Calibrating agent based models to real data affects the timing and speed of the energy transition

Oreane. Y. Edelenbosch*[a,b,1], Luciana Miu[c,1], Julia Sachs[c], Adam Hawkes[c], Massimo Tavoni[a,b]

[a] Department of Management and Economics, Politecnico di Milan, Via Lambruschini 4/B, Milan

[b] RFF-CMCC European Institute on Economics and the Environment, Via Bergognone, 34, Milan

[c] Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, UK

[1] These authors contributed equally to this work.

Abstract

The building sector has a high potential to reduce energy consumption. Achieving this depends on household's choices, which are known to be highly heterogeneous. Agent-based models are tools used to describe energy choices, but require data demanding calibration. Here we combine a novel, cross-country European household-level survey -including socio-demographic characteristics, energy-saving habits, energy-saving investments, and metered household electricity consumption- with a global agent based energy choice model. Cluster analysis reveals that households who demand and consume energy in very similar ways cannot easily be mapped to standardly used socio-demographic classes or attitudes. However, the data also shows interesting patterns both between and within the clusters. Most noticeably, income, consistently, has the largest effect on demand, dwelling efficiency and energy-saving investments. Dwelling improvement potential also incentivizes energy efficiency investments. We use this cluster analysis to calibrate agents of the residential sector of an agent based model, including also the within cluster variations and uncertainty. Including these various sources of heterogeneity affects the timing and speed of the transition, two key indicators in the context of climate change mitigation. The results reinforce the need for grounding agent-based models in real data, to take real advantage of their capabilities and contribute to a better understanding of energy transition dynamics.

Acknowledgements

OE and MT acknowledge financial support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 154 336155 - project COBHAM "The role of consumer behaviour and heterogeneity in the integrated assessment of energy and climate policies".

1. Introduction

Energy use in residential buildings accounts for approx. 25% of the EU's total final energy consumption (EC, 2018), and has been the focus of numerous emissions reduction programmes, such as the Energy Labelling Directive and the Energy Performance of Buildings Directive. While the potential for reducing residential energy use is substantial (Berardi, 2017), energy consumption in buildings relies heavily on the residents behavior, depending on attitudes, habits and consumption patterns. These behaviors attribute to sizeable variation in households energy consumption and introduces heterogeneity into the likelihood of efficient technology adoption and other solutions to achieve low energy buildings (Hong, Taylor-Lange, D'Oca, Yan, & Corgnati, 2016).

Research efforts to better understand this heterogeneous behavior and its drivers are therefore central to successfully implementing residential energy efficiency policies and programmes (Lopes, Antunes, & Martins, 2015). There is a vast amount of research that has examined residential energy-saving behavior, employing a range of disciplinary perspectives ranging from the more economics-focused research frameworks to theories which incorporate social and contextual factors (Wilson & Dowlatabadi, 2007). There is consensus that the long-term, deliberative behaviors of investing in residential energy-saving technologies are also affected by numerous non-economic drivers and barriers (Klöckner & Nayum, 2017; Hesselink & Chappin, 2019; Wilson, Crane, & Chryssochoidis, 2014).

Decision-makers in the energy policy domain are often relying on energy system models, and integrated assessment models (IAMs), that are used to evaluate long-term pathways of climate change mitigation and conduct global reviews of the energy system (Tavoni et al., 2015). The heterogeneity of user behavior and energy technology investments is difficult to capture in these long-term models (McCollum et al. 2017), that focus on high-level trends (Krey, 2014), and are therefore less suited for the representation of individual choices. This lack of this representation can lead to several challenges, including inaccurate estimation of demand-side reduction potential (Sovacool et al., 2015), failure to capture interaction such as social influence effects (Pettifor, Wilson, McCollum, & Edelenbosch, 2017), and difficulties in evaluating the impact of behavioural policies (Mundaca, Neij, Worrell, & McNeil, 2010).

The challenges in representing heterogeneous behaviour could be addressed through the use of agent-based models (ABMs), which can capture the individuality, interactions and varied drivers of decision making (Farmer & Foley, 2009)(An, 2012). Rai & Henry (2016) therefore make a compelling argument that for a better understanding of the complexities of energy choices affecting climate mitigation strategies ABMs should be used to model energy demand projections. Rai & Robinson (2015), however, point out that there are two key challenges when developing ABMs: There is a lack of empirical data to appropriately initialize, verify and validate the models (data challenge), and second: Often behavioral decision rules applied in the models are introduced in an ad hoc fashion (theoretical challenge) (Durlauf, 2012).

In this paper, we address the first challenge of using empirical data to appropriately define ABMs capable of modelling realistic behavioural heterogeneity within IAMs. With a focus on residential energy efficiency, our research asks two main questions: 1) Can empirical data reveal customer profiles adequate for defining agents in a residential energy ABM? and 2) Can we use this data to identify drivers of energy-saving investment behaviours and incorporate them into the ABM? To address these questions, we use a unique, cross-country European dataset of household-level survey responses, including socio-demographic characteristics, attitudes, energy-saving habits and energy-saving investments, and metered household electricity consumption. Through cluster-analysis this dataset is used to explore persistent, heterogenous patterns and drivers of energy-saving behaviour; and crucially, the availability of metered electricity usage data allows agent typologies to be related to *actual* consumption levels. Moreover, the diversity of questions asked in the survey allows to distinguish between service demand and energy efficiency contributing to household energy consumption.

This data is applied to the Residential Building Simulation Module (RBSM) ABM, part of the new MUSE ® IAM (Sachs, Meng, Giarola, & Hawkes, 2019), designed to represent household level decision-making affecting technology choices and investment, at a regional scale. The RBSM is a bottom-up, technology-rich model, characterizing 70 different residential energy technologies, based on unit technology cost, efficiencies, lifetime and emissions. The model framework is based on the Theory of Planned Behaviour (Ajzen 2001), and agents' energy-saving investment behaviour is represented as a step-wise decision-making process for selecting and implementing energy-saving technologies.

By using empirical data as the starting point for defining agent typologies and investment behaviours, the aim of this research is to better represent the heterogeneity in user energy choices in the RSBM, and then examine how this effects the high-level trends of technology adoption and energy consumption. We find that the data allows to define clusters of household that act in similar in terms of energy consuming indicators. While patterns between these cluster in terms of socio-demographics and preferences can be seen, the data also shows a high level of within cluster variation. Including these various sources of heterogeneity in the ABM model a significantly different timing and rate of implementation of energy-saving technologies by households is found, compared to the case where conventional patterns of demand, consumption and investment are assumed. In the long term the new empirically grounded ABM does not show different levels up technology uptake, but particularly affects the development of the transition.

2. Analysis

The data used in this study was collected via two large-scale surveys: PENNY, conducted in Italy, Switzerland and the Netherlands, and COBHAM, conducted in Italy. 6,138 responses were recorded, containing information on socio-demographic and socio-psychological characteristics, dwelling and household characteristics, technologies and energy services used, and their metered electricity consumption[1]. More information on the survey can be found in the Supplementary Materials.

To describe the energy consumption characteristics of respondents, the survey responses were used to construct a number of "energy efficiency indicators". Firstly, lighting and appliance electricity demand were estimated for each respondent household based on their dwelling characteristics and survey responses related to service level (e.g. floorspace and number of appliances, see Supplementary Methods). These two indicators are in the paper referred to as lighting and appliance service demand, and can be seen as the expected electricity use for that level of service. Secondly, the sum of the lighting and appliance service demand was compared to the actual metered electricity consumption to produce two energy efficiency indicators associated with each respondent:

- The "efficiency gap": the difference between metered electricity consumption and the appliance and lighting service demand (the values can be positive or negative). The values were grouped into 5 quintiles, ranging from very efficient dwellings (negative

[1] There were a large number of missing responses for metered electricity consumption in the Netherlands, leading to an under-representation of data from this country.

efficiency gap, high absolute value) to very inefficient dwellings (positive efficiency gap, high absolute value)[2].

- The "relative savings potential": the proportion of the efficiency gap compared to the appliance and lighting service demand, as an indicator of how close a household is to consuming as much electricity as expected based on its dwelling characteristics, and also an indicator of the potential energy cost savings. The values were grouped into 5 quintiles, ranging from very high savings potential (i.e. households with the largest positive efficiency gaps, as a proportion of their estimated electricity demand) to very low savings potential (i.e. households with the largest negative efficiency gaps, as a proportion of their estimated electricity demand).

In addition to these electricity-related indicators, an environmental preferences index was also constructed, as a simple aggregate of respondents' scoring of questions on environmental value, morality, identity and injunctive norms.

The survey responses were clustered based on the lighting service demand, appliance service demand and the efficiency gap (k-means clustering with Jaccard dissimilarity measure). Because the sum of these three clustering variables is equal to the metered electricity consumption, the clustering grouped the data based on the existence of similar patterns in electricity consumption, as well as lighting and appliance service demand and efficiency gaps. These groups were then analysed for within- and between-group differences in socio-demographic and socio-psychological characteristics, energy literacy, metered electricity consumption, energy-saving habits and energy-saving investments. The variables which showed the strongest between-cluster differences were also fitted to regression models, to determine whether they were significantly affected by other variables. Supplementary Table 1 provides an overview of all variables analysed.

The resulting clusters were used as a basis for defining and characterizing agents in the RBSM, and projecting the uptake of residential technologies in the EU-18 residential sector, to the year 2050. The extension of the model scope from the survey data to the residential building sector of the EU-18[3] region is clearly an approximation which could be improved by using globally available datasets. The goal of this study was not to produce accurate model projections, but rather to analyse how the use of empirically-based clusters and assumptions on investment drivers affects the projection of technology uptake by the RBSM.

The age, income and household size distributions within each cluster were linked with data from the European Union Statistics on Income and Living Conditions (EU-SILC, 2019), to determine the overall share of the European population represented by each agent. Each agent was linked to different rules for searching for and deciding to invest in residential energy technologies, based on the relationship between household characteristics and investment behaviour identified in the cluster analysis. The drivers of energy-saving investments identified in the survey data were used to make assumptions about what objectives agents would seek to fulfil as they make decisions on investing in energy-saving measures. The main objectives used in the RBSM ABM are:

- Capital cost – agents seek to invest in technologies of a certain capital cost based on their income constraints and risk preferences;

[2] Note that those household with electric heating or cooking (which were a small number of respondent households) were taken out of the data sample.

[3] EU-18 refers to EU-18: Austria, Belgium, Czech Republic, France, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, Netherlands, Poland, Portugal, Slovak Republic, Spain, UK, Slovenia, Estonia (IEA, 2017)

- Equivalized annual cost (EAC) (the annual cost of owning, operating, and maintaining an asset over its entire life) – agents will seek to adjust their EAC based on their income constraints and savings potential;
- Fuel consumption cost – agents will seek to reduce their fuel consumption cost based on their savings potential;
- Efficiency – agents will seek to adjust the efficiency of their energy use for non-economic reasons
- Emissions – agents will seek to adjust their emissions levels based on their preferences such as environmental awareness.

A more detailed description of the RSBM can be found in the Supplementary Materials.

To test this empirically-based ABM, in Section 3.2 we compare the outcomes of using the original, macro-economically-driven agent parameterization that uses links between age, income level, occupation and housing data but to determine agent groups without considering information about energy usage, described in Sachs et al., 2019, to the new cluster-driven model. The cluster-driven model simulates 5 agent clones for the 5 sub-ranges of estimated electricity demand in each cluster, and each agent clone was duplicated to indicate whether their investment was in a new residential building or an existing one (retrofit). The ABM code presented in Sachs et al., 2019 was extended to incorporate this variation. All scenarios were run with the average carbon tax of the Energy Modelling Forum-27 450ppm full-tech scenario (Kriegler et al., 2014). The case study does not directly account for the changing carbon intensity of electricity, but rather considers the influence of the carbon price on the electricity price as a proxy.

3. Results and discussion

3.1 Description of clusters

The clustering produced an optimal solution of 10 clusters, containing 4,874 responses. Out of these 10 clusters, 5 were of suitable size for further analysis[4] and are presented in Table 1. Most variables were significantly different between clusters (Table 2[5]), apart from household size (although significant at $p<0.1$), and the environmental preference index. This was confirmed by post-hoc tests of difference. The differences between clusters were strongest for the absolute and relative efficiency gaps. The main characteristics of respondents grouped in each cluster can be seen in Figures 1, 2 and 3, and are described below.

Table 1. Size of clusters identified in the optimal cluster solution

Cluster	**Number of survey responses**			
	Italy	**Switzerland**	**Netherlands**	**Total**
1	643	125	39	807
2	150	147	21	318
3	319	86	40	445
4	1204	70	34	1308
5	1538	116	36	1690

[4] Detail the criteria for considering them too small.

[5] The p-values are reported for the Kruskal-Wallis tests that accounted for ties in the data.

Table 2. Significance of differences between the 5 clusters.

Variable	**Significant differences between clusters?**	**Variable**	**Significant differences between clusters?**
Income class	Y	Efficiency gap	Y
Age range	Y	Relative savings potential	Y
Education level	Y	Energy-saving habits: frequency of switching lights off	Y
Household size	N (p=0.051)	Energy-saving habits: frequency of unplugging appliances	Y
Environmental preference index	N (p=0.46)	Respondents' value of wealth	Y
Energy literacy	Y	Dwelling floor area	Y
Metered electricity consumption	Y	Length of residence in dwelling	Y
Estimated lighting electricity demand	Y	Risk preferences	Y
Estimated appliance electricity demand	Y	Level of energy-saving investments	Y

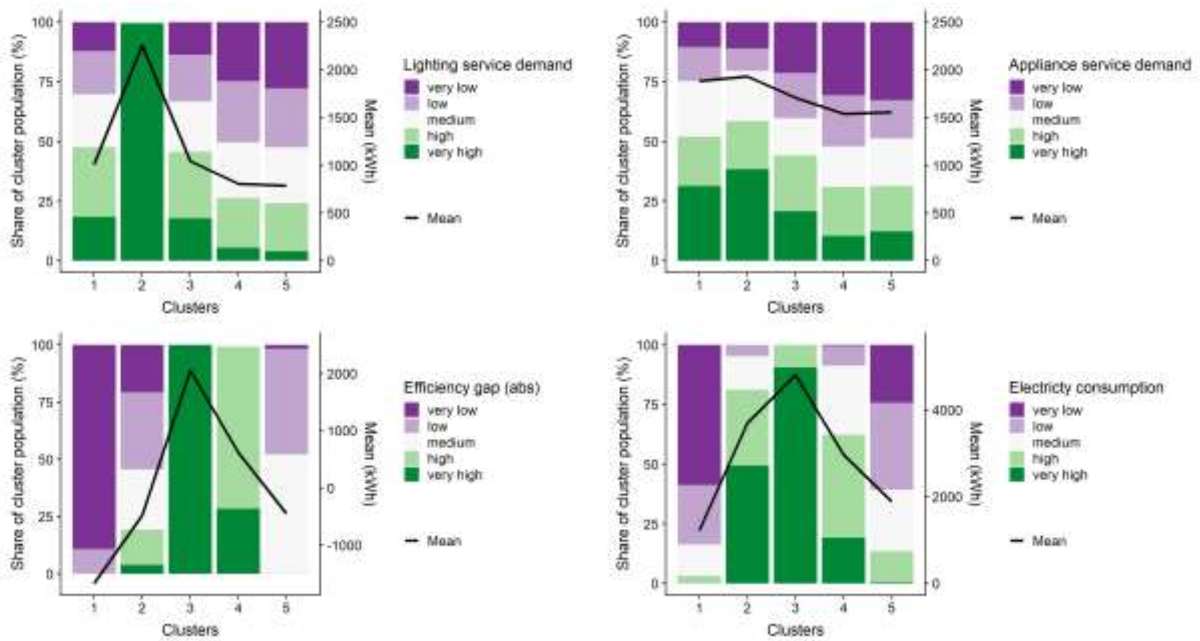


Figure 1. Key characteristics of energy consumption of the 5 clusters.

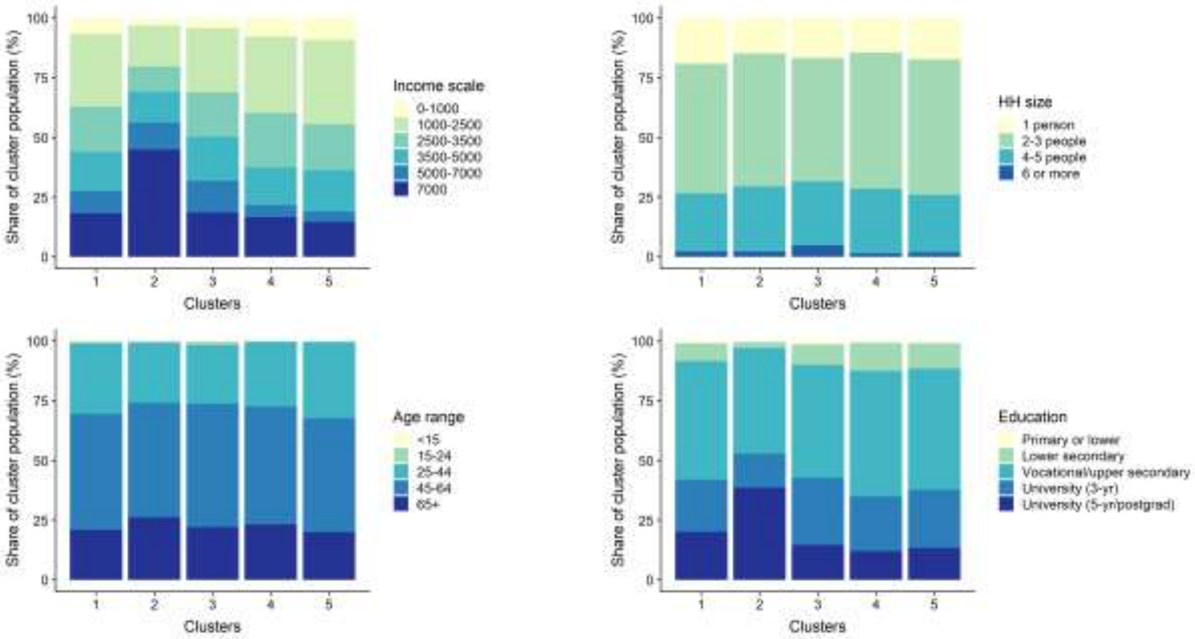


Figure 2. Socio-demographic indicators in the 5 clusters.

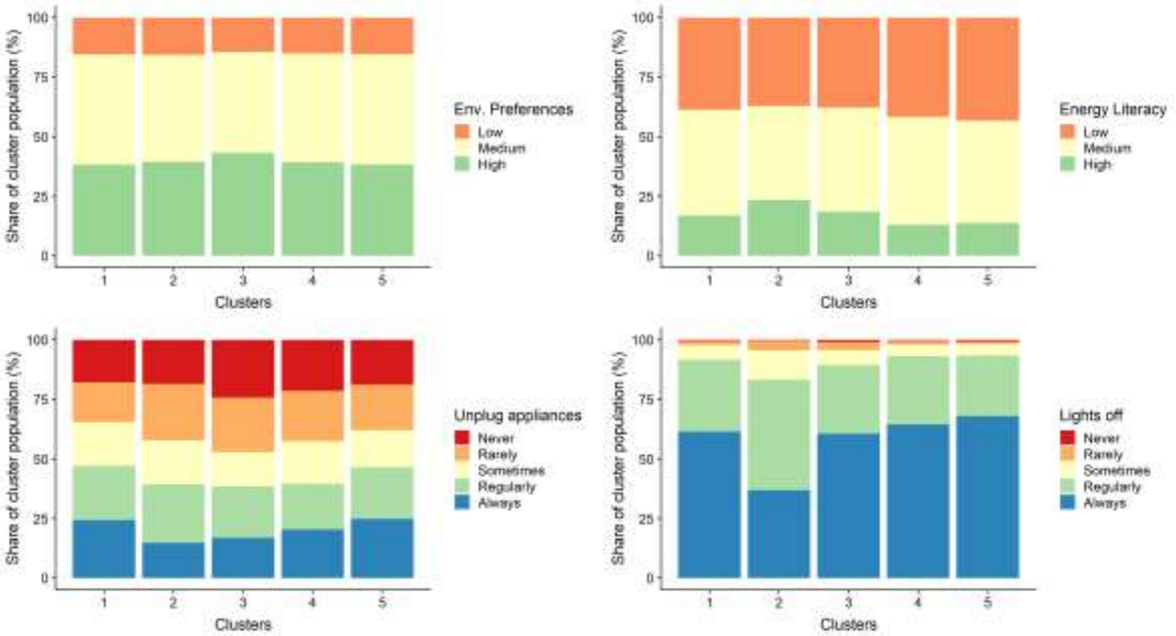


Figure 3. The environmental preference, energy literacy and energy efficient behaviour of the 5 clusters.

Cluster 1 is *comfortably efficient*, living in very efficient (86%) or efficient (14%) households, despite having a higher appliance service demand than other clusters. Households in this group have already achieved significant savings, and there is little room for further improvement, with annual electricity consumption being on average 58% less than expected demand. Members of this group are relatively young, live in slightly smaller households and have the lowest residence times. They have good energy-saving habits, in particular being the best at unplugging their appliances, despite having slightly lower environmental preferences.

Cluster 2 is a *well-off, medium-efficient* group, with a range of household efficiency gaps, the highest lighting and appliance electricity demand of all groups and potential savings of up to 26% (compared to the expected use) in inefficient households, which represent 24% of the group. Members have higher education, strikingly higher income levels and are slightly older, more energy literate and have slightly worse energy-saving habits, than other clusters. They have slightly larger household sizes, live in slightly larger dwellings, and value wealth highly.

Cluster 3 is an *inefficient high-consumption* group, whose members live exclusively in very inefficient households, consume the most electricity of all groups and have potential savings of 64% on average. They have the largest household sizes, are fairly well-educated and have high environmental preferences, but have the highest proportion of group members who never unplug their appliances to save energy.

Cluster 4 is an *inefficient low-consumption* group, whose members live in inefficient (57%) and very inefficient (39%) households with average potential energy savings of 28%, have the very low lighting and appliance electricity demand and medium levels of electricity consumption. They live in the smallest households and dwellings of all groups, are relatively low-educated and have longer residency times than most other groups.

Cluster 5, the *resource-constrained and medium-efficient,* is also a low-consuming cluster, where the majority lives in efficient households (92%) or in inefficient households with potential savings of less than 4.5% to consume electricity according to their expected demand. Members of this group are have the lowest income, are slightly less energy literate and live the smallest dwellings of all groups, despite having a similar household size distribution to the comfortably efficient group.

Analysing the efficiency gap for different groups can reveal how much households are over- or under-consuming relative to their expected electricity demand. For example, a group of households with very low efficiency gaps is already very efficient (e.g. cluster 1, which consumes on average 1,546 kWh less than its expected annual electricity demand); a group of households with very high efficiency gaps is very inefficient (e.g. cluster 3, which consumes on average 1,737 kWh more than its expected annual electricity demand), and thus has to make substantial savings in order to reduce its electricity consumption to expected levels. The efficiency gap thus describes the "improvement effort" of inefficient households and the "accomplished savings" of efficient households.

On the other hand, analysing the relative savings can reveal how much of a difference reducing this efficiency gap could make to a household: two inefficient households requiring the same improvement effort per kWh may perceive the resulting savings as vastly different if their expected electricity needs are different. For example, despite households in clusters 2 and 5 having similar average efficiency gaps (on average -483 and -449 kWh/year, respectively), relative to their electricity demand these gaps are different – lower for cluster 2, whose richer, larger households demand more electricity, and higher for cluster 5, whose poorer households living in small dwellings demand less.

An analysis was conducted to determine whether the "improvement effort" or "accomplished savings", in conjunction with the relative savings, had an effect on the energy-saving investments made in each cluster. The analysis could only be performed on 4 out of the 5 original clusters (clusters 1, 3, 4 and 5), due to the lack of investment data for the well-off medium-efficient group.[6]

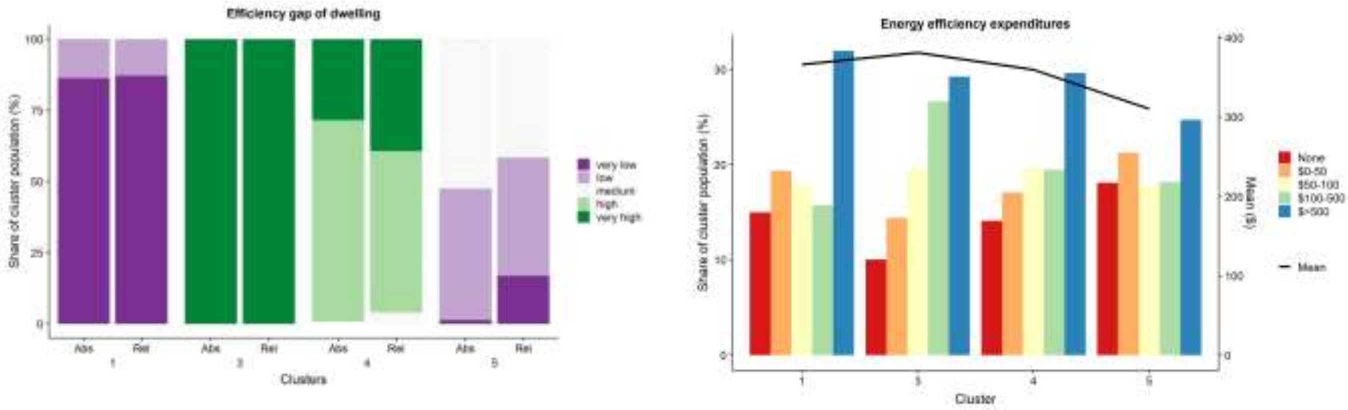


Figure 4. Relative efficiency gap compared to the absolute energy efficiency for the five clusters. Only responses from the COBHAM survey are used (left). Energy efficiency expenditures per cluster (right).

As shown in Figure 4, the differences in energy-saving investment levels between groups are less noticeable than those in efficiency gaps or relative savings, indicating that investment is based on more factors. These two variables do demonstrate an effect on investment in the two least efficient clusters: clusters 3 and 4, which invested on average \$312/year and \$299/year, respectively, and around a third of whose members invested at the highest level of over \$500/year. However, there are noticeable differences in the distribution of investment levels in each cluster, with cluster 4 having fewer respondents investing at the mid-levels than would be expected when considering the significant proportion (61%) that can achieve substantial relative savings with less effort than those in cluster 3[7]. This is due to significant interaction effects between the efficiency gap and relative savings ($p<0.05$, ordered probit regression model). These differences do not affect cluster 3, but they do affect cluster 4: in the highest category of relative savings, investments of over \$500/year occur even if the efficiency gap is very high (thus requiring substantial effort), but in the second-highest category of relative savings, investments of over \$500/year are 62% less likely to occur if the efficiency gap is high, than if it is very high. Without these interaction effects, the proportion of investors in the highest-investing categories, and likely the average investment, would likely be higher in cluster 4. This slightly counter-intuitive finding may be explained by the fact that these respondents with high potential savings and high efficiency gaps (61% of cluster 4), have lower overall electricity consumption levels than their counterparts with very high potential savings, and may therefore perceive their relative savings less clearly.

Other factors also demonstrate clear effects on respondents' investments. In cluster 3, larger households (2-3 person and 4-5 person, which make up 80% of the group) are less likely to invest significantly in energy-saving measures than single-person households ($p<0.05$ ordered probit regression). In cluster 1, which has already accomplished its savings but still invests a very high amount, investments were negatively affected by increasing environmental preferences[8] and positively affected by increasing the energy-saving habit of unplugging appliances[9] ($p<0.05$, ordered probit regression). Furthermore, despite still having potential to

[6] This was because most responses in this group were obtained from the survey that did not collect data on energy-saving investments.
[7] Those with high relative savings and high, but not very high, efficiency gaps
[8] Above the threshold of medium-low environmental preferences.
[9] Above the level of "sometimes" unplugging appliances.

save energy, cluster 5 made the least investments in energy-saving measures, and significantly less than other groups. This was due to the significant positive effect of income on investments, at the level of the entire survey population: increasing income levels led to an increase in investment, apart from when respondents entered the highest income level ($p<0.05$, ordered probit regression model). With cluster 5 containing the lowest-income respondents overall, an income constraint on this group becomes apparent in addition to its relatively low potential for savings. This potential income constraint is also manifested when comparing the other clusters: between clusters 3 and 4, the higher-income group is also the higher-investing group. However, income does not seem to constrain investment by respondents in cluster 1, a relatively low-income which invests a high amount regardless.

The cluster analysis of survey data demonstrates several points. Firstly, there is a large variation in socio-demographic indicators, energy-saving habits and energy-saving investments within groups with similar demand and consumption profiles. Secondly, there was still heterogeneity in how respondents within the different clusters decided to make energy-saving investments. These reinforces previous findings on the range and diversity of factors affecting residential energy demand, consumption and energy-saving investments (Hesselink & Chappin, 2019;Trotta, 2018;Klöckner & Nayum, 2017;Klöckner & Nayum, 2017). However, despite this variation, the cluster analysis identified several patterns for the energy-saving investments of different clusters. At population level, income generally played a significant role in incentivising investments, in line with a range of existing research (Nair, Gustavsson, & Mahapatra, 2010). Additional drivers on energy-saving investments also had clear effects within and between the different groups, indicating that purely income-driven investment decisions are likely to be unrealistic representations of actual uptake of residential energy-saving measures (Wilson et al., 2014). The efficiency gap and relative savings achievable by a household partially drove investment, similar to the findings of Hrovatin & Zorić (2018) and Nakamura (2016) who outline the importance of physical building characteristics and dwelling efficiency in driving energy-saving investments. On the other hand, there were significant interaction effects between a household's efficiency gap and relative savings, while income and, within certain clusters, household size and energy-saving behaviour, also contribute to investment levels, confirming the complexity of investment decision-making processes and the heterogeneity of drivers for these processes. In the next section, we describe the translation of this heterogeneity in the MUSE ® RBSM, and discuss the differences in model projections that this may introduce.

3.2 Application of findings to the RBSM ABM

The clusters described above were used to define 4 groups of agents with different efficiency gaps and drivers for investment in new energy technologies. As shown above, energy-saving investment was driven by a variety of factors, including income, which were used to make assumptions about the objectives that drive agents, and the constraints that block them, when they invest in energy-saving measures (Table 3).

- Cluster 1 had already achieved significant savings, but still invested substantially, driven by good energy-saving habits. We therefore assume that agents associated with cluster 1 are motivated by the desire to improve their efficiency, and are therefore assigned "efficiency" and "emissions" objectives. Although they invest significantly more than their income would suggest, their relatively low income levels may constrain them from adopting more expensive energy-saving technologies in the future. This behaviour is captured through an upper constraint on "capital cost" of an asset to only allow investments in technologies within a certain price range.

- Cluster 3 had very high potential to improve its efficiency and benefit from large relative savings, and investments were helped by having high income. However, having high improvement potential was not sufficient to incentivise investment from larger households. We therefore assume that some agents associated with cluster 3 will be motivated by "efficiency", "fuel consumption costs" and "equivalized annual costs" objectives. Since people within this cluster show the highest income, the required initial investment is assumed to only have a minor role where the total lifetime cost, EAC, is more likely to be taken into account.
- Cluster 4 also had high potential to improve its efficiency and benefit from savings, but less so than cluster 3. Its overall investment was reduced by the under-investment from respondents with high efficiency gaps and high relative savings potential, who already had low electricity consumption. Some of agents associated with cluster 4 were therefore assumed to be motivated by the "fuel consumption costs", "EAC" and "capital cost" objectives, but only if their electricity consumption was high.
- Cluster 5 had little potential to further improve its efficiency, having already achieved significant savings. Most agents were therefore assumed to be driven by the desire to be more efficient, and some by the desire to reduce their energy costs, and are therefore assigned "efficiency" and "fuel consumption" objectives, where the potential for savings still exists. Regardless of their drivers, agents associated with this cluster were strongly constrained by their low income, which caused the lowest investment level of all groups. Thus, an upper constraint on the maximum amount of initial investment is integrated.

Risk preferences and energy literacy were also used as a proxy for agents' openness to new technologies, which defined their rules for searching for new energy technologies and the desired maturity level when deciding to make an energy-saving investment (Table 3).

Table 3. Translating the cluster findings in to agent objectives

Cluster	**Objectives of agents**	**Constraints**	**Openness to new technologies**
Cluster 1 – AGENT 1	Emissions Efficiency	Potentially constrained by capital cost	High – open to invest in new technologies
Cluster 3 – AGENT 2	Efficiency Fuel consumption costs Equivalized annual cost	Not constrained by capital cost	Low – prefers to invest in mature technologies (10% maturity threshold)
Cluster 4 - AGENT 3	Fuel consumption costs (partially) Equivalized annual cost Capital cost	Potentially constrained by lower electricity consumption	Neutral – neither takes high risks with new technologies nor prefers to invest in mature ones (5% maturity threshold)
Cluster 5 – AGENT 4	Efficiency Fuel consumption costs	Highly constrained by capital cost	Neutral- neither takes high risks with new technologies nor prefers to invest in mature ones (5% maturity threshold)

These assumptions on objectives and constraints were used to define the 4 agents within the RBSM for the EU-18 region, and project the uptake of heating and lighting technologies in this region, up to the year 2050. The key changes to the model that define the new agents are the share of the population represented by an agent, the demand for energy service, the technology maturity threshold and constraints on the budget. The assumption here is that the energy consumption patterns seen for lighting and appliances will also be reflected in other end-uses of energy for residential buildings. Figures 5 show how the projected uptake of heating and lighting technologies changes when using the agents derived from the cluster analysis, versus a model driven exclusively by macro-economic indicators. For the uptake of lighting technologies, the effect of using cluster-based agents is to increase the pace of uptake of LED lighting, as well as extend the use of conventional lighting until 2050, where efficient lighting bulbs dominate in the macro-economically driven model. A similar technology landscape can be seen for heating in 2050, where heat pumps have been adopted and replaced conventional boilers; however, in the cluster-based model, initially heat pumps have a smaller share, but then experience a more distinct transition phase with a rapid uptake in the period 2025-2035. The macro-economic model, in contrast, shows a more gradual transition.

These different diffusion patterns are an effect of basing the assumptions on the socio-demographic characteristics of an agent and consequently the share of the population presented by the agent, its financial limitations and openness to new technologies on the empirical cluster findings, rather than on broader assumptions on the correlation of investor behavior and macro-economic indicators. For example, in the macro-economically driven model, the assumption that all high-income, well-educated agents within a certain age group tend to adopt energy-efficient technologies leads to an uptake of heat pumps in all high-income groups, and therefore a more gradual diffusion to other groups over time. In the cluster-driven model, the heterogeneity in investment drivers across agents (also within high income groups), affected by amongst others the role of efficiency gap and relative savings, means that uptake of heat pumps will start earlier and has the potential to then grow fast.

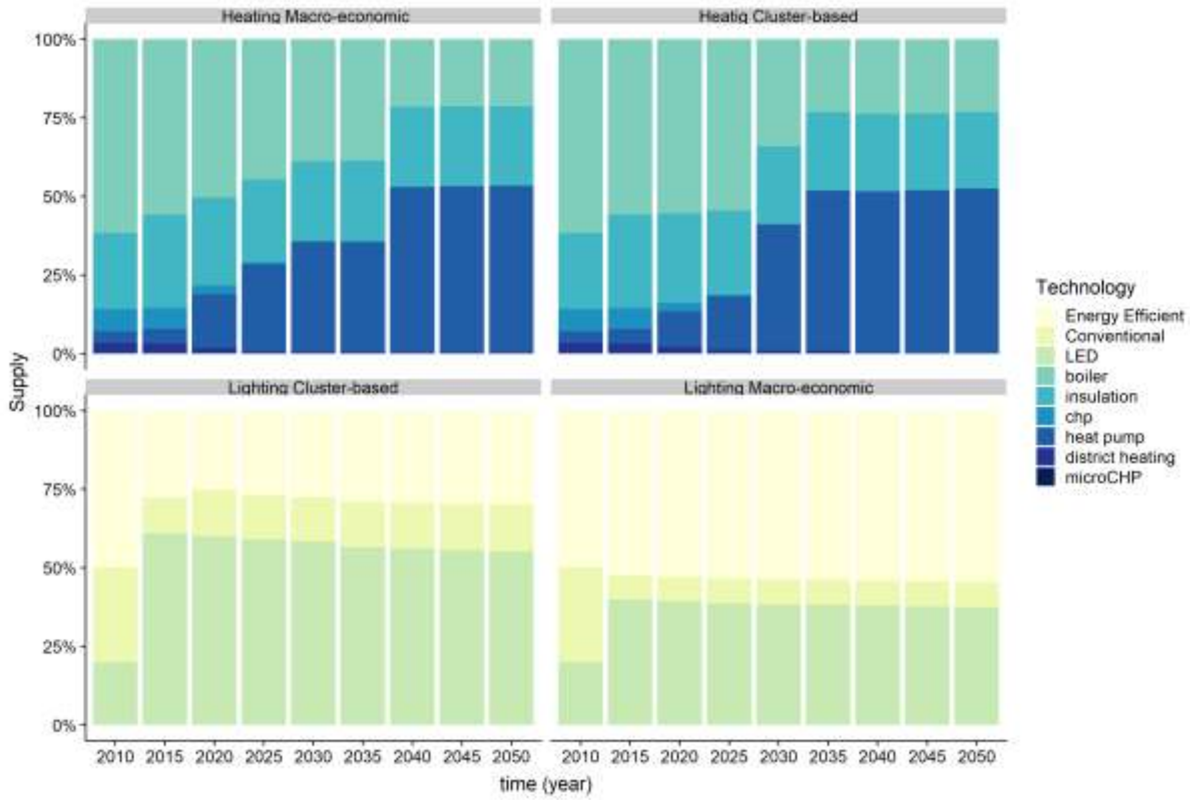
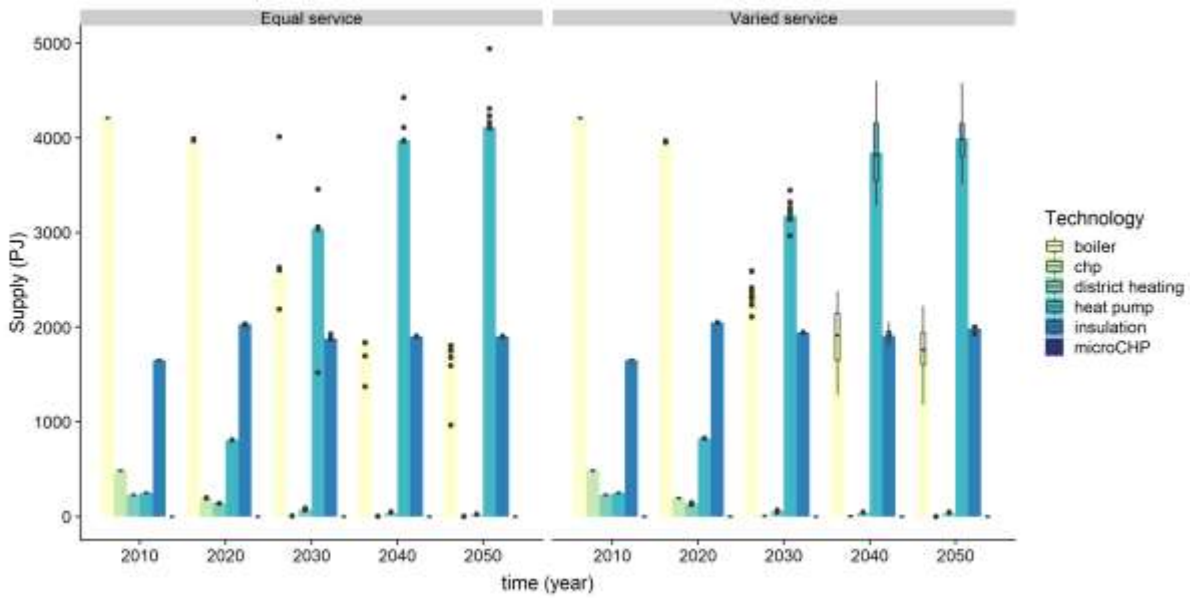


Figure 5. Lighting and heating technology penetration in the model formulation with 1) cluster-based approach including different service levels 2) macro-economic driven approach (from left to right).

We also examined the effect of adding (1) within-cluster variation in service demand, by creating "clones" for each agent, with the same socio-demographic profile but different demand levels; and (2) stochastics around the parameters of agents' decision heuristics (e.g. weights assigned to the agents' objectives when deciding whether or not to invest in a technology) to capture the heterogeneity in decision making of an agent. The five different electricity demand levels are modelled using a scaling factor based on the electricity consumed in each specific level, compared to the total average energy consumption across all clusters. To capture the within-cluster variation, each of the agent 'clones' a certain share of the in cluster population belonging to one demand level. The stochastics are implemented by multiplying the decision heuristic with a scaler drawn from a normal distribution with mean one and variance of 20%. Every time the decision process is carried out, a random value from this distribution is taken and multiplied to the decision heuristic.

Figure 6 shows the mean and standard deviation of the installed stock of lighting and space heating technologies, over 100 simulations, both for the model with within-cluster variation in service demand, and the one without. These figures show between the two models, clear variations in the uptake of heating and lighting technologies. In particular, there are a higher number of extreme scenarios, identified as outliers, during the transition phase around 2030 and a higher variation in the total uptake towards 2050. An increasing carbon price is assumed, which means that low-carbon technologies become more competitive around 2030, and a small change in the investment heuristics leads to the adoptions of different technologies and thus different transition patterns. Due to the introduction of varied service levels, agents with less demand will be less incentivized to adopt high efficient technologies quickly, following the lower energy saving potential. This is obvious for the uptake of lighting technologies, where LED lighting is taken up more rapidly in the varied service level scenario, and high variations in the equal service level scenario.

For the uptake of heating technologies, assuming equal service levels leads to an increased number of outliers but a more stable uptake trend for heat pumps and boilers, whereas incorporating varied service demand leads to more uncertainty in the uptake of these technologies towards 2050.



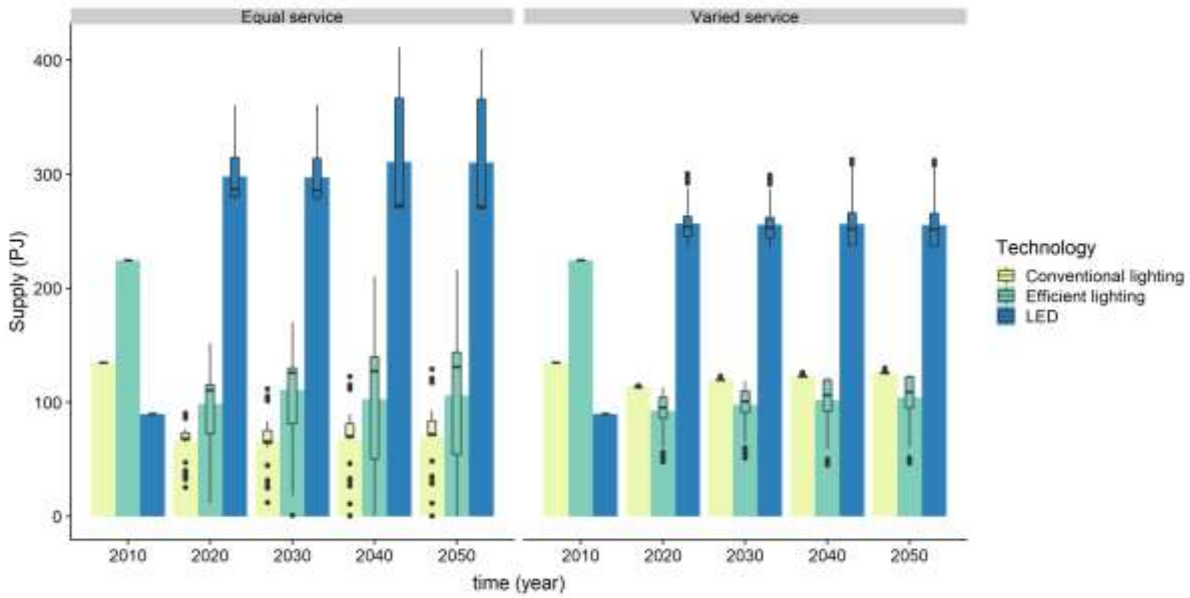


Figure 6. Heating and lighting technology supply for the equal service scenario compared to the varied service scenario. The figure shows the mean supply across the 100 runs (indicated by the bar), and the spread in technology uptake (indicated by the boxplot).

The results of this agent-based modelling show that the cluster-based modelling of investor agents introduces differences in the projected uptake of heating and lighting technologies. These differences are a reflection of the heterogeneity in service level consumption profiles. By profiling agents according to their improvement potential, the cluster-driven ABM parameterization was less driven by socio-economic factors, and more by agents' demand profiles, showing greater heterogeneity in technology uptake when within-group demand differences are considered. Ameli & Brandt (2015) state that models that fail to include this behaviour-based heterogeneity have been found to under-estimate investment in energy-saving technologies. In this study, we find that compared to the macro-economically driven model, the cluster-based models in the long term show similar levels of efficient technology uptake but show different transition pathways of how to get there (Figure 6). The inclusion of non-socio-economic drivers in the cluster-based model, under an increasing carbon price, can lead to significant shifts towards energy-saving technologies over short periods, at uncertain points in time. By adding uncertainty around the parameters of agent decision heuristics, further variation in the speed of the uptake of heating and lighting technologies appears. Therefore, the uncertainty surrounding heterogeneous decision-making appears to affect the penetration rates less than the timing and speed of the transition development, which are both crucial aspects in the context of climate change mitigation.

4. Conclusions

The impact of behavioural factors on user investments in energy-saving measures introduces heterogeneity in energy technology choices and affects the dynamics of the residential energy efficiency transition. Agent based models (ABMs) are powerful tools for representing this heterogeneity, but are challenged by the need for empirical grounding of agent definitions and behavioural parameters. In this study, we use a cross-country survey to identify patterns and drivers of energy-saving investments. By estimating household-level electricity demand and comparing it to metered electricity consumption, we have defined measures of dwelling energy efficiency, which we used to conduct a cluster analysis and partition the large survey population into groups with characteristic demand and consumption profiles. These groups formed the basis for deeper investigation into socio-demographical characteristics, dwelling improvement potential and investment behaviour.

The cluster analysis showed significant variation in energy consumption and dwelling efficiency between groups, but much less in socio-demographic characteristics, indicating respondents who demand and consume energy in very similar ways, cannot easily be mapped to socio-demographic classes. There are, however, interesting patterns both between and within the clusters. Income, consistently, is the biggest driver that affects demand, dwelling efficiency and energy-saving investments. Dwelling improvement potential also plays an important role, incentivising energy efficiency investments in households where they might have a significant impact. However, there is interference from other socio-demographic and psycho-social characteristics such as household size, age and, in specific groups, environmental preferences and energy-saving habits. By translating these drivers and user characteristics into ABM objectives and agent typologies, respectively, we have empirically grounded the MUSE ® RBSM and produced a different outlook for the EU-18 residential heating and lighting sectors, including more heterogeneity in the uptake of energy technologies and thus providing a more accurate representation of investment in energy-saving measures. A systematic treatment of uncertainty, as done here by including stochastics in several model parameterizations, is required in order to reflect on the variety and diversity in the observed energy choices.

Further research can address several limitations in this study. Firstly, longer time-serie data is required to understand how dwelling improvement potential, environmental and risk preferences, energy literacy and energy-saving habits may change over time (Friege, 2016). Secondly, while this research focused on several countries within Europe it would be interesting to perform similar test in more countries within and outside of Europe. In addition, the empirical work of this study focussed on lighting and appliances, while ideally energy consumption data on space heating would be included. Moreover, several other factors could be incorporated for improved analysis of investment behaviour: perceived consumer effectiveness – particularly in the case of studies based on self-reported investment (Armitage & Conner, 2001) actual energy savings investment data and contextual influences such as household characteristics, home tenure and property characteristics (Wilson et al., 2014).

Capturing heterogeneity in long-term models is complex and strenuous also due to the large data requirements. However, in the current age of big data collection, the potential of characterising household energy consumption combined with questionnaires to clarify household motivations, as was done in this study, is promising. This study demonstrates a modelling approach to link empirical data to ABM modelling. The results show that including this heterogeneity in long-term projections can affect transition dynamics, which is surrounded by large uncertainties affecting directly the impacts of energy policy.

References

Ameli, N., & Brandt, N. (2015). *What Impedes Household Investment in Energy Efficiency and Renewable Energy?* https://doi.org/10.1787/5js1j15g2f8n-en

An, L. (2012). Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling*, *229*, 25–36. https://doi.org/10.1016/j.ecolmodel.2011.07.010

Armitage, C. J., & Conner, M. (2001). Efficacy of the Theory of Planned Behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471–499. https://doi.org/10.1348/014466601164939

Berardi, U. (2017). A cross-country comparison of the building energy consumptions and their trends. *Resources, Conservation and Recycling*, *123*, 230–241.

Durlauf, S. N. (2012). Complexity, economics, and public policy. *Politics, Philosophy & Economics*, *11*(1), 45–75. https://doi.org/10.1177/1470594X11434625

EC. (2018). Eurostat database. Retrieved from https://ec.europa.eu/eurostat/data/database

EU-SILC. (2019). European Union Statistics on Income and Living Conditions. 2008 Module.

Farmer, J. D., & Foley, D. (2009). The Economy Needs Agent Based Modelling. *Nature*, *460*(August), 685–686. https://doi.org/doi:10.1038/460685a

Friege, J. (2016). Increasing homeowners' insulation activity in Germany: An empirically grounded agent-based model analysis. *Energy and Buildings*, *128*, 756–771. https://doi.org/10.1016/J.ENBUILD.2016.07.042

Hesselink, L. X. W., & Chappin, E. J. L. (2019). Adoption of energy efficient technologies by households--Barriers, policies and agent-based modelling studies. *Renewable and Sustainable Energy Reviews*, *99*, 29–41.

Hong, T., Taylor-Lange, S. C., D'Oca, S., Yan, D., & Corgnati, S. P. (2016). Advances in research and applications of energy-related occupant behavior in buildings. *Energy and Buildings*, *116*, 694–702.

Hrovatin, N., & Zorić, J. (2018). Determinants of energy-efficient home retrofits in Slovenia: The role of information sources. *Energy and Buildings*, *180*, 42–50. https://doi.org/10.1016/J.ENBUILD.2018.09.029

IEA. (2017). *Energy Technology Perspectives*. Paris, France.

Klöckner, C. A., & Nayum, A. (2017). Psychological and structural facilitators and barriers to energy upgrades of the privately owned building stock. *Energy*, *140*, 1005–1017. https://doi.org/10.1016/J.ENERGY.2017.09.016

Krey, V. (2014). Global energy-climate scenarios and models: A review. *Wiley Interdisciplinary Reviews: Energy and Environment*, *3*(4), 363–383. https://doi.org/10.1002/wene.98

Kriegler, E., Weyant, J. P., Blanford, G. J., Krey, V., Clarke, L., Edmonds, J., … others. (2014). The role of technology for achieving climate policy objectives: overview of the EMF 27 study on global technology and climate policy strategies. *Climatic Change*, *123*(3–4), 353–367.

Lopes, M. A. R., Antunes, C. H., & Martins, N. (2015). Towards more effective behavioural energy policy: An integrative modelling approach to residential energy consumption in

Europe. *Energy Research & Social Science*, *7*, 84–98.

Mundaca, L., Neij, L., Worrell, E., & McNeil, M. (2010). Evaluating energy efficiency policies with energy-economy models. *Annual Review of Environment and Resources*, *35*, 305–344.

Nair, G., Gustavsson, L., & Mahapatra, K. (2010). Factors influencing energy efficiency investments in existing Swedish residential buildings. *Energy Policy*, *38*(6), 2956–2963. https://doi.org/10.1016/J.ENPOL.2010.01.033

Nakamura, E. (2016). Electricity saving behavior of households by making efforts, replacing appliances, and renovations: empirical analysis using a multivariate ordered probit model. *International Journal of Consumer Studies*, *40*(6), 675–684. https://doi.org/10.1111/ijcs.12292

Pettifor, H., Wilson, C., McCollum, D., & Edelenbosch, O. Y. (2017). Modelling social influence and cultural variation in global low-carbon vehicle transitions. *Global Environmental Change*, *47*. https://doi.org/10.1016/j.gloenvcha.2017.09.008

Rai, V., & Henry, A. D. (2016). Agent-based modelling of consumer energy choices. *Nature Climate Change*, *6*(6), 556–562. https://doi.org/10.1038/nclimate2967

Rai, V., & Robinson, S. A. (2015). Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental factors. *Environmental Modelling and Software*, *70*, 163–177. https://doi.org/10.1016/j.envsoft.2015.04.014

Sachs, J., Meng, Y., Giarola, S., & Hawkes, A. (2019). An agent-based model for energy investment decisions in the residential sector. *Energy*.

Sovacool, B. K., Ryan, S. E., Stern, P. C., Janda, K., Rochlin, G., Spreng, D., … Lutzenhiser, L. (2015). Integrating social science in energy research. *Energy Research & Social Science*, *6*, 95–99.

Tavoni, M., Kriegler, E., Riahi, K., Van Vuuren, D. P., Aboumahboub, T., Bowen, A., … others. (2015). Post-2020 climate agreements in the major economies assessed in the light of global models. *Nature Climate Change*, *5*(2), 119.

Trotta, G. (2018). Factors affecting energy-saving behaviours and energy efficiency investments in British households. *Energy Policy*, *114*, 529–539. https://doi.org/10.1016/J.ENPOL.2017.12.042

Wilson, C., Crane, L., & Chryssochoidis, G. (2014). Why do people decide to renovate their homes to improve energy efficiency? Retrieved from http://www.tyndall.ac.uk/sites/default/files/publications/twp160.pdf

Wilson, C., & Dowlatabadi, H. (2007). Models of Decision Making and Residential Energy Use. *Annual Review of Environment and Resources*, *32*(1), 169–203. https://doi.org/10.1146/annurev.energy.32.053006.141137