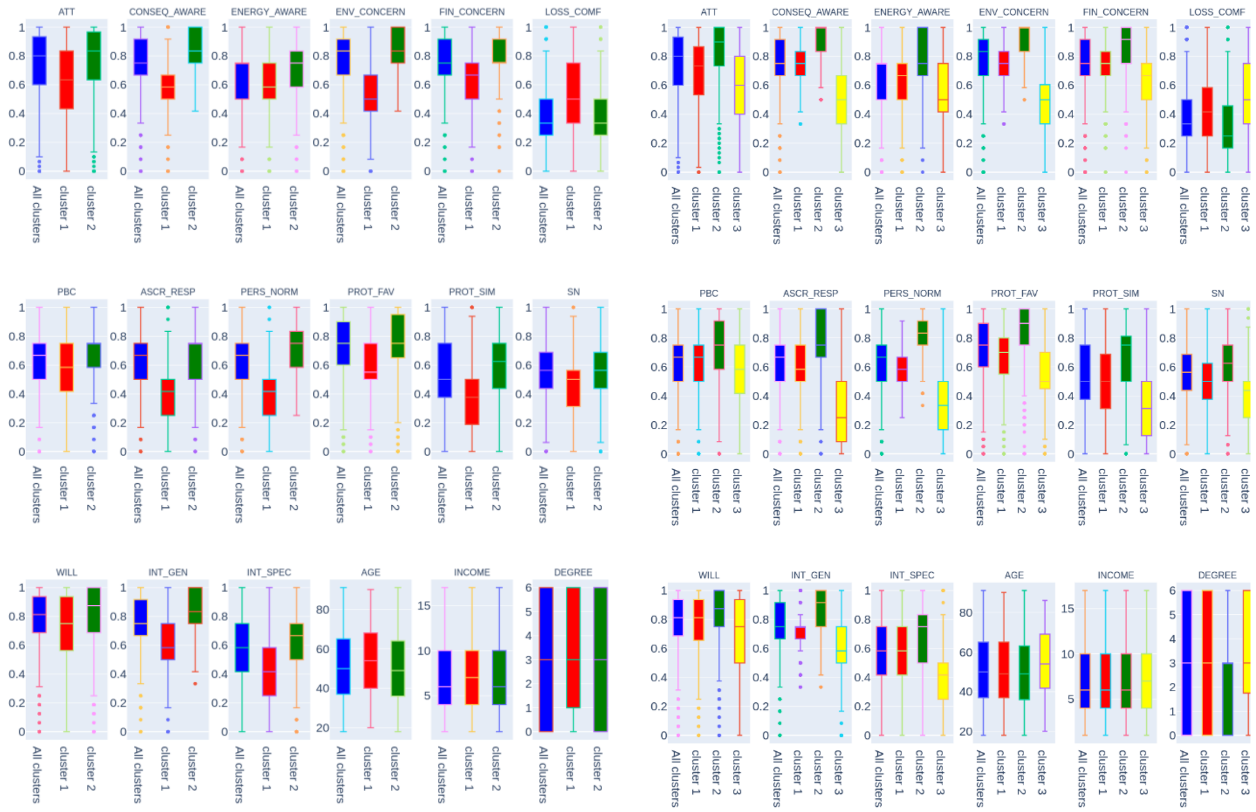
Table 4: Feature sets of size 3, 4, 5 achieving top Hopkins statistic scores.

Feature set size	**5**	**4**	**3**
Hopkins statistic value range	**0.65-0.79**	**0.71-0. 0.86**	**0.8-0.94**
Feature sets with top score	'CONSEQ_AWARE, ENV_CONCERN, ASCR_RESP, PERS_NORM, INT_GEN': 0.79 'CONSEQ_AWARE, ENV_CONCERN, FIN_CONCERN, PERS_NORM, INT_GEN': 0.79 'CONSEQ_AWARE, ENV_CONCERN, FIN_CONCERN, ASCR_RESP, INT_GEN': 0.79 'CONSEQ_AWARE, ENV_CONCERN, FIN_CONCERN, ASCR_RESP, PERS_NORM': 0.79 'CONSEQ_AWARE, ENERGY_AWARE, ENV_CONCERN, ASCR_RESP, PERS_NORM': 0.79 'CONSEQ_AWARE, ENERGY_AWARE, ENV_CONCERN, ASCR_RESP, INT_GEN': 0.79 'CONSEQ_AWARE, ENERGY_AWARE,	'CONSEQ_AWARE, ENV_CONCERN, ASCR_RESP, INT_GEN': 0.86 'CONSEQ_AWARE, ENV_CONCERN, FIN_CONCERN, INT_GEN': 0.85 'CONSEQ_AWARE, ENV_CONCERN, PERS_NORM, INT_GEN': 0.85 'CONSEQ_AWARE, ENV_CONCERN, FIN_CONCERN, ASCR_RESP': 0.85 'ENV_CONCERN, FIN_CONCERN, PERS_NORM, INT_GEN': 0.85 'CONSEQ_AWARE, ENV_CONCERN, ASCR_RESP, PERS_NORM': 0.85	'FIN_CONCERN, INT_SPEC, INT_GEN': 0.94 'ENV_CONCERN, INT_SPEC, INT_GEN': 0.93 'CONSEQ_AWARE, ENERGY_AWARE, INT_GEN': 0.93 'CONSEQ_AWARE, ENV_CONCERN, ASCR_RESP': 0.93 'ENV_CONCERN, PERS_NORM, INT_GEN': 0.93 'CONSEQ_AWARE, ENV_CONCERN, INT_GEN': 0.93 'CONSEQ_AWARE, PBC, INT_SPEC': 0.93 'ENV_CONCERN, FIN_CONCERN, INT_GEN': 0.93

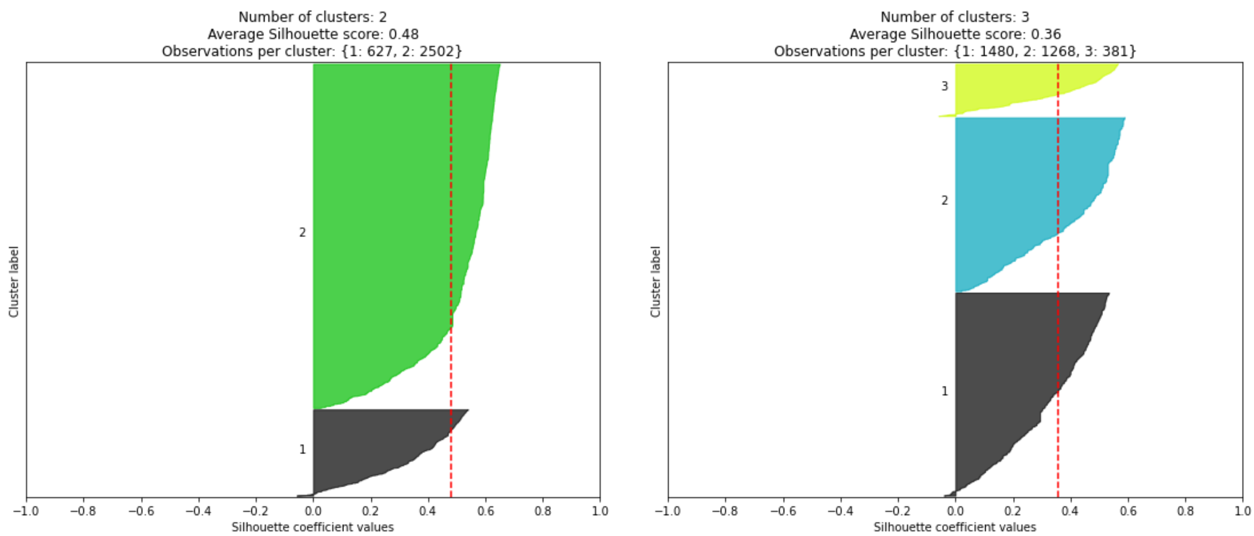
	ENV_CONCERN, FIN_CONCERN, INT_GEN': 0.79	'ENV_CONCERN, FIN_CONCERN, ASCR_RESP, INT_GEN': 0.84	'ENV_CONCERN, ASCR_RESP, INT_GEN': 0.93
	'CONSEQ_AWARE, ENV_CONCERN, ASCR_RESP, WILL, INT_GEN': 0.78	'ENV_CONCERN, ASCR_RESP, PERS_NORM, INT_GEN': 0.84	'ENERGY_AWARE, FIN_CONCERN, INT_GEN': 0.93
	'CONSEQ_AWARE, ENERGY_AWARE, ENV_CONCERN, PERS_NORM, INT_GEN': 0.78	'CONSEQ_AWARE, ENV_CONCERN, FIN_CONCERN, PERS_NORM': 0.84	
	'CONSEQ_AWARE, ENV_CONCERN, FIN_CONCERN, WILL, INT_GEN': 0.78	'CONSEQ_AWARE, ENERGY_AWARE, ENV_CONCERN, INT_GEN': 0.84	

Experiment results: We have experimented with several of the feature sets listed in Table 4. We present and discuss below one of those experiments that carried out clustering on the {CONSEQ_AWARE, ENV_CONCERN, ASCR_RESP, PERS_NORM} feature subset. The Hopkins statistic value equals 0.85 for this feature subset, whose features strongly point to important motivating factors for energy-saving behaviour. The results of experimentation with other subsets are similar to what we report below for this subset and the conclusions drawn here apply for those experiments as well.

Figure 27 shows the cluster-average scores of the resulting clusters in the 15 features under clustering structures of 2 and 3 clusters, while Figure 28 does the same when we demand structures of 4 and 5 clusters. The figures also depict the distribution of Silhouette values of all clustered observations in each case.



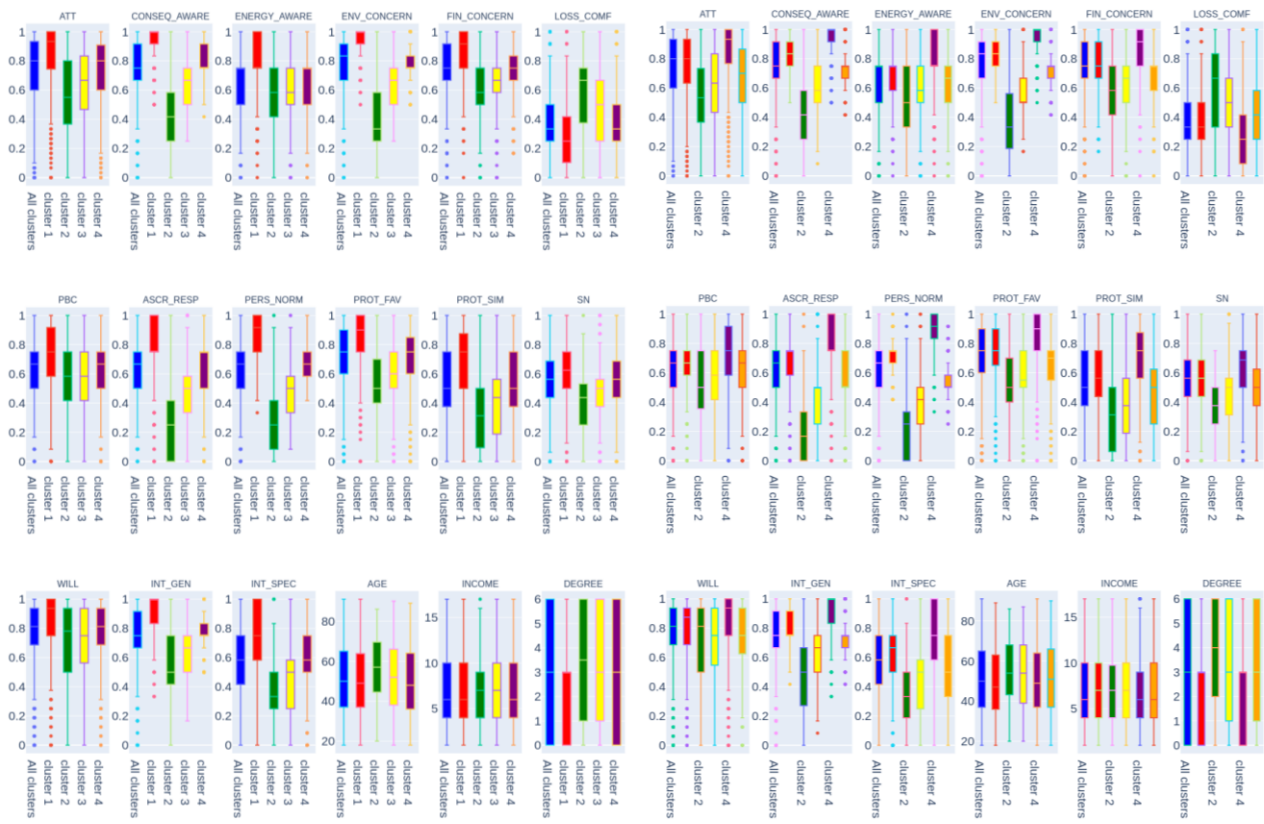
(a) Cluster-average scores in 15 features and 3 socio-demographic variables



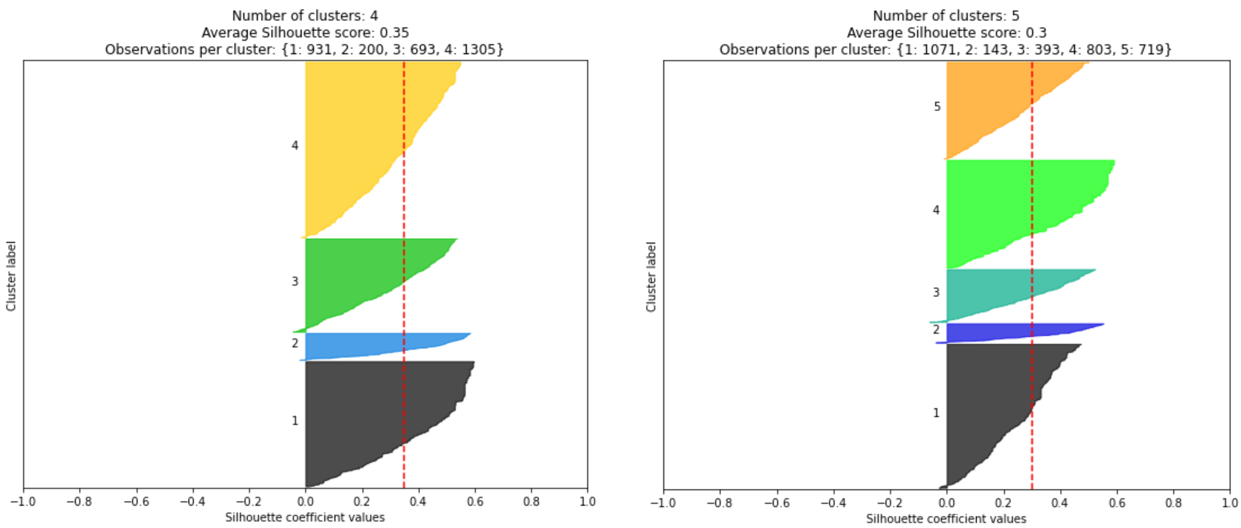
(b) Silhouette values

2 clusters 3 clusters

Figure 27: Box plots of cluster-based scores in the 15 features as well as their age, income and educational degree characteristics: intervention-assisted feature selection, k-means with Euclidean distance, number of clusters equals 2 (left column) and 3 (right column)



(a) Cluster-average scores in 15 features and 3 socio-demographic variables



(b) Silhouette values

4 clusters 5 clusters

Figure 28: Box plots of cluster-based scores in the 15 features as well as their age, income and educational degree characteristics: intervention-assisted feature selection, k-means with Euclidean distance, number of clusters equals 4 (left column) and 5 (right column)

In terms of fitness, the clustering structures perform well: with very few exceptions, observations have positive silhouette values, implying that they fit better in the group they were assigned to by clustering rather than any other group. The average values over all observations range from 0.34 (for 5 clusters) up to 0.47 (for only 2 clusters).

Looking at the cluster sizes, we avoid the extreme cases of very large or very small clusters, as shown in Table 5. Only in the case of 5 clusters, we get one cluster (cluster 2) with size marginally smaller than the 5% of the overall observations.

Table 5: Sizes of clusters emerging from the clustering experiment when k-means generates 2, 3, 4 and 5 clusters.

Number of clusters	Cluster 1 size	Cluster 2 size	Cluster 3 size	Cluster 4 size	Cluster 5 size	Silhouette value
2	627	2502				0.48
3	1480	1268	381			0.36
4	931	200	693	1305		0.35
5	1071	143	393	803	719	0.3

How informative are these clustering structures? Looking into the 2-cluster structure in Fig. 27, we can remark that there is one cluster (cluster 2) that scores consistently higher in all 15 features than the other (cluster 1). When we add one cluster to come up with a 3-cluster structure, we get:

- **one "good" cluster (cluster 2) that outperforms the other two in all 15 features;**
- **one "bad" cluster (cluster 3) with the worst scores in all 15 features; and**
- **a third cluster (cluster 1), which positions in between the other two in that it consistently demonstrates intermediate scores in all 15 features. This cluster is a quite large one (1268 users), drawing a significant portion of users who were assigned to the two "extreme" (the good and the bad) clusters in the 2-cluster structure.**

When we look into the 4-cluster structure, we can still identify one "good" and one "bad" cluster (clusters 1 and 2, respectively) together with two more clusters: cluster 3 that ranks consistently 3rd in all 15 features and cluster 4 that ranks consistently 2nd in all 15 features! Predictably, this trend of proportional scaling of feature scores from cluster to cluster pertains to the 5-cluster structure as well. Thereby, the top-performing cluster is cluster 4, followed by clusters 1, 5, 3 and 2, which is the worst-performing one in all 15 features.

Summarizing, this clustering experiment yields clustering structures that perform acceptably in terms of fitness and balance but demonstrate a very particular pattern in the way the cluster-based scores rank in the different features: the rankings of a given cluster in all features are identical. This pattern of strongly correlated score rankings is in agreement with the quite high positive pairwise feature correlations reported in Figure 18 in section 4.2 but **does not provide us with much information as to which intervention would be most appropriate for each energy consumer group**. In positive terms, if we attempted to interpret the information conveyed by such clustering structure, we could conclude that:

- There are energy consumers who demonstrate strong intentions to engage in energy-saving behaviour, motivated by a combination of factors (concern about the environment, interest in savings related to heating arrangements, good awareness of how to save energy and a strong sense of moral obligation to save energy). No intervention is particularly needed for those consumers other than preserving their positive attitude towards energy-saving.
- Then, there are one or more groups of energy consumers who demonstrate these characteristics in varying and highly correlated levels of intensity. These groups of consumers do not exhibit some particular behavioural traits (e.g., concern for the environment or the financial implications of energy consumption) to a distinctly higher extent than other groups, so that they could be targeted through specific interventions. We could essentially try the same (any) interventions, albeit at higher intensity, for each group, as far as the specific intervention allows this regulation (e.g., tuning the frequency of real-time feedback on the financial or environmental footprint of a specific energy-saving action).

We have carried out the cluster analysis with different parameterization of its components (clustering algorithm, distance measure), as described in section 5.2.2. The same pattern has been consistently observed in all these experiments (refer to Annex II for more such experiments).

5.2.3.2.2 Clustering with feature transformation – PCA

Principal Component Analysis (PCA) is a common pre-processing technique for clustering analysis. It transforms the original feature space (features and their variation range) to a lower-dimension feature space made up of weighted combinations of the original features called principal components (PCs). We can then project the data on this space of fewer dimensions while preserving as much of the data's variation as possible. In PCA, the number of principal components that are used for this transformation is determined by the amount of the variance in the data they can preserve (interchangeably: explain).

We report two experiments involving transformation of the original feature space.

PCA with 15-dimensional original feature space and 4-dimensional transformed feature space

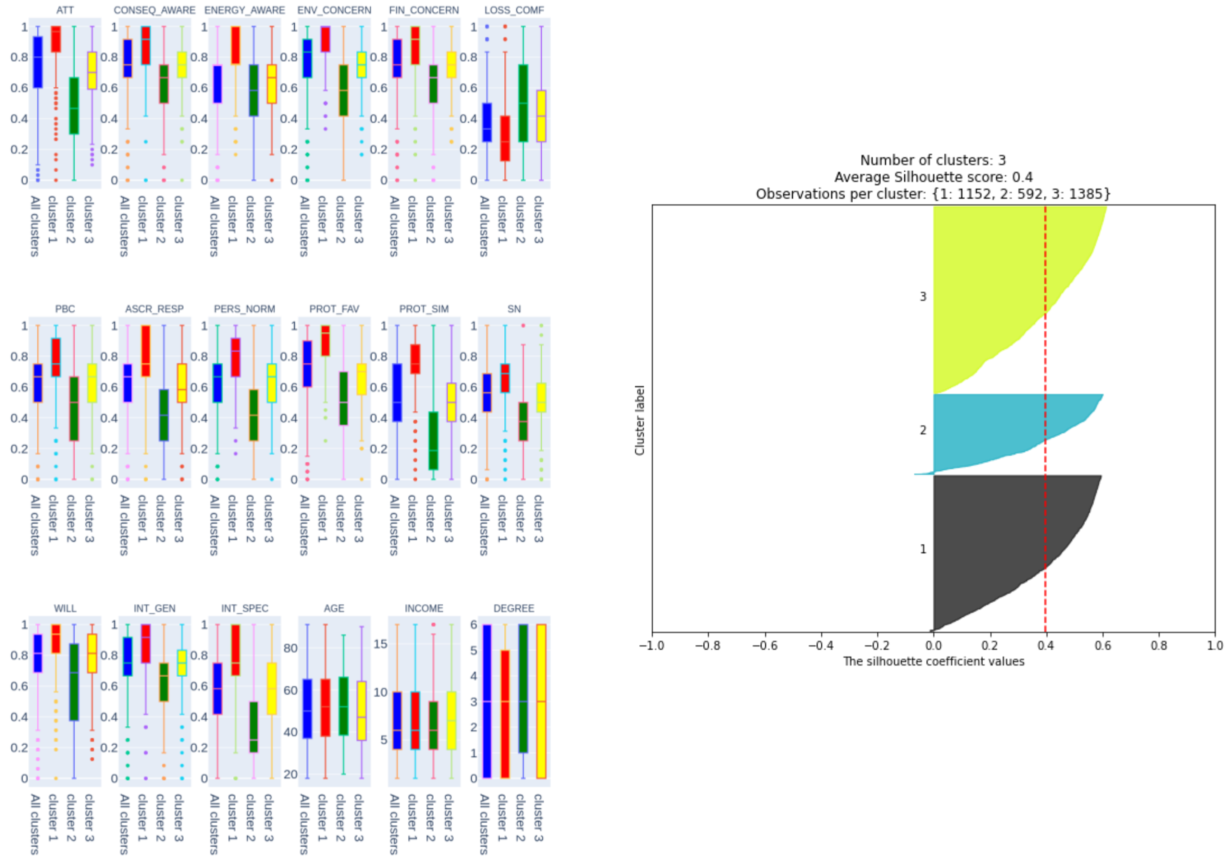
The parameterization of this experiment was as follows:

Feature transformation: Original feature space of 15 features; transformed feature space of dimension 4, the four PCs explaining 62% of the variance in data.

Clustering algorithm: k-means. The Euclidean distance is employed as the distance metric and the number of clusters is given as an argument to the algorithm.

Clustering fitness assessment: Silhouette score

Figure 29 plots the per cluster scores in the 15 features under a 3-cluster structure. The results with 4-cluster and 5-cluster structures are similar and follow the same pattern as the previous intervention-assisted effort with the original feature set.



(a) Cluster-average scores in 15 features and 3 socio-demographic variables

(b) Silhouette values

Figure 29: Box plots of cluster-based scores in the 15 features as well as their age, income and educational degree characteristics (left) and cluster silhouette values (right): feature transformation, k-means with Euclidean distance, number of clusters = 3

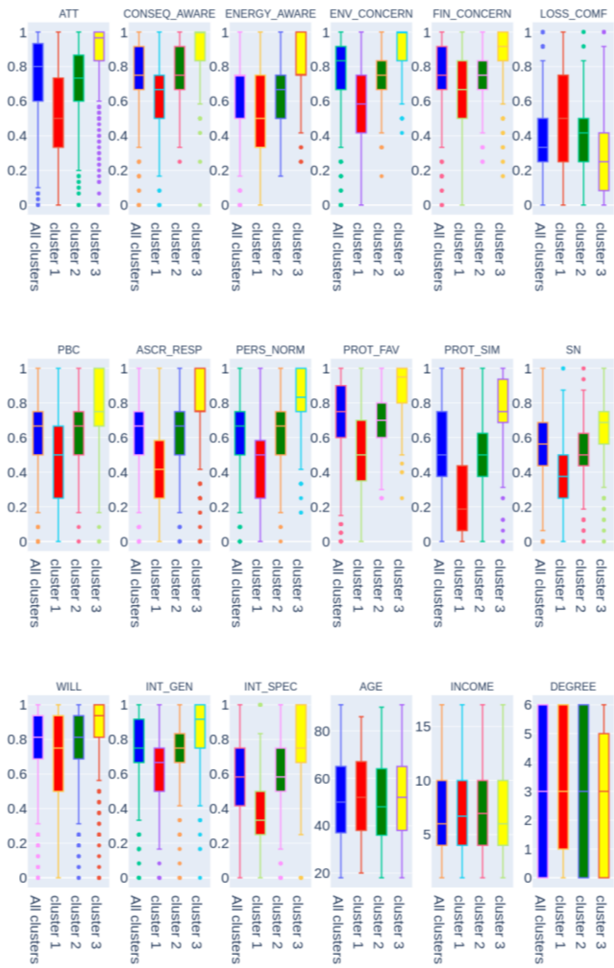
PCA with 10-dimensional original feature space and var_{ex} used to determine the dimension of the transformed feature space

Feature transformation: Original feature space of 10 features {ASCR_RESP, CONSEQ_AWARE, ENERGY_AWARE, ENV_CONCERN, FIN_CONCERN, LOSS_COMF, PBC, PROT_FAV, PROT_SIM, SN} which are the first-level or independent model variables as they appear in the behavioural research model (Figure 1). The dimension of the transformed feature space results from the requirement var_{ex} = 60%

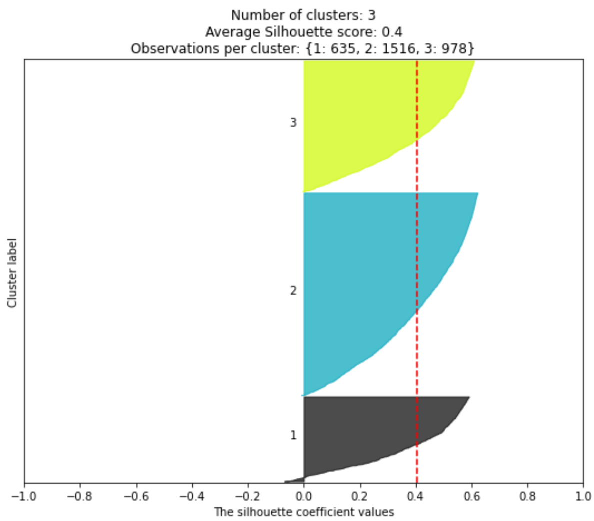
Clustering algorithm: k-means. The Euclidean distance is employed as the distance metric and the number of clusters is given as an argument to the algorithm.

Clustering fitness assessment: Silhouette score

Figure 30 plots the per cluster scores in the 15 features under a 3-cluster structure. The results with 4-cluster and 5-cluster structures are similar and follow the same pattern as in the other experiments discussed so far.



(a) Cluster-average scores in 15 features and 3 socio-demographic variables



(b) Silhouette values

Figure 30: Box plots of cluster-based scores in the 15 features as well as their age, income and educational degree characteristics (left) and cluster silhouette values (right): feature transformation, k-means with Euclidean distance, number of clusters = 3

We observe again that the same pattern of identical cluster-based score rankings across constructs persists in both scenarios where the PCA is applied. Looking at how the percentage explained variance scales with the number of components in Table 5.4, we can see that a single PC can already explain more than 40% of variance. This is another symptom of the significant correlation that is evidenced between the features, reflecting the highly interrelated responses of the survey users to the different constructs.

Table 6: The number of principal components needed per explained variance threshold.

Explained Variance	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Principal Components	1	1	1	2	4	5	8	11

...

Hence, the combination of clustering with feature transformation yields clustering structures that resemble those we obtained with feature selection. We now turn into the second approach we outlined in the beginning of section 0 that specifies energy consumer classes as a combination of conditions that the respondents' scores in the different features should satisfy.

5.3. Segmentation of energy consumers with *a priori* class specification

As already mentioned, this approach is methodologically distinct from clustering analysis. The input is the same, namely the scores of respondents in the variables measured by the survey. However, the groups of users (hereafter called "classes") are not emerging in automated manner as the result of clustering. We rather specify beforehand certain classes (interchangeably: profiles) of users and then seek to classify the survey respondents into one of those. The specification of classes is intervention-driven: we define them so that it is clear which intervention(s) is(are) applicable to them.

When comparing this approach to the cluster analysis, we can remark the following:

- With clustering, the groups that emerge need to be a posteriori analysed and characterized to find their defining characteristics. Depending on this cluster characterization, the selection of proper interventions for them may become a difficult task (as seen in sections 5.1.2-5.1.4). On the contrary, with this intervention-based class definition, we know from the beginning which interventions are proper for each class.
- Clustering yields disjoint user groups, whereby each survey respondent (user) is uniquely assigned to a single cluster. With the a priori specification of classes, several users may find themselves satisfying the requirements/criteria of more than one classes; we call them "multihomed" users. This is not necessarily a problem in our case since this implies the appropriateness of more than one intervention for the specific users. If, however, it is necessary to have one-to-one mapping of users to classes, i.e., due to the cost of carrying out interventions, we need an additional processing step to choose one of those classes for each of those users.
- Likewise, there may be users who are not falling under any of those classes. Again, we need a way to find out how we are going to treat those users and which intervention, if any, applies to them.

We can identify three steps in the way this intervention-based segmentation of energy consumers proceeds:

5.3.1. Step 1: Specification of energy consumer classes

Each energy consumer class is characterized by the following components:

- the class features, namely the features that participate in the class specification. As class features can serve any subset of the original 15 features.
- the range (interval) of acceptable scores in the class features. In other words, each class feature comes up with a class-specific acceptance interval.

To classify a user into a given class, her scores in all class features should lie within the corresponding acceptance interval. Assuming normalized scores in [0,1], the class-specific acceptance interval for a feature *v* may be of three types[4]:

- [thr, 1]: the user's score in feature v should be higher than some threshold value thr.
- [0, thr]: the user's score in feature v should be lower than some threshold value thr.
- [thr_1, thr_2], $thr_1 < thr_2$: the user's score in feature v should lie within an interval [thr_1, thr_2].

In principle, these threshold values are feature-dependent.

The specification of classes is a heuristic exercise combining the analysis of the situation at hand and experience-based intuition. In our case, the iterative process has led us to specify **6 classes of energy consumers**. One of them represents ideal consumers who consistently score high in all the 15 features in Figure 25. In the specification of the other five classes, we try to identify distinct features (variables) that (a) on the one hand, appear to be justifying lower intentions to adopt energy-saving behaviour; and, (b) on the other hand, prescribe specific type of interventions for strengthening these intentions.

The six classes are the following:

1. Environmentally conscious and well-informed energy consumers: These energy consumers resemble the "idealistic energy savers" in (Sütterlin, B., Brunner, T. A., & Siegrist, M, 2011). They combine the high concern about the environment with good knowledge about the climate change problem, its context and its consequences, together with a strong sense of personal responsibility for action against it. Energy saving sets a favourable paradigm for them and the intentions to engage into energy-saving activities, with respect to heating but also overall, are very strong.

In terms of features and feature acceptance intervals, this class demands:

---

[4] The shape of acceptance intervals does not change if we consider original scores in the Likert scale. For instance, the first type would become [thr, 5] for scores in the 5-Likert scale or [thr,7] in the 7-Likert scale, with the threshold numbers appropriately scaled in each case.

Feature	CONSEQ_ AWARE	ENV_CONCERN	ASCR_RESP	PROT_FAV	INT_SPEC	INT_GEN
Acceptance interval type	[thr_{11}, 1]	[thr_{12}, 1]	[thr_{13}, 1]	[thr_{14}, 1]	[thr_{15}, 1]	[thr_{16}, 1]

where all thresholds equal "high" values to capture the consistently excellent scores of this class's users in all class features.

At intervention level, these are users whose existing interest and behaviour need to be preserved through regular reminders of the environmental issues and the importance of energy-saving behaviour. Such reminders could be operationalized through educational material (e.g., documentaries), brief letters exposing the energy situation (similar to what humanitarian aid NGOs send to their members) but also general-purpose marketing campaigns in social media.

2. Concerned but comfort-oriented energy consumers. In principle, users in this class demonstrate clear intentions for acting in an energy-friendly manner. However, energy-saving behaviour with respect to heating in particular, implies compromises that appear not to be acceptable for them, such as setting the thermostat at lower temperature and wearing more clothes to make up for it. The characteristic that opens the way to interventions is that users in this class are highly concerned about the monetary cost involved in higher energy consumption.

In terms of features and feature acceptance intervals, this class demands:

Feature	LOSS_COMF	FIN_CONCERN	INT_SPEC	INT_GEN
Acceptance interval type	[thr_{21}, 1]	[thr_{22}, 1]	[0, thr_{23}]	[thr_{24}, 1]

where:

- thr_{21} equals a middle-to-high value to embrace users with better-than-average score in the loss comfort feature;
- thr_{22} is a high value demanding from class users high scores in the FIN_CONCERN feature;
- thr_{23} also equals a middle-to-high value, to capture intentions about heating-related energy saving that clearly lag behind those in the first class;
- thr_{24} takes a high value to ensure that users in this class have strong intentions for saving energy, in principle and as far as this does not threaten their comfort.

For those users, the financial burden of excessive consumption may outweigh or, at least, balance the discomfort that heating-related energy-saving activities bring about. Hence, candidate interventions for this class include real-time feedback messages that remind her about the additional cost (projected at yearly or multi-annual basis) resulting from her actions (e.g., raising the thermostat target temperature or wasting hot water in the bathroom).

3. Concerned but lacking awareness energy consumers. Much as loss of comfort serves as a barrier for energy-saving behaviour in the 2nd class, energy awareness is the Achilles' heel of energy consumers in this third class. Whereas they are concerned about the environment, they acknowledge the risks for it and they are willing to undertake their share of responsibility in this matter, they miss the practical knowledge that would strengthen their intention to adopt ideal energy-saving behaviour.

In terms of features and feature acceptance intervals, this class demands the conditions:

Feature	CONSEQ_AWARE	ENERGY_AWARE	ENV_CONCERN	INT_GEN
Acceptance interval type	$[thr_{31}, 1]$	$[0, thr_{32}]$	$[thr_{33}, 1]$	$[0, thr_{34}]$

where:

- thr_{31} and thr_{33} equal middle-to-high values, denoting higher-than-average scores in the CONSEQ_AWARE and ENV_CONCERN features;
- thr_{32} equals a "low" value so that the acceptance interval for the Energy Awareness feature includes users with distinctly lower scores than in other classes;
- thr_{34} is a middle-to-high value, to reflect that the generic intentions of users in this class lag behind the "star" 1st class.

There are two types of interventions that are candidate for this class of users. The first one, more appropriate for the "lazy" users who are not willing to invest effort in learning tips and secrets to save energy, is the use of energy(heating)-friendly *defaults* in the operation of devices. The second one, more appropriate for those who want to learn new things, is the use of tips, either online, when users take some energy-related action (e.g., changing the thermostat setting) or offline.

4. Materialistic energy consumers escaping their personal responsibility. This class includes energy consumers combining lower than average energy-saving intentions with a low anticipation of personal responsibility to act and high concern for the financial implications of energy-saving activities on the monthly bills.

In terms of features and feature acceptance intervals, this class prescribes the following:

Feature	ASCR_RESP	FIN_CONCERN	INT_GEN
Acceptance interval type	$[0, thr_{41}]$	$[thr_{42}, 1]$	$[0, thr_{43}]$

where:

- thr_{41} is a low-to-middle score so that only users with low scores in ASCR_RESP are included in the class;
- thr_{42} is a high value demanding from class users high scores in the FIN_CONCERN feature; and

- thr_{43} is a middle-to-high value, to reflect that the generic intentions of users in this class lag behind the top 1st class, playing the same role that thr_{33} plays in the specification of the 3rd class

The motivating idea for the specification of this class is that the financial concerns of users are the characteristic that can be targeted by an intervention intending to counterbalance the missing sense of personal responsibility for energy-saving behaviour. The lack of personal responsibility is a strong barrier to energy-saving behaviour. Targeting this directly rather than through the proxy of financial concerns calls for larger-scale interventions at the level of educational system.

5. Prone to social influence energy consumers. Energy consumers in this class state low intentions for heating-related energy saving behaviour but, contrary to the well-being driven consumers, they exhibit distinctly higher than average scores in the Subjective Norm variable, implying that they are "vulnerable" to interventions that try to leverage the social pressure effect.

In terms of features and feature acceptance intervals, this class requires the following conditions:

Feature	SN	INT_SPEC	INT_GEN
Acceptance interval type	$[thr_{51}, 1]$	$[0, thr_{52}]$	$[0, thr_{53}]$

where:

- thr_{51} is a middle-to-high score that lower-bounds the SN acceptance interval so that the scores of users classified into the 5th class are distinctly higher-than-average;
- thr_{52} and thr_{53} values should be high serving for 5th class users the same purpose that thr_{33} and thr_{43} did for the 3rd and 4th class, respectively

The idea is that interventions of the social comparison type are most appropriate for this class of users.

6. Indifferent energy consumers. The defining feature for this class, which appears to be serving as a barrier towards strong intentions for energy-saving, is the low perception of behavioural control. In this case, this is more related to perceived self-efficacy, i.e., if the users' belief that they have the capacity to engage in activities related to energy-saving and really have an impact (Bandura, 1991). Energy consumers in this class do not really identify with the prototype of energy-saver.

In terms of features and feature acceptance intervals, this class requires the following conditions:

Feature	PBC	PROT_SIM	INT_SPEC	INT_GEN
Acceptance interval type	[0, thr61]	[0, thr62]	[0, thr63]	[0, thr64]

where:

- thr_{61} is a low-to-middle score that upper-bounds the PBC acceptance interval so that the scores of users classified into the 6th class possess distinctly lower scores in the PBC feature;
- thr_{62} is also a low-to-middle score which sets an upper bound for the score of this class users in the PROT_SIM feature;
- thr_{63} and thr_{64} are values denoting moderate intentions for both general-purpose and heating-related energy-saving.

On the intervention side, facilitating nudges are applicable for this class. These may involve practical tips about energy conservation or use of defaults to make energy-saving alternatives more salient.

The six energy-consumer classes are parameterized. Their parameters are the threshold values that determine the class feature acceptance intervals. As it can be seen in the class specifications there are four kinds of thresholds (we assume normalized score values):

- high values left bounding feature acceptance intervals (type 1 thresholds). They could lie anywhere in the [0.7,0.85] interval. As such a threshold grows higher, fewer users satisfy the respective feature acceptance interval.
- high values right bounding feature acceptance intervals (type 2 thresholds). Plausible range for them is the interval [0.65, 0.75]. The higher this type of threshold, the more users score within the feature acceptance interval.
- moderate values left bounding feature acceptance intervals (type 3 thresholds). They take values in [0.45, 0.65]. The higher the value the fewer the users whose score lies in the respective feature acceptance interval.
- moderate values right bounding feature acceptance intervals (type 4 thresholds). They take values in [0.3, 0.5]. As thresholds of this type grow larger, the number of users satisfying the score condition increases.

In general, there are two possibilities for defining the thresholds that are relevant to each feature and each class:

- **Common thresholds.** The first one is to set common thresholds for all features. This means that there are four parameters to be defined, say thr1, thr2, thr3 and thr4, respectively, for the thresholds of type 1 to 4, respectively. The 24 threshold parameters of the six classes are mapped to these four parameters as shown.

Table 7: Threshold parameters to be defined under the common thresholds approach

Class feature threshold	Threshold type
thr_{11} - thr_{16}, thr_{22}, thr_{24}, thr_{42}	thr_1
thr_{23}, thr_{34}, thr_{43}, thr_{52}, thr_{53}, thr_{63}, thr_{64}	thr_2
thr_{21}, thr_{31}, thr_{33}, thr_{51}	thr_3
thr_{32}, thr_{41}, thr_{61}, thr_{62}	thr_4

NUDGE
Nudging consumers towards energy efficiency through behavioural science

- **Feature-specific thresholds.** Set different threshold(s) for each feature *j*, irrespectively of the class it relates to. This implies the determination of 1-4 thresholds for each one of the features that appear one or more times in the classes' specification. The threshold parameters that would need to be defined in this case are listed in the table below.

Table 8 : Threshold parameters to be defined when thresholds are assumed to be feature-specific

Feature	Relevant thresholds
CONSEQ_AWARE	thr_1,thr_3
ENERGY_AWARE	thr_4
ENV_CONCERN	thr_1, thr_3
FIN_CONCERN	thr_1
LOSS_COMFORT	thr_3
PBC	thr_4
ASCR_RESP	thr_1, thr_4
PROT_FAV	thr_1
PROT_SIM	thr_4
SN	thr_3
INT_SPEC	thr_1, thr_2
INT_GEN	thr_1, thr_2

Namely, we would need to find the values of 17 different parameters for the 12 features that are involved in the class specification process.

5.3.2. Step 2: Assignment of survey respondents to energy-consumer classes

The assignment of users to the six classes is carried out simultaneously with the parameterization of the six classes. The latter is the goal of an optimization problem, hereafter abbreviated as (OPT).

Given the specification of classes, **the objective of (OPT) is to determine the values of thresholds that end up minimizing the number of users who are not assigned to any of the six energy consumer classes**. For this optimization problem:

- Its decision variables are the parameter values and may be 4 or 17, depending on whether we work with common or feature-specific thresholds (see section 5.3.1). With normalized score values, these decision variables are continuous variables in [0,1]; if scores are not normalized, they are continuous variables in [1,5] ([1,7] in the case of ATT).
- Besides the allowed range of values for each parameter, there are two types of constraints related to:
    - *The order of the four types of thresh*olds, thr_1-thr_4. Namely, for any feasible problem solution it should hold that $thr_1 \geq thr_2 \geq thr_3 \geq thr_4$.
    - *The size of individual classes.* We demand to derive classes with at least 5% of the overall sample (around 156 users).

One additional decision relates to whether we will work with common or feature-specific thresholds. The first option implies lower computational complexity: we need to find 4 instead of 17 values. In the second case, we gain in terms of higher flexibility to optimize the objective at the expense of processing overhead. Moreover, we would need more effort to reason about the choice of different scores per variable to denote similar qualitative levels (e.g., high, low, moderate).

In what follows, we proceed with the first option. Under common thresholds, the six classes are parameterised as shown in Table 9.

Table 9: Specification of energy-consumer classes based on common thresholds

Class 1	**CONSEQ_AWARE** ≥ thr_1 AND **ENV_CONCERN**≥ thr_1 AND **ASCR_RESP** ≥ thr_1 AND **PROT_FAV** ≥ thr_1 AND **INT_SPEC** ≥ thr_1 AND **INT_GEN** ≥ thr_1
Class 2	**LOSS_COMF** ≥ thr_3 AND **FIN_CONCERN** ≥ thr_1 AND **INT_SPEC** ≤ thr_2 AND **INT_GEN** ≥ thr_1
Class 3	**CONSEQ_AWARE** ≥ thr_3 AND **ENERGY_AWARE** ≤ thr_4 AND **ENV_CONCERN**≥ thr_3 AND **INT_GEN** ≤ thr_2
Class 4	**ASCR_RESP** ≤ thr_4 AND **FIN_CONCERN** ≥ thr_1 AND **INT_GEN** ≤ thr_2
Class 5	**SN** ≥ thr_3 AND **INT_SPEC** ≤ thr_2 AND **INT_GEN** ≤ thr_2
Class 6	**PBC** ≤ thr_4 AND **PROT_SIM** ≤ thr_4 AND **INT_SPEC** ≤ thr_2 AND **INT_GEN** ≤ thr_2

We then carry out an exhaustive enumeration over all possible combinations of values that parameters thr_1-thr_4 can assume in the ranges [0.7,0.85], [0.65,0.75], [0.45,0.65] and [0.3, 0.5], respectively, in steps of 0.02. Table 10 summarizes the derived optimal solution.

Table 10: Optimal threshold parameterization (after exhaustive enumeration)

thr_1	**thr_2**	**thr_3**	**thr_4**	**Users classified in at least one class**	**class 1**	**class 2**	**class 3**	**class 4**	**class 5**	**class 6**
0.75	0.75	0.5	0.5	2132	529	400	440	259	392	112

Therefore, 2132 users can be directly assigned to one or more classes through this process for optimal threshold values $thr_1 = thr_2 = 0.75$ and $thr_3 = thr_4 = 0.5$, when feature scores are normalized ($thr_1 = thr_2 = 4$ and $thr_3 = thr_4 = 3$ when feature scores are measured in the 1-5 Likert scale).

Figure 31 plots the number of classes to which each of these 2132 users belongs. We can see that the majority (1180, 55.27%) of the users can be assigned to one class, whereas the rest are multihomed. In particular, 643 (30.16%) users satisfy the specifications of 2 classes, 262 users (12.27%) could be classified into 3 classes, 43 (2.01%) users could belong to 4 classes and another 4 (~0.2%) of them could even fit into 5 classes.

The number of users who are at first eligible for each class are 529, 477, 507, 425, 1041 and 465 users, respectively. To assign each user to only one class, we can randomly rank the classes and

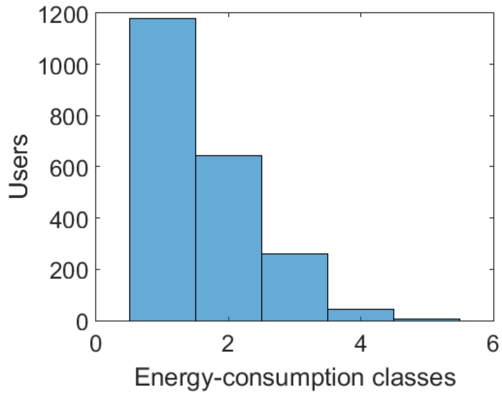


Figure 31: Distribution of classes a user is eligible for under the optimal parameterization of the six classes

then parse them sequentially to eliminate each time users who have earlier been assigned to another class. For instance, if we rank classes in the order 1-2-3-4-5-6, we come up with 529 (24.81%), 400 (18.76%), 440 (20.64%), 259 (12.15%), 392 (18.39%), and 112 (5.25%) users in each of the six classes, respectively.

Figure 32 plots how each class of users scores in each of the main 13+2 variables of interest in the survey, considering the responses of those 2132 users who have been assigned to one of the six energy-saving behaviour classes.

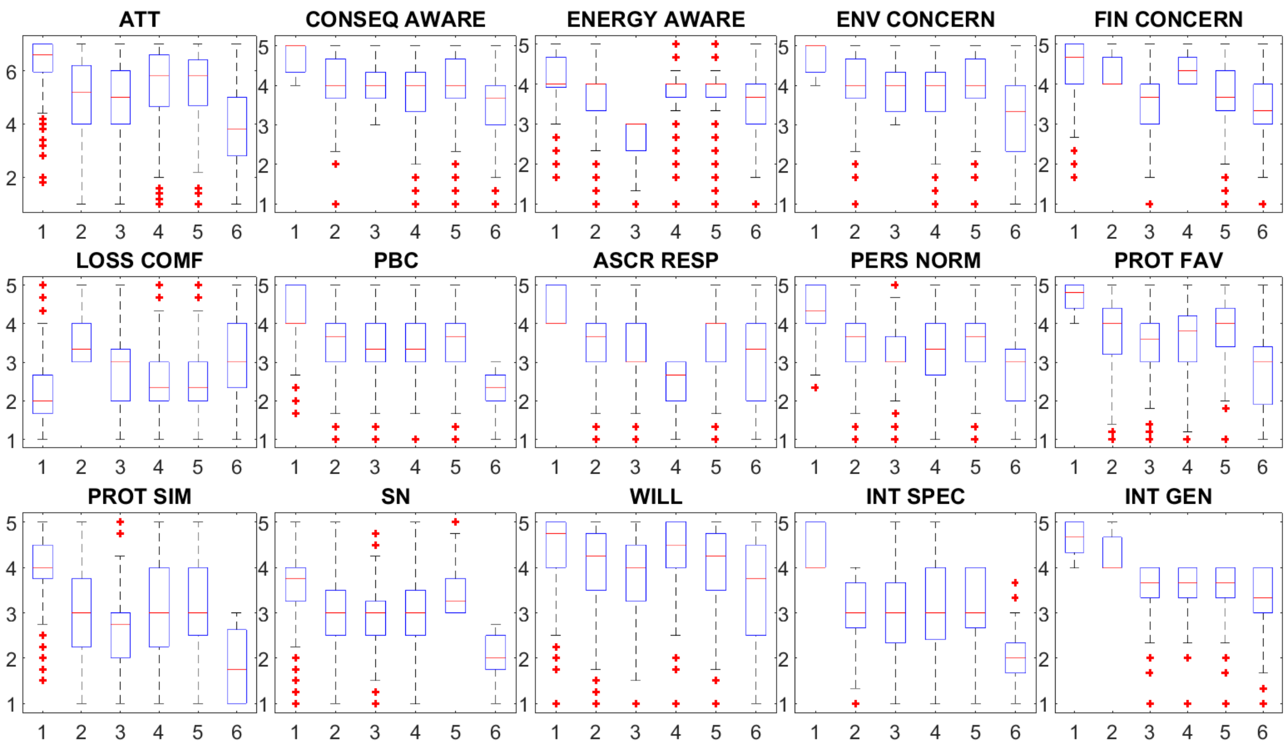


Figure 32: Box plots of class scores (y-axis) in the main 15 variables measured in the survey. The class score in a variable equals the average score of users assigned to that class (x-axis) in the particular variable, measured on the 1-5 Likert scale (exception: ATT, which is measured on the 1-7 Likert scale)

5.3.3. Step 3: Assignment of *remaining* survey respondents to energy-consumer classes

By the end of the second step of the process, 2132 users have been identified as members of a unique class. What about the remaining 3129-2132 = 997 users?

The way we handle them is as follows:

- First, for each class we compute the average scores over all users assigned to the class, for each of the 15 variables of Figure 32. This could be viewed as the score of the class centroid.
- Then, for each of those 997 users, we compute its normalized distances from all five class centroids. These normalized distances take into account only the variables that are involved in the specification of each class. They are computed as the 'cityblock' distances (also: Manhattan or taxicab distances, see section 5.2.2.4) between the centroids and the sample at hand. The normalization consists in dividing the distance metric over the number of variables involved in the class specification and used in the computation of the distance.
- Finally, we assign the user to the nearest class, i.e., the class, whose centroid lies closest to the user.

After this step is executed, the 6 classes include 917 (29.31%), 733 (23.43%), 497 (15.88%), 311 (9.94%), 499 (15.95%) and 172 (5.49%) users, respectively. Table 11 summarizes how the sizes of the six classes evolve through the three steps of the overall process.

Table 11: Evolution of sizes of six classes during steps 2 and 3 of the user classification process

Class size evolution through the process	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Class 5**	**Class 6**	**Sum**
After identifying which users are eligible for which class	529	477	507	425	1041	465	3444
After all 2132 users are assigned to a class	529	400	440	259	392	112	2132
After the remaining 1071 users are assigned to the "closest" class	917	733	497	311	499	172	3129

5.3.4. Characterization of energy consumer segments

5.3.4.1 Beliefs, attitudes, norms

Figure 33 plots how the six energy consumer classes of users score in each of the 15 features corresponding to psycho-social constructs. When we compare Fig. 33 (after the assignment of the 997 users) to Fig. 32 (at the end of step 2), we note a few more outlier observations, as expected, but, otherwise, the class median scores and their relative rank remain practically intact.

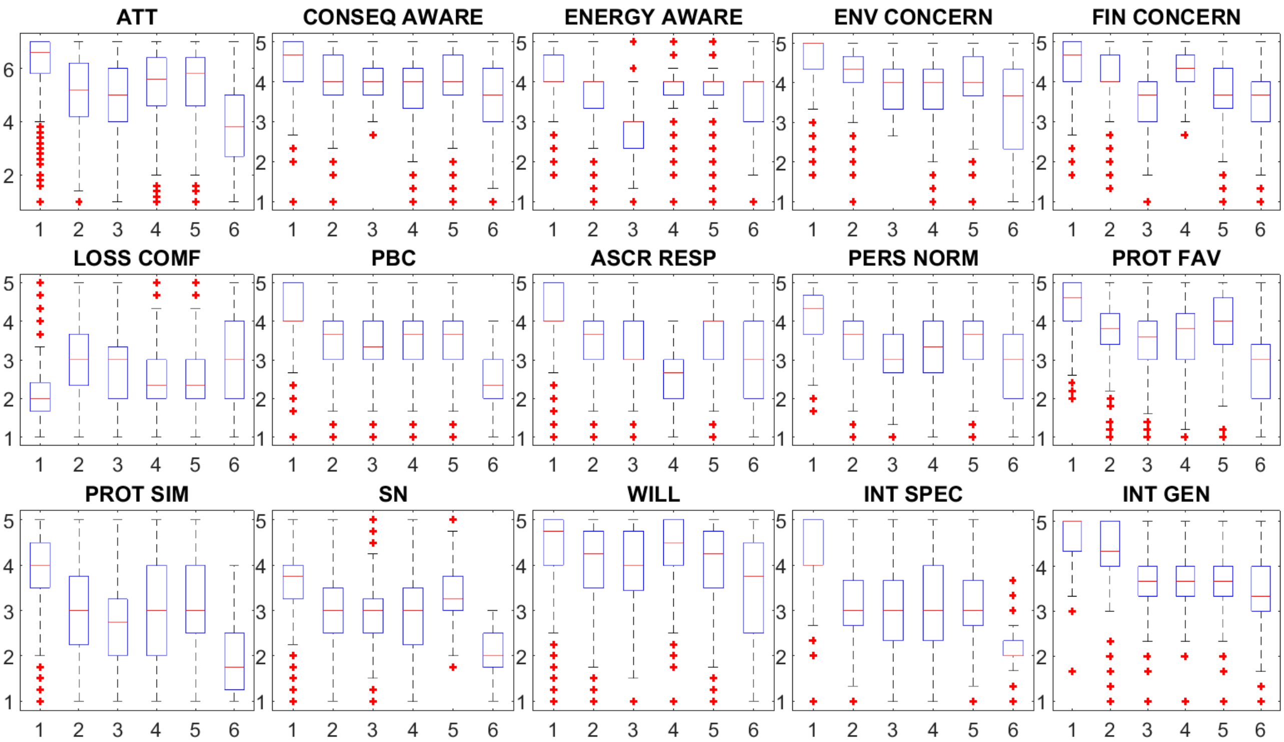


Figure 33: Box plots of class scores in the main 15 variables measured in the survey. The class score in a variable equals the average score of users assigned to the class in the particular variable.

Environmentally conscious and well-informed energy consumers indeed represent a benchmark in terms of energy-saving behaviour. They score high not only in the class feature set but also in other constructs that have been identified as important motivators for energy-saving behaviour. Besides being concerned about the environmental matters, and aware about the consequences of irresponsible energy-related behaviour, they are well informed about ways to save energy. They bear a strong sense of personal responsibility for acting in energy-friendly manner and the possibility of sacrificing some of their comfort to do so does not stand as a barrier to act this way.

Concerned but comfort-oriented energy consumers form a very distinct segment of energy consumers. Their overall intention to adopt an energy-saving behaviour is high (in fact: the 2nd highest after the 1st benchmark class) and this is supported by high concern about the environment and good understanding of the risks involved in energy-wasting ways. Nevertheless, the intentions of these consumers are clearly weaker when the question is about energy-saving with respect to heating and cooling. Namely, the possible sacrifice of comfort that might result from tolerating a slightly lower temperature as thermostat's setting in winter or a higher one during summer appears to be much less tolerable for this segment of energy consumers than any other.

Concerned but lacking awareness energy consumers are one of the four energy-consumer segments (ref. Figure 33, INT_SPEC and INT_GEN box-plots), whose stated intentions to save energy can be strengthened, both specifically with respect to heating and, more generally, with respect to other energy-consuming activities (use of appliances, kitchen, lighting). The lack of knowledge about practical ways to save energy serves as barrier for an unconditionally positive attitude towards energy-saving. The latter exists, even in less profound way than in the first benchmark segment.

Materialistic energy consumers escaping personal responsibility is the second energy consumer segment that lags in overall energy-saving intentions. Neither concern about the environment, nor knowledge about ways to save energy are missing in their case. Yet, whereas they claim awareness of the consequences that increasing energy demand bears for the environment and the society, they do not accept their own share of responsibility to act on this. On the other hand, and this gives some hope for their treatment, they demonstrate high concern for the height of their energy bills and the monetary fingerprint of energy-saving activities.

Prone to social influence energy consumers attribute high value to the fact that people they deem important in their lives approve and support energy-saving, which sets a strongly favourable behavioural prototype. Hence, this form of indirect social pressure serves as facilitator of energy-saving in their case.

Indifferent energy consumers are users demonstrating profoundly low intentions for energy-saving. They doubt their own capacity to adopt energy-saving behaviour as well as the impact this can have on the energy-saving challenge, shaping their overall attitude towards energy-saving on the negative. They are nowhere close to the energy-saver prototype (which they do not find favourable anyway) and they do not perceive social pressure to adopt energy-saving behaviour.

5.3.4.2 Socio-demographic characteristics

5.3.4.2.1 Gender

Overall, the classes do not exhibit significant differentiation with respect to gender distribution. For three classes, the Environmentally conscious and well-informed energy consumers, the Concerned but comfort-oriented energy consumers, and the Indifferentenergy consumers, the portion of females and males is approximately identical with the one in the overall dataset (48.86% and 51.14%, respectively). For the other three classes, males are marginally overrepresented, as shown in Figure 34. The two-sample t-test for the proportion of males in each of the three clusters and the overall dataset are marginally rejected (57.42% males, $p = 0.024$ for the Concerned but lacking awareness energy consumers, 56.14% males with $p=0.035$ for the Materialistic energy consumers escaping their personal responsibility and 53.85% males with $p=0.035$ for the Prone to social influence energy consumers).

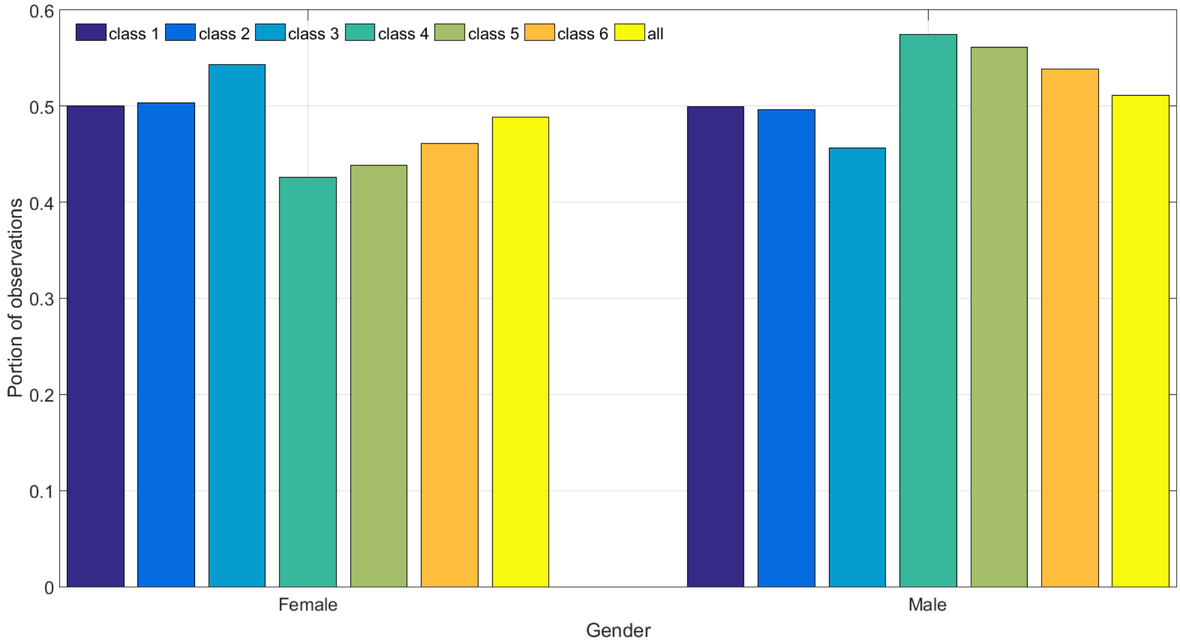


Figure 34: Gender representation in the 6 energy consumer classes and the overall dataset.

5.3.4.2.3 Age

The within-class age distributions deviate from the one in the total dataset. At 5% significance level the two-sample Kolmogorov-Smirnov test rejects the hypothesis that the within-class age distribution and the overall (average age M = 50.41) are identical (samples of the same underlying distribution) for all energy consumer segments except for the *Concerned but comfort-oriented energy consumers (*M =49.02 , *p* = 0.09*)*: strongly for the *Environmentally conscious and well-informed energy consumers* (M=48.74, *p*=0.004), the *Materialistic energy consumers escaping personal responsibility* (M=*55.52,* p=0.000), and the *Prone to social influence energy consumers (*M=53.49, *p*=0.004) and marginally for the *Concerned but lacking awareness energy consumers (*M=48.36, *p*=0.027*)* and the *Indifferent energy consumers (*M=53.49, *p*=0.04*).*

Figure 35 shows that *Materialistic energy consumers escaping personal responsibility* tend to be older, with more than half of them exceeding the age of 57 years, whereas the *Prone to social influence energy consumers* exhibit similar mass concentration in the interval 44-69 years old, prevailing in the ages 44-56. The lack of self-confidence is also prevalent in middle and high ages, 7 out of 10 users in this class being older than 44 years. On the other hand, almost half the *Concerned but lacking awareness energy consumers* are found in the two youngest groups [18-43], implying that educating the younger generations about energy-saving should remain high in the list of possible interventions.

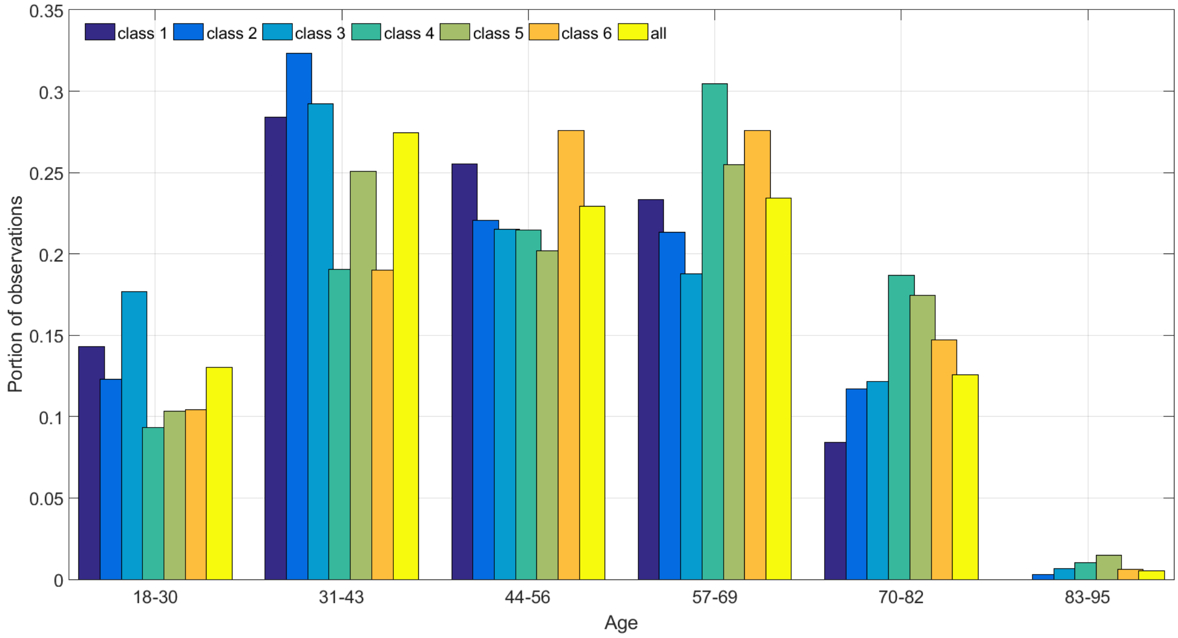


Figure 35: Age distribution within the six energy consumer classes and the overall dataset (minimum age = 18 years and maximum age =95 years in the dataset).

5.3.4.2.3 Education degree

As explained in section 2.2.5, the education level of respondents is measured as an ordinal number on a scale of 0 to 6: None, Primary education, Lower secondary education, Upper secondary education, Bachelor's or equivalent level, Master's or equivalent level and Doctoral or equivalent level.

Applying the two-sample Kolmogorov-Smirnov test at 5% significance level to assess the hypothesis that the cluster-level distributions of education's degree is identical to the one of the overall sample (average M = 3.972):

- the hypothesis is rejected for the *Environmentally conscious and well-informed energy consumers* (M=4.21, p<0.0001), the *Materialistic energy consumers escaping personal responsibility* (M=3.54, p<0.0001) and the *Indifferent energy consumers* (p=0.02). Note that the average educational status of the 1st well-behaving cluster is noticeably higher than the average status in the overall sample, whereas it is the other way round (noticeably lower educational status) with the *Materialistic energy consumers escaping personal responsibility* and the *Indifferent energy consumers* (p=0.02).
- the hypothesis cannot be rejected for the other three classes, i.e., the *Concerned but comfort-oriented energy consumers* (M =4.03, p = 0.22), the *Concerned but lacking awareness energy consumers* (M=3.82, p=0.058), the *Prone to social influence energy consumers* (M=3.93, p=0.99).

Looking at Figure 36, it is noteworthy making the following remarks:

- 3 out of 4 *Environmentally conscious and well-informed energy consumers* have at least a Bachelor's level degree. The respective proportions are 2 out of 3 for the *Concerned but comfort-oriented energy consumers* and 3 out of 5 for the *Prone to social influence energy consumers.*
- On the other extreme, more than half of the *Materialistic energy consumers escaping personal responsibility* and 1 out of 2 *Indifferent energy consumers* have not obtained a degree from a higher education institution. This obviously relates to the fact that these two energy consumer groups involve the oldest (on-average) consumers.

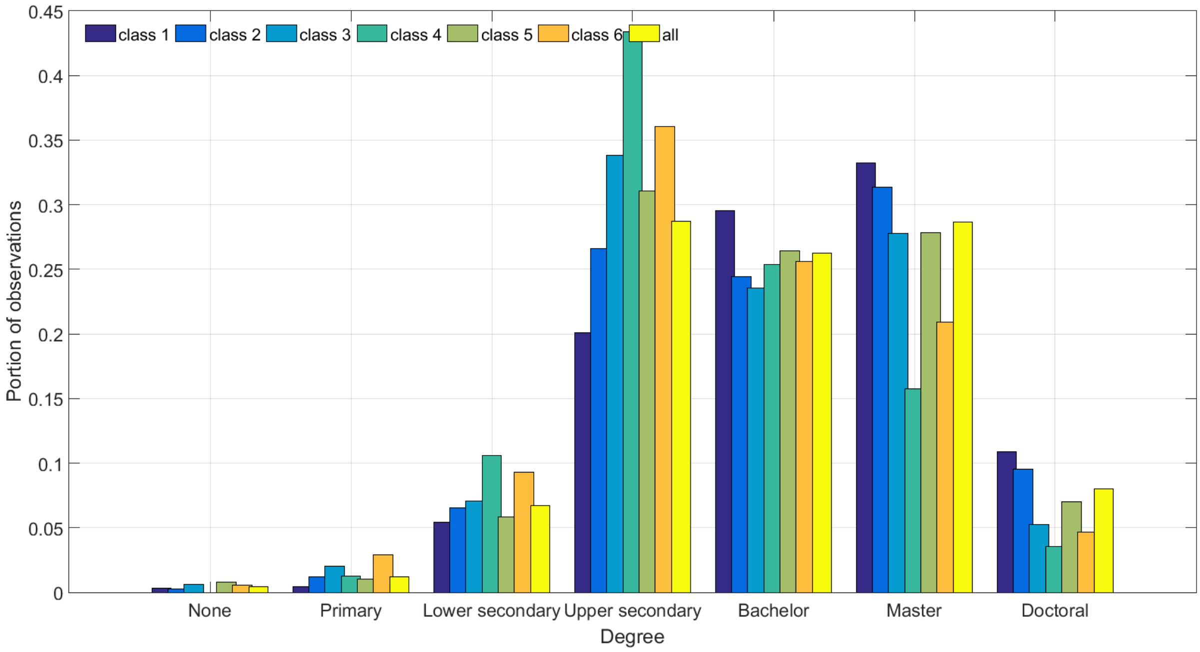


Figure 36: Education degree representation in the 6 energy consumer classes and the overall dataset.

5.3.5. On identifying interventions for the six energy consumer segments

In what follows, we summarize which interventions of the nudging type are deemed appropriate for each of the six energy consumer classes. In doing that, we expand on the discussion in section 5.3.1, where the energy-consumer classes were specified in view of these interventions, leveraging additional information about these classes from the characterization in section 5.3.4 and some more insights drawn in this section. The main reference for this task is the categorization of nudges and their types in NUDGE Deliverable 2.1 titled "*Design document of nudging interventions per pilot*", where the main characteristics as well reservations about these interventions are elaborated.

Environmentally conscious and well-informed energy consumers

This class of consumers is the one in minimum need of an intervention treatment. The main requirement is to keep them sensitized about energy-saving and this can happen through simple reinforcement nudges such as regularly providing them with information about energy-saving and the positive consequences of their behavior.

Key points	(Nudge) intervention type	Description
High scores in all features	Reinforcement	*Feedback & awareness:* keep the interest warm through regular but sparse information about energy-saving (selected notifications, regular marketing campaigns)

Concerned but comfort-oriented energy consumers

For this energy consumer class, the interventions can target their concern about the financial implications of energy consumption. Otherwise, it is not straightforward to cope with their concerns about losses in terms of personal comfort. Note that the overall energy-saving intentions of this group are strong but the required compromises in terms of comfort serve as barrier for heating/cooling-related energy-saving, in particular.

Key points	(Nudge) intervention type	Description
Strong concern about comfort & financial implication of energy-saving	Confronting	*Reminding of consequences:* prompt the user to consider the consequences of an action e.g., increasing the target temperature of the thermostat or the air-conditioning, insisting on the extra cost it incurs. It could be the net increase of the energy bill, projecting the impact of the action at monthly/annual level.

Concerned but lacking awareness energy consumers

This group of users can be nudged in two ways: either by securing default operational conditions that favor energy-saving, essentially bypassing the missing know-how barrier, or by trying to (gradually) render it obsolete by gradually educating and training people in the optimal energy-saving behaviors.

Key points	(Nudge) intervention type	Description
Concern about the environment, awareness of consequences but lack of know-how to practically save energy	Facilitating	*Default:* Turn energy-friendly operational settings of devices (thermostat, air conditioning equipment) into defaults, to save the user from the "burden" of learning what is appropriate and what is not.

	Reinforcement	*Just-in-time prompts and tips:* Provide the user with tips and recommendations exactly upon the time she mingles with devices' settings that have an impact on energy consumption.

Materialistic energy consumers escaping personal responsibility

This is another "difficult" group in the sense that the key barrier to its energy-saving behavior cannot be treated, at least in obvious and generic manner, by interventions of the nudging type. Nevertheless, nudging can exploit their sensitivity to the financial aspects of energy-related behavior and focus on the possible direct monetary savings that are feasible with energy-saving.

Key points	(Nudge) intervention type	Description
Concern about the environment, awareness of consequences but lack of know-how to practically save energy	Confronting	*Reminding of consequences:* prompt the user to consider the consequences of e.g., increasing the target temperature of the thermostat or the air-conditioning, insisting on the extra cost it incurs. It could be the net increase of the bill, projecting the impact of the action at monthly/annual level.

Prone to social influence energy consumers

The main goal with this group of consumes is to exploit their vulnerability to norms and social pressure. The idea of exposing users to social comparison is one of the most explored ones in experiments around nudging, including those related to energy-friendly behavior. Goal setting programs, on the other hand, are viewed as smart ways to elicit consumers' commitment to save upon what they consume.

Key points	(Nudge) intervention type	Description
Strong sense of subjective norms, average scores-no distinct differentiation in other features	Social influence	*Enabling social comparison:* leverage different means (from written text and diagrams printed on a paper to online social platforms and dynamic query response systems) to facilitate the comparison with other peers (friends, neighbors, consumers of similar demographic characteristics).
	Social influence	*Goal setting & commitment:* get the consumers to sign a formal commitment to reduce the energy they consume, many times in return of some (non-monetary) reward.

An additional check to carry out when planning such interventions is the extent to which this group of users is willing to share data about their consumption and make it visible to other peers (e.g., friends, neighbors, or even more openly), also implying their processing from commercial entities directly related to the energy provision (providers, distributors) or third parties.

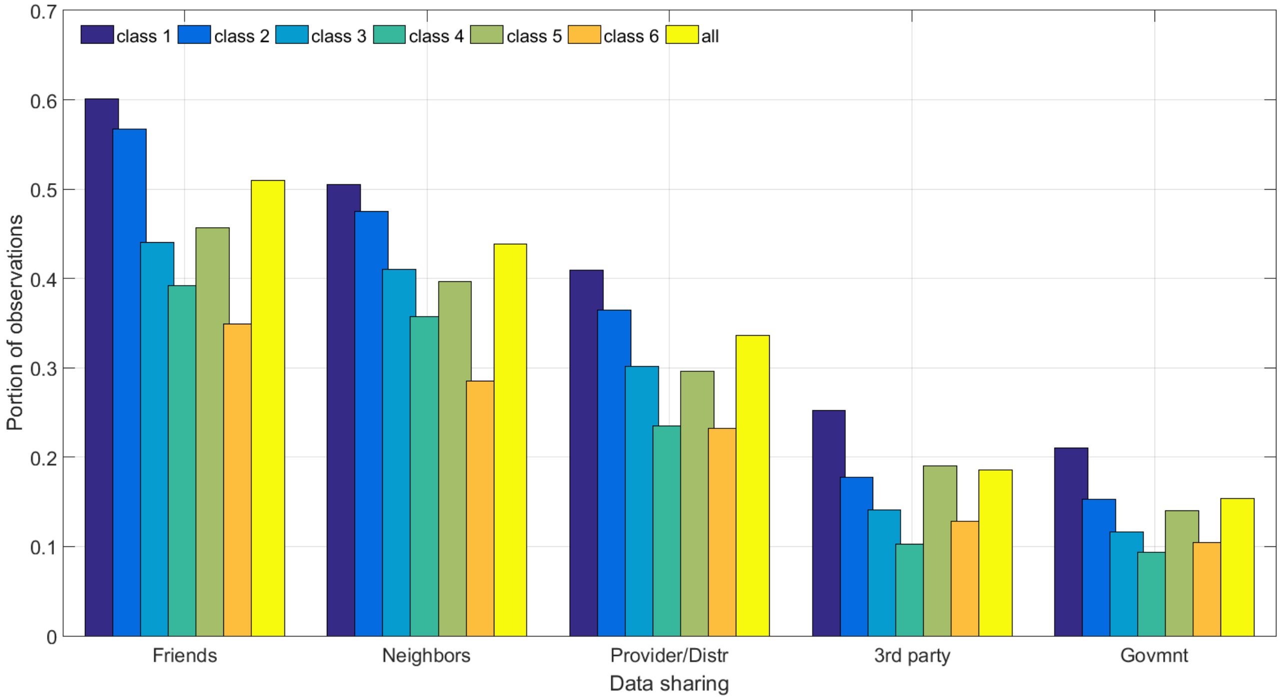


Figure 37: Willingness within energy consumer classes and overall to share data about their monthly energy usage

Figure 37 suggests that additional barriers exist, this time to operationalize the aforementioned interventions, e.g., additional effort has to be invested for collecting data from this group of users. Fewer than half of them (44.06%) are not willing to share monthly consumption data even with friends of theirs, let alone with neighbors (41.05%) or the energy provider/distributor (30.18%). Notably, even across the *Environmentally conscious and well-informed energy consumers*, which emerges as the group with the comparatively stronger disposition to share monthly consumption data, a good 40% resists sharing data even with their friends. Equally interesting is the fact that people trust their governments less than the private sector (providers/distributors, third-party entities) on this matter.

Indifferent energy consumers

This is another difficult-to-treat group. The original idea while specifying this class was to isolate users who perceive their self-efficacy to be low. The characterization of the group showed that these users also share, on average, the lowest levels of environmental concern and energy awareness and the lowest pressure from norms of any kind. Hence, much work may be needed with this group of energy consumers on multiple fronts.

Key points	(Nudge) intervention type	Description
Low perception of self-efficacy and possible impact of personal action, low concern and awareness about environmental matters.	Facilitating	*Default:* Turn energy-friendly operational settings of devices (thermostat, air conditioning equipment) into defaults, to save the user from the "burden" of learning what is appropriate and what is not.
	Reinforcement	*Feedback & awareness:* use tips, notifications, marketing campaigns, to sensivitize this group of users and overcome their reservations about the efficacy of their behavior.
	Reinforcement	*Hedonic goal* : stress the big picture and the impact on big things, possibly with some exaggeration, to render energy-saving a goal.

5.4. Conclusions

We have carried out a segmentation study of energy consumers relying on self-reports of their behaviors, beliefs, norms, and attitudes. The ultimate goal of this segmentation was to identify consumer segments with loose intentions to save energy and facilitate the selection of proper interventions that could amend this. As mentioned in section 1, the survey measured, for the first time, a large number of psychosocial constructs (15) that draw their origins into three distinct theoretical models of human behavior. One expectation about this "opening" of the measured construct space, which provides richer information about energy consumers at the psychosocial level, is that it could better serve the goals of the segmentation study.

The findings of our study are two-fold in this respect. On the one hand, we could specify six energy consumer classes (i.e., profiles of energy consumers) that combine in distinct manner the aforementioned constructs. These six classes range all the way from Environmentally conscious and well informed energy consumers down to Indifferent energy consumers and, with the exception of the former, they embody characteristics that serve as barriers towards energy-saving but also others that motivate certain interventions for them. On the other hand, standard clustering techniques used in literature for automated segmentation of observations (i.e., energy consumers) to groups with similar characteristics were not of much help in our case. Despite our experimentation with a range of clustering algorithms and trying many different parameterizations, including feature selection and transformation techniques, the obtained clustering structures routinely shared the same pattern: there would always be one cluster that scored top in all 15 constructs (features), one that would score second best in all features, one that would score third best in all features and so on. Namely, no groups with profound differentiation across the 15 features could be identified this way. This is why in section 5.3 we followed an alternative yet plausible approach, which has a starting point the interventions at hand and *a priori* *manually* specified energy consumer groups that make the intervention selection intuitive. This is how the six energy consumer classes emerged.

An exhaustive experimentation with more clustering techniques (e.g., density-based or spectral clustering techniques) would be needed before concluding whether clustering techniques as a whole could be or could not be of help in our case. The reply to this question has independent theoretical interest and it probably touches on the theoretical foundations of clustering that determine how it works and what kind of clustering structures it generates depending on the data it is applied to.

The results of our study will be used in later stages of the project in different ways. For instance, they will inform the preparation of the co-creation workshops that will take place with pilot participants. They give some concrete hints about possible energy consumer behaviors that will be met in the pilot trials but also interventions that are a priori relevant for them. Then, the subject of these co-creation sessions will be the operationalization of these interventions in ways that better fit the user needs and particular characteristics. Likewise, the study findings about the psychosocial variables that differentiate distinct consumer segments are of direct interest to stakeholders such as consumer associations and energy providers, both within and outside the project, in addressing their members/customers. Finally, the study adds to the volume of energy consumer segmentation studies in the scientific literature (see Annex III), giving new insights to the problem, including methodological ones.

Nudging consumers towards energy efficiency through behavioural science

6. Bibliography

NUDGE Deliverable D2.1,Design document of nudging interventions per pilot, April 2021

Abrahamse, W. (2007). *Energy conservation through behavioral change : Examining the eff ectiveness of a tailor-made approach*.

Abrahamse, W., & Steg, L. (2011). Factors Related to Household Energy Use and Intention to Reduce It: The Role of Psychological and Socio-Demographic Variables. *Human Ecology Review*, *18*(1), 30–40.

Abrahamse, W., Steg, L., Gifford, R., & Vlek, C. (2009). Factors influencing car use for commuting and the intention to reduce it: A question of self-interest or morality? *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(4), 317–324. https://doi.org/10.1016/j.trf.2009.04.004

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2007). The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. *Journal of Environmental Psychology*, *27*(4), 265–276. https://doi.org/10.1016/j.jenvp.2007.08.002

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211.

Ajzen, I. (2002). Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior. *Journal of Applied Social Psychology*, *32*(4), 665–683. https://doi.org/10.1111/j.1559-1816.2002.tb00236.x

Caraban, A., Karapanos, E., Gonçalves, D., & Campos, P. (2019). 23 Ways to Nudge: A review of technology-mediated nudging in human-computer interaction. *Conference on Human Factors in Computing Systems - Proceedings*, (*May*). https://doi.org/10.1145/3290605.3300733

Chen, C. fei, Xu, X., & Day, J. K. (2017). Thermal comfort or money saving? Exploring intentions to conserve energy among low-income households in the United States. *Energy Research and Social Science*, *26*, 61–71. https://doi.org/10.1016/j.erss.2017.01.009

Dianshu, F., Sovacool, B. K., & Vu, K. (2010). The barriers to energy efficiency in China: Assessing household electricity savings and consumer behavior in Liaoning Province. *Energy Policy*, *38*(2), 1202–1209. https://doi.org/10.1016/j.enpol.2009.11.012

Elliott, M. A., McCartan, R., Brewster, S. E., Coyle, D., Emerson, L., & Gibson, K. (2017). An application of the prototype willingness model to drivers' speeding behaviour. *European Journal of Social Psychology*, *47*(6), 735–747. https://doi.org/10.1002/ejsp.2268

Engelken, M., Römer, B., Drescher, M., & Welpe, I. (2018). Why homeowners strive for energy self-supply and how policy makers can influence them. *Energy Policy*, *117*, 423–433. https://doi.org/10.1016/j.enpol.2018.02.026

Fornara, F., Pattitoni, P., Mura, M., & Strazzera, E. (2016). Predicting intention to improve household energy efficiency: The role of value-belief-norm theory, normative and informational influence, and specific attitude. *Journal of Environmental Psychology*, *45*, 1–10. https://doi.org/10.1016/j.jenvp.2015.11.001

Frater, J., Kuijer, R., & Kingham, S. (2017). Why adolescents don't bicycle to school: Does the prototype/willingness model augment the theory of planned behaviour to explain intentions? *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 250–259. https://doi.org/10.1016/j.trf.2017.03.005

Gadenne, D., Sharma, B., Kerr, D., & Smith, T. (2011). The influence of consumers' environmental beliefs and attitudes on energy saving behaviours. *Energy Policy*, *39*(12), 7684–7694. https://doi.org/10.1016/j.enpol.2011.09.002

Gao, L., Wang, S., Li, J., & Li, H. (2017). Application of the extended theory of planned behavior to understand individual's energy saving behavior in workplaces. *Resources, Conservation and Recycling*, *127*(August), 107–113. https://doi.org/10.1016/j.resconrec.2017.08.030

Gerrard, M., Gibbons, F. X., Houlihan, A. E., Stock, M. L., & Pomery, E. A. (2008). A dual-process approach to health risk decision making: The prototype willingness model. *Developmental Review*, *28*(1), 29–61. https://doi.org/10.1016/j.dr.2007.10.001

Gibbons, F. X., Stock, M. L., & Gerrard, M. (2020). The Prototype-Willingness Model. In *The Wiley Encyclopedia of Health Psychology* (pp. 517–527). Wiley. https://doi.org/10.1002/9781119057840.ch102

Guo, Z., Zhou, K., Zhang, C., Lu, X., Chen, W., & Yang, S. (2018). Residential electricity consumption behavior: Influencing factors, related theories and intervention strategies. *Renewable and Sustainable Energy Reviews*, *81*(January 2017), 399–412. https://doi.org/10.1016/j.rser.2017.07.046

Han, M. S., & Cudjoe, D. (2020). Determinants of energy-saving behavior of urban residents: Evidence from Myanmar. *Energy Policy*, *140*(October 2019), 111405. https://doi.org/10.1016/j.enpol.2020.111405

Hopkins, B. and Skellam, J.G. (1954) A New Method for Determining the Type of Distribution of Plant Individuals. Annals of Botany, 18, 213-227. https://doi.org/10.1093/oxfordjournals.aob.a083391

Karlin, B., Davis, N., Sanguinetti, A., Gamble, K., Kirkby, D., & Stokols, D. (2014). Dimensions of Conservation: Exploring Differences Among Energy Behaviors. *Environment and Behavior*, *46*(4), 423–452. https://doi.org/10.1177/0013916512467532

Kilbourne, W., & Pickett, G. (2008). How materialism affects environmental beliefs, concern, and environmentally responsible behavior. *Journal of Business Research*, *61*(9), 885–893. https://doi.org/10.1016/j.jbusres.2007.09.016

Lind, H. B., Nordfjærn, T., Jørgensen, S. H., & Rundmo, T. (2015). The value-belief-norm theory, personal norms and sustainable travel mode choice in urban areas. *Journal of Environmental Psychology*, *44*, 119–125. https://doi.org/10.1016/j.jenvp.2015.06.001

Long, J. E. (1993). An econometric analysis of residential expenditures on energy conservation and renewable energy sources. *Energy Economics*, *15*(4), 232–238. https://doi.org/10.1016/0140-9883(93)90012-G

Olsen, S. O., Heide, M., Dopico, D. C., & Toften, K. (2008). Explaining intention to consume a new fish product: A cross-generational and cross-cultural comparison. *Food Quality and Preference*, *19*(7), 618–627. https://doi.org/10.1016/j.foodqual.2008.04.007

Perugini, M., & Bagozzi, R. P. (2001). The role of desires and anticipated emotions in goal-directed behaviours: Broadening and deepening the theory of planned behaviour. *British Journal of Social Psychology*, *40*(1), 79–98. https://doi.org/10.1348/014466601164704

Proudlove, R., Finch, S., & Thomas, S. (2020). Factors influencing intention to invest in a community owned renewable energy initiative in Queensland, Australia. *Energy Policy*, *140*(March), 111441. https://doi.org/10.1016/j.enpol.2020.111441

Ratliff, K. A., Howell, J. L., & Redford, L. (2017). Attitudes toward the prototypical environmentalist predict environmentally friendly behavior. *Journal of Environmental Psychology*, *51*, 132–140. https://doi.org/10.1016/j.jenvp.2017.03.009

Rivis, A., Sheeran, P., & Armitage, C. J. (2006). Augmenting the theory of planned behaviour with the prototype/willingness model: Predictive validity of actor versus abstainer prototypes for adolescents' health-protective and health-risk intentions. *British Journal of Health Psychology*, *11*(3), 483–500.

https://doi.org/10.1348/135910705X70327

Song, Y., Zhao, C., & Zhang, M. (2019). Does haze pollution promote the consumption of energy-saving appliances in China? An empirical study based on norm activation model. *Resources, Conservation and Recycling*, *145*(February), 220–229. https://doi.org/10.1016/j.resconrec.2019.02.041

Staats, H., Harland, P., & Wilke, H. A. M. (2004). *EFFECTING DURABLE CHANGE A Team Approach to Improve Environmental Behavior in the Household*, *36*(3), 341–367. https://doi.org/10.1177/0013916503260163

Steg, L., Dreijerink, L., & Abrahamse, W. (2005). Factors influencing the acceptability of energy policies: A test of VBN theory. *Journal of Environmental Psychology*, *25*(4), 415–425. https://doi.org/10.1016/j.jenvp.2005.08.003

Stern, P. (2000). Toward a Coherent Theory of Environmentally Significant Behavior. *Journal of Social Issues*, *56*(3), 407–424.

Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., & Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. *Human Ecology Review*, *6*(2), 81–97.

Sütterlin, B., Brunner, T. A., & Siegrist, M. (2011). Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics. *Energy Policy*, *39*(12), 8137–8152. https://doi.org/10.1016/j.enpol.2011.10.008

Taufique, K. M. R., & Vaithianathan, S. (2018). A fresh look at understanding Green consumer behavior among young urban Indian consumers through the lens of Theory of Planned Behavior. *Journal of Cleaner Production*, *183*, 46–55. https://doi.org/10.1016/j.jclepro.2018.02.097

Thøgersen, J., & Ölander, F. (2002). Human values and the emergence of a sustainable consumption pattern: A panel study. *Journal of Economic Psychology*, *23*(5), 605–630. https://doi.org/10.1016/S0167-4870(02)00120-4

Tonglet, M., Phillips, P. S., & Read, A. D. (2004). Using the Theory of Planned Behaviour to investigate the determinants of recycling behaviour: a case study from Brixworth, UK. *Resources, Conservation and Recycling*, *41*(3), 191–214. https://doi.org/10.1016/j.resconrec.2003.11.001

Van Gool, E., Van Ouytsel, J., Ponnet, K., & Walrave, M. (2015). To share or not to share? Adolescents' self-disclosure about peer relationships on Facebook: An application of the Prototype Willingness Model. *Computers in Human Behavior*, *44*, 230–239. https://doi.org/10.1016/j.chb.2014.11.036

Van Raaij, F. W., & Verhallen, T. M. M. (1983). a Behavioral Model of Residential Energy Use @Bullet. *Journal of Economic Psychology*, *3*, 39–63.

Wang, Z., Zhang, B., & Li, G. (2014). Determinants of energy-saving behavioral intention among residents in Beijing: Extending the theory of planned behavior. *Journal of Renewable and Sustainable Energy*, *6*(5), 1–18. https://doi.org/10.1063/1.4898363

Wang, Z., Zhang, B., Yin, J., & Zhang, Y. (2011). Determinants and policy implications for household electricity - saving behaviour : Evidence from Beijing China. *Energy Policy*, *39*(6), 3550–3557.

Webb, D., Soutar, G. N., Mazzarol, T., & Saldaris, P. (2013). Self-determination theory and consumer behavioural change: Evidence fromahousehold energy-saving behaviour study. *Journal of Environmental Psychology*, *35*, 59–66. https://doi.org/10.1016/j.jenvp.2013.04.003

Winett, R. A., Leckliter, I. N., Chinn, D. E., Stahl, B., & Love, S. Q. (1985). Effects of television modeling on residential energy conservation. *Journal of Applied Behavior Analysis*, *18*(1), 33–44.

Wittenberg, I., Blöbaum, A., & Matthies, E. (2018). Environmental motivations for energy use in PV households: Proposal of a modified norm activation model for the specific context of PV households. *Journal of Environmental Psychology*, *55*, 110–120. https://doi.org/10.1016/j.jenvp.2018.01.002

Yoo, S., Eom, J., & Han, I. (2020). Factors driving consumer involvement in energy consumption and energy-efficient purchasing behavior: Evidence from Korean residential buildings. *Sustainability (Switzerland)*, *12*(14), 1–20. https://doi.org/10.3390/su12145573

NUDGE
Nudging consumers towards energy efficiency through behavioural science

Annexes

Annex I. Linear Regression models

Table 12 details specific intent, while Table 13 looks at general intent. In both cases, we first present our socio-demographic variables (model 1a and 2a). This is followed by our three theory of planned behaviour variables (model 1b and 2b). The third model in the series examines our remaining attitudinal variables (model1c and model 2c). Finally, we present the complete analysis, containing all variables (model 1d and 2d).

Table 12: Linear regression with specific intent as outcome

	Dependent variable:	*Specific intent to reduce heating*		
	Model 1a	Model 1b	Model 1c	Model 1d
Age	-0.003^{**}			-0.005^{***}
Region (1)				
Northern Europe	-0.37^{***}			-0.10
Southern Europe	0.32^{***}			0.06
Western Europe	0.02			-0.12^{*}
Degree (2)				
Upper secondary	0.01			-0.06
Bachelor	0.08			-0.07
Master	0.10			-0.09
Doctor	0.12			-0.12^{*}
Attitudinal Factors				
Attitude		0.16^{***}		0.14^{***}
Perceived behavioural control		0.49^{***}		0.47^{***}
Subjective Norms		0.33^{***}		0.28^{***}
Financial Concern			0.16^{***}	0.10^{***}
Loss of comfort			-0.18^{***}	-0.10^{***}
Energy knowledge			-0.01	-0.05^{***}
Environmental concern			0.09^{**}	-0.01
Awareness of consequences			0.07^{*}	0.01
Ascription of responsibility			0.004	-0.02
Personal norms			0.29^{***}	0.11^{***}
Constant	3.35^{***}	-0.33^{***}	1.50^{***}	0.07
Observations	3,131	3,131	3,131	3,131
R^2	0.04	0.57	0.24	0.62
Adjusted R^2	0.04	0.57	0.23	0.62
Residual Std. Error	1.00 (df = 3122)	0.67 (df = 3127)	0.90 (df = 3123)	0.63 (df = 3112)
F Statistic	16.08^{***} (df = 8; 3122)	$1{,}399.52^{***}$ (df = 3; 3127)	137.33^{***} (df = 7; 3123)	281.77^{***} (df = 18; 3112)

Notes:
1: Reference category is Eastern Europe
2: Reference category is Lower education
$^{*}p^{**}p^{***}p<0.001$

Table 13: Linear regression with general intent as outcome

		Dependent variable:	General intent to reduce consumption	
	Model 2a	Model 2b	Model 2c	Model 2d
Age	-0.003***			-0.01***
Region (1)				
Northern Europe	-0.36***			-0.21***
Southern Europe	0.11*			-0.04
Western Europe	-0.25***			-0.22***
Degree (2)				
Upper secondary	0.03			-0.04
Bachelor	0.15**			0.02
Master	0.20***			0.03
Doctor	0.21**			0.01
Attitudinal Factors				
Attitude		0.04***		-0.02*
Perceived behavioural control		0.16***		0.11***
Social Norms		0.13***		0.06***
Financial Concern			0.19***	0.20***
Loss of comfort			-0.08***	-0.07***
Energy knowledge			0.04*	0.05***
Environmental concern			0.19***	0.14***
Awareness of consequences			0.02	0.03
Ascription of responsibility			0.04*	0.01
Personal norms			0.10***	0.06***
Constant	4.23***	2.91***	2.07***	2.25***
Observations	3,131	3,131	3,131	3,131
R^2	0.10	0.12	0.28	0.34
Adjusted R^2	0.09	0.12	0.28	0.34
Residual Std. Error	0.70 (df = 3122)	0.70 (df = 3127)	0.63 (df = 3123)	0.60 (df = 3112)
F Statistic	41.20*** (df = 8; 3122)	136.74*** (df = 3; 3127)	171.89*** (df = 7; 3123)	88.93*** (df = 18; 3112)

Notes:
1: Reference category is Eastern Europe
2: Reference category is Lower education
*p**p***p<0.001

Annex II. Complementary material on clustering results

We report below some more experiments, with different parameterization, attempting to come up with a segmentation of energy consumers using standard clustering techniques.

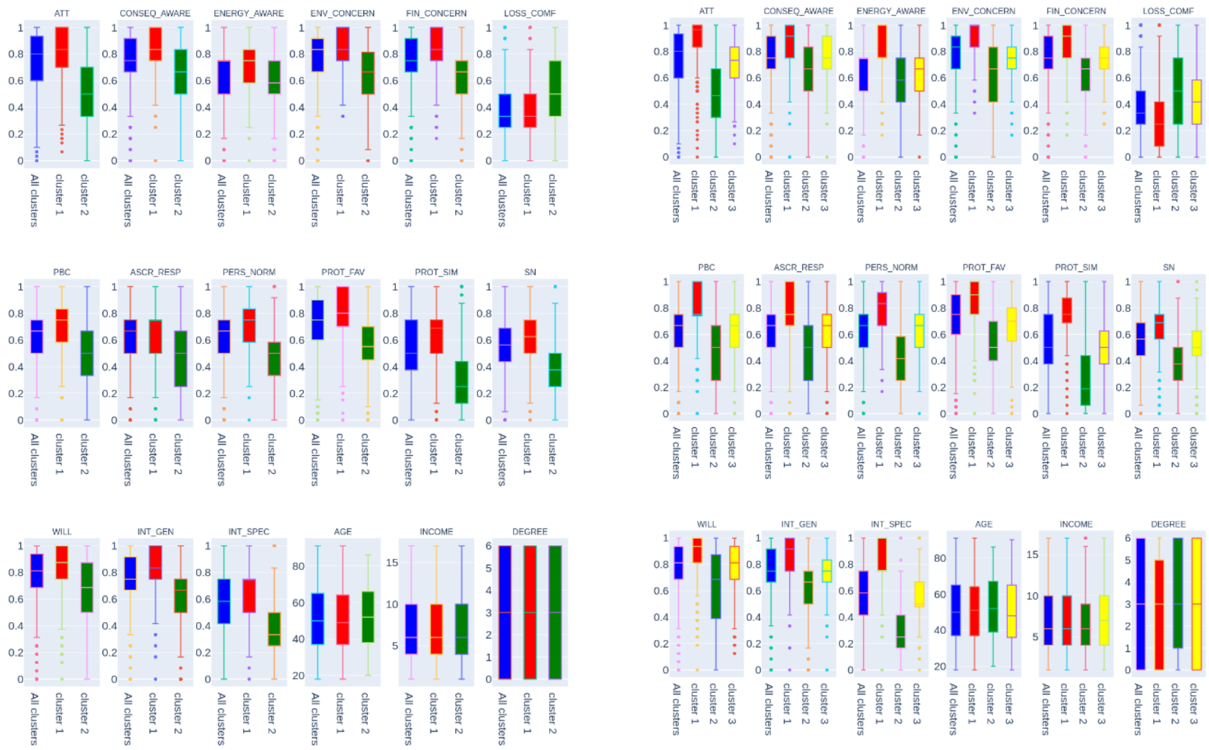
AII.A Clustering with full feature set (15 features – no feature selection)

The parameterization of this experiment is as follows:

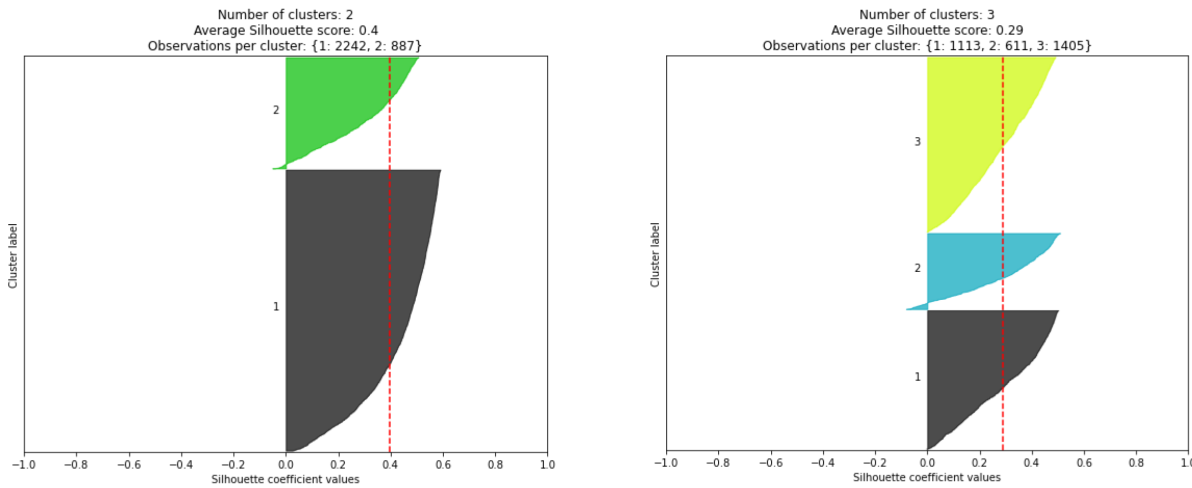
Feature selection: No feature selection is applied. The full set of 15 features (see section 5) is used.

Clustering algorithm: k-means. The Euclidean distance is employed as the distance metric and the number of clusters is given as an argument to the algorithm.

Clustering fitness assessment: Silhouette score.

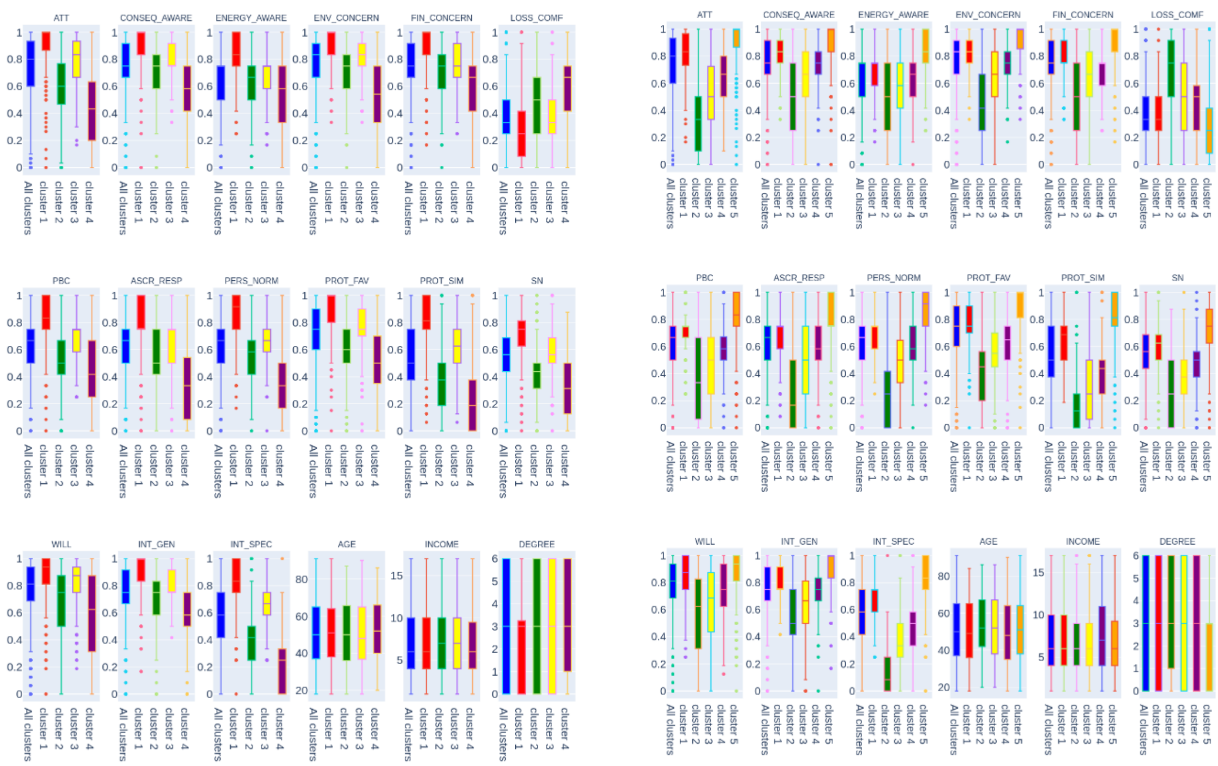


(a) Cluster-average scores in 15 features and 3 socio-demographic variables

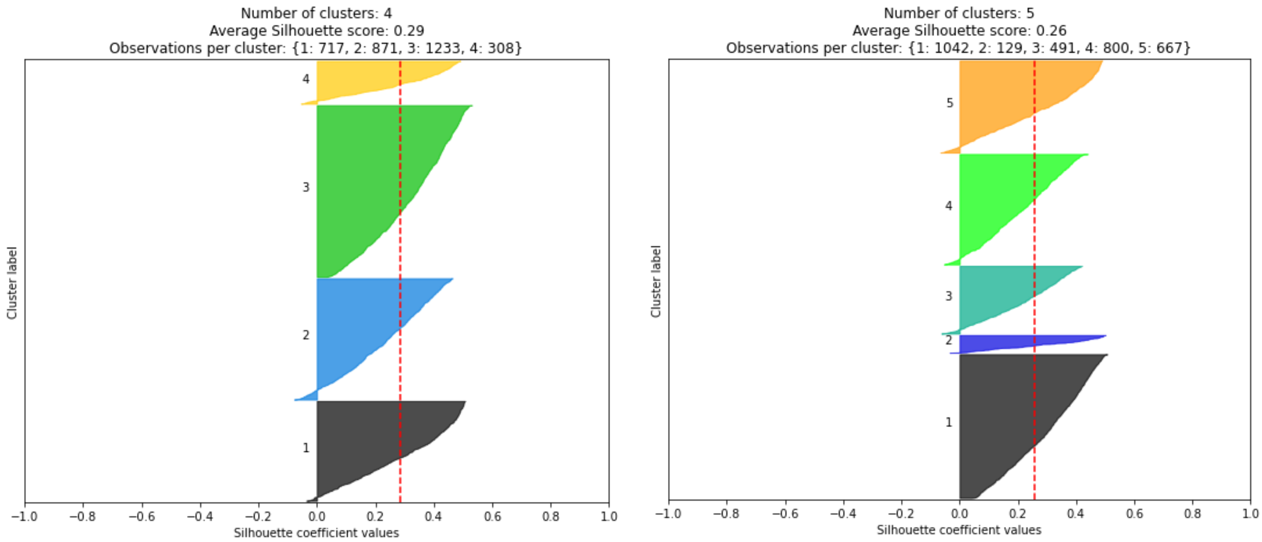


(b) Silhouette values

Figure 38: Box plots of cluster-based scores in the 15 features as well as their age, income and educational degree characteristics: number of clusters equals 2 (left column) and 3 (right column)



(a) Cluster-average scores in 15 features and 3 socio-demographic variables



(b) Silhouette values

Figure 39: Box plots of cluster-based scores in the 15 features as well as their age, income and educational degree characteristics: number of clusters equals 4 (left column) and 5 (right column)

The quality of the resulting clustering structure can be summarized as follows:

Fitness: The Silhouette scores are not improved in any tangible way; they are rather worse than those in section 5.2.3.2.1, obtained with the same clustering algorithm and parameters but with a reduced subset of features, for all possible numbers of clusters.

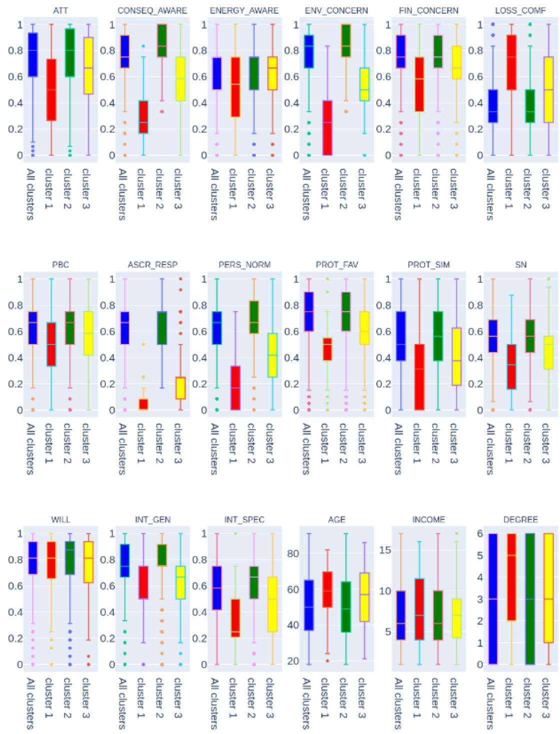
Balance: The sizes of the resulting clusters mostly satisfy the rule of thumb requirement (no cluster should have fewer than 5% of the total observations). Only the experiment with 5 clusters, results in one cluster (cluster 2) that violates this rule.

Structure: The results of this experiment follow the same pattern reported in cluster analysis (section 5.2.3.2.1, Figure 27 and Figure 28). The ranking of clusters across features is consistently observed, as Figure 38 and Figure 39**Error! Reference source not found.** for 2,3,4 and 5 clusters, respectively.

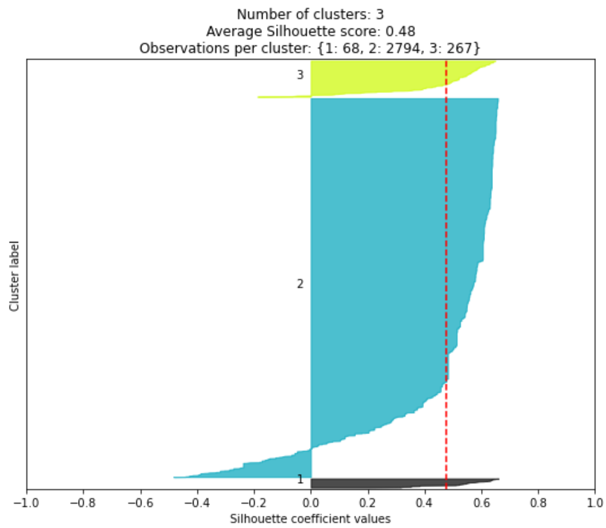
AII.B Agglomerative clustering and standard feature selection

The parameterization of this experiment is as follows:

Feature selection: standard, based on the Hopkins statistic. The best combination of 4 features is selected such that the Hopkins score is maximized, ignoring the interventions. The selected features set is {CONSEQ_AWARE, ENV_CONCERN, ASCR_RESP, INT_GEN}.



(a) Cluster-average scores in 15 features and 3 socio-demographic variables



(b) Silhouette values

Figure 40: Box plots of cluster-based scores in the 15 features as well as their age, income and educational degree characteristics (left) and cluster silhouette values (right): number of clusters = 3

Clustering algorithm: Agglomerative. The Manhattan distance is employed as the distance metric and the number of clusters is given as an argument to the algorithm.

Clustering fitness assessment: Silhouette score.

Figure 40 plots the per cluster scores in the 15 features and 3 socio-demographic variables under a 3-cluster structure. The results with 4-cluster and 5-cluster structures are similar and follow the same pattern as in the other experiments discussed so far.

Fitness: The Silhouette score in this experiment is higher than all the experiments with 3 cluster structures reported so far, scoring 0.48.

Balance: The balance of the clusters is less satisfactory, given that there is one cluster (cluster 2) gathering the vast majority of the observations, while another cluster (cluster 1) has very few observations violating the minimum rule of thumb of at least 5% of observations per cluster.

Structure: The results of this experiment follow the same pattern reported in cluster analysis (section 5.2.3.2.1, Figures 27, 28). The ranking of clusters across features is observed here, as well. The green cluster (cluster 2) is the one to score higher in the features, with the yellow one (cluster 3) being the second in hierarchy and red cluster (cluster 1) exhibiting the worst scores.

Annex III. Brief survey of literature segmentation of energy consumers

1. The energy efficiency behaviour of individuals in large organisations: A case study of a major UK infrastructure operator (Zierler, Wehrmeyer, & Murphy, 2017)

Study/scope: Survey study about energy-efficient behavior in general (use of appliances, heating, chargers etc). Exploratory factor analysis study.

Participants: mid- and high-salary employees of the UK national rail company

Survey structure: 38 question items on 5-point Likert scale (not clear how they were chosen, i.e., with which constructs in mind); 9 additional statements represent (stated) behavior (dependent variable); 5 statements about demographic variables.

Methodology: PCA with varimax rotation and Kaiser normalization (in SPSS). Three criteria used for extracting constructs (factors): eigenvalue >1, at least three question items with loadings > 0.5 and Cronbach alpha > 0.6 over all items with loadings > 0.4

Ten constructs/factors are derived. 35 out of 38 items are uniquely mapped to only one construct (two items load high on two questions). Eventually, six out of the 10 constructs (with Cronback alpha> 0.6) are retained for cluster analysis: Technology Adoption Norms, Benefit Evaluation, Energy Intentions, Goal Flexibility, Energy Awareness and Energy Self-Efficacy

Clustering results: The input to the clustering task are the scores of participants to the 6 constructs that came out of the PCA. They apply a two-level method with log-distance as metric. 5 clusters emerged:

'Technological Sceptic' group (20.9%): low scores for Energy Self-Efficacy, Benefit Evaluation, and Energy Intentions, and no particular high scores. Interpreted as a group who neither feel able nor willing to save energy, and cannot see the economic or environmental benefits to the company of doing so. The only cluster which groups together both low Energy Intentions and low Benefit Evaluation, suggesting this as a key defining feature for this cluster.

'Efficiency-Aware' group (26.3%): scored very highly for Energy Awareness, and somewhat high for Energy Self-Efficacy, with no particularly low scores. This cluster identifies individuals with the highest awareness of energy efficiency campaigns who feel that energy savings are relatively easy for them, but not necessarily those with the highest intention to do so. It perhaps represents those with the best (perceived) access to information.

'Barrier Sensitive' group (n=139, 22.1%) score highly for Benefit Evaluation and Energy Intentions but have low scores for Energy Awareness, Energy Self-Efficacy, Technology Adoption Norms and

Goal Flexibility. This grouping of Factors suggests a personal intention to save energy and a high level of support for energy efficiency measures, but may be held back by a perception that the rest of the organisation needs to adopt technologies faster, and that their personal efforts to save energy will therefore have minimal effect.

'Organisational Barriers' (n=96, 15.3%) score particularly high for Energy Self-Efficacy and fairly high for Benefit Evaluation, but exhibit low scores for Energy Intentions, Energy Awareness, and Goal Flexibility. Of all the clusters, this group had the lowest overall intention to save energy in future, but the highest perceived ease of doing so at a personal level, particularly in economic terms. This suggests that this group may perceive conflicts in desired performance goals as a reason for not pursuing energy efficiency efforts within the business.

'Benefit Sceptic' (15.4%) have high scores for Technology Adoption Norms, Energy Intentions, Energy Self-Efficacy, and Goal Flexibility, and low scores for Benefit Evaluation, and Energy Awareness. The exceptionally high score for Technology Adoption Norms suggests that this group receives the highest perceived technological support from the company, but the low Benefit Evaluation score implies that they are not necessarily in agreement that energy efficiency is a worthwhile use of company resources.

Comments/thoughts

(a) A possible correspondence between the factors derived by PCA and the constructs in our study is:
Technology Adoption Norms ⇔ not clear mapping, partial correspondence to Social Norms
Benefit Evaluation ⇔ Environmental Concern AND Financial Concern AND Consequence Awareness
Energy Intentions ⇔ Attitude
Goal Flexibility ⇔ No mapping
Energy Awareness ⇔ Energy Awareness
Energy Self-Efficacy ⇔ Perceived Behavioral control

(b) One variable is used for Benefit Evaluation, including both Financial Concern and Environmental Concern.

2. *Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioural characteristics* (Soetterlin, Brunner, & Siegrist, 2011)

Study/scope: Study based on emailed questionnaires; The aim was to derive a segmentation of customers that is suited to mapping measures/interventions/marketing campaigns.

Participants: Random sample of Swiss households was emailed. 1506 returned out of 3200+ sent out, 214 with missing values, a total of 1292 responses were subsequently processed.

Survey structure: organized into seven sections addressing energy-saving efforts; motives; acceptance of policies; energy beliefs; general energy-related attitudes; energy-awareness; and socio-demographic variables. Several items score in scale 1-5, some others are binary 0-1.
They distinguish and measure curtailment behavior (related to habits and everyday behavior) and once-done action (such as purchasing energy-efficient products) in various domains (food, housing, mobility). They also question the acceptance of different policy measures.

Methodology: The clustering features (variables) are 17, including both what we could call independent (factors influencing behavior) and dependent variables (actual energy-related behavior). With regard to psycho-social variables, they adopt the VBN model constructs (ascription of responsibility, awareness of consequences and personal norm) and add to it factors such as social norms, personal comfort, perceived self-efficacy and personal efficacy (reminiscent of perceived behavior control) and perceived response efficacy (belief that a specific measure will have a positive outcome). They also (directly) measure financial concerns and environmental concerns, as well as the perceived social pressure for a total of 11 psychosocial variable. They use hierarchical clustering (agglomerative) with the Euclidean distance serving as the distance metric and Ward's method for driving the cluster-merging process. The fitness is measured through the clustering coefficient.

Clustering results: They end up with six clusters, 5.3% is the smallest one (around 60 consumers) and 26.4% the largest (around 330 users).

'Idealistic Energy Savers' (15.6%): They score top at 9 out of 11 psycho-social constructs and come only 3rd in the financial concern variable and 6th in perceived social pressure.
They show the most energy-saving efforts based on both curtailment and energy efficiency behavior and fully accept policy measures in terms of sales and use regulations.

'Selfless inconsequent energy savers' (26.4%): On the one hand, they state high acceptance of policy regulations, pronounced awareness of consequences, and believe in consumers' energy-saving actions; on the other hand, their energy-saving efforts, in particular with respect to curtailment in food domain and energy efficiency in the housing domain, seem rather inconsequential. Their median scores in the 8 pure psychosocial constructs rank always 2nd or 3rd. They are 3rd and 4th in financial and environmental concern and second last in perceived social pressure.

'Thrifty energy savers' (14%) Thrifty energy-savers highly engage in energy-saving efforts as long as they involve no financial disadvantages. Accordingly, they disapprove policy measures based on sales or regulations that are associated with additional financial efforts. Their energy-saving efforts are, in general, rather extrinsically motivated, since besides financial considerations they also

experience the most social pressure to engage in energy-saving behavior. The group scores in the psycho-social variables rank in positions 2-4 (3 times 2nd, 3 times 3rd, 2 times 4th).
They have top scores in financial concern and perceived social pressure (1st) and environmental concern (2nd).

'Materialistic energy consumers' (25.1%) Fewer energy-saving efforts, especially in the domains of mobility and food. Energy-saving actions based on energy efficiency measures in the housing domains, however, are considerably pronounced. Policy measures with possible financial consequences are less accepted. If they engage in energy-saving behavior, this is mainly due to financial considerations.
They hold rank no 4 in almost all variables (they are 2nd in financial concern and 3rd in personal efficacy)

'Convenience-oriented indifferent energy consumers' (5.3%) Convenience-oriented indifferent energy consumers are least likely to engage in energy-saving actions. They largely ignore the fact that the increase in energy consumption and its consequences constitute a serious problem for society, and they neither feel jointly responsible for the present energy situation, nor have energy consciousness anchored in their personal norms. Their behavior is less driven by financial considerations than by concerns regarding personal comfort and convenience. Restrictive political regulations and interference are strongly disapproved of.They have the worst score in all constructs (8+3) and the lowest score in stated behavior.

'Problem-aware well-being-oriented energy consumers' (13.6%) Not eager to engage in energy-saving actions. Their awareness of consequences is rather pronounced and they believe that their energy-saving efforts can make a difference. However, they still do not feel obliged to avoid unnecessary energy. Furthermore, they consider their ability to perform energy-saving behaviors as rather limited. A possible loss of comfort and convenience constitutes a barrier to their engagement in energy-saving efforts, but on the other hand, they perceive a certain social pressure to save energy. Second worst replies to most variables except for consequences of awareness (3rd) and perceived social pressure (2nd).

The proposed interventions per cluster are:
Idealistic energy savers ⇒ no special measure needed, just keep their interest warm through reminding them the energy situation, related problems etc.

Selfless inconsequent energy savers ⇒ more emphasis on trustworthiness of information and sources, information at sales/action points through brochures and expert info.

Thrifty energy savers ⇒ lowest income and with financial concerns, information campaigns stressing the financial consequences of energy-saving measures

Note: This class of users is defined/characterized with respect to its behavior across different domains (food, transportation, energy) and depends on the different reported behaviors in each of them.

Materialistic energy consumers ⇒ financial incentives, subsidies, rewards, campaign that energy savings and quality of life are not incompatible.

Convenience oriented indifferent consumers ⇒ Evoke curiosity and address their desire for pleasure and novel experiences

Comments/thoughts
(a) The study includes/measures the VBN variables we have in our model plus a set of variables that we have covered under TPB and its antecedents, even without explicit reference to the TPB model. There are no variables from the Prototype Willingness model.

(b) The clusters exhibit a trend towards uniform ranking across all psycho-social variables, at least in the 8 basic ones. The Perceived social pressure (~SN) and the Financial Concern, mainly, break this uniformity.

3. Are you an energy saver at home? The personality insights of household energy conservation behaviors based on theory of planned behavior (Liu, et al., 2021)

Study/scope: Survey study aiming to explore the effects of Big Five personality traits on the energy-saving behavior of residents based on the extended theory of planned behavior (TPB).
Participants: 1119 survey participants in Xi'an, China

Survey structure: 1119 valid survey responses out of 2276 participants. Demographic variables included age, education, income, marital status, and occupation. The questionnaire employed in this study has four parts: (1) typical household energy-saving behaviors, (2) psychological factors affecting energy-saving behaviors, (3) five-factor personality traits, and (4) socio-demographic information.

Methodology: k-means used to cluster participants into four groups: i) positives, ii) temperates, iii) conservatives, iv) introverts

Three steps take place as follows:
i) SEM to exam the validity of the theoretical framework.

ii) Cluster the respondents using hierarchical k-means according to personality characteristics. Number of clusters determined using Ward's method and squared Euclidean distance.

iii) Group analysis to compare the psychological and behavioral characteristics across clusters. ANOVA takes place to determine whether TPB factors and household energy-saving behavior have significant differences based on the cluster analysis results. Then, a comparison of TPB attributes and household energy-saving behaviors is executed to find out the differentia of each cluster. Secondly, the study conducts SEM on analysing the psycholog-ical and behavioral patterns of each cluster's respondents and compares the characteristics of each cluster with the overall model to explore the behavioral features of different clusters.

Clustering results: The following five personality attributes have been used for clustering: Extra-version, Agreeableness, Conscientiousness, Neuroticism and Openness. 4 clusters emerged:

Positive people (n=169) have the highest score in most personality components: conscientiousness (i.e., 1.534), agreeableness (i.e., 1.407), openness (i.e., 1.375), and extra- version (i.e., 0.850). They have the lowest neuroticism score (i.e., 1.607), which suggests their highest emotional stability. Positive re-spondents have higher self-discipline and friendliness than others, and they usually have no persistent negative emotions.

2. Temperate people's (n=229) Big Five Personality Traits' overall performance is relatively balanced, and their personality trait scores are higher than the average. The extent of neuroticism (i.e., 0.755) and extraversion (i.e., 0.555) is relatively high within the Big Five personality traits. The Temperate individual is highly perceptive and likes to communicate with others actively, but their emotions are reacting strongly with life experience changes.

3. Con-servative people (n=391) have a high score on neuroticism (i.e., 0.117), while its other Big Five personality traits are below average. For example, openness (i.e., − 0.946) and conscientiousness (i.e., − 0.748) are lower than average. It seems that conservative people get used to the routine and reservation and are easily influenced by the external envi-ronment.

4. The cluster of introverted (n=330) has higher neuroticism (i.e., 0.160) and openness (i.e., 0.143) and lower agreeableness (i.e., − 0.519) and extraversion (i.e., − 0.711). Therefore, people in the cluster seem to have passive emotion regulation and response but are willing to accept new perspectives.

Comments/thoughts: This study confirms the differences in terms of energy saving behaviors across the four groups of energy consumers, according to their personality characteristics. An interesting outcome of the clustering analysis is that the pattern of ranking of clusters across features reported in our cluster analysis (section 5.2.3.2.1), appears in this study as well (Figures 6,7).

4. What makes people seal the green power deal? — Customer segmentation based on choice experiment in Germany (Tabu, Hille, & Wüstenhagen, 2014)

Study/scope: This study aims at identifying the drivers that motivate consumers to subscribe to green electricity tariffs. The main goal of the research is to provide an estimation of the size of the market that could potentially be reached with green electricity products. A great focus has been given to determine the factors that differentiate the consumers that have already subscribed to a green electricity tariff (Adopters) and those that display strong preferences towards a green electricity product without having completed a subscription (Potential Adopters).

Survey structure: Based on the 4968 experimental choices of a sample of 414 German consumers, different consumer segments were identified based on their preferences for different electricity product attributes. The survey data were retrieved by a previous research. Each participant received a series of 12 choice tasks involving comparisons of different electricity products with varying levels of attributes. Each question includes a choice of electricity products depending on 7 different characteristics. The variables included socio-demographics (age, gender, income, education, household size, etc.) and behavioral questions related with practical examples of energy saving awareness and willingness to adopt aiming at eliminating the hypothetical bias of providing more generic and common answers to the questions.

Methodology: A latent class analysis is carried out, with the range of clusters ranging from two to five. To determine the best model fit, four criteria are mainly used; Percent Certainty (Pct Cert), Consistent Akaike Information Criterion (CAIC), Chi-square and Relative Chi-square. One segment was excluded and formed a cluster prior to the latent class analysis, as described below. The rest of the consumers were separated into four more clusters, forming a set of five clusters in total.

Clustering Results: The consumers are categorized into 5 clusters as follows:

- Adopters (n=29): 29 respondents were already subscribed to a green electricity tariff, and therefore, characterized as Adopters.
- Potential Adopters: These profiles are described as Potential Adopters based on their preference for green electricity products. They can be further categorized as:
- Truly Greens (n=117):Truly Greens demonstrate similar product attribute preferences to Adopters.
- Price Sensitive Greens (n=78): These consumers put some emphasis on the electricity mix, but mainly take into account the monthly electricity costs.
- Local Patriots (n=108): Local Patriots show the strongest preferences for local electricity generation compared to all other clusters. Members of this group also consider the monthly electricity costs to be the most important product attribute.
- Likely Non-Adopters (n=82): This segment is the most price-sensitive and takes into consideration the cost of the monthly electricity product. The members of this group are the least likely to choose to buy green electricity due to its higher price.

A further analysis is carried out to evaluate whether significant differences appear across different clusters based on socio-demographic characteristics. Some of them (gender, age, household net income and household size) were similarly distributed across the five clusters except from the level of education (Adopters were higher educated). In addition, the average net income of the Truly Greens was significantly higher than that of Price Sensitive Greens and Local Patriots. With regard to psychological characteristics, segments with a high preference for electricity mixes sourced from renewable energy are characterized by a higher degree of concern for issues related to climate change. Potential Adopters also believe that science and technology will improve many environmental issues without requiring changes in our ways of living.

Comments/thoughts: This study does not directly examine the energy saving behaviour of the consumers. It mainly focuses on their willingness to subscribe to green electricity tariffs, a factor that combines attributes related to both their environment and financial concerns in order to change their behaviour. The findings are very interesting, by confirming the previous literature and providing some new characteristics that seem to play an important role in order to promote energy efficient policies in the future.

Annex IV. Statistical abbreviations

Table 14: Abbreviation table for statistical terms used throughout the document

Abbr.	Statistics' term
CI	Confidence interval
F	F-test or Fisher's F ratio
ANOVA	Analysis of variance
SE	Standard error
R^2	Regression
R	Pearson's product moment correlation coefficient
M	Mean
OR	Odds ratio
alpha, α	Significance level of a hypothesis test (also type I error rate).
n	Sample size
p	p value / value of statistical significance

Annex V. Survey dissemination strategy

We present below some of the material that was used to contact different stakeholders for the prupose of raising awareness about the survey.

A5.1 Sample email for contacting energy sector stakeholders

Dear Sir/Madam,
I hope this email finds you well.

I am writing to you on behalf of the Italian consumer organization Cittadinanzattiva - Active Citizenship Network.

As you might already know, we are currently engaged in a European initiative focused on energy transition and the active role of citizens in achieving a low-carbon future across Europe.

The project, funded by the EU H2020 program, is called "NUDging consumers towards enerGy Efficiency through behavioral science" (NUDGE). Its main goal is to assess and fully unleash the potential of behavioral interventions towards achieving higher energy efficiency, paving the way to the generalized use of such interventions as a worthy addition to the policy-making toolbox

At this stage, we are engaged in the dissemination of an online survey that will allow us to learn how European citizens consume and save energy in their households. Consequently, and given your commitment to energy issues, it would be great if you could participate in the survey!

The survey is available online in 14 different languages.
Click this link, select the language you prefer, and join the survey now!

To find the questionnaires in all the available languages, you can click here.

Thank you in advance for your help!

Best regards,

Manuela Amadori

A5.2 Flyers prepared for raising awareness about the survey

Flyer 1

Flyer 2

Annex VI. Complete Survey

NUDGE - Profiling survey

Start of Block: FILTER

INTRO Dear participant,

Ghent University and imec are studying energy consumption within households and opinions on energy issues. This project is supported by the European Union and aims to inform energy policy.

Filling out this survey takes about 20 minutes. Several questions ask about your personal opinion. Keep in mind that no right or wrong answers exist. Only your opinion matters. Also realize that your participation is completely voluntarily and you can stop whenever you want without having to justify.

Thank you for your participation!
Questions or remarks? Please contact Peter.Conradie@imec.be
The full list of partners in this project include:
- Institute for Climate and Policy (IEECP, NL) - Athens University of Economics and Business (AUEB, GR) BEEGY GmbH (DE) Institute of Science and Innovation in Mechanical and Industrial Engineering (INEGI, PT) Cittadinanzattiva (IT) MVV Energie (DE) DOMX (GR) Spring-Stof (BE) Fraunhofer (DE) Interuniversity Microelectronics Centre (IMEC, BE) Zelena Energetska Zadruga (ZEZ, HR)

PrivacyStatement Before starting, we would like to give you some information on the data we collect, the purpose of the study, and your rights as participant of this study.

Privacy statement
imec vzw is responsible for the processing of your Personal data. They process your feedback and data in accordance with the General Data Protection Regulation (EU) 216/679.

1. What data do we collect?
We process the following personal data from you:
• Identification data: year of birth, gender, country of residence, family composition
• Questions about your subjective estimate of the energy consumption in your household and your opinion energy consumption in general. We use the following validated scales: Theory of Planned Behavior, Prototype Willingness Model, Value Belief Norm Theory, Environmental Concern, Financial Concern, Energy Awareness, and Loss of Comfort.

2. Why do we collect the data?

All data collected is used in the context of the EU project NUDGE (Grant agreement ID: 957012). We use the data to gain insight into your attitudes about energy, your perception of energy consumption within your household and related socio-economic factors. For example, we try to profile European households according to their energy consumption and we can support policy makers in decision making

3. Who will have access to your personal data?
We would also like to give an explanation about the persons, companies or organizations that will have access to your personal data in order to avoid misunderstandings.
• Our own researchers within the research group imec-mict-UGent research group who are participating in this project.
• For the survey, we work with Qualtrics, survey software. We only share personal data with organizations outside the European Economic Area (EEA) if this is necessary to comply with our legal obligations or to ensure our services. Third parties are always screened in advance to ensure that your personal data is handled with the necessary care. The necessary security measures are implemented in accordance with Articles 44-50 of the GDPR.

4. Legal basis for the processing
Data is processed on the basis of consent. However, you always have the right to withdraw this consent. When you withdraw your consent, we will delete your personal data from our database.

5. Your rights
You have the following rights when processing your personal data:
• the right to access your personal data;
• the right to correct your personal data;
• the right to delete your personal data;
• the right to portability of your personal data;
• the right to restriction of processing; and
• the right to withdraw your consent.
You can easily do this by e-mail to Peter.Conradie@imec.be or by post to Interdisciplinary imec research group for Media, Innovation and Communication Technologies, De Krook, Miriam Makebaplein 1, 9000 Gent. You can also ask to have your data deleted from the. To do so, we ask that present your identity card to ensure that we do not delete and / or modify data without you wanting to do so yourself.

Finally, you also have the right to lodge a complaint about how your data is treated with the Belgian supervisory authority responsible for enforcing data protection law:
Data protection authority (GBA)
Drukpersstraat 35, 1000 Brussels
Tel. +32 2 274 48 00
e-mail: contact@apd-gba.be
Website: www.dataprotectionauthority.be

6. Technical and organizational measures and duration of storage
imec has taken appropriate technical and organizational measures, which are incorporated in internal information security documents, to protect your personal data.
Your collected personal data will be stored up to 5 years after the last collection.

7. Contact
If you have any questions about how imec uses your data, please contact:
https://www.imec-int.com/en/privacy-statement
I have read the privacy statement and agree (1)
I disagree (2)

FILTER_HEAD Are you one of the persons in your residence who pays the energy bills? (e.g. electricity bill, gas charges)
No (1)
Yes (2)

FILTER_BIRTHYEAR What is your birth year (e.g., 1981)?

____

End of Block: FILTER

Start of Block: Housing characteristics

HOUSE_INTRO First, we would like to ask you some questions about your residence. If you live at multiple residences, we are interested in the residence you most lived and slept last winter.

FILTER_COUNTRY What is your country of residence?
Belgium (1)
Bulgaria (16)
Croatia (5)
France (9)
Germany (3)
Greece (4)

Italy (6)
Latvia (12)
Lithuania (13)
The Netherlands (2)
Portugal (7)
Romania (10)
Slovakia (11)
Slovenia (15)
Spain (14)
Other (8) ____

HOUSE_TYPE Which of the following best describes your principal residence?
Single-family detached house (1)
Semi-detached house (house with two separate entrances) (2)
Terraced house (row house) (3)
Apartment in a multi-family house (4)
Other (5) ____

HOUSE_PROPERTY Which of the following best describes the property type of your principal residence?
Owned residence (1)
Rented residence (2)
Living in a free residence (possibly payment in kind) (3)

HOUSE_SURFACE What is the living area of your house/apartment in terms of square meters? Please round to the nearest multiple of 10 (e.g., 30, 330).
less than 20 (1)
Living area (m²) (2) ____
400 or more (3)
I don't know (4)

HOUSE_RENOV Since it was built, did your house undergo any major energy-saving renovations? If yes, when did the last one occur? (e.g. renovation of windows, facade or roof, change or renovation of space or water heating system)
No (1)
Yes, before 2005 (2)
Yes, in the period 2005-2009 (3)
Yes, in the period 2010-2014 (4)
Yes, in the period 2015-2019 (5)
Yes, major renovations are taking place from 2020 (6)
Yes, but I don't know the exact year (7)
I don't know (8)

X→

HOUSE_HEATING Which energy source do you primarily rely on for space heating?
Fuel oil (1)
Wood (2)
Pellets (3)
District Heating (4)
Heat Pump (5)
Solar (6)
Gas (7)
Electricity (8)
Other (9) ____
I don't know (10)

X→

HOUSE_HEATING_SYSTEM How do you heat the following rooms? Multiple answers are possible.

	Central heating radiator(s) (1)	Electrical radiator(s) (2)	Underfloor heating (3)	Convector(s) (4)	Stove (5)	Other (6)	I don't know (8)	I don't heat this room (7)
Kitchen (HOUSE_HEATING_KITCHEN)								

Living room (HOUSE_HEATING_LIVING)
(master) Bedroom (HOUSE_HEATING_BED)
(master) Bathroom (HOUSE_HEATING_BATH)

X→

HOUSE_RENEW Which renewable energy systems do you have or energy-efficient measurements have been taken? Multiple answers are possible.
⊗None (1)
Solar panels (2)
Solar water heating (3)
Heat pump(s) (e.g., based on air source or ground source) (4)
Biomass heating system (e.g., wood pellet stove, biomass boiler) (5)
Other (7) ____
⊗I don't know (8)

HOUSE_PV Would you be interested in installing a photovoltaic (PV) system?
Yes (1)
No (2)
I already have installed a PV system (or in the process of installation) (4)

HOUSE_PV_REFUSE What is currently preventing you from using solar panels? Multiple answers are possible.
Lack of information (1)
Insufficient financial resources (2)
Non-ideal conditions of my rooftop (technical: unfavorable orientation, rooftop cover) (3)
Low return on investment (4)
Administration (5)
Low level of interest (6)
Other (7) ____

EPC Estimate the Energy Performance Certificate (EPC) of your property on the color scale below. A property with a green EPC label is very energy-efficient with low energy costs, and on the other side a property with a red EPC label is not at all energy-efficient with high energy costs.

HOUSE_ENER_TRACKING Which of the following devices do you have in your household that track energy usage in real-time? multiple answers are possible.
A smart meter consists of a digital display and communication ability that allows the meter to be read remotely. An energy management system provides real-time energy monitoring and allows to control and automate energy flows.

⊗I don't have one (1)

Smart meter (2)

Energy management system (3)

Other (4) ____

⊗I don't know (5)

HOUSE_THERMOSTAT Which type(s) of thermostat do you have to adjust the temperature setting at your residence? Multiple answers are possible.
With a manual thermostat the temperature setting will stay at a certain point unless it is manually changed. A programmable thermostat allows to set a consistent temperature setting for week and weekend days. A smart thermostat learns and predicts your preferences.

⊗I don't have one (1)

Manual or non-programmable thermostat (2)

Programmable thermostat (3)

Smart thermostat (4)

Programmable thermostatic radiator valves (TRVs) or zone control to set different temperatures in individual rooms (5)

Other (6) ____

⊗I don't know (7)

End of Block: Housing characteristics

Start of Block: Energy saving behavior

SAVBEH_INTRO The following questions address daily activities that might take place in your household to save energy. For your convenience these activities are grouped by four thematic areas; activities related to space heating, activities that take place in the kitchen, activities that take place in the bathroom, and doing laundry and other activities. If an activity is not possible in your household (e.g., because you don't have a particular appliance), please indicate with 'not applicable'.

X→

SAVBEH01 How regularly do you perform these activities in your daily life that concern space heating to save energy?

	Not applicable (0)	Never (before) (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)
Turning heating off while airconditioning is on (1)						
Closing windows when heating is on (2)						
Keeping the doors closed to unheated areas in winter (3)						
Closing curtains and/or blinds to prevent heat loss in winter and heat gain in summer (4)						
Wearing more clothes instead of turning the heating up (5)						
Lowering daytime/nighttime thermostat setting (6)						

Turning off heating when absent (7)
Turning down temperature in unused rooms (8)

SAVBEH02 How regularly do you perform these activities in your daily life that take place in the bathroom to save energy?

	Not applicable (0)	Never (before) (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)
Reducing hot water temperature in thermostat settings (1)						
Reducing the number of baths/showers per week (2)						
Turning off tap when soaping up/cleaning teeth (3)						
Preferring a shower over bathing (4)						
Reducing showering time (5)						

SAVBEH03 How regularly do you perform these activities in your daily life that take place in the kitchen to save energy?

	Not applicable (0)	Never (before) (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)
Filling the kettle only with the needed amount of water before boiling (1)						
Cooking with pots covered (2)						
Only using dishwasher when fully loaded (3)						
Turning off tap when washing dishes (4)						
Using energy-saving programme (e.g. eco-mode) of dishwasher (5)						
Defrosting of freezer to remove icing (6)						
Optimizing temperature set point of cold appliances, such as refrigerator and freezer to prevent freezing of interior of the appliance (7)						

SAVBEH04 How regularly do you perform these activities in your daily life to save energy?

	Not applicable (0)	Never (before) (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)
Switching on electric devices when sun is shining as PV production is						

high (SAVBEH04_8)
Charging my electric vehicle when sun is shining as PV production is high (SAVBEH04_9)
Only using washing machine when fully loaded (SAVBEH04_1)
Frequently doing laundry at lower temperature, i.e. 40°C instead of 60°C (SAVBEH04_2)
Only using tumble drier when fully loaded (SAVBEH04_3)
Using a clothes line rather than a tumble drier (SAVBEH04_4)
Turning off all unnecessary appliances completely when not in use (not in stand-by) (SAVBEH04_5)
Switch of the TV when no-

one is watching (SAVBEH04_6)	
Turning off lights when leaving a room (SAVBEH04_7)	

X→

PERC_IMPACT Imagine a ladder with 9 steps, with people on the first step who live not at all energy conscious and have relative high energy bills, and on the highest step, the ninth, people who live very energy conscious and have relative low energy bills. Where are you at the moment?

9 = very energy conscious and relative low energy bills (9)
8 (8)
7 (7)
6 (6)
5 (5)
4 (4)
3 (3)
2 (2)
1 = not at all energy conscious and relative high energy bills (1)

X→

INT_GEN To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
I intend to save energy at home (1)					
I want to save energy at home (2)					
There is a chance that I					

save energy at home (3)	

End of Block: Energy saving behavior

Start of Block: Pschychological statements

ATT_INTRO
You are almost halfway through the survey! The next section consists of a number of statements that address your personal opinion on energy saving in the household and energy issues in general.

Please indicate to what extent you agree with the following statements. We are interested in your personal opinion; there are no right or wrong answers.
Some statements are rather similar, but please answer all questions.

X→

ATT Indicate which answer applies to you.

For me, saving energy by lowering the temperature setting in winter is ...

	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	7 (7)	
Useless								Useful
Disadvantageous								Advantageous
Foolish								Wise
Ineffective								Effective
Dull								Interesting

X→

SN To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
Most people who are important in my life would approve that I save energy by lowering the temperature setting in winter (1)					
People who are important to me expect that I save energy by lowering the temperature setting in winter (2)					
I think most people who are important in my life would not mind that I save energy by lowering the temperature setting in winter (3)					
Most people who are important in my life save energy by lowering the temperature setting in winter (4)					

PBC To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
I have the capabilities to save energy by lowering the temperature setting in winter (PBC_1)					
If I would want it, I could save energy by lowering the temperature setting in winter (PBC_2)					
If it were entirely up to me, I am confident that I could save energy by lowering the temperature setting in winter (PBC_5)					

INT_SPEC To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)

I intend to save energy by lowering the temperature setting in winter (1)
I want to save energy by lowering the temperature setting in winter (2)
There is a chance that I save energy by lowering the temperature setting in winter (3)

PERC_BEH In the last winter, how often did you save energy by lowering the temperature setting?

never (1)

a few times (2)

a number of times, but less than half the days (3)

on about half the days (4)

most days (5)

almost every day (6)

every day (7)

PROT_FAV We want to know what you think about someone who saves energy by lowering the temperature setting in winter. We don't mean anyone in particular, just someone with a similar income who does or would do this. Can you indicate to what extent you think he/she has the following properties

	Not at all (1)	(2)	(3)	(4)	Totally (5)

conscious (PROT_FAV_1)
progressive (PROT_FAV_3)
smart (PROT_FAV_4)
green (PROT_FAV_5)
responsible (PROT_FAV_6)

PROT_SIM Think about a person that saves energy by lowering the temperature setting in winter.

Do you resemble the typical person who saves energy by lowering the temperature setting in winter? (1)	No (1)	Rather no (2)	Neither yes nor no (3)	Rather yes (4)	Yes (5)
How similar or different are you to the type of person who saves energy by lowering the temperature setting in winter? (2)	Different (1)	Rather different (2)	Neither similar nor different (3)	Rather similar (4)	Similar (5)
I am comparable to the typical person who saves energy	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)

by lowering the temperature setting in winter. (3)					
To what extent are you like the typical person who saves energy by lowering the temperature setting in winter? (4)	Not at all alike (1)	Slightly alike (2)	Somewhat alike (3)	Moderately alike (4)	Extremely alike (5)

WILL How likely are the following situations during winter?

	Extremely unlikely (1)	Unlikely (2)	Neither likely nor unlikely (3)	Likely (4)	Extremely likely (5)
You lower the temperature setting in all unused rooms when you are at home all day. (1)					
You lower the temperature setting when you leave home. (2)					
You keep the doors closed to prevent heat loss. (3)					
You go to sleep and you lower the					

temperature setting. (4)

ENV_CONCERN To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
I am very concerned about the environment (ENV_CONCERN_1)					
I would be willing to reduce my energy consumption to help protect the environment (ENV_CONCERN_3)					
Major political change is necessary to protect the natural environment (ENV_CONCERN_4)					

FIN_CONCERN To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
I pay attention to energy-saving tips to reduce my electricity bills (FIN_CONCERN_1)					

I keep track of my (monthly) electricity bills (FIN_CONCERN_2)	
I am motivated to keep my (monthly) electricity costs under a reasonable amount (FIN_CONCERN_3)	

ENERGY_AWARE To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
I know energy-saving methods well (5)					
I know much about the energy-saving tips of daily life (6)					
I feel knowledgeable about saving energy (7)					

LOSS_COMF To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)

Energy conservation means I have to live less comfortably (LOSS_COMF_5)	
My quality of life will decrease when I reduce my energy use (LOSS_COMF_6)	
To me, energy-saving behavior entails losses of comfort that are too high (LOSS_COMF_8)	

X→

PREVBEH Indicate your opinion on the following statements or insert the most appropriate behaviour.

	Reduced a lot (1)	Reduced (2)	Haven't changed (3)	Increased (4)	Increased a lot (5)
My efforts to save energy by lowering the temperature setting in winter over the past half year have ... (PREVBEH_1)					
My efforts to save energy by lowering the temperature setting in winter over the past four					

weeks have .... (PREVBEH_3)	

PREVBEH Indicate your opinion on the following statements or insert the most appropriate behaviour.

Because of lowering the temperature setting in winter ...

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
... my energy consumption has fallen a lot over the past half year. (PREVBEH_2)					
... my energy consumption has fallen a lot over the past four weeks. (PREVBEH_4)					

CONSEQ_AWARE To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
Energy conservation contributes to a reduction of global warming (CONSEQ_AWARE_2)					

The increasing energy demand is a serious problem for our society (CONSEQ_AWARE_4)
The increasing shortage of energy sources is a serious problem for our society (CONSEQ_AWARE_5)

ASCR_RESP To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
I take joint responsibility for the depletion of energy resources (ASCR_RESP_1)					
I feel jointly responsible for the greenhouse effect (ASCR_RESP_2)					
I take joint responsibility for environmental problems (ASCR_RESP_3)					

PERS_NORM To what extent do you agree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
I feel morally obliged to reduce my energy use, regardless of what other people do (4)					
I feel guilty when I use a lot of energy (5)					
I feel good about myself when I do not use a lot of energy (6)					

End of Block: Pschychological statements

Start of Block: Potential of future technologies

DATA_INTEREST
Energy platforms are able to reveal certain activities that occur within a dwelling depending on its accuracy. This ranges from average monthly energy usage to real-time usage of particular appliances. Imagine that such an energy platform is present at your household, to what extent are you interested to get more insight in the following energy consumption levels of your household?

An energy platform is a smart thermostat or energy management system that provides real-time energy monitoring and allows to control and automate energy flows.

	None at all (1)	A little (2)	A moderate amount (3)	A lot (4)	A great deal (5)
Average monthly energy usage (1)					

Daily property occupancy (patterns) (2)
Active occupancy (i.e., residents are awake or asleep) (3)
Energy type (i.e., gas and electricity) (4)
Purpose of energy usage (e.g., space heating, hot water preparation) (5)
Usage of major appliances (e.g., washing machine) (6)
Real-time usage of many appliances (7)

DATA_SHARING Imagine that such an energy platform is present at your household, which energy consumption data are you willing to share with whom? Multiple answers are possible.

An energy platform is a smart thermostat or energy management system that provides real-time energy monitoring and allows to control and automate energy flows.
Your family members are all people that live at the same address.
A neighborhood consists of the local residents who live geographically nearby.

An energy provider (or supplier) provides electricity, gas, etc. to homes and businesses, and is the company that you pay your energy bill to.
An energy distributor (or utility) owns and maintains the infrastructure, such as power lines and

meters that get electricity to the homes and businesses.
Third parties, such as technical service companies, that provide maintenance services for key home appliances (e.g., space/water preparation boilers, water heaters).
The government sets the (regional) energy policy.

	Not shareable (1)	Family members (2)	Neighborhood (7)	Energy provider (3)	Energy distributor (4)	Third parties (5)	Government (6)
Monthly energy usagereveals average energy usage (1)							
Daily energy usage reveals property occupancy (patterns) (2)							
Half-hourly energy usage reveals active occupancy (i.e., residents are awake or asleep) (3)							
Every minute identifies usage of major appliances (e.g., washing machine) (4)							
Every second identifies usage of							

many appliances (5)

End of Block: Potential of future technologies

Start of Block: Socio-demographics

SOCIODEMO_INTRO You are almost at the end of the survey! In this last part we are interested in you as a person and the household you live in.

GENDER What is your gender, as indicated on your national ID or passport?
Female (1)
Male (2)
Other (3)

HOUSEHOLD Which of the following best describes your household type?
Single person (1)
Single parent with 1 or more children (2)
Couple, without children (3)
Couple, with 1 or more children (4)
Living with parents (5)
Non-family household (e.g., co-housing friends) (6)

FAMILY At the end of year 2020, how many people of the following age groups lived in your residence, including yourself?

This includes all people that live at the same address and fall under your household's financial responsibility, including children and young adults who might live elsewhere during the week (e.g., boarding school, shared custody)

None									10 or more	
0	1	2	3	4	5	6	7	8	9	10

Children younger than 14 years (1)	
Children/Teenagers between 14 and 19 years (2)	
Adults between 20 to 64 years (3)	
Adults from 65 years (4)	

DEGREE What is the highest educational degree you have completed?
None (1)
Primary education (2)
Lower secondary education (3)
Upper secondary education (4)
Bachelor's or equivalent level (5)
Master's or equivalent level (6)
Doctoral or equivalent level (7)

CAREERSTATUS What is your principal career status?
Employed (full time) (1)
Employed (part time) (2)
Self-employed / Freelancer (3)
Student / Trainee (4)
House-wife / House-husband (5)
Seeking work (6)
Temporary leave (e.g., sick or maternity leave, career break) (7)
Unable to work due to long term illness or disability (8)
Retired (9)
Other (10) ____

INCOME What was the average monthly total net income of your household in 2020 (in Euros)? Think about all net incomes of all your household members. Count all incomes together, for example wages, benefits for unemployment, retirement fees, etc.?
Below 501 (1)

501 - 1000 (2)
1001-1500 (3)
1501-2000 (4)
2001-2500 (5)
2501-3000 (6)
3001-3500 (7)
3501-4000 (8)
4001-4500 (9)
4501-5000 (10)
5001-5500 (11)
5501-6000 (12)
6001-6500 (13)
6501-7000 (14)
Above 7000 (15)
No answer (16)
I don't know (17)

X→

INCOME_COHOUSING What was your average monthly total net income in 2020 (in Euros)? Count all incomes together, for example wages, benefits for unemployment, retirement fees, etc.?
Below 501 (1)
501 - 1000 (2)
1001-1500 (3)
1501-2000 (4)
2001-2500 (5)
2501-3000 (6)
3001-3500 (7)
3501-4000 (8)
4001-4500 (9)
4501-5000 (10)
5001-5500 (11)
5501-6000 (12)
6001-6500 (13)
6501-7000 (14)
Above 7000 (15)
No answer (16)
I don't know (17)

ZEZ_TXT At the Green Energy Cooperative (ZEZ), the leading Croatian organization in the field of civic energy, we are doing everything in our power to make it easier for citizens to get solar energy. One of the products of this is the Solar Club, a Facebook community of solar enthusiasts and current and future owners of small solar power plants. Completely free and without obligations, a member of the Solar Club can be any person who: Wants to follow current information on solar innovations in Croatia and the world, Wants to know more about the support of the Green Energy Cooperative in the installation and use of a small solar power plant, Wants to be an active part of the solarization movement in Croatia! For more information, go here.

REMARK Should you have any remarks or questions, please leave them below.

____

____

____

____

____

End of Block: Socio-demographics