Application of cumulative prospect theory in understanding energy retrofit decision
A study of homeowners in the Netherlands

Ebrahimigharehbaghi, Shima; Qian, Queena K.; de Vries, Gerdien; Visscher, Henk J.

DOI
10.1016/j.enbuild.2022.111958

Publication date
2022

Document Version
Final published version

Published in
Energy and Buildings

Energy & Buildings 261 (2022) 111958

Contents lists available at ScienceDirect

Energy & Buildings

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

Application of cumulative prospect theory in understanding energy retrofit decision: A study of homeowners in the Netherlands

Shima Ebrahimigharehbaghi [a,*], Queena K. Qian [a], Gerdien de Vries [b], Henk J. Visscher [a]

[a] Delft University of Technology, Faculty of Architecture & the Built Environment, Julianalaan 134, Delft, The Netherlands
[b] Delft University of Technology, Faculty of Technology, Policy and Management, Jaffalaan 5, Delft, The Netherlands

A R T I C L E I N F O

Article history:
Received 14 December 2021
Revised 6 February 2022
Accepted 15 February 2022
Available online 19 February 2022

Keywords:
Energy retrofit
Cumulative prospect theory
Expected utility theory
Cognitive bias
Insulation
Double-glazing

A B S T R A C T

Retrofitting residential buildings can help mitigate the effects of climate change. Cognitive biases are systematic deviations from rationality in decision making and can lead to inaction, delay, and uncertain decisions. Understanding the cognitive biases involved in residential renovation decisions and developing interventions to overcome them can help increase residential renovation rates. Despite their importance, few studies have examined the impact of cognitive biases on energy retrofits. The question addressed in this study is: "Can accounting for cognitive biases improve the prediction of homeowners' actual investment decisions, and how can the outcomes be used to recommend potential behavioural interventions?". Expected Utility Theory (EUT) and Cumulative Prospect Theory (CPT) are compared to evaluate which model(s) more accurately describes actual decision-making behaviour regarding energy retrofits. The EUT assumes a rational decision maker. The CPT is a quantitative model that assumes a decision-maker operating under risk and uncertainty and subject to the cognitive biases of reference dependence, loss aversion, decreasing sensitivity, and probability weighting. The influences of cognitive biases on energy retrofit decisions can be quantified if the relative performance of CPT versus EUT is more accurate. The data for these analyses come from housing surveys conducted in the Netherlands in 2012 and 2018, which also collected data on energy modules. 2,784 and 2,878 homeowners were surveyed, respectively. The model is validated by estimating the coefficients of EUT and CPT and identifying the similarities and differences between the results of the two datasets. Before estimating the parameters, four household clusters are identified using grey relational analysis and the K-Means cluster. For the first time, the EUT and CPT parameters are estimated for four clusters and two energy retrofits, double glazing and insulation, using a genetic algorithm because the equations are nonlinear. The results confirm that CPT provides a better description of the actual decision behaviour than EUT using the two previously established initial values of Layard et al. (2008) and Häckel et al. (2017) as well as the parameters estimated by the genetic algorithm. In the latter case, CPT correctly predicts the decisions of 86% of homeowners to renovate their homes to be energy efficient or not. EUT, on the other hand, overestimates the number of decisions to renovate: it incorrectly predicts retrofit for 52% of homeowners who did not renovate for energy efficiency reasons. Using the estimated parameters of CPT, the cognitive biases of reference dependence, loss aversion, diminishing sensitivity, and probability weighting can be clearly seen for different target groups. The groups with the highest average incomes and house values had the highest loss risk aversion parameters. These households invested more in installing insulation and double glazing.

1. Introduction

Two hundred countries have agreed to the Paris Climate Agreement, which states that global temperatures should be less than 2 °C and ideally less than 1.5°C above the pre-industrial era baseline. Energy inefficient buildings account for about 75% of the EU building stock. EU countries need to double their retrofit rates if they are to achieve the energy and climate targets set by the EU Commission. In the Netherlands, the housing stock consumes a large amount of natural gas, which accounts for almost 71% of the country's total energy consumption. The Dutch government has therefore set a target to eliminate natural gas as an energy source in this sector by 2050. The technical and financial solutions to improve the rates are in place, but homeowners are not using them as much as expected.

* Corresponding author.
E-mail address: s.ebrahimigharehbaghi@tudelft.nl (S. Ebrahimigharehbaghi).

https://doi.org/10.1016/j.enbuild.2022.111958
0378-7788/

Recent research suggests that cognitive biases are important factors influencing home investment behaviour [78,14,20]. Cognitive biases are systematic deviations from rationality in decision making and can lead to inaction, delay, and unstable decisions. Understanding the cognitive biases involved in retrofit decisions and developing interventions to overcome them can help increase retrofit rates. Often, homeowners choose not to renovate due to risk aversion, which explains why energy retrofit rates remain low despite energy targets and policy interventions. In order to encourage homeowners to renovate, it is important to understand their behaviour in terms of their preferences, expectations, experiences and especially their cognitive biases during the retrofit decision process. Despite their importance, few studies have examined the impact of cognitive biases on energy retrofits.

Expected utility theory (EUT) and various variants of prospect theory (PT) are probably the most widely used models for evaluating decisions under risk. EUT assumes a perfectly rational individual who maximises the highest expected utility. In contrast, Cumulative Prospect Theory (CPT) describes what occurs in a situation rather than what should occur. This theory considers the influencing factors that lead to less optimal, less rational decisions. CPT offers potential explanations for many cognitive biases and allows for the quantitative study of these biases. CPT considers the cognitive biases of reference dependence, loss aversion, diminishing sensitivity, and probability weighting [64,71,43]. Policy makers can benefit from quantifying cognitive biases because the effectiveness of policies can be subsequently analysed given the presence of cognitive biases.

In a very recent study, Rockstuhl et al. [65] mentioned that volatile future energy cost savings is one of the main barriers to implementing energy retrofits. The authors applied the EUT to investigate the investment decisions using the averaged data of German commercial buildings. The highest optimal investment amount is achieved when decision makers aim to maximise energy cost savings compared to the investment-only perspective (considering expected wealth in the future, as well as related perceived risk). The use of CPT has the potential to improve the results of this study by accounting for cognitive biases. The impacts of behavioural biases on energy efficiency investments were investigated by Häckel et al. [33]. The authors used the original specifications by Tversky and Kahneman [71] and compared the EUT and CPT results based on different scenarios.

Few studies have examined energy efficiency investments using CPT [30,33]. Empirical investigations of the parameters of CPT have not been conducted for energy efficiency investments in previous studies. This study aims to investigate the impact of cognitive biases on energy retrofit investment decisions. To achieve this goal, the models with and without cognitive biases are compared. The questions to be answered by our study are as follow: (a) Whether CPT describes the actual decision-making behaviour more accurate compared to EUT in the context of energy efficiency investments?, (b) Which cognitive biases significantly determine the Dutch homeowners' behaviours towards energy efficiency investments?, (c) Whether CPT parameters vary for different groups of individuals and types of energy efficiency investments?, and (d) How can the results of CPT be used to recommend potential behavioural interventions for promoting the energy efficiency renovations in the Dutch owner-occupied sector?.

We are the first to empirically estimate the parameters of CPT. The parameters of EUT and CPT are estimated from actual data to avoid potential problems associated with assumed responses, such as not accounting for cognitive biases [75,37]. The energy modules of the 2012 and 2018 Dutch household surveys are used to investigate the parameters of CPT in terms of predicting the actual behaviour of homeowners in their decisions to renovate or not. The approach used in this study is innovative in several ways: (a) the EUT and CPT parameters are estimated for 2,784 and 2,878 homeowners, respectively; previous studies have examined only one building type and individual homeowners (e.g., [33]); (b) homeowners are grouped based on building and household characteristics so that different parameters are estimated for each group of households; and (c) two types of energy retrofits, insulation and double-glazing, are examined from actual data using the 2012 and 2018 energy modules.

This article is organised as follows: cognitive biases and behavioural interventions are discussed in Section 2. Section 3 describes EUT and CPT. Section 4 explains the data sets and research methodology. The results of the analyses, discussion, and conclusions are presented in Sections 5–7, respectively.

2. Review of earlier studies on cognitive biases and behavioural interventions in the energy efficiency literature

2.1. Overview of cognitive biases in the energy efficiency literature

The energy efficiency gap shows the difference between the theoretical potential energy efficiency and the actual achieved energy efficiency. Neoclassical theory explains the existence of the energy efficiency gap through market failures, environmental externalities or imperfect information. In contrast, behavioural economics attributes energy efficiency gaps to systemic biases, such as high uncertainty about future energy savings [7]. The determinants of the energy efficiency gap are examined in classical, institutional and behavioural economics literature by Gillingham and Palmer [28]. These determinants are shown in Table 1. Häckel et al. [33] compared expected utility theory and prospect theory, following the work done by Mayer; Greene [54,32]. Based on cumulative prospect theory, the behavioural biases of reference dependence, loss aversion, diminishing sensitivity/risk aversion, and probability weighting/uncertainty explain the energy efficiency gap. Furthermore, the energy efficiency gap is determined by high sunk costs and uncertainty of energy prices. More importantly, loss aversion influences investment decisions drastically, compared to other cognitive biases.

The barriers from individual, organisational, and institutional perspectives are evaluated for the construction of green buildings in the United States by Hoffman and Henn [38]. Two main theories of behavioural economics are adopted: (1) bounded rationality: individuals are restricted in their ability to achieve pure rationality; and (2) heuristics thinking: individuals rely on simplifying strategies, which cause a wide variety of decision-making biases. For individuals, the following barriers are considered: (a) over-discounting the future; (b) ego-centrism; (c) positive illusions; (d) presumed associations; (e) mythical fixed-pie bias; and (f) environmental literacy. The authors proposed the following strategies to overcome the decision-making biases: issue framing, targeting the right demographic, education, structural and incentive change, compensating risk, green building standard improvements, and tax reform.

Klotz et al. [45] examined the *anchoring effect* on the energy performance goals of commercial buildings in the United States. Three surveys were conducted. The first four questions asked about benefits and incentives for energy consumption reduction. In each of these surveys, identical questions were asked, but with different energy consumption reductions, thereby creating different anchors. One survey arranged an anchor of 90% energy consumption reduction over standard practice; one arranged a 30% anchor; and one set no anchor. At the end of the surveys, participants exposed to different anchors were asked to set an energy efficiency target for a new project. Participants exposed to either the 90% energy consumption reduction anchor or no anchor set higher tar-

Table 1
Determinants of the energy efficiency gap [28].

Category	Factors influencing the energy efficiency gap			
Behavioural anomalities and failures	Non_standard preferences: self-control problems, reference - dependent preferences	Non_standard beliefs: systematic incorrect beliefs about the future	Non_standard decision making: limited attention, framing, sub-optimal decision heuristics	
Market Failures	Imperfect information, regulatory failures	Principal-agent issues	Credit constraints	Learning by using: no evidence for energy efficiency technologies
Other reasons	Transaction Costs, Uncertainty	Consumer Heterogeneity	Rebound effect	because of engineering calculation, e.g. not including the interactions between different investments

gets. Therefore, building rating systems that only support incremental energy improvements may accidentally generate low anchors and, thereby, discourage more advanced energy performance goals that are both technically and economically feasible. In an analytical framework study, Klotz et al. [44] indicated that cognitive biases are one of the main hindrances to achieving sustainability targets of commercial buildings in the United States. Professional bias and group thinking bias were specifically identified in the study. *Motivational framing*, e.g. achieving a healthier neighbourhood by using less greenhouse gases, changes behaviour more than *sacrifice framing*, e.g. getting used to driving less, turning off the lights and reducing the heat. These conclusions were based on a survey among 1,000 householders in Ontario, Canada [27]. Taranu and Verbeeck [68] highlighted the role of both rational and heuristic thinking in explaining pro-environmental behaviour. The results verified that homeowners' positive arguments in favour of energy retrofit are mostly rational, and that negative arguments are mainly heuristic. In a very recent study by Good [30], a behavioural economics model was developed to evaluate the impact of behavioural biases on reducing energy consumption for a demand-responsive electricity system. Among biases, *the endowment effect and the time-discounting* were considered. These biases influence the demand-response provision, particularly when the demand of an entity is high.

2.2. Overview of behavioural interventions in the energy efficiency literature

The household behaviours can be influenced by behavioural interventions and nudge tools. The current study also focuses on proposing potential behavioural interventions using the results of CPT for different target groups of households and buildings. Osbaldiston and Schott [58] and Abrahamse et al. [2] reviewed articles and provided a list of interventions targeting the householder's behaviours with regard to energy. They categorised them into: (1) convenience: easy and prompt interventions; (2) information: information on justifying behaviours and guidance on changing behaviours; (3) monitoring: feedback and rewards; and (4) social influence: social modelling, cognitive dissonance, commitment and setting goals. Many of these interventions change the context in which the behaviour takes place, e.g. smart meters provide live information about the current and accumulated energy consumption of a household. Context change is the core pillar of the 'nudge' tools. [1] These tools are generally similar to behavioural interventions [36].

Thaler and Sunstein [70] used the term 'nudge' and defined it as 'any aspect of the choice architecture that predictably alters people's behaviour, without forbidding any options or significantly changing their economic incentives' [p.6]. A nudge can also be seen as a tool to modify people's choices, without removing or changing the number of choices. For example, changing the default setting leads to different choices by households. In an experiment, two default choices of 'green' and 'grey' utility electricity providers were suggested to two groups of households. People in the green utility group were more likely to choose this default option, compared to other groups [1]. These types of behavioural interventions and nudges are not widely covered in ongoing studies.

Frederiks et al. [25] extensively reviewed cognitive biases and behavioural anomalities, in predicting household behaviours with regard to energy consumption. The most prevalent biases were status quo bias, loss and risk aversion, sunk-cost effects, temporal and spatial discounting, and availability bias. Additionally, psychological factors such as normative social influence, intrinsic and extrinsic rewards, and trust significantly change household behaviours with regard to energy use. See also [57]. In addition to the most frequent biases, more effective policies in terms of energy consumption were attributed to each bias. Table 2 shows the behavioural biases, definitions, and associated policies. In similar studies, taking account of behavioural biases when designing effective environmental and climate change policies is recognised as being crucial [31,72]. Providing transparent information stimulates energy saving behaviour among households, and giving feedback can considerably reduce the energy bills of households [5]. Dietz et al. [16] proposed an integrated framework from economic, engineering, behavioural and social science, for designing energy policies that aim to increase the energy efficiency of the residential sector. Households use cognitive shortcuts and different mental considerations in making their decisions. Energy policies need to concentrate on the decisions with the highest impact on energy consumption, by the highest number of capable households with the highest probability of making changes.

2.3. Review of energy-efficiency literature on the main influencing factors and barriers of energy retrofits

Previous studies are reviewed to investigate the importance of identifying the barriers for specific groups of households. According to the literature review [76,10,19], the household characteristics, socio-demographics, property characteristics, and salient events (e.g. moving house) determine EER decisions. For instance, in a study of eight European countries, considerable differences in adoption of energy-efficiency technologies were recognised, among different income groups of households. The willingness to pay is considerably lower for the lowest income groups and all types of energy-efficient technologies [66]. In another study, the effectiveness of subsidy programs was investigated by Lihtmaa et al. [51]. Subsidies were assigned equally for all residents. However, an unequal distribution occurred on a regional basis. Low-performing regions gained a lower proportion of national subsidies, leading to the inequalities between regions increasing further, over and above current socio-economic differences.

Abreu et al. [3] emphasised the importance of designing specific policies for different groups of homeowners, with regard to

[1] 'Libertarian paternalism'. Nudge tools are policies designed to encourage individuals toward better choices without restricting their freedom [35].

Table 2
A list of biases, definitions and the policy implications influencing the energy consumption.

Biases	Definition	Policy implications
Status quo bias and defaults	people are not willing to change and prefer to go with the flow of default settings, even where other options may have better outcomes.	Applying the energy related practices with easy and effortless changes to the default settings, e.g. introducing the energy efficient option as default rather than encouraging them to choose energy efficient option between others.
Satisficing	Applying only the effort needed to achieve a satisfactory rather than an optimal result	Inessential complexity and sensory overburden need to be avoided by framing messages in a clear, concise and comprehensible format.
Be loss averse	Considering losses more with the same size gains,	Emphasizing on the cost/ loss reductions of using energy efficiency measures rather than energy savings
Be risk averse	People are more likely to engage in a risky behaviour to avoid a certain loss rather than to engage in a similarly risky behaviour to obtain a comparable gain.	Focusing on low-risk, safe, stable, and secure energy saving measures and investments
Sunk cost effects	After purchasing appliances, people insist to use them even if better choices become available.	Reduce the importance of old energy efficient investments and emphasize on the costs of any inefficient investments
Temporal discounting/spatial discounting	Less valuable further away in time/space. Avoiding on expenses on energy efficiency appliances if the benefits are further away in the future.	Emphasise to the longer-term payoffs of energy consumption
Conform to social norms	The behaviours and attitudes of other people always influence peoples behaviours, such as herd behaviour, the Bandwagon effect.	Formulate energy-saving practices socially desirable behaviour
Be motivated by rewards and incentives	The incentives lead to more behavioural responses.	Use non-monetary rewards such as praise, recognition and social approval
Free-riding effect	Tendency to contributing less for public good when possible and believing that others are contributing less.	Making a group and showing the participations of other people in energy saving activities
Trust	A trustworthy professional is an effective source to influence the decision-making process.	Providing information that originates from a high-credibility source (e.g., public service commission)
Availability bias	People usually use the available information	Specifying the well-publicised popular energy-saving behaviours and favourable to consumers

energy-saving retrofits. In addition to household and building characteristics, daily activities and social practices are identified as important influencing factors. The authors focused on different age groups of homeowners in single-family dwellings. They examined which home-related activities and social practices drive this group of households to conduct energy-efficiency retrofits. The authors concluded that younger homeowners appeared to be more environmentally conscious and implemented 'little-by-little' energy retrofits. The motivational factors for older groups must be stronger, despite their higher incomes. The use of framing is also recognised: when energy-related retrofits are linked with other aspects of the home, such as aesthetics and indoor comfort, the likelihood of those retrofits being undertaken increases. In a study of German homeowners, installation of insulation was assessed, to examine the effect of policy interventions for this group of households [26]. For this type of energy-efficiency retrofit, the policies focused on wall insulation. Furthermore, the homeowners' decisions on energy-efficiency retrofits were highly dependent on their financial resources, age and attitude towards insulation, as well as the structural conditions of the walls. Compelling new homeowners to insulate their walls within the first year can potentially increase the total insulation rate by up to 40%.

Another group of studies focused on multi-family dwellings [17,69,9,11,61]. Dodoo et al. [17] analysed the cost-effectiveness of various energy-efficiency measures, such as insulation, improved windows, or a glass-enclosed balcony, for a typical 1970s multi-family buildings in Sweden. The results indicated that the highest energy saving for a single measure is achieved by improved windows. Furthermore, the cost-effectiveness showed sensitivity to both the real discount rate and energy price growth. The energy-saving potential of deep energy-efficiency retrofits, such as various types of insulation and improved energy-efficient windows and doors, were evaluated for Swedish multi-storey residential building of the 1970s by Tettey and Gustavsson [69]. Energy savings for space heating were significantly increased by the use of energy-efficient windows and doors, balanced ventilation with heat recovery (VHR), and additional insulation to external walls. The benefits of energy-efficiency s, such as insulation, window glazing, and district heating for individuals and national government, were investigated for apartment buildings constructed during the 1970s and 1980s in Estonia [61]. The authors used the net present value (NPV) method to calculate the economic benefits of energy-efficiency s, following the European commission's cost optimality methodology [22]. The authors concluded that energy-efficiency s contributed considerably to economic benefits for both individuals and national governments. These economic benefits would be even higher, if one could place a monetary figure on non-energy benefits whose economic value is difficult to calculate. Similarly, Bonakdar et al. [9] investigated the cost-optimum level of building fabric elements, of extra insulation thicknesses for considered opaque elements, and different U-values for new windows in a multi-storey Swedish residential building. A variety of different economic outcomes was assessed, by including different discount rates and energy price growth rates. However, the results were not particularly sensitive to changes in the lifespan, to figures of 40, 50 or 60 years. Brown et al. [11] evaluated the economic, indoor environmental quality (IEQ), and environmental aspects of energy-efficiency packages for a Swedish multi-family building. A base case, and two packages with higher initial investment costs and higher levels of energy-efficiency s, were defined for each building. Based on the results, the packages that reduced the energy demand considerably (50% energy reduction) have a higher life-cycle cost. Hence, higher initial investment costs for multi-family dwellings are essential to achieving national and international energy efficiency goals.

Another group of studies evaluated the energy-efficiency for single-family dwellings [40,53]. In a study for the Nordic countries (i.e. Denmark, Sweden, Norway, Finland), the behavioural, economic, and market-related hindrances to promoting energy-efficiency s of single-family detached houses built before 1980 were analysed. These dwellings are expected to have substantial primary energy saving potential [53]. The identified barriers were:

lack of need; lack of regulatory requirements on the energy standard of a renovated building; insufficient information; lack of knowledge/awareness about the energy-efficiency measures and the energy and non-energy benefits; lack of trust; uncertainties among financiers and end-users regarding the energy saving levels; difficulty in measuring the monetary values of non-energy benefits, such as improving quality of life; no agreement on the suitable discount rate; difficulty in controlling the occupants' energy use behaviours; and, finally, difficulty in predicting future energy prices. In the current study, different target groups of households, such as by age or type of dwelling, are investigated, to identify their cognitive biases and recommend potential behavioural interventions for each group of homeowners. In a very recent study, Cristino et al. [13] examined the barriers to energy retrofits in Brazil. The first main identified category of barriers was related to governmental and financial aspects. Residents believe that the government is responsible for the implementation of energy retrofits in the country. The importance of behavioural barriers was also explored. Many households were reluctant to invest because they lacked concrete information about the benefits of energy efficient technologies.

3. Expected utility theory and cumulative prospect theory in explaining the individual decision-making processes regarding the energy retrofits

Two psychologist, Amos Tverskey and Daniel Kahneman, developed 'prospect theory' which explains the relation of preferences with regard to attitudes to gains and losses. They won the Nobel memorial prize in economic sciences' for developing this theory and comparing their cognitive models of decision-making under risk and uncertainty to economic models of rational behaviour [71,43]. This study also uses cumulative prospect theory (i.e., in which a decision-maker unintelligently confronts risk and uncertainty) and compares this theory with expected utility theory (i.e., in which a decision-maker who intelligently confronts risk). By comparing these models, appropriate model/s that predict the homeowners' decisions about energy retrofits can be examined. The following subsections describe the EUT and CPT theories.

3.1. Expected Utility Theory (EUT)

Standard neoclassical theory assumes that individuals behave and make decisions rationally under risk by maximising their expected utility. The formal representation of decision-making under risk is as follows:

$$EUT = \sum_{i=1}^{n} p(X_i) \cdot u(\varphi_i) \tag{1}$$

Where n shows the number of payoffs, X_i indicates the payoffs, $p(X_i)$ presents the probability of payoffs X_i and $u(\varphi_i)$ indicates the individual utility of total wealth φ_i. The total utility is calculated based on the initial wealth and the payoffs (φ_0+X_i). The utility function is defined based on Eqs.2 and 3. This utility function indicates the constant risk aversion by individuals [33,49].

$$u(\varphi_i) = \begin{cases} \frac{1}{1-\theta} \cdot (\varphi_i^{1-\theta} - 1) \quad for \quad \theta \neq 1 & (a) \\ ln(\varphi_i) \quad for \quad \theta = 1 & (b) \end{cases} \tag{2}$$

When the EUT is higher with the payoffs. The investment increases the utility and must be implemented.

3.2. Cumulative Prospect Theory (CPT)

Few studies have investigated energy-efficiency investments using behavioural economic theory. Previous work mainly used this model and simulated the behaviour by making assumptions regarding the parameters in the model [30,33]. Häckel et al. [33] suggested empirical investigations of the parameters of CPT as subjects for future research. The current study aims to examine and verify the CPT parameters empirically. The cumulative prospect theory advances the prospect theory by modifying the possible error of first- order stochastic dominance. Furthermore, CPT enables comparison with EUT. CPT defines a value function that depends on the differences in the payoffs. CPT offers advantages in the quantification of many cognitive biases. CPT mainly covers four cognitive biases as presented in Table 3:

Two functions are defined for positive and negative differences [71,33,64]. The value functions are as follows:

$$v(\Delta x_i) = \begin{cases} (\Delta x_i)\alpha \quad \Delta x_i \geqslant 0 \quad (a) \\ -\lambda(\Delta x_i)\beta \quad \Delta x_i \leqslant 0 \quad (b) \end{cases} \tag{3}$$

α and β>0 (usually $\alpha, \beta \leqslant$1) indicate the curve for the positive and negative payoffs, respectively. The λ>0 parameter shows the loss aversion, i.e., weighting the loss more than equal gain. For instance, λ equal to 2 means that individual perceives loss twice more than gain.

Diminishing sensitivity (i.e., risk avoidance in gain situations and risk seeking in loss situations) are included in CPT. This bias is applied in the model by objective probabilities instead of weighting with their subjective values.

$$w(p(\Delta x_i)) = \begin{cases} \frac{p(\Delta x_i)^{\gamma}}{(p(\Delta x)^{\gamma} + (1-(p(\Delta x_i))^{\gamma})^{1/\gamma}} \quad for \quad \Delta x_i \geqslant 0 \\ \frac{p(\Delta x_i)^{\delta}}{(p(\Delta x)^{\delta} + (1-(p(\Delta x_i))^{\delta})^{1/\delta}} \quad for \quad \Delta x_i \leqslant 0 \end{cases} \tag{4}$$

Different studies validate this functional form [48,29]. The weighting function uses the objective probabilities $p(\Delta x_i)$ to the subjective/perceived probabilities $w(p(\Delta x_i)$. Two parameters of δ and γ control the curve of the value function. Diminishing sensitivity is incorporated in the model by the parameters of γ and $\delta \in$ (0,1]. The higher sensitivities are expected for individuals with lower values of γ and δ. Furthermore, larger values for loss δ are expected compared to gain γ as specified by [71]. CPT differs in weighting cumulative probabilities compared to PT by weighting single probabilities ($p(\Delta x_i)$). CPT applies weighting to the cumulative probability distribution. There are a few steps in calculating the weight π_i (1) ranking the payoffs (ascending), (2) using the probabilities of each payoff, (3) using the right function for positive and negative payoffs, and (4) calculating the differences in neighbouring probability weightings. The weighting for the positive payoffs depends on the probabilities of the payoffs being at least as good as the payoff i and the higher payoffs compared to payoff i. The weighting for negative payoffs depends on the weighted probabilities of the payoffs being at least as bad as the payoff i and the weighted probabilities of the payoffs being worse than payoff i. Eqs.8 and 9 show the decision weight π_i.

$$\pi_i = \begin{cases} w(p(\Delta x_i) + \ldots + p(\Delta x_n)) - w(p(\Delta x_{i+1}) + \ldots + p(\Delta x_n)) \quad for \quad \Delta x_i \geqslant 0 \\ w(p(\Delta x_1) + \ldots + p(\Delta x_i)) - w(p(\Delta x_1) + \ldots + p(\Delta x_{i-1})) \quad for \quad \Delta x_i \leqslant 0 \end{cases} \tag{5}$$

The equations are valid for:

$$\Delta x_i \geqslant 0 \quad \text{for} \quad k+1 \leqslant i \leqslant n-1, and$$
$$\Delta x_i \leqslant 0 \quad \text{for} \quad 2 \leqslant i \leqslant k.$$

Where k indicates the number of negative payoffs. CPT is calculated by multiplying the decision weight (π_i) and value function ($v(\Delta x_i)$)

Table 3
Cognitive biases of cumulative prospect theory.

Cognitive bias	Definition
Reference dependence	Individual decision-making depends on the difference between the changes in the utility of the current wealth with their reference point or status quo (usually in the past).
Loss aversion	Individuals perceive the value of loss higher compared to the same value of gain.
Diminishing sensitivity	Generally speaking, people prefer to avoid risk given the prospect of a positive outcome (i.e., gain), but the reverse is true given the prospect of negative outcomes (i.e., loss).
Probability weighting	Individuals use the probabilities of the outcomes instead of statistical probabilities and place a lower weight on the average payoffs (the centre of the distribution), but a higher weight on events with low probabilities (the tails of the distribution).

$$CPT = \sum_{i=1}^{n} \pi_{\mathrm{i}} \cdot v(\Delta x_{\mathrm{i}}) \tag{6}$$

A CPT more and less than 0 indicates a favourable and unfavourable decision, respectively. The values of CPT are not in monetary terms. Although the CPT is originally formulated for one period, the value function can be extended for the long term as well. To achieve this, the status quo, aggregation of future outcomes, and consideration of the time value of money need to be considered. This study uses a multi-period NPV including the status quo (the initial wealth of the individual) to extend the one-period CPT function to the multi-period CPT similar to studies done by Häckel et al. [33,61,22,9]. Therefore, Δx_{i} is replaced by NPV in all the formulas (Eq. 11). Energy-efficiency investments are similar to any other types of investment. An initial financial cost and usually uncertain outcomes are the components of energy- efficiency investments. Therefore, in this case, Net Present value (NPV) is used to evaluate the energy-efficiency investments over the long term. If the total present value of an energy-efficiency investment is higher than the initial investment costs, people might invest in it. The mathematical form of the NPV is:

$$NPV = -I_0 + \sum_{n=0}^{T} \frac{CF_{\mathrm{t}}}{(1+r)_{\mathrm{t}}} \tag{7}$$

Where I_0 is the initial investment. T is the lifetime of the energy-efficiency investments. r is the indicator of the time value of money discounted by the interest rate. The following formula is used to calculate the CF_t in each period:

$$CF_{\mathrm{t}} = P_{\mathrm{t}} \cdot C_{\mathrm{t}} \cdot \varepsilon + UCB_{\mathrm{t}} \tag{8}$$

Where P_t is the stochastic price per source of energy (for instance, for gas it is kWh) for period t. C_t is the energy consumption for period t, and ε indicates the percentages of energy savings per source of energy. UCB_t shows costs and benefits that are difficult to measure and not observable, such as time and effort expended to find reliable information or contractors (i.e., transaction costs) [20,21]. Regarding the benefits, energy saving is the main and observable benefit of implementing energy-saving measures. This benefit might be calculated more easily compared to other benefits such as enhanced quality of life.

4. Methodology

4.1. Database

The data of the main variables defined in EUT and CPT models is collected from the Netherlands Household Survey Energy modules 2012 and 2018. This dataset is released every 5–6 years. Table 4 shows the datasets of this study.

4.2. Methods of analyses

Building and household characteristics significantly influence the energy retrofits according to the previous studies [76,78,6,20,21]. In this subsection, the clustering method is explained. The cluster of households is defined using building and household characteristics. Before finalising the main influencing factors for clustering the data, more variables were used to define different clusters. However, the identified numbers of observations per cluster are not very well distributed. A sufficient amount of data per cluster is required for examining the parameters of EUT and CPT. Therefore, these variables are removed one by one to evaluate the distribution of data per cluster of households. Based on this investigation, it was found that the variables with so many missing values cause a very uneven distribution of the numbers of observations for different clusters. The main building and household characteristics are as follows: (a) building characteristics including building types (multi and single family houses); type of single-family dwellings (detached, two under one roof, middle houses, row houses); type of multi-family dwellings (flat and maisonette); number of rooms; construction period; type of heating system (gas boiler, block or neighbourhood heating, district heating, etc.). (b) household characteristics: number of households; whether relocated or not in the last two years; age; income; and education.

4.2.1. Cluster analysis

The purpose of clustering is to group observations into the classes or clusters, so that objects in the same group have high similarity, and objects in other groups are not alike. 'n' buildings and 'n' households are in the datasets that are called 'instances'. 'm' attributes, i.e. specific characteristics, are defined for each instance. Each instance is assigned to a group. One important step needs to be conducted before the main cluster analysis. The degrees of influence of these attributes differ. For instance, the construction period

Table 4
Data sets for the estimation of EUT and CPT parameters.

Datasets	Number of respondents	Variables
Energy module 2012	2784	Time series of energy consumption (2004–2010), energy labels, energy savings in the dwelling stock, building and household characteristics, energy efficiency in the past five years (i.e. insulation and double glazing), planning for energy retrofit in the coming two years, investment costs, the house values. The data regarding the expectation of households are collected using a survey among almost 5000 of buildings.
Energy module 2018	2787	The only difference with energy module 2012 is that the time-series of energy consumption is not included in this new version. - Energy consumption 2018
Eurostat	_	The initial values of the gas and electricity prices in the Netherlands.
Milieu Centraal	_	The initial values for energy saving percentages, initial investment costs.

is expected to influence energy retrofit considerably more than other factors, such as the number of rooms in the buildings. Therefore, a weighting system is needed to consider the importance of different attributes in the clustering process. Furthermore, the values of attributes need to be normalised to prevent two important miscalculations in grouping the instances: (a) most of the attributes have different units of measurement, and the differences between units influence the quality and accuracy of clustering; (b) this weighting system prevents data with large ranges from having more weight than attributes with smaller ranges.

Grey relational analysis (GRA) is conducted to normalise and to identify the weights of different attributes [47,79,50,81]. First, the min–max normalisation is conducted to make the scale of attributes comparable and within the same range. This method of normalisation retains the relationship between the initial data since it performs a linear normalisation. In this study, the new range is defined between [0,1].

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \left(X'_{\max} - X'_{\min}\right) + X'_{\min} \tag{9}$$

For categorical variables with meaningful orders, it is essential to make an order and then to assign the values between [0,1].

$$X' = \frac{rank_{\mathrm{i}} - 1}{rank_{\max} - 1} \tag{10}$$

Based on geometrical mathematics, grey relational analysis (GRA) is performed to identify the grey relational grades and a grey relational order (i.e., the rank of grey relational grades). These values can show the primary values between the influencing factors and the target variable, i.e. the decision. As mentioned, the influencing factors are defined within two categories of attributes: building and household characteristics. The grey relational grades indicate the degree of the influencing factors on the energy retrofit decision. The advantages of this method are its simplicity and its lack of assumptions regarding the type of probability distributions of the attributes [79. X and Y indicate the influencing factors and the energy retrofit decision, respectively. To calculate the GRA, these steps are followed: 1) normalisation of the data, 2) calculate the grey relational coefficients using Eq. 15 (normally alpha = 0.5):

$$\xi_{\mathrm{i}}(k) = \frac{\underset{i}{min}\underset{k}{min}|y(k) - X_{\mathrm{i}}(k)| + \alpha \underset{i}{max}\underset{k}{max}|y(k) - X_{\mathrm{i}}(k)|}{|y(k) - X_{\mathrm{i}}(k)| + \alpha \underset{i}{max}\underset{k}{max}|y(k) - X_{\mathrm{i}}(k)|} \tag{11}$$

3) compute the grey relational grades 4) rank the obtained grey relational grades, so that the grey relational order can be identified. The grey relational grades are used to weight related attributes in cluster analysis. This grade is between [0,1].

K-Means clustering is used to group similar instances within one group [74,4,55,56]. In this method, no target variable is predicted, i.e., an unsupervised learning problem. Each cluster should have different features: (1) all instances should be very much alike. The sum of squares of distances of each instance from the centroid of a cluster, also called the 'intra cluster distance', are calculated. In this regard, lower values result in a better cluster; (2) the instances in one cluster should be as distinct as possible from other clusters. The 'inter cluster distance' is calculated to indicate the distances between clusters. After calculating these two values, the Dunn index is calculated using Eq. (16). The Dunn index is applied. The values of this index need to be maximised, and a higher value of the Dunn index indicates better clustering. The K-means clustering is an algorithm for minimising the sum of distances between the instances in a cluster with their corresponding cluster centroid [18,60,52].

$$Dunn_i = \frac{min(Inter - cluster - distance)}{max(Intra - cluster - distance)} \tag{12}$$

A few criteria are used to rely on the clusters by the K-means algorithm: centroids of new clusters do not change in the new iteration; instances stay in the same cluster; and finally, the maximum number of iterations is obtained. It is a challenge to achieve the appropriate size of different clusters in terms of scale. If a cluster is too large, a cluster analysis can be conducted for this specific cluster to make several clusters out of it. Another challenge is when the densities of different clusters differ. Again, the k-means clustering algorithm and the use of a higher number of clusters can be applied to solve this issue. The elbow method is used to determine the optimal number of clusters for each data set. In the elbow method, data clustering is performed several times, in each attempted data is clustered in predefined number of clusters. Then, the sum of squared distanced of each data point from the centre of the corresponding cluster is plotted as a function of the number of clusters. The resulting plot should have a shape of an arm where the number of clusters at the location of the elbow will correspond to the highest Dunn index and indicated the optimal number of clusters [80,8,46].

4.2.2. Calculation of the main components of EUT and CPT models

Net present values of energy efficiency investments

Energy saving depends on gas and electricity prices. Future gas and electricity prices are sources of uncertainty in the NPV model. Therefore, different NPVs are calculated for the same values of energy consumption and energy saving percentages for different paths of energy prices. First, these NPVs are computed for each household per type of energy saving measure in the samples (2,784 and 2,878 instances). The NPVs are used as the inputs for calculating the values of the EUT and CPT.

Predicting the energy prices using "Geometric Brownian Motion"

Energy prices are a source of uncertainty for energy retrofit decisions. As with Häckel et al. [33]; Postali and Picchetti [62], energy prices are simulated using Geometric Brownian Motion (GBM). The main reason for using this method is the characteristics of energy prices, chiefly the uncertainty of predicting their increase over time. Extended periods of low and high energy costs are both apparent. The GBM contains two important parameters: the long term average of (μ) and the degree of randomness surrounding this average (σ).

$$\delta P_{\mathrm{t}} = \mu P_{\mathrm{t}}(\mathrm{dt} = 1\mathrm{year}) + \sigma P_{\mathrm{t}} \mathrm{d}W_{\mathrm{t}} \tag{13}$$

Where P_{t} is the gas price, μ is the average trend of the gas price, σ is the randomness or volatility of gas prices, and W is the Brownian Motion. The Brownian Motion is the random part of the equation. The W is the result of using the actual continuous-time stochastic process, known also as the Wiener process. W has the standard normal distribution, $W \sim N(0,1)$. Each W is calculated using a standard random variable z by the square root of the time changes. We forecast gas prices on a yearly basis because these are usually decided annually. We calculate 50 energy price paths for each building. A higher number of energy price simulations was not possible, due to computational burdens in terms of time and the capacity of the computer. Häckel et al. [33] estimated the energy prices for 10,000 simulation runs per year. However, that analysis was performed for only one type of building. In this paper, analyses were conducted for 2,784 and 2,878 buildings. Therefore, it was not possible to create a greater number of energy price simulations.

Probability of each NPV

For deriving EUT and CPT, the probability of each NPV needs to be estimated. For this purpose, a kernel density estimator (KDE) is employed. This probability density estimator has advantages over other estimators, such as normal distribution. A density estimator is an algorithm which aims to model the probability distribution that generated a dataset. A histogram is a widely-used density estimator for one-dimensional data. The data is divided into different ranges and the number of data is calculated for each range. The advantage of KDE is that it is more precise in terms of estimation of probabilities. The histogram calculates the probabilities for a range and block of data. KDE instead estimates the probability for each point in the dataset. The results are more robust in estimating the actual data characteristics compared to a histogram density estimator [34,59,33]. For KDE, a Gaussian kernel is used, which gives more accurate results on the shape of data distribution, and the results are changed less due to the differences in sampling [2]. One input parameter of KDE is the bandwidth. The bandwidth is estimated for each house using the optimisation method, which is incorporated in the Scikit-Learn library in Python 3.[3]

Energy consumption and energy saving

In the energy module 2012 dataset, energy consumption is provided for 2784 houses and for seven years before collection of the dataset (2004–2010). The data is the official energy consumption for each house, and is not taken from a survey. The average values for energy consumption are calculated. These average values are used as indicators of energy consumption for each household. The energy module 2018 contains the energy consumption for the year 2018. Therefore, the energy consumption for the year 2018 is considered to be the reference for energy consumption of these dwellings. For energy saving, the percentages of energy saving per measure, i.e. insulation, and double glazing, are collected using reliable sources such as Milieu Centraal, and Netherlands Enterprise Agency (RVO). Similar to Häckel et al. [33], UCB = 0 is assumed, i.e., the unobserved benefits and costs compensate for each other. (See Fig. 1).

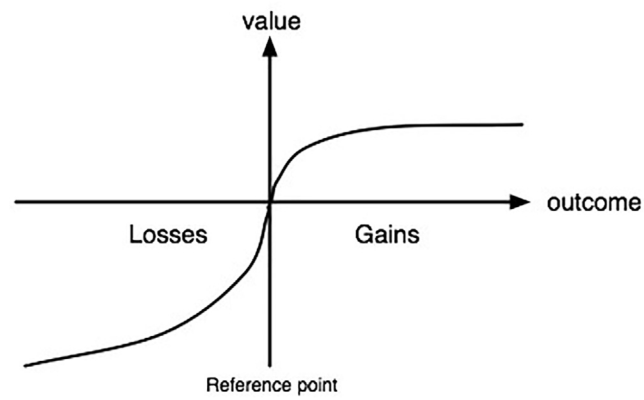


Fig. 1. Schematic presentation of the value function for cumulative prospect theory.

5. Results

5.1. Cluster analysis

To achieve more meaningful clusters, weight factor are used for different attributes based on their grey relational grades. The grey relational grade shows the relative importance of factors in determination of the outcome. Therefore, factors with higher grey relational grades should be given more weights in formation of clusters. The results of grey relational grades are presented in Table 5. For energy module 2012, building characteristics: type of single family/multi-family dwellings, year of construction, and type of heating system, show a higher correlation to decisions compared to household characteristics. Regarding household characteristics, the number of people in the households and relocation in the last two years are important factors. For energy module 2018, the type of building has the highest degree of correlation with investment decisions, followed by the household composition (i.e. one person, a family with or without children, etc.) (Tables 5 and 6).

The K-means clustering method is used to cluster similar instances in one group. Considering the number of instances and their distribution, the k-mean method is more suitable for the dataset used in this publication. The algorithm disregards the missing values, and the calculation is completely based on the actual observations. The cluster analysis is conducted from one to thirty clusters. Therefore, the population per cluster and the distribution of instances per cluster can be depicted. The elbow method is used to determine the optimal number of clusters for each data set. In the elbow method, data clustering is performed several times, in each attempted data is clustered in predefined number of clusters. Then, the sum of squared distanced of each data point from the centre of the corresponding cluster is plotted as a function the number of clusters. The resulting plot should have a shape of an arm where the number of clusters at the location of the elbow will correspond to the highest Dunn index and indicated the optimal number of clusters [80,8,46]. Using the elbow method, the number of clusters equal to 4 has the highest Dunn index for two datasets of the energy modules 2012 and 2018. Furthermore, this number of clusters contains the most appropriate population per cluster, as well as distribution of instances across different clusters. (See Fig. 2).

The clustering analysis is performed with an in–house python script. The number of observations is shown in Fig. 3. The optimal values of the squared sum of the cluster analysis is equal to 0.19 and 0.27 using the energy modules 2012 and 2018, respectively. Later, the characteristics of these clusters are investigated. The cluster analysis is an unsupervised learning process, and these clusters are grouped without connection to any type of target variable.

After clustering all data point into the optimal number of clusters, in Tables 7 and 9 the number of conducted energy retrofit measures are compared in different clusters per type of retrofit. Tables 7–10 indicate the characteristics of different clusters using the average values. It can be seen that cluster number 2 in module 2012 and cluster number 8 in module 2018 have the highest number of implemented energy retrofits. These clusters have highest average income and value of the house among other clusters. The occupants of these clusters are around 55 to 60 years old. In terms of construction year, the average year is equal to 1988 and 1969 in clusters 2 and 8, respectively. Conversely, the cluster with the lowest number of installed energy-saving measures has the lowest house values and household incomes.

This study focuses on households that with only one or no energy-saving measure for estimating the EUT and CPT. The main reason for this choice is that it is not possible to separate the effect of various energy efficiency measures in case of multiple energy retrofits.

5.2. Variables in calculating the Net Present Values (NPVs)

NVP includes the effects of all costs and benefits throughout the life time of a measure. Most of the cost is paid upfront while the benefits are accumulated later on. To be able to calculate the ben-

[2] we used sklearn package in python 3 because of its flexibility and efficiency. This package can estimate KDE in multiple dimensions with one of six kernels and one of a couple dozen distance metrics. Because KDE can be fairly computationally intensive, the Scikit-Learn estimator uses a tree-based algorithm under the hood and can trade off computation time for accuracy using the atol (absolute tolerance) and rtol (relative tolerance) parameters.

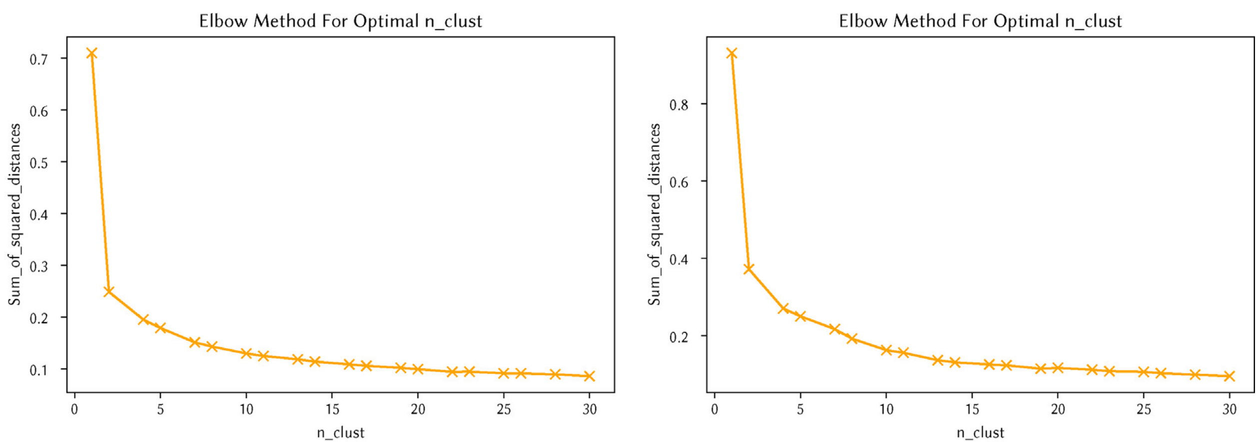
[3] GridSearchCV implements a 'fit' and a 'score' method. It also implements 'predict', 'predict_proba', 'decision_function', 'transform' and 'inverse_transform' if they are implemented in the estimator used. The parameters of the estimator used to apply these methods are optimized by crossvalidated grid-Ssearch over a parameter grid.

Table 5
Grey relational grades for each attribute using the energy module 2012.

Target variable	Number of people	Household relocated	Age	Income	Type of house (S/M)	Type of multi-family house	Construction year
Energy efficiency decision	0.6014	0.6176	0.5750	0.5203	0.6537	0.7345	0.6252
	Type of heating system	Building type, e.g. detached houses	Gas	Electricity	Number of rooms	Type of single family house	
	0.6216	0.5722	0.5323	0.5576	0.5222	0.5605	

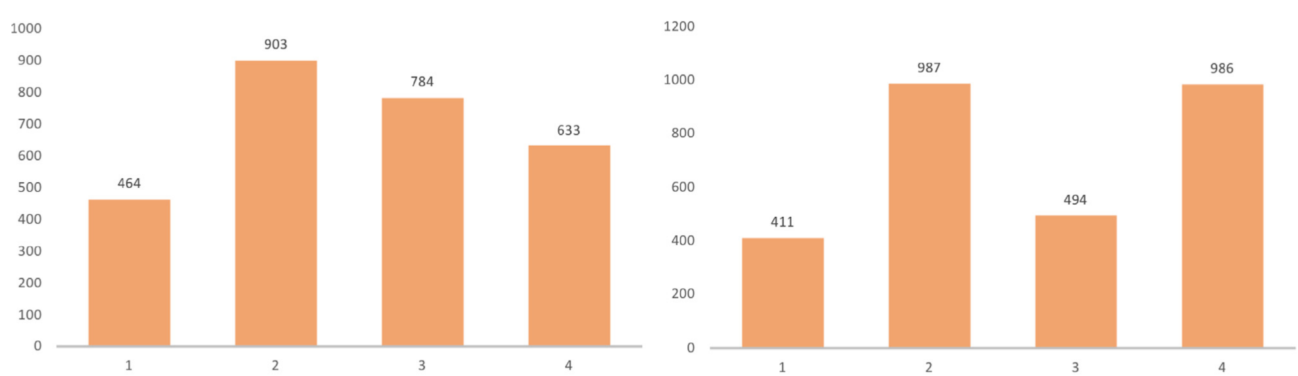
Table 6
Grey relational grades for each attribute based on the energy module 2018.

Target variable	Number of people	Household relocated	Age	Income	Type of house (S/M)	Type of multi-family house	Construction year
Energy efficiency decision	0.5479	0.5780	0.5433	0.5879	0.5631	0.6964	0.5384
Household composition	Building type	Gas	Electricity	Number of rooms	Type of single family house		
	0.6905	0.5507	0.5052	0.5235	0.5378	0.6044	



(a) The energy module 2012 (b) The energy module 2018

Fig. 2. Elbow Method for the optimal number of clusters.



(a) The energy module 2012 (b) The energy module 2018

Fig. 3. Population per cluster.

Table 7
Different characteristics of four clusters - National Household Survey energy module 2012.

Cluster	Insulation	Boiler	Double glazing	PV_panel	Number_energy_saving	House_value	Income
1	25	104	68	1	198	217,445	47,708
2	160	329	223	38	750	353,921	65,922
3	112	265	205	14	596	262,960	56,292
4	136	188	171	15	510	240,183	58,932

Table 8
Different characteristics of four clusters - the energy module 2012.

Cluster	Number_of_people	building year	EI	Type of house	Number_of_rooms	Relocated	Age
1	1.7	1994	1.81	multi_family	3.4	0.153	54.5
2	2.6	1988	1.84	single_family	5.4	0.016	59.8
3	2.3	1984	1.89	single_family	4.9	0.005	60.7
4	3.3	2004	1.71	single_family	4.9	0.210	38.0

Table 9
Different characteristics of four clusters - the energy module 2018.

Cluster	Insulation	Boiler	Double glazing	PV_panel	Number_energy_saving	House_value	Income
5	79	121	79	67	346	256,437	43,182
6	232	311	247	223	1013	263,199	69,418
7	72	136	72	19	299	257,196	58,097
8	258	317	243	285	1103	370,734	84,051

Table 10
Different characteristics of four clusters - the energy module 2018.

Cluster	Number_of_people	building year	EI	Type of house	Number_of_rooms	Relocated	Age
5	1.08	1964	1.71	single_family	4.7	0.087	56
6	2.79	1972	1.66	single_family	5.1	0.129	48.5
7	1.64	1959	1.68	multi_family	3.5	0.224	51.5
8	2.6	1969	1.6	single_family	5.7	0.068	55

efits of each energy efficiency measure in terms of monetary savings by consuming less energy, the energy consumption in case of no energy saving measure, the expected percentages of energy saving per type of energy-saving measure, and the future energy prices must be known.

5.2.1. Predicting gas prices

Knowing the long-term price trend μ , and price volatility σ, different possible trajectories can be estimated gas price in the future using geometric Brownian motion. The values for the μ = 5.7%, and σ = 17% are obtained from historic data from the Eurostat dataset. This study aims for the long-term evaluation of investment decisions; therefore, energy prices are estimated for 30 years. Gas prices are simulated 50 times per year. Fig. 4 shows 10 examples of gas price predictions over 30 years. Given bi-yearly data, 60 instances are generated to compute data for 30 years. (See Fig. 5).

5.2.2. Energy-saving measures

Table 11 shows the total numbers and percentages of installed insulation and double glazing for houses with one energy saving measure. To calculate the energy saving for each measure, the expected energy saving is required. The expected energy savings are collected from reliable data sources such as the Netherlands Environmental centre and the Netherlands Enterprise Agency (RVO). From this source, the expected energy savings of insulation and double glazing are 12% and 14%, respectively. These percentages are multiplied by average gas consumption to calculate the average energy saving per type of energy-saving measure. (See Table 12).

5.2.3. Energy efficiency investment

In the dataset, households provide information on the investment costs of energy-saving measures. The information is not available for all households. The average investment cost per cluster of households is computed in place of the missing household cost. The outliers, such as investment costs less than 100 euros, were removed. Previous studies identify high upfront costs as one of the main barriers to conducting energy retrofits [21,39,33]. Therefore, the investment costs, including subsidies and no-subsidies, are used to test the importance of these influencing factors. Another reason for calculating with both subsidies and no-subsidies is to examine the cognitive bias of the reference dependence.

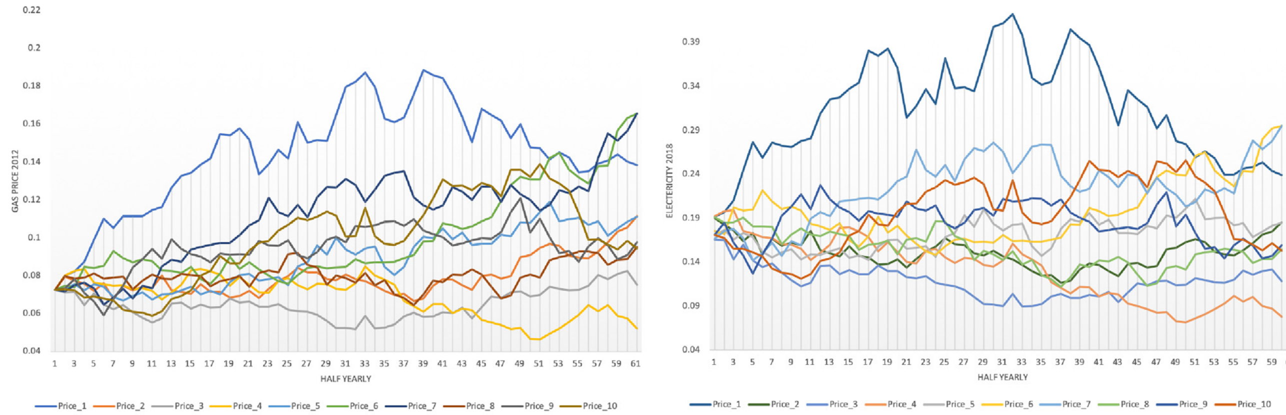
5.3. Expected utility theory (EUT) and cumulative prospect theory (CPT)

This study aims to estimate the input parameters of the EUT and CPT models to more accurately predict individual energy efficiency investment decisions. First, the initial input parameters are applied following previous studies [49,33,71,64]. The EUT and CPT parameters are calculated using the Eqs. (1–10) and (11–17), respectively. The initial values are presented in Table 13.

The EUT and CPT parameters are calculated for 2,779 and 2,878 homeowners of the two datasets using the initial parameter values. The results show that the CPT predicts the decisions of 86% of homeowners correctly. However, the EUT overestimates decisions and shows a positive value for 1441 homeowners who did not invest in any type of energy-saving measure. In the following subsections, the parameters of EUT and CPT per each cluster are identified. The differences between the predicted and actual percentages of households that made energy retrofits are minimised to estimate these parameters.

5.4. Identification of the parameters of expected utility theory (EUT) and cumulative prospect theory (CPT)

Genetic algorithm is used to estimate EUT and CPT parameters. The goal of this optimisation is to minimise the deviation between the retrofit rate estimated by EUT and CPT and actual retrofit rate obtained from the data set by changing the parameters. Estimating the values of all these parameters necessitates high computational times resulting from the complexity and non-linearity of the CPT parameters as well as the need to calculate 50 different energy price scenarios. The values of $\beta, \theta, \gamma, \delta$, and λ are calculated. To reduce the computational times, boundaries are defined for these parameters. The possible range for different parameters were identified using trial and error to narrow down the domain space. The



(a) the Netherlands Household Survey Energy module 2012 (b) the Netherlands Household Survey Energy module 2018

Fig. 4. Population per cluster.

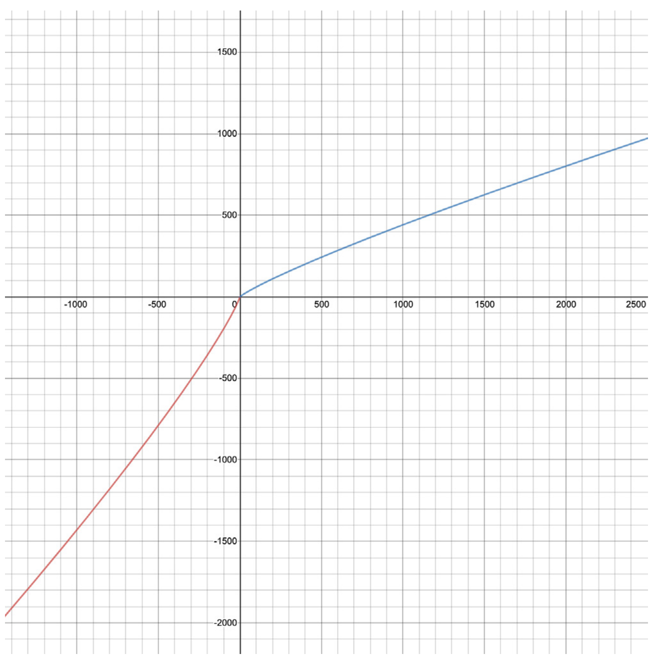


Fig. 5. the value function of CPT for cluster 3, installation of double-glazing, and the investment costs with subsidies.

maximum number of generations is constrained as well. An initial population of 100 with 10 generations appear to be sufficient to achieve sufficient results. increasing the number of generation above 10 and population size about 100 resulted in minimal improvement in the goodness of the objective function, therefore, these values are selected. The optimisations are conducted for four clusters of households and two types of energy-saving measures, i.e. insulation and double-glazing. The percentage of households that made positive decisions to conduct energy-saving measures is the indicator in estimating the parameters for each cluster and per type of energy-saving measure. Two reference points (RPs) are used to estimate the parameters and to examine the reference dependency of preferences proposed by Tversky and Kahneman [71] ($NPV = I_0 + \sum_{n=0}^{T} \frac{CF_t}{(1+r)^t} + RP$). Following this formula, the investment costs can be decreased or increased to include different reference points. Here, the investment costs with subsidies and without subsidies are included to evaluate the cognitive bias of reference dependence. The subsidy data is collected from the Milieu Centraal website. Tables 14 and 15 show the optimisation parameters for EUT and CPT models using the energy modules 2012 and 2018.

5.4.1. Comparing the expected utility theory (EUT) and cumulative prospect theory parameters (CPT) and overall results of CPT parameters

As mentioned, the EUT has a utility function that includes the wealth level and shows the constant relative risk aversion (CRRA) for each individual. The values of CRRA (i.e. θ) are calculated using the genetic algorithm method. The objective function is similar to CPT. Therefore, the differences between the actual percentages and the predicted ones are minimised. Based on the results and similar to previous studies [64,23,33,71], CPT captures the actual behaviours reasonably well for the majority of the clusters, for two energy-saving measures, as well as, two reference points with and without subsidies. As mentioned by Rieger et al. [64], the CPT's more accurate parameters are not solely due to the fact that CPT contains a large number of parameters. The main reason for better performance of CPT in comparison with EUT in predicting the rate of retrofit is due to inclusion of reflection effect and probability weighting. The reflection effect is another term for risk-seeking behaviour for loss and risk-averse behaviour for gain. Probability weighting refers to the fact that the demand of high gain to repay

Table 11
Total numbers and percentages of energy-saving measures per cluster from the energy module 2012 and 2018 (households with one energy-saving measure).

2012					2018				
	Insulation		Double glazing			Insulation		Double glazing	
name	numbers	percentages	numbers	percentages	name	numbers	percentages	numbers	percentages
cluster 1	2	0.43	26	5.60	cluster 5	24	5.84	21	5.11
cluster 2	24	2.66	67	7.42	cluster 6	57	5.77	68	6.89
cluster 3	24	3.06	74	9.44	cluster 7	22	4.45	24	4.86
cluster 4	25	3.95	50	7.90	cluster 8	55	5.58	40	4.06

Table 12
The average investment costs for different types of energy-saving measures and per cluster of households (Euro)-the energy modules 2012 and 2018

2012			2018		
Cluster	Insulation	Double-glazing	Cluster	Insulation	Double-glazing
cluster 1	1875	2750	cluster 5	2589	2061
cluster 2	1250	1200	cluster 6	2277	2670
cluster 3	1950	3500	cluster 7	1300	2000
cluster 4	2500	2267	cluster 8	2358	3111

Table 13
Initial values for input parameters.

Input parameters	θ	α	β	γ	δ	λ
Values	1.26	0.88	0.88	0.61	0.69	2.25

Table 14
Cluster-level CPT estimates and Mean Error using the energy module 2012.

	CPT												EUT	
Reference	Subsidies						No_subsidies						Subsidies	
	β	γ	δ	λ	θ	dev	β	γ	δ	λ	θ	dev	θ	dev
							Insulation							
cluster 1	0.33	0.71	0.71	1.37	9.94	0	0.24	0.92	0.07	1.81	3.53	2.48	1.84	21.43
cluster 2	0.68	0.11	0.05	4.85	7.28	4.23	0.21	0.45	0.46	6.35	0.32	2.08	26.59	3.28
cluster 3	0.28	0.34	0.41	2.86	8.16	12.46	1.31	0.87	0.56	8.77	8.56	6.1	2.50	32.13
cluster 4	1.11	0.01	0.31	0.52	7.82	9.19	0.76	0.78	0.03	7.12	9.42	3.67	2.48	37.80
							Double-glazing							
cluster 1	0.26	0.57	0.39	4.45	0.78	0.29	0.53	0.93	0.24	3.22	0.02	10.43	0.78	5.79
cluster 2	0.03	0.31	0.71	2.26	1.60	0.43	0.06	0.54	0.24	7.05	0.21	3.00	2.49	17.17
cluster 3	0.86	0.10	0.48	3.76	0.14	0.61	0.04	0.44	0.41	6.91	0.65	0.61	3.08	19.94
cluster 4	0.30	0.55	0.91	6.44	1.33	1.07	0.38	0.53	0.02	7.61	7.54	4.64	3.06	21.35

Table 15
Cluster-level CPT estimates and Mean Error using the energy module 2018.

	CPT												EUT	
	Subsidies						No_subsidies						Subsidies	
	β	γ	δ	λ	θ	dev	β	γ	δ	λ	θ	dev	θ	dev
							Insulation							
cluster 5	1.53	0.60	0.83	0.29	1.12	0.0	0.36	0.35	0.14	7.26	0.31	0.93	0.62	21.49
cluster 6	1.03	0.98	0.61	5.16	0.35	1.78	0.91	0.68	0.03	7.55	6.05	0.25	1.02	16.01
cluster 7	1.40	0.85	0.03	3.35	0.91	0.33	0.02	0.94	0.06	4.29	7.11	6.62	1.90	3.64
cluster 8	0.76	0.77	0.18	3.27	3.51	2.68	0.41	0.95	0.47	4.13	0.28	0	0.42	28.69
							Double-glazing							
cluster 5	1.14	0.96	0.70	4.62	0.35	1.78	0.05	0.91	0.45	6.24	0.63	1.78	1.37	16.11
cluster 6	0.11	0.51	0.67	9.54	0.70	0.24	0.20	0.45	0.23	5.73	0.21	9.11	0.79	8.87
cluster 7	0.37	0.64	0.43	0.65	0.30	1.97	0.37	0.71	0.21	4.28	0.25	7.57	1.89	3.29
cluster 8	0.01	0.71	0.52	2.34	6.45	0.56	0.64	0.83	0.06	4.20	8.41	0.28	1.32	22.35

loss with equal chances cannot be described by a reasonable degree of risk aversion in the EUT model [63].

Evaluating the parameters of CPT determines several interesting results for the overall clusters as shown in Table 16:

Next, the parameters of CPT models are interpreted per cluster and per type of energy-saving measure, i.e. insulation and double-glazing, using the energy modules 2012 and 2018.

5.4.2. Results of the CPT's parameters for each cluster using the Netherlands Household Survey Energy module 2012

The parameters of different clusters using the energy module 2012 are explained in the following:

Insulation—subsidies Cluster 2 shows high risk-seeking behaviour (β=0.68) of this group in losses compared to other clusters (similar to the proposed parameter value by [71]). Furthermore, the loss-aversion parameter is the highest as well (λ=4.85). Therefore, the individuals in this cluster prefer to accept higher risk to prevent future losses or to gain in future compare with other clusters. this results in higher rate for implementation of energy retrofits in this cluster compared to others (Tables 7 and 8). The level of average incomes and house values are the highest, as well. In contrast, cluster 4 shows the least risk-seeking behaviour in losses for insulation (β=1.11). This group is the youngest with newly constructed buildings (average year of construction of 2004). Apparently, people with higher income and more expensive houses are more likely to invest in energy retrofits compared to the young and low income group.

Clusters 2 and 4 have the lowest values for the parameters γ and δ, respectively. Therefore, cluster 2 (δ=0.05) would be highly concerned regarding the low probability of losses. In contrast, cluster 4 (γ=0.01) is more concerned about the small probability of gains with the second-highest number of installation of insulation (Tables 7 and 8). Clusters 2 and 4 overweight the small probability of losses and gains, respectively. Cluster 1 has the highest and

Table 16
Overall interpretation of CPT parameters.

Parameters	Interpretations
$\gamma, \delta, \lambda, \theta$	Changing the reference point has a statistically significantly influence on the coefficient values of γ and δ. No-significance impacts are identified for the coefficients of λ, θ (t-test, P$\leqslant$0).
β	The parameter β (i.e. $0 < \beta \leqslant 1$) shows the convexity of the value functions in 81% of CPT parameter estimations. This indicates the risk-seeking behaviours for loss on energy efficiency investments.
θ	The parameter of θ (i.e. $\geqslant 0$) indicates the concavity of the value functions in all cases. Namely, it shows the risk-averse behaviours of individuals in gain regarding energy efficiency investments.
β, θ	The average β (0.52) is smaller than θ (3.25). The differences between these two parameters are statistically significant. Therefore, asymmetric value functions in gain and loss are identified. It means that in loss, the risk-seeking behaviour increase considerably with more losses. However, the rate of risk-aversion behaviour does not increase considerably with more gains from energy efficiency investments (Fig. 5).
γ and δ	Studies in decision making indicate that decision-makers do not weight rare events according to their actual probability chances of happening. Instead, small probability events are inclined to be overweighted for two reasons: (1) decision-makers may overestimate the chance that rare events happen; and (2) small probabilities are overweighted in terms of their impact on decisions. Considering these two reasons, rare events are given greater psychological weight in our minds than actual weight [12,71]. The obtained values for γ and δ are between 0 and 1. Based on these values, the overweighting of small probabilities is confirmed for both gain and loss.
λ	The estimated values for λ for almost all and except for two cases are greater than 1. This indicates that the majority of homeowners consider losses more important compared to gains. For instance, the maximum value of λ is equal to 9.54. Therefore, on average, the individuals of this cluster perceive losses almost 10 times more important than gains.
Reflection principle	The results of the CPT parameters are tested by the reflection principle. Reflection principle means that there should be no correlations between the estimated parameters of CPT modelas stated by Kahneman [41]. The bi-variate correlations of the five CPT parameters are tested using the Pearson correlation and Superman's rho tests. The results show that no-significant correlations are identified. Therefore, the interpretation of coefficients using the CPT model is not influenced, and these coefficients can be interpreted independently.

equal values of the parameters $\gamma = 0.71$ and $\delta = 0.71$. This group overweights the small probability of losses and gains less than other clusters and behaves consistently in the domain of losses and gains. Cluster 1 contains multi-family dwellings with less than two household members. This result indicates when multiple families are living in the same building they will probably make decisions collectively. Therefore, they will cancel out individual tendencies to overweight probability of losses or gains.

Double-glazing—subsidies The parameters of the CPT model for cluster 3 are estimated similarly to the original parameter values (i.e. β=0.86 and θ=0.86) proposed by Kahneman and Tversky [42]. The individuals of this cluster are less risk averse in gains but risk seeking in losses for installation of double-glazing. In addition, cluster 3 has the lowest γ (=0.10) and mean error (0.14) compared to other clusters. This could be due to the fact that buildings in this cluster are mainly from the construction period of 1984 and households are about 61 year old. Therefore, most of the houses are lacking double glazing in the original construction and as elderly people have a preference for higher indoor temperatures they are likely to install double glazing. Cluster 4 has the highest loss-aversion coefficient for the installation of double-glazing (low number of double-glazing installations). Cluster 1 has the lowest δ (=0.39) for double-glazing, which means that this group would be highly concerned about the small probability of losses. This cluster also has the lowest number of installations of double-glazing compared to other groups.

Insulation—no-subsidies The reference point is changed in the NPV formula by including the investment costs of insulation without subsidies. Different values for five parameters of CPT are identified. In this new optimisation, the lowest mean error is calculated for cluster 2, similar to the RP with subsidies. The inclusion of subsidies makes the proposed optimisation model more realistic as subsidies could have significant effect in decision towards implementation of energy efficiency measures. As excepted inclusion of subsidies in the reference point improved the goodness of the objective function 50% in almost all cases. The highest values of β equal to 1.31 and 0.76 are calculated for clusters 3 and 4. Therefore, these two groups are risk seeking in the domain of losses for insulation. For these two groups, the values of loss aversion parameters are the highest, as well. Namely, clusters 3 and 4 consider losses more than 8.77 and 7.12 as important as gains.

Double-glazing—no-subsidies Compared to the insulation, the mean errors are increased when the RP includes the investment costs without subsidies. In addition, three θs (out of four) are $0 \leqslant \theta \leqslant 1$. This range of θ makes the value functions of CPT more comparable to the one proposed by Kahneman and Tversky [42]. The results for the parameter γ are not changed in terms of interpretation. However, the order of magnitude has increased.

5.4.3. Results of CPT's parameters for each cluster using the energy module 2018

The results of CPT parameters for different clusters are explained using the Netherlands National Household Survey Energy module 2018.

Insulation—subsidies The mean errors are smaller using the energy module 2018. The parameter β shows risk-seeking behaviour and risk aversion in gains. For all clusters, the θ follows the original curvature proposed by Kahneman and Tversky [42]. By comparing the parameter β, cluster 8 has the lowest β_8=0.76. This indicates the highest risk-seeking behaviours of cluster 8 in losses for insulation compared to other clusters. Cluster 6 has the second-lowest value for parameter β. In terms of dwelling and household characteristics, cluster 8 has the highest number of installed insulation, highest average income, and average house values compared to the other clusters. Furthermore, cluster 6 ranks second for all these attributes. Furthermore, cluster 6 has the lowest value for θ, which means this group has the least risk aversion to gains. Considering the household characteristics (Table 10), cluster 6 has the youngest average compared to other clusters. Regarding the parameter γ, the value of this parameter for cluster 6 is the highest. Therefore, this group would be less concerned about the small probability of gains for insulation. At the same time, the loss-aversion coefficient (λ=5.16) is the highest compared to other clusters. Furthermore, cluster 6 also has the lowest value of θ. Hence, this group is the lowest risk-averse in gains. Based on Tables 9 and 10, cluster 6 ranks second in terms of insulation, average income, and house value compared to other clusters. Clusters 7 and 8 have the lowest δ_7=0.03 and δ_8=0.18, respectively. Therefore, these groups would be highly concerned about the small probability of losses.

Double-glazing—subsidies For both datasets, double-glazing with investment costs including subsidies has the lowest mean error compared to other combinations of investment costs with no subsidies. Regarding the parameter results, clusters 6, 7, and 8 are risk-seeking in preventing losses. Based on the values of θ, clusters 5, 6, and 7 confirm the curvature proposed by Kahneman and Tversky

[42]. Among these clusters, cluster 7 with $\theta = 0.30$ is less risk-averse in gains. This cluster belongs to multi-family dwellings with the highest probability of relocation over the last two years compared to other clusters. The highest value of parameter λ is equal to 9.54, which implies that individuals of cluster 6 perceive losses to be more than 9 times as important as gains. Based on Table 9, this group's number of double-glazing installations (=247) is the highest compared to other clusters. Regarding the probability weighting function parameters, cluster 5 has the highest values of $\gamma = 0.96$ and $\delta = 0.70$, which indicates that these individuals are less concerned about the small probability of losses and gains from double-glazing installation. This cluster has the lowest number of double-glazing installations in the past 5 years (Table 9).

Insulation-no—subsidies The mean errors of clusters 5,6, and 8 are declined when the reference point is set to the investment costs without subsidies. In this case, cluster 6 and 5 have the highest loss aversion parameters of $\lambda = 7.55$ and $\lambda = 7.26$ compared to other clusters, respectively. The values of θ for clusters 5 and 8 are equal to 0.31 and 0.28, respectively. This implies that these groups tend to be risk averse in the gains for installation of insulation. Cluster 6 has the highest $\beta = 0.91$ and $\lambda = 7.55$ values, which indicates less risk seeking in losses and high loss aversion in insulation installation. Based on Table 10, this cluster ranks second in insulation installation. The highest values of parameter γ are identified for clusters 7 and 8 (0.94 and 0.95, respectively). In addition to this, cluster 8 has the highest values of δ, which indicates these people would be less concerned about the small probability of losses. This high-income group of households has the highest rate of installed insulation in the past five years (Table 9). *Double-glazing—no-subsidies* Similar to the results using the energy module 2012, the mean errors are increased for clusters 6 and 7 compared to the previous RP. Again, the values of the parameter θ of cluster 5, 6, and 7 are $0 \leqslant \theta \leqslant 1$, which confirm the proposed shape of value functions by Kahneman and Tversky [42]. The parameter values of β for all clusters conform to the proposed convexity as well (risk-seeking for loss). Cluster 8 has the lowest parameter value of $\delta = 0.06$, which implies this group would be highly concerned about small probabilities of loss. Based on Table 10, this group has the highest average income, highest average house value, and the second most installations of double-glazing.

6. Discussion

This study has applied quantitative methods to examine the impacts of cognitive biases on energy efficiency investment decisions. It has compared expected utility theory (EUT) and cumulative prospect theory (CPT) and evaluating their potential to predict and explain decision-makers' behaviours. EUT assumes rational decision-making under risk, an assumption CPT disputes by explaining actual behaviours. According to CPT, decision-makers display different cognitive biases: reference dependency, loss aversion, diminishing sensitivity, and probability weighting. These agents generally behave asymmetrically, to their loss and gain.

6.1. Comparing the performance of expected utility theory (EUT) and cumulative prospect theory (CPT) in predicting the decision-makers behaviours

This study has demonstrated CPT's superiority in explaining renovators' decision-making behaviours. EUT is useful in a minority of cases, but CPT can predict these cases as well [64,23,15], while also providing deeper insights into qualitative and quantitative studies on energy efficiency [33,30,25,28,77]. CPT's explanatory strength derives from its consideration of cognitive biases, such as the reflection/framing effect (i.e. risk-seeking in loss and risk aversion in gain), probability weighting (over/underweighting small/average probabilities), and loss aversion [64,33,15]. The evidence of CPT's superiority is as follows:

1. Using the initial parameters from previous research [49,33], EUT overestimated the actual decisions of approximately 50% of homeowners. However, CPT predicted the decisions of 86% of individual homeowners accurately.
2. CPT determined homeowners' risk-seeking behaviours for losses for 81% of the total number of groups. Furthermore, homeowners' risk-averse behaviours of homeowners were identified for all cases, confirming the function proposed by Tversky and Kahneman [71]. Furthermore, CPT verified overweighting of small probabilities based on estimations for the corresponding parameters (i.e. γ and δ). For the majority of the groups, the loss aversion factors were considerably greater than those proposed by Tversky and Kahneman [71].
3. CPT's mean errors were in many cases smaller than EUT's.

Cognitive bias of reference dependence

Modifying the reference points significantly influences the parameter values of γ and δ. Therefore, the results of this study are somewhat similar to the quantitative study by Häckel et al. [33], which found the determination of a reference point to be very important. Reference dependence's importance is stated in previous qualitative studies on energy efficiency (e.g. [25,28,77]). For cases involving insulation, using the reference point of no-subsidies slightly improved CPT's results. The reverse was identified for cases involving double-glazing; namely, the reference point including subsidies resulted in closer predictions of homeowners' actual behaviours. The conclusions were the same using both datasets.

6.2. Identifying and comparing cognitive biases for different groups regarding the installation of insulation and double-glazing

This study mainly contributes to extant knowledge by empirically examining CPT's parameters for each group of homeowners. The cognitive biases of reference dependency, loss aversion, diminishing sensitivity, and probability weighting were quantified for four clusters of homeowners, using household and building characteristics, as well as the probability of relocation in the past two years. The energy retrofits of insulation and double-glazing were also investigated. Rieger et al. [64] estimated the CPT parameters for different groups using an international survey of 53 countries. The author investigated the risk preferences of a large number of undergraduate students, as shown through hypothetical choices in a predefined set of lotteries. This is the first study in the field of energy efficiency that empirically investigates CPT parameters for clusters of homeowners.

Table 17 presents the main findings of this study. The main clusters' highlights are illustrated for different types of energy saving measures, using different reference points using the energy modules 2012 and 2018. Based on the investigation of CPT parameters, the study identified the importance of risk-seeking in loss, concern about small probabilities and loss aversion factors:

1. For insulation, the households that invested more in installing insulation were also more risk-seeking in loss and highly concerned about the small probabilities of loss and gain, as well as being highly loss-averse. These households often had the highest average income and house values. For cluster 4 of energy module 2012, this group was the youngest and had

Table 17
Identifying and comparing CPT's parameters among different clusters and energy saving measures (i.e. Insulation and Double-glazing).

	2012		2018	
	CPT parameters	household and building characteristics	CPT parameters	household and building characteristics
insulation-subsidies	Cluster 2 and 4: (1) highest values for the β, i.e. the highest value for risk-seeking behaviour in loss, (2) lowest values for γ and δ, i.e. highly concerned about small probabilities of loss and gain - cluster 2: has the highest loss aversion parameter = 4.85	Highest number of installed insulation, single-family, highest number of rooms highest average income and highest average house value cluster 4: higher probability of relocation in the past 2 years	cluster 8: (1) highest value of $0 \leqslant \beta \leqslant 1$ proposed by [71], other clusters have β values more than 1. (2) the lowest δ, i.e. highly concerned about small probabilities of loss cluster 6: highest loss aversion parameter = 5.16, highest value of γ, i.e., least risk aversion in gain	the highest number of installed insulation, highest income and house values single-family highest number of room average construction year: 1970s
double-glazing-subsidies	Cluster 3: (1) highest value of β, i.e. highest values for risk-seeking behaviour, (2) the lowest value for γ, i.e. highly concerned about small probabilities of gain			
	Second ranking of installed double-glazing, second-ranking of house value average construction-year: 1984	cluster 6: highest loss aversion parameter = 9.54	the highest number of installed double-glazing, highest income and house values	
	Cluster 1: the highest value of γ, not highly concerned about small probabilities of gain	the lowest ranking of installed double-glazing	cluster 3: lowest loss aversion parameter	the lowest number of installed double-glazing, highest income and house values
insulation-no subsidies	- cluster 2: (1) high loss aversion = 6.35, (2) the value of γ is less than δ, which is exactly following the pattern proposed by [71]. cluster 2 concerned more about the small probabilities of loss compared to gain. - cluster 4: results remain the same: (1) highest $0 \leqslant \beta \leq 1$, i.e. risk-seeking behaviour in loss, (2) high loss aversion = 7.12.	- highest number of installed insulation, income and house value - second highest number of installed insulation, income, and house value	cluster 6: results remain the same: (1) highest β, i.e. risk seeking behaviour in loss, (2) highest loss aversion = 7.55	the second-highest number of installed insulation, income, and house value
double-glazing- no subsidies	cluster 2 and 3: (1) high loss aversion parameters (7.05 and 6.91, respectively), cluster 3: (2) more risk-averse in gain	cluster 2 and 3: first and second-ranking of the number of installed double-glazing cluster 2: highest income, house value cluster 3: highest house value	cluster 8: (1) highest value of $0<\beta<1$ proposed by [71]. (2) lowest δ, i.e. highly concerned about small probabilities of loss cluster 6: (1) highest loss aversion parameter = 5.73, (2) lowest value of δ=0.23, i.e. highly concerned about small probabilities of losses	the highest number of installed double- glazing, highest income and house values

Table 18
Behavioural interventions using the identified cognitive biases.

Cognitive bias	Behavioral interventions
Loss aversion	The loss-averse people were recognised in different clusters. Furthermore, the groups of homeowners with the most installed energy saving measures were often risk seekers in loss. For these people who are highly loss-averse, highlighting the cost/loss reductions from using energy retrofits can be more effective than emphasising the benefits of energy saving measures, as also mentioned by Frederiks et al. [25].
Risk averse in gain	The majority of homeowners were risk-averse in gain. That is, they would rather engage in a risky behaviour to avoid certain loss than to engage in a similarly risky behaviour to obtain a comparable gain [71]. Promoting low-risk and secure energy retrofit might be more persuasive for risk-averse homeowners [25].
Social influence	Multi-family dwellings often invest less compared to single-family dwellings. This might be due to barriers preventing an agreement among the homeowners for conducting energy retrofits. This group can be motivated by other people's attitudes; for instance, a trusted, well-informed neighbour can explain the benefits of energy efficiency retrofits to other neighbours successfully. Furthermore, formulating energy retrofits as a socially desirable behaviour can increase the probability of other people conducting retrofits. The importance of these behavioural interventions are specified by [24,25].

the highest probability of relocation. The results remained the same upon modifying the reference point. In this case, the groups of people who invested more were extremely loss-averse as well.

2. For double-glazing, the highly loss-averse people had invested more in double-glazing than other groups. For cluster 8 of energy module 2018, homeowners who invested the highest amount in double-glazing were risk-seekers in loss and highly concerned about the small probabilities of loss. These groups always had high house values and were often the highest average income group. For instance, cluster 3 had the second highest house values. However, in terms of income, they were ranked third.

6.3. Insights for behavioural interventions using the identified cognitive biases

This study did not evaluate the impact of behavioural intervention. However, based on the results, potential behavioural interventions can be identified. Table 17 presents the cognitive biases in energy retrofit decisions regarding insulation and double-glazing for homeowners. Using this table and the results of Sections 5.4.2 and 5.4.3., the identified cognitive biases and potential behavioural interventions can be proposed as presented in Table 18.

7. Conclusion and policy implications

The current study contributes to the identification of cognitive biases in energy retrofits by developing a theoretical framework and conducting empirical analyses of homeowners' retrofit decisions in the Netherlands. From the theoretical perspective, models and approaches incorporating specific cognitive biases have been presented, including prospect theory as proposed by Tversky and Kahneman [71] (considering the cognitive biases of isolation effect, certainty effect, endowment effect, and reflection effect); and the theory of moral sentiments proposed by Smith [67] (social norm, social approval and status). The study has also reviewed current studies on cognitive biases to identify the main known cognitive biases in the field of energy efficiency. The identified cognitive biases include status quo bias/default setting, loss aversion, risk aversion, availability bias, and sunk cost fallacy.

This study compared expected utility theory (EUT), which assumes a rational decision-maker, to cumulative prospect theory (CPT), which assumes the influence of risk and uncertainty, when studying homeowners' retrofit decisions. EUT as developed by Von Neumann and Morgenstern [73] and CPT as proposed by Tversky and Kahneman [71] were presented. For the empirical application, CPT was applied for investigating cognitive biases in homeowners' energy retrofits. These cognitive biases included: (a) reference dependence/status quo bias/default setting (b) diminishing sensitivity/reflection effect/framing effect/certainty effect (i.e. different behaviours for gains vs. losses), (c) loss aversion, and (d) probability weighting. Furthermore, the cognitive biases were investigated for four homogeneous groups of individuals, as well as two types of energy retrofits, i.e. insulation and double-glazing. The differences and similarities of cognitive biases for different groups were then evaluated. Finally, potential behavioural interventions for each cluster of individuals biases were proposed.

Overall, there is evidence of reference dependency, reflection effect, loss aversion, and probability weighting. CPT was considerably better than EUT at predicting the energy efficiency decision behaviours for four clusters and two types of energy retrofits (insulation and double-glazing). Furthermore, changing the reference points significantly influenced the parameter values of the probability weighting function (i.e. γ and δ). This indicates the importance of status quo bias in individuals' decision-making. For the reflection effect, individuals' risk-seeking in losses and risk aversion in gains were also identified as significant. Furthermore, diminishing sensitivity in losses was less compared to gains, since the average of β (for negative outcomes) was less than θ (for positive outcomes). Based on CPT, people overweigh the small probability of both gains and losses. For this purpose, the corresponding parameters of CPT were $0 < \gamma, \delta \leqslant 1$. These ranges of parameter values were estimated for four clusters of individuals. Finally, people prevented losses significantly. The maximum loss aversion parameter is equal to 9.54 for energy saving investment, which is almost 5 times more than the estimated value by Tversky and Kahneman [71].

The groups with highest average income and house values in the National Household Surveys of 2012 and 2018 showed highest risk-seeking parameter for losses. These groups installed the highest number of insulation and overweighted the small probabilities of losses. In data for both years, the youngest group of individuals were among the least risk-averse in gains. However, the correlation between age and installation of insulation is not clear. The average income and house values are significantly more important in determining the decision towards installing insulation. For double-glazing, similar conclusions could be drawn using the 2012 and 2018 datasets (cluster 3 of 2012, cluster 6 of 2018). The risk-seeker individuals for losses, who also overweighted the small probabilities of losses/gains more than other groups, installed more double-glazing compared to other groups (second-ranking). In 2018 dataset, the cluster with the lowest amount of installed double-glazing, showed less risk-seeking behaviours in losses. Similar conclusions are achieved for the reference point no-subsidies. Our findings show that cost/loss reductions for installing energy retrofits can be more effective, compared to promoting energy retrofits by their advantages and benefits.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This project is financially supported by the MMIP 3–4 scheme of the Ministry of Economic Affairs & Climate and the Ministry of the Interior & Kingdom Relations (In Dutch: Dit project wordt uitgevoerd met ondersteuning vanuit de MMIP 3–4 regeling van het Ministerie van Economische Zaken & Klimaat en het Ministerie van Binnenlandse Zaken & Koninkrijksrelaties.).

References

[1] Abrahamse, W., Schuitema, G., 2020. Psychology and energy conservation: Contributions from theory and practice. In: Energy and Behaviour. Elsevier, pp. 19–44..

[2] W. Abrahamse, L. Steg, C. Vlek, T. Rothengatter, A review of intervention studies aimed at household energy conservation, Journal of environmental psychology 25 (3) (2005) 273–291.

[3] M.I. Abreu, R.A. de Oliveira, J. Lopes, Younger vs. older homeowners in building energy-related renovations: Learning from the portuguese case, Energy Reports 6 (2020) 159–164.

[4] M. Ashouri, F. Haghighat, B.C. Fung, A. Lazrak, H. Yoshino, Development of building energy saving advisory: A data mining approach, Energy and Buildings 172 (2018) 139–151.

[5] I. Ayres, S. Raseman, A. Shih, Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage, The Journal of Law, Economics, and Organization 29 (5) (2013) 992–1022.

[6] Baginski, J.P., Weber, C., 2017. A Consumer Decision-Making Process? Unfolding Energy Efficiency Decisions of German Owner-Occupiers. HEMF Working Paper 08..

[7] N.C. Barberis, Thirty years of prospect theory in economics: A review and assessment, Journal of Economic Perspectives 27 (1) (2013) 173–196.

[8] P. Bholowalia, A. Kumar, Ebk-means: A clustering technique based on elbow method and k-means in wsn, International Journal of Computer Applications 105 (9) (2014).

[9] F. Bonakdar, A. Dodoo, L. Gustavsson, Cost-optimum analysis of building fabric renovation in a swedish multi-story residential building, Energy and buildings 84 (2014) 662–673.

[10] G. Bravo, G. Pardalis, K. Mahapatra, B. Mainali, Physical vs. aesthetic renovations: Learning from swedish house owners, Buildings 9 (1) (2019) 12.

[11] N.W. Brown, T. Malmqvist, W. Bai, M. Molinari, Sustainability assessment of renovation packages for increased energy efficiency for multi-family buildings in sweden, Building and Environment 61 (2013) 140–148.

[12] Burns, Z., Chiu, A., Wu, G., 2010. Overweighting of small probabilities. Wiley encyclopedia of operations research and management science..

[13] T. Cristino, A.F. Neto, F. Wurtz, B. Delinchant, Barriers to the adoption of energy-efficient technologies in the building sector: A survey of brazil, Energy and Buildings 252 (2021) 111452.

[14] G. De Vries, M. Rietkerk, R. Kooger, The hassle factor as a psychological barrier to a green home, Journal of Consumer Policy (2019) 1–8.

[15] S. Dhami, A. Al-Nowaihi, Why do people pay taxes? prospect theory versus expected utility theory, Journal of Economic Behavior & Organization 64 (1) (2007) 171–192.

[16] T. Dietz, P.C. Stern, E.U. Weber, Reducing carbon-based energy consumption through changes in household behavior, Daedalus 142 (1) (2013) 78–89.

[17] A. Dodoo, L. Gustavsson, U.Y. Tettey, Final energy savings and cost-effectiveness of deep energy renovation of a multi-storey residential building, Energy 135 (2017) 563–576.

[18] J.C. Dunn, Well-separated clusters and optimal fuzzy partitions, Journal of cybernetics 4 (1) (1974) 95–104.

[19] S. Ebrahimigharehbaghi, Q.K. Qian, G. de Vries, H.J. Visscher, Identification of the most critical behaviour-influencing factors in the decision-making processes of dutch homeowners: installations of double-glazing, insulation, solar pv-panel, and sustainable heating system, in: Building research and information, 2021.

[20] S. Ebrahimigharehbaghi, Q.K. Qian, F.M. Meijer, H.J. Visscher, Unravelling dutch homeowners' behaviour towards energy efficiency renovations: What drives and hinders their decision-making?, Energy policy 129 (2019) 546–561

[21] S. Ebrahimigharehbaghi, Q.K. Qian, F.M. Meijer, H.J. Visscher, Transaction costs as a barrier in the renovation decision-making process: A study of homeowners in the netherlands, Energy and Buildings 109849 (2020).

[22] European Commissions, 2012. Commission Delegated Regulation (EU) No 244/2012 of 16 January 2012 Supplementing Directive 2010/31/EU of the European Parliament and of the Council on the Energy Performance of Buildings by Establishing a Comparative Methodology Framework for Calculating Cost-optimal Levels of Minimum Energy Performance Requirements for Buildings and Building Elements. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:081:0018:0036:EN:PDF..

[23] H. Fehr-Duda, M. De Gennaro, R. Schubert, Gender, financial risk, and probability weights, Theory and decision 60 (2) (2006) 283–313.

[24] D.C. Feldman, The development and enforcement of group norms, Academy of management review 9 (1) (1984) 47–53.

[25] E.R. Frederiks, K. Stenner, E.V. Hobman, Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour, Renewable and Sustainable Energy Reviews 41 (2015) 1385–1394.

[26] J. Friege, Increasing homeowners' insulation activity in germany: An empirically grounded agent-based model analysis, Energy and Buildings 128 (2016) 756–771.

[27] R. Gifford, L.A. Comeau, Message framing influences perceived climate change competence, engagement, and behavioral intentions, Global Environmental Change 21 (4) (2011) 1301–1307.

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

[29] R. Gonzalez, G. Wu, On the shape of the probability weighting function, Cognitive psychology 38 (1) (1999) 129–166.

[30] N. Good, Using behavioural economic theory in modelling of demand response, Applied energy 239 (2019) 107–116.

[31] J.M. Gowdy, Behavioral economics and climate change policy, Journal of Economic Behavior & Organization 68 (3–4) (2008) 632–644.

[32] D.L. Greene, Uncertainty, loss aversion, and markets for energy efficiency, Energy Economics 33 (4) (2011) 608–616.

[33] B. Häckel, S. Pfosser, T. Tränkler, Explaining the energy efficiency gap-expected utility theory versus cumulative prospect theory, Energy Policy 111 (2017) 414–426.

[34] Q. Han, S. Ma, T. Wang, F. Chu, Kernel density estimation model for wind speed probability distribution with applicability to wind energy assessment in china, Renewable and Sustainable Energy Reviews 115 (2019) 109387.

[35] D.M. Hausman, B. Welch, Debate: To nudge or not to nudge, Journal of Political Philosophy 18 (1) (2010) 123–136.

[36] Heiskanen, E., Matschoss, K., Laakso, S., Apajalahti, E.-L., 2020. A critical review of energy behaviour change: The influence of context. In: Energy and Behaviour. Elsevier, pp. 391–417..

[37] R.L. Hicks, A comparison of stated and revealed preference methods for fisheries management, Tech. rep. (2002).

[38] A.J. Hoffman, R. Henn, Overcoming the social and psychological barriers to green building, Organization & Environment 21 (4) (2008) 390–419.

[39] A.J. Hope, A. Booth, Attitudes and behaviours of private sector landlords towards the energy efficiency of tenanted homes, Energy Policy 75 (2014) 369–378.

[40] Jakob, M., et al., 2007. The drivers of and barriers to energy efficiency in renovation decisions of single-family home-owners. Center for Energy Policy and Economics CEPE, Department of Management, Technology and Economics, ETH Zurich, Switzerland.< http://www. cepe. ethz. ch/publications/workingPapers/CEPE_WP56. pdf>[22 March 2010]..

[41] D. Kahneman, Maps of bounded rationality: Psychology for behavioral economics, American economic review 93 (5) (2003) 1449–1475.

[42] D. Kahneman, A. Tversky, Prospect theory: An analysis of decision under risk, Econometrica 47 (2) (1979) 263–291.

[43] D. Kahneman, A. Tversky, Prospect theory: An analysis of decision under risk, in: Handbook of the fundamentals of financial decision making: Part I, World Scientific, 2013, pp. 99–127.

[44] L. Klotz, Cognitive biases in energy decisions during the planning, design, and construction of commercial buildings in the united states: an analytical framework and research needs, Energy Efficiency 4 (2) (2011) 271–284.

[45] L. Klotz, D. Mack, B. Klapthor, C. Tunstall, J. Harrison, Unintended anchors: Building rating systems and energy performance goals for us buildings, Energy Policy 38 (7) (2010) 3557–3566.

[46] J.-H. Ko, D.-S. Kong, J.-H. Huh, Baseline building energy modeling of cluster inverse model by using daily energy consumption in office buildings, Energy and Buildings 140 (2017) 317–323.

[47] Y. Kuo, T. Yang, G.-W. Huang, The use of grey relational analysis in solving multiple attribute decision-making problems, Computers & industrial engineering 55 (1) (2008) 80–93.

[48] P.K. Lattimore, J.R. Baker, A.D. Witte, The influence of probability on risky choice: A parametric examination, Journal of Economic Behavior & Organization 17 (3) (1992) 377–400.

[49] R. Layard, G. Mayraz, S. Nickell, The marginal utility of income, Journal of Public Economics 92 (8–9) (2008) 1846–1857.

[50] W.-S. Lee, Y.-C. Lin, Evaluating and ranking energy performance of office buildings using grey relational analysis, Energy 36 (5) (2011) 2551–2556.

[51] L. Lihtmaa, D.B. Hess, K. Leetmaa, Intersection of the global climate agenda with regional development: Unequal distribution of energy efficiency-based renovation subsidies for apartment buildings, Energy Policy 119 (2018) 327–338.

[52] Z. Ma, R. Yan, N. Nord, A variation focused cluster analysis strategy to identify typical daily heating load profiles of higher education buildings, Energy 134 (2017) 90–102.

[53] K. Mahapatra, L. Gustavsson, T. Haavik, S. Aabrekk, S. Svendsen, L. Vanhoutteghem, S. Paiho, M. Ala-Juusela, Business models for full service energy renovation of single-family houses in nordic countries, Applied energy 112 (2013) 1558–1565.

[54] P.C. Mayer, Electricity conservation: Consumer rationality versus prospect theory, Contemporary Economic Policy 13 (2) (1995) 109–118.
[55] F. McLoughlin, A. Duffy, M. Conlon, A clustering approach to domestic electricity load profile characterisation using smart metering data, Applied energy 141 (2015) 190–199.
[56] H. Naganathan, W.O. Chong, X. Chen, Building energy modeling (bem) using clustering algorithms and semi-supervised machine learning approaches, Automation in Construction 72 (2016) 187–194.
[57] S. Organ, D. Proverbs, G. Squires, Motivations for energy efficiency refurbishment in owner-occupied housing, Structural Survey (2013).
[58] R. Osbaldiston, J.P. Schott, Environmental sustainability and behavioral science: Meta-analysis of proenvironmental behavior experiments, Environment and Behavior 44 (2) (2012) 257–299.
[59] T.B. Ouarda, C. Charron, J.-Y. Shin, P.R. Marpu, A.H. Al-Mandoos, M.H. Al-Tamimi, H. Ghedira, T. Al Hosary, Probability distributions of wind speed in the uae, Energy conversion and management 93 (2015) 414–434.
[60] S. Papadopoulos, B. Bonczak, C.E. Kontokosta, Pattern recognition in building energy performance over time using energy benchmarking data, Applied Energy 221 (2018) 576–586.
[61] E. Pikas, J. Kurnitski, R. Liias, M. Thalfeldt, Quantification of economic benefits of renovation of apartment buildings as a basis for cost optimal 2030 energy efficiency strategies, Energy and Buildings 86 (2015) 151–160.
[62] F.A. Postali, P. Picchetti, Geometric brownian motion and structural breaks in oil prices: a quantitative analysis, Energy Economics 28 (4) (2006) 506–522.
[63] M. Rabin, Risk aversion and expected-utility theory: A calibration theorem, in: Handbook of the fundamentals of financial decision making: Part I, World Scientific, 2013, pp. 241–252.
[64] M.O. Rieger, M. Wang, T. Hens, Estimating cumulative prospect theory parameters from an international survey, Theory and Decision 82 (4) (2017) 567–596.
[65] S. Rockstuhl, S. Wenninger, C. Wiethe, B. Häckel, Understanding the risk perception of energy efficiency investments: Investment perspective vs. energy bill perspective, Energy Policy 159 (2021) 112616.
[66] J. Schleich, Energy efficient technology adoption in low-income households in the european union–what is the evidence?, Energy Policy 125 (2019) 196–206
[67] Smith, A., 1822. The theory of moral sentiments. Vol. 1. J. Richardson..
[68] V. Taranu, G. Verbeeck, Overview of dual process behavioural models and their implications on decision-making of private dwellers regarding deep energy renovation, in: WBC16 Proceedings: Volume II Environmental Opportunities and challenges Constructing Commitment and Acknowledging Human Experiences, TUT–Tampere University of Technology, 2016, pp. 591–603.
[69] U.Y.A. Tettey, L. Gustavsson, Energy savings and overheating risk of deep energy renovation of a multi-storey residential building in a cold climate under climate change, Energy 202 (2020) 117578.
[70] R.H. Thaler, C.R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness, Penguin, 2009.
[71] A. Tversky, D. Kahneman, Advances in prospect theory: Cumulative representation of uncertainty, Journal of Risk and uncertainty 5 (4) (1992) 297–323.
[72] L. Venkatachalam, Behavioral economics for environmental policy, Ecological economics 67 (4) (2008) 640–645.
[73] Von Neumann, J., Morgenstern, O., 1947. Theory of games and economic behavior, 2nd rev..
[74] E. Wang, Benchmarking whole-building energy performance with multi-criteria technique for order preference by similarity to ideal solution using a selective objective-weighting approach, Applied Energy 146 (2015) 92–103.
[75] M. Wardman, A comparison of revealed preference and stated preference models of travel behaviour, Journal of transport economics and policy (1988) 71–91.
[76] C. Wilson, L. Crane, G. Chryssochoidis, Why do homeowners renovate energy efficiently? contrasting perspectives and implications for policy, Energy Research & Social Science 7 (2015) 12–22.
[77] C. Wilson, H. Dowlatabadi, Models of decision making and residential energy use, Annual review of environment and resources 32 (2007).
[78] C. Wilson, H. Pettifor, G. Chryssochoidis, Quantitative modelling of why and how homeowners decide to renovate energy efficiently, Applied energy 212 (2018) 1333–1344.
[79] G.Y. Yun, K. Steemers, Behavioural, physical and socio-economic factors in household cooling energy consumption, Applied Energy 88 (6) (2011) 2191–2200.
[80] S. Zhan, Z. Liu, A. Chong, D. Yan, Building categorization revisited: A clustering-based approach to using smart meter data for building energy benchmarking, Applied Energy 269 (2020) 114920.
[81] G. Zheng, Y. Jing, H. Huang, Y. Gao, Application of improved grey relational projection method to evaluate sustainable building envelope performance, Applied Energy 87 (2) (2010) 710–720.