Aalborg Universitet

Variances in residential heating consumption – Importance of building characteristics and occupants analysed by movers and stayers

van den Brom, Paula; Hansen, Anders Rhiger; Gram-Hanssen, Kirsten; Meijer, Arjen; Visscher, Henk

[Link to publication from Aalborg University](#)

Variances in residential heating consumption – Importance of building characteristics and occupants analysed by movers and stayers

Paula van den Brom[a,*], Anders Rhiger Hansen[b], Kirsten Gram-Hanssen[b], Arjen Meijer[a], Henk Visscher[a]

[a] Delft University of Technology, Faculty of Architecture and the Built Environment, OTB-Research for the Built Environment, P.O. Box 5043, 2600 GA Delft, the Netherlands
[b] Danish Building Research Institute, Aalborg University, A.C. Meyers Vænge 15, DK-2450 Copenhagen SV, Denmark

HIGHLIGHTS

- Longitudinal analysis of heating consumption data from the Netherlands and Denmark.
- Comparing variances in consumption of 887,685 moving and staying households.
- Occupants and buildings each explain half of the variation in heating consumption.
- The relative influence of occupants differs for different building characteristics.

ARTICLE INFO

ABSTRACT

It is commonly accepted that occupants have a significant influence on the variation in residential heating consumption. However, the scale of that influence lacks empirical investigation. The aim of this study was to distinguish which part of the variance in actual residential heating consumption can be attributed to the occupants, and which part to the building itself. This was achieved by applying and extending a method suggested by Sonderegger in 1978, using updated and significantly improved data from two different countries: the Netherlands and Denmark. These data contain different types of heating supply systems (district heating and natural gas) and different housing forms (multi and single-family social housing, and private detached single-family houses). For the studied databases, the results indicate that approximately 50% of the variance in heating consumption between houses can be explained by differences related to occupants. The other 50% can be explained by the characteristics of the building itself and other physical parameters, which are often not taken into account in simulation models of heat transmission within buildings. Additional analyses indicate that the relative influence of occupants on heating consumption differs depending on the building characteristics of the dwelling. For example, the influence of occupants is larger when the building is more energy efficient. Based on the research results, it can be concluded that it is unrealistic to aim for a building simulation model that perfectly projects residential heating consumption for individual cases. However, creating building simulation models and occupant consumption profiles that accurately represent average residential heating consumption should be possible.

1. Introduction

Household energy consumption is estimated to be responsible for approximately 26% of the total energy consumption in Europe [1]. Therefore, policymakers see a large potential for energy savings in this sector. However, previous studies have indicated that thermal renovations often result in lower energy savings than expected [2]. This discrepancy between actual and theoretical savings is caused (among other factors) by the energy performance gap (EPG), which is the discrepancy between actual and calculated energy consumption of a household. The EPG illustrates that it is not possible to explain residential energy consumption by solely relying on building simulation models [3]. Several studies have also demonstrated that residential energy consumption varies largely due to the characteristics of the

Nomenclature

DHW	domestic hot water
$Adjheat_t$	standardised heat consumption year t [kWh]
$heat_t$	annual heat consumption year t [kWh]
avg_heat$_{2010}$	average annual heat consumption of the year 2010 [kWh]
avg_heat$_{2010}$	average annual heat consumption of the year 2015 [kWh]
Cv	Coefficient of variance
Sd	Standard deviation
$Nheat_t$	normalised heat consumption year t [kWh]
c_t	constant, result from linear regression year t
b_t	coefficient, result from linear regression year t
LRC	logarithm of relative heat consumption [-]
Var_{max}	maximum variance
σ_t^2	variance year t
$\sigma^2[LRC^{stayers}]$	variance in heating consumption of 'stayers': due to changes in heating consumption of the same occupants over time (SO)
$\sigma^2[LRC^{movers}]$	variance in heating consumption of 'movers' due to

changes in heating consumption of the same occupants over time (SO) **and** variance due to changes in heating consumption due to new occupants moving into the house (NO)

$\sigma^2[Var_{max}]$ maximum variance in heating consumption, when everything is different compared to the previous period. Due to changes in heating consumption of the same occupants over time (SO) **and** changes in heating consumption due to new occupants moving into the house (NO) **and** change of physical characteristics that are not taken into account in the linear regression model (Ph)

SO	changes in heating consumption over time of the same occupants [%]
NO	changes in heating consumption due to new occupants moving into the house [%]
Ph	Physical characteristics that are not taken into account in the linear regression analyses [%]
AB	building characteristics that were available in the database and are taken into account in the linear regression [%]

occupants as indicators of behavioural patterns [4–6]. For example incomes in England were found to be positively correlated with the actual energy consumption in a household [5] and a larger number of household members also results in higher energy consumption, but it decreases the energy consumption per person [6]. Age is found to be the most determining indirect effect on heating [4].

Based on previous studies, it is expected that occupants play an important role in this EPG, but the scale of this role is unclear [7]. Some researchers even expect the occupant role to be more important than the role of building characteristics [8,9]. Sonderegger [10] was one of the first who attempted to define the extent to which occupants are responsible for the variance in energy consumption among similar houses, by studying *movers* (houses with changed occupants) and *stayers* (houses with the same occupant over time). Accordingly, Sonderegger compared the variance in energy consumption of houses with movers and houses with stayers. The aim of his method was to define the extent to which the variance in residential energy consumption was related to either occupants or building characteristics.

This study applies Sonderegger's method to two significantly larger and more diverse datasets from the Netherlands and Denmark. This means that our data contains almost one million houses and households, compared to the 200 similar houses in Sonderegger's study. This comparative design enables a stronger generalisability of the results, which is seldom seen in quantitative energy consumption studies. Because many researchers found a relation between building characteristics and occupant behaviour, the analyses are extended by studying whether the influence of occupant behaviour depends on the building characteristics.

By doing this, the importance of the role of occupants for understanding variation in energy consumption among households is indicated, and the interaction of different types of building characteristics with the behaviour of occupants is shown. Knowing how much of the variance in energy consumption is caused by occupants enables a better insight in how to interpret the energy consumption results and how much variance in energy consumption can be expected due to variation in occupant behaviour. The results also indicate over which range the energy simulation can expected to be assumed to be correct. Further, the paper will show which part of the variance can be explained by the physical characteristics that are not taken into account in the energy simulation.

This paper first reviews research studies investigating the influence of the occupant on residential energy consumption. This section is

followed by an explanation of the data used for this study, an explanation of Sonderegger's method, and how this method is adapted to make it suitable for our datasets. Then, the results of the analysis are presented. In the discussion section, the authors consider both the advantages and disadvantages of the adapted method and the data used. Finally, conclusions are drawn in the final section.

2. Literature review

Many researchers have already investigated variations in residential energy consumption in similar dwellings, and sought to explain the reasons for the variance in energy consumption among similar dwellings. In this literature review, an overview of studies on this topic is provided, and the research results, applied methods, and type and origins of the data are discussed. The aim of this review is to indicate current knowledge about the influence of occupants on building-related energy consumption and to define how this study could contribute to further insights.

The literature for this review was selected based on the following conditions: First, the aim of the research must include a better understanding of residential energy consumption and the influence of occupants; Second, the research must be based on measured data/post-occupancy data. This means that studies using simulated data were excluded from this literature review. The reasoning behind this that the use of simulated data is a simplification of reality, and therefore does not reflect the complexity of actual energy consumption. Finally, only references from academic journal papers are used.

2.1. Comparing results

Table 1 shows a summary of the literature review, and the first column lists the aims of the study. Although the aims of the studies appear similar, the results and conclusions vary. All studies concluded that occupants and their behaviour play a significant role in the amount of residential energy consumption. However, the amount of the impact is different across the studies, with some claiming that occupants are the most influential factor. For example, Steemers and Yun [5] found that the roles of occupant behaviour and socio-economic factors are the most important components for determining residential energy consumption. According to their research, the physical characteristics of dwellings (such as construction year, type and floor area) are less important. However, it should be taken into account that they also

considered that the type of heating and/or cooling system and its control to be a decision of the occupant, and thus a behavioural factor.

Other studies concluded that the building characteristics are the principal determining factor for residential energy consumption. For example, Guerra Santin et al. [11] found that 42% of residential energy consumption can be determined by the building characteristics, and only 4.2% by occupant characteristics. In this study, it has to be taken into consideration that Guerra Santin et al. [11] used the linear regression to determine those percentages with the building characteristics, and subsequently added the occupant characteristics. Therefore, they did not consider possible relationships between occupant behaviour and building characteristics. These results might have been different if they had started with the occupant characteristics. Huebner et al. [12] found that building characteristics account for approximately 39% of the variability in energy consumption, socio-demographic factors are 24%, heating behaviour is 14%, and attitudes and other behaviour account for only 5%. However, a combined model including all predictors explains only 44% of all variability. Sonderegger [10] found that 54% of the variance in energy consumption among similar buildings could be explained by "obvious building characteristics", 15% by the change of occupants, 17% by lifestyle, and 13% by house-related quality differences. The obvious building characteristics referred to by Sonderegger include for example the number of bedrooms, which he takes into account by applying a regression analysis. House related quality differences are the physical characteristics of the house that are not considered in the regression model, for example, if a tree blocks the solar radiation. Further, Brounen et al. [13] found that residential heating consumption is primarily determined by the building characteristics, such as its construction year or type.

Other studies found the same (or almost the same) impact level of building and occupant characteristics on residential energy consumption. For example, Verhallen and Raaij [14] discovered that household behaviour explains 26% of residential energy consumption, and house characteristics explain 24%. They also found an interaction between building characteristics and residential energy consumption. As an illustration, house insulation has a positive effect because people tend to lower their thermostat settings more often, and they are more likely to open their windows more frequently. Similarly, a recent study [15] investigated how occupant behaviour is related to building characteristics (including heating and ventilation installations and building year). Gill et al. [16] found that energy efficiency behaviour accounts for 51% of the variance in heat consumption between dwellings. However, they explicitly state that behaviour is not claimed to be the dominant factor.

Several aspects can explain why the conclusions differ although the aim of the studies is similar. For example, the sample size and the level of detail of the collected data differ significantly between studies. Comparing the research of Spataru et al. [17] and the study of Brounen et al. [13] similar aims can be ascertained, but the data and focus of the researchers are completely different. The first used highly detailed monitoring data from a single house, while the latter used a large but more aggregated database containing information of one million houses and their occupants. Unavoidably, this results in different types of research and different research results.

In addition, the starting point of the researcher (and the definition of the influence of the occupant on residential energy consumption) can mean that those studies with similar aims arrive at different conclusions. For example, all studies indicated that occupants have a significant influence on residential energy consumption. However, there is discussion about the magnitude of this influence, and whether it is more influential than, building characteristics. One of the reasons for these different research results is the different starting point of the research. Some researchers take the house and its physical characteristics as a starting point [18], while others focus on the occupant. Here, they assume the occupant chooses the house and therefore the influence of this choice is part of the influence of the occupant on residential energy

consumption [8]. Often, when the first starting point is applied, the building characteristics seem to be more important. Conversely, when the second starting point is applied, occupant influence appears to be more important. Several studies have indicated an awareness of these direct and indirect effects [5,19,20]. For example: Steemers and Yun [5] demonstrated that behavioural, physical and socio-economic parameters have direct and indirect influence of energy use; and Estiri [20] showed that household characteristics have almost the same impact on building energy consumption as building characteristics, if not only their direct effect but also their indirect effects are taken into account.

2.2. Occupant characteristics

Many of the studies use occupant characteristics to indicate the influence of the role of occupants on residential energy consumption. The main reason for this is that occupant characteristics are easier to collect than (for example) detailed behavioural indicators, and they are available for a higher number of households. As several studies suggest that occupant characteristics indicate occupant behaviour, it also appears a sensible approach. Several occupant characteristics are found to correlate with actual energy consumption. The strongest and most frequently-mentioned correlations are those between the number of occupants [4,12,18–24], and income [5,12,19,20,22,25].

2.3. Statistical methods

While the studies have differences in data and focus, their statistical methods are similar. Almost all studies use cross-sectional statistical analysis[1] techniques, with the majority using linear regression or multiple linear regression analysis. Within studies on the impact of prices on residential energy consumption, panel data are more frequently used [26,27]. In our literature review, only the study of Sonderegger [15] makes use of longitudinal/panel data.[2] In his research, 205 similar houses were monitored for 3 years (1971–1973). The resulting data included energy consumption figures, building characteristics, and which occupants were living the house during the monitored years. The research is based on the assumption that if the occupants remain the same, energy consumption will be more constant over time than if they move and are replaced by other occupants.

Conducting energy consumption research can benefit significantly from longitudinal data and the accompanying statistical data analysis techniques. In the past, many studies used data from similar houses to compare the influence of the occupant on residential energy consumption. However, no houses are exactly similar, owing to different locations and layouts. Therefore, longitudinal data and the accompanying statistical data analysis techniques are highly beneficial for conducting energy consumption research. For example, multiple houses over time can be monitored, and the direct influence of the building characteristics can be excluded from the analysis because these factors remain the same (assuming that the house is not renovated). This presents significant potential for evaluating the effect of policy changes, newly installed technologies and renovations.

2.4. Conclusions of the literature review

Based on this literature study, it can be concluded that determining the effect of the occupant behaviour on residential energy consumption is highly dependent on the boundaries that the researcher set for the term occupant influence. The results of determining the influence of occupants on residential energy consumption varied from 4.2% to more than 50%. Furthermore, if longitudinal data are available then the

---

[1] *Cross sectional data is data of many different subjects at the same point of time.*
[2] *Longitudinal/panel data is data of many different subject that are followed over multiple points in time.*

Table 1
Literature overview of studies that aim to get a better insight into residential energy consumption and the reason for the variance in energy consumption among similar dwellings.

Aim/research question	Data, type, country	Method	Conclusion	Ref.
To determine the factors responsible for the remaining 46% variation that cannot be explained by conventional factors	*Twin rivers project, 248 townhouses *Monthly electric and natural gas meter readings *UK	Regression, three-factor multiplicative model	54% of the variance is explained by obvious building characteristics, change of occupants explains 15%, lifestyle explains 17% and persistent house-related quality differences explain 13%	[10]
To determine the factors that determine energy use for home heating are investigated in this study	*145 similar houses 79 with standard insulation and 78 with superior insulation *Natural gas meters, 4 moments in time *The Netherlands	Factor analysis	* Home characteristics, special circumstances, and sociodemographic together explain 58% of the energy use variance. * Household behaviour alone explains 26% * Home characteristics alone 24% * Special circumstances alone explain 11% of the energy-use variance	[14]
To determine to what extent consumer behaviour influences space heating energy demand and test the linear approach describing space heating energy demand by means of a simple linear dependence on climate (heating degree days) and the thermal quality of a building (heat load)	* 400 households * Data on energy consumption (without electricity demand for appliances) by fuel type are available for at least 1 year, in most cases for 2 or 3 years. sociological, and structural data * Austria	Service factor analysis	* The result of this investigation provides evidence of a rebound-effect of about 15 to 30% due to building retrofit	[28]
To determine to what extent energy performance is determined by interactions between occupants, behaviour and buildings systems, as well as building and climate characteristics establish	*3358 housing units for heating and 2718 housing units for cooling climate * Actual energy consumption data for heating and cooling and building, occupant behaviour and socioeconomic characteristics data * 50 states in US	Regression models and path analysis	* Climate and building characteristics alone are insufficient as determinants of energy demand * Most significant parameter is climate. Second is a set of parameters related to occupant behaviour, specifically in terms of the choices made about heating and cooling systems and their control * Less important than might be expected are some physical characteristics of the dwellings	[5]
To gain greater insight into the effect of occupant behaviour on energy consumption for space heating by determining its effect on the variation of energy consumption in dwellings while controlling for building characteristics	* 15,000 interview-based survey * 3 years of heating (gas consumption data) including household characteristics and use of the dwelling * The Netherlands	ANOVA & multiple regression analysis	* Building characteristics determine 42% of the energy use in a dwelling * Adding occupant characteristics and behaviour increases the explanation factor with 4.2%	[11]
To determine the direct, indirect, and total impacts of household and building characteristics on residential energy consumption	* Microdata from the 13th Residential Energy Consumption Survey (RECS) * Total household energy consumption * US	Structural equation modelling	* The direct impact of household characteristics on residential energy consumption is significantly smaller than the indirect impact *Taking both direct and indirect impact into account the total impact of households on energy consumption is only slightly smaller than that of building characteristics.	[20]
Understanding the spectrum of residential energy consumption	* Residential Energy Consumption Survey (RECS) public use microdata set * Total household energy consumption * US	Quantile regression analysis	* Results show that housing size matters for space conditioning * Housing type has a more nuanced impact. * Some, not all, types of multifamily housing offer almost as much savings as a reduction in housing area by 100 m², compared to single-family houses	[24]
Identifying the key determinants and effects of occupants' behaviour on energy use for space heating	* 313 household * Annual gas consumption * The Netherlands	Pearson correlation samples t-test, ANOVA, Chi-square regression model	* Interaction between occupant behaviour and building characteristics are found * Occupant behaviour (indirect and direct) can predict 11,9% of the variation in energy use	[18]
To evaluate the relationships between occupancy and energy usage, as well to diagnose the performance and energy efficiency	* 1 house, one family was extensively monitored * Energy consumption for heating * UK		* In order to reach the 2050 target to reduce carbon emissions by 80%, the behaviour of the occupant is increasingly important, being responsible for the energy consumption in the building	[17]
The contribution of behaviours to actual performance	* 26 similar dwellings * Domestic electricity heat and water consumption and occupant behaviour * UK	Linear regression	* Energy-efficiency behaviours account for 51% of the variance in heat consumption in dwellings * 37% of the variance in electricity consumption can be explained by energy behaviour * And 11% of the variance in water consumption can be explained by energy behaviour	[16]

(continued on next page)

Table 1 (*continued*)

Aim/research question	Data, type, country	Method	Conclusion	Ref.
To identify the influences of the occupant behaviour on the building energy consumption.	* Annual building energy use intensity (EUI) 2003 * Annual energy consumption * Japan	Cluster analysis, Grey relational analysis	* Weather conditions significantly influenced occupant behaviour, thereby impacting building energy consumption * Households tended to maintain their lifestyles, and the level of their general indoor activities associated with these end-use loads did not fluctuate widely from month to month	[29]
To determine if energy efficiency of appliances and houses or user behaviour is the more important	* 50,000 households * Meter readings heating and electricity consumption, socio-economic information on their inhabitants, building information * Denmark	Regression and literature study	* User behaviour is at least as important as the efficiency of technology when explaining households' energy consumption in Denmark	[8]
Determining the extent to which the use of gas and electricity is determined by the technical specification of dwellings as compared to the demographic characteristics of the residents	* 3,000,000 Dutch homes and their occupants* annual gas and electricity consumption* the Netherlands	Regression	* Residential gas consumption is determined principally by structural dwellings characteristics, such as the vintage, building type, and characteristics of the dwelling * While electricity consumption varies more directly with household composition, in particular income and family composition	[13]
To determine the impact of occupants on residential energy consumption in China	* 642 surveys related to behaviour and energy use in winter and 838 surveys in summer* household energy data building and occupant characteristics and behaviour* China, Hangzhou	Bivariate correlation, path, and multiple linear regression analysis	* Household socio-economic and behaviour variables are able to explain 28.8% of the variation in heating and cooling energy consumption	[21]
To what extent different types of variables (building factors, socio-demographics, attitudes and self-reported behaviours) explain annualized energy consumption in residential buildings	* data from a sample of 924 English households collected in 2011/12* annual energy consumption* England	Lasso regression	* Building variables on their own explained about 39% of the variability in energy consumption * Socio-demographic variables 24% * Heating behaviour 14% * Attitudes &other behaviours only 5% * A combined model encompassing all predictors explained only 44% of all variability, indicating a significant extent of multicollinearity between predictors	[12]
Socio-cultural differences in heat consumption	* household data and building characteristics data* households' annual heat consumption for space heating and heating of hot water* Denmark	Regression	* Households' heat consumption levels vary across social groups * Social groups indicate differences in heating-consuming habits * The results of the paper indicate that around one-third of the impact of educational and income differences between households on heat consumption are due to differences in heat-consuming habits (direct effect), whereas the rest, two thirds, are due to differences in households and houses (indirect effects)	[19]
To provide a better understanding of the main determinants of residential energy consumption in order to guide energy policymaking	* Survey data 36,000 occupants, national housing survey * Household energy consumption *France	Multiple correspondence analysis and ascending hierarchical classifications, OLS regression	*Energy prices were the most important factors determining domestic energy consumption * Occupant characteristics significantly affect domestic energy use	[30]

research should benefit from its possibilities. Further, most studies on the influence of occupants on residential energy consumption are based on one dataset from one country or region. Moreover, the literature review indicates that all studies acknowledge that occupants affect actual energy consumption but the degree of influence varies between the studies. A lack of large databases and detailed building and occupant data makes it difficult to establish a constant value or even a range for such influence, since many of the previous studies have been conducted on small databases.

3. Data

Two databases are used in this study: one with data from Dutch houses and households and one from Danish houses and households. This section explains the two datasets and how they are used in this study. The first part explains the Dutch database and the second part the Danish database.

3.1. Dutch data

The Dutch data originate from two different sources. The first one is the SHAERE database, which is a database from Dutch social housing organisations in the Netherlands. It is primarily used to monitor energy efficiency and contains 60% of the Dutch social housing stock. Of the total housing stock, social housing stock in the Netherlands is relatively large compared to other countries, accounting for 30%. This means the database contains a significant share of all houses in the Netherlands. Within these houses in the database, 46.9% are single family houses and 53.1% are multi-family houses. For single-family houses, the vast majority are terraced. The database contains most of the input variables that are used to calculate the energy performance of houses, the energy performance certificate, and predicted energy consumption per house for six years (2010–2015). This dataset is combined with actual annual energy consumption data from Statistics Netherlands. Energy consumption data are considered private (sensitive information); therefore, it is only allowed to publish the results on an aggregated level. Apart from actual energy consumption data the Statistics Netherlands database also contains occupant characteristics data (such as income, number of household members, and employment status).

Approximately 95% of Dutch households use gas as a heating source for their house [31]. In countries such as the Netherlands and Denmark, energy for heating constitutes the main energy demand of a house. Further, energy consumption for heating has the highest energy performance gap. Therefore, only houses that use gas as a heating source are studied. This enables us to distinguish energy consumed for heating and domestic hot water (and sometimes cooking) on one side and energy consumed for electrical appliances on the other side. Because domestic hot water is on average a relatively small part of the gas consumption of Dutch houses from now gas consumption will be referred to as the energy used for heating. However the amount of gas consumption for domestic hot water is significant (in the Netherlands

on average 16%) and therefore it is important that the reader should be aware that this is included in the term "heating consumption" [32]. Energy supply companies in the Netherlands are only obligated to report actual energy consumption every three years. If the data is not reported, energy consumption data of the previous year is used and therefore all cases with exactly the same gas consumption as the previous year are deleted (approximately 15% of the total amount of cases). It is assumed highly unlikely that a household would use precisely the same amount of gas every year.

Houses with collective installation systems are deleted from the database because the Dutch statistical experts expressed doubts about the quality of this data. Further, because the databases that we use are relatively large, there is an increased probability of them recording unrealistic values that might affect the results. To avoid possible bias of those unrealistic values and errors biasing the results, the highest and lowest 1% of household energy consumption (kWh) and area (m2) are removed for each year in the analysis. Because the relative energy consumption is used in this study (explained in section 5 energy consumption 2015/energy consumption 2010), cases with a relative consumption higher than 12 were deleted. This is because some extreme values were found that are highly unlikely and yet have a significant influence on the mean (891 cases), so they can be considered outliers. For this analysis, it is important that the building characteristics are constant. Therefore, dwellings with changed building characteristics (such as renovations or administrative corrections) are deleted (approximately 30% of the cases). Finally, only cases that had at least an energy consumption record for the years 2010 and 2015, and a theoretical energy consumption record for at least one year are taken into account. After filtering, data on 375,382 houses remained.

3.2. Danish data

The Danish data came from two sources. Data on building and household characteristics were taken from Statistics Denmark's administrative registers, which covers the full population. These were, merged with data on household energy consumption for space heating and hot water from the Danish Building and Dwelling Register (BBR), which is part of the Danish Ministry of Taxation. Heat supply utilities in Denmark are required by law to submit household energy consumption data to BBR, who subsequently compile and prepare data for research and other purposes. The administrative data from Statistics Denmark is accessible in anonymised form through an online server.

The data are registered on housing units. Therefore, the used data on energy consumption is from single-family detached houses that are individually metered to avoid uncertainties about which households the consumption relates to. Single-family houses are the predominant type of housing in Denmark, accounting for 44% of the housing stock in 2014 (Statistics Denmark). Further, in the Danish sample, 92.57% of the houses are owner-occupied. Data for houses with an individual heat supply (for example oil-fired boiler) has some uncertainties regarding the periodisation of yearly energy consumption because it is not clear at

Table 2

Variables used in the regression model as building characteristics for Danish dataset.

| Variable name	Variable description	Categories
Gas	Heating supply: natural gas or district heating?	0 = District heating; 1 = Natural gas
Area	Heated area (m2)	Continuous
Rooms	Number of rooms	Count
Woodstove	Do the house have a woodstove or fireplace?	1 = yes
Attic floor	Do the house have an attic floor?	1 = yes
Basement	Do the house have a basement?	1 = yes
Roof	Roof material	1 = fibrecement; 2 = cement stone; 3 = tile, 4 = other material
Exteriorwall	Exterior wall material	1 = Bricks; 2 = Wood; 3 = Concrete; 4 = Other material
Buildingperiod7a	Building period in 7 categories	1 = < 1938; 2 = 1938–1960; 3 = 1961–1972; 4 = 1973–1978; 5 = 1979–1998; 6 = 1999–2006; 7 = > 2006
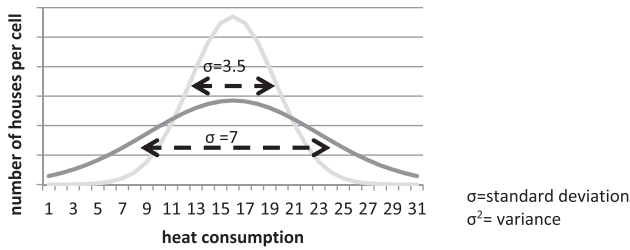
Fig. 1. Fictive normal distributions to show the effect of data with the same mean but a different variance.

what time the fuel is used. Therefore, data is restricted to houses supplied with district heating or gas, which together supplied 78% of Danish households in 2015 (Statistics Denmark). By law, all households in Denmark have individual metering of their energy consumption, independently if the supply is by gas or by district heating. By restricting the study to households supplied with district heating, or a gas supply that has registered heat consumption, the data covers approximately 64% of all single-family detached houses in Denmark. It is not possible to distinguish between energy used for space-heating and domestic hot water, but it is estimated that space-heating accounts for approximately 80%, while the remainder is for domestic hot water [33]. However, in newer houses the percentage attributes to space heating might be lower due to their higher energy efficiency. To mitigate the risk of unrealistic values and errors biasing the results, the highest and lowest 1% of household energy consumption (kWh) and areas (m^2) are removed for each year in the analysis. Moreover, the sample was restricted to domestic housing, not for business. Further, if the house had no registered occupants, its data were removed from the sample. Taken together, this removed approximately 17% of the observations. Finally, 1,425 observations were removed because their consumption in 2015 was more than five times higher than in 2010. Also 27,547 observations were removed because they did not have the same building characteristics registered in 2010 and 2015. After filtering, data of 512,393 houses remained.

Table 2 shows the variables used in the regression as building characteristics.

4. Method

This section explains the method used in this study, which is based on the method proposed by Sonderegger [10]. This method is based on the difference in variance between movers and stayers. Therefore, this methodology section starts by describing how movers and stayers are identified. This is followed by an explanation of Sonderegger's method, which describes step-by-step how the method was applied, and how it was made applicable for our data. This description also explains why the variance in relative heat consumption instead of the average relative heat consumption is studied (heat consumption 2015 divided by heat consumption 2010). Further, it should be mentioned that when heating consumption is referred to in the text, this also includes energy consumption for domestic hot water. This is included because the amount of energy consumed for hot water is relatively small compared to energy used for heating (approximately 20%) [33]. Energy for *Domestic Hot Water* (DHW) is, compared to energy for heating, less dependent on the technical characteristics of a building. The amount of energy consumption for DHW will be relatively large for energy-efficient buildings compared to relatively energy-inefficient buildings, because the energy demand for heating is in energy-efficient buildings is lower than in energy-inefficient buildings, while the domestic hot water demand is not influenced by the energy-efficiency of the building. This is something to be aware of because it allows for possible bias.

4.1. Identifying movers and stayers

To identify movers and stayers in the databases, it was determined whether the reference person in a household stayed the same or changed between 2010 and 2015. For the Dutch case, the reference person of a house is already identified in Statistics Netherlands data. For the Danish case, the oldest person in the house is selected as the reference person (if two people have exactly the same age, one is randomly chosen). This method could cause some bias because it is possible that the reference person will leave the house but the others will stay (or the other way around). However, given the large size of the datasets, this is considered acceptable, and so the authors do not expect those cases to influence the results significantly.

4.2. Method description

The starting point of Sonderegger's method is the assumption that the heat consumptions of two different time periods will have a higher correlation for houses with the same occupant than for houses with different occupants, because occupants continue to have the same behaviour over time. To investigate this, a comparison is made of the variance in relative heat consumption of a group of houses where occupants remained the same (stayers) and a group where occupants changed (movers). The <u>variance</u> of relative heat consumption and not the <u>mean</u> is chosen for study, because the variance shows how far the relative heat consumption of different cases is distributed. A large variance would mean that the spread of the relative heating consumption is wide, whereas a small variance would mean the opposite (Fig. 1)

The analyses used heat consumption data from 2010 and 2015. To make the heat consumption of those two years comparable, a standardization method is applied: the heat consumption of 2015 is multiplied by the ratio of the means of the years 2010 and 2015 (Eq. (1)), Doing this ensures the removal of variances in heat consumption due to weather and other external factors.

$$adjheat_{2015} = heat_{2015}\left(\frac{avg\_heat_{2010}}{avg\_heat_{2015}}\right) \tag{1}$$

$adjheat_{2015}$ = standardised heat consumption 2015
avg_heat_{2010} = average annual heat consumption 2010
avg_heat_{2015} = average heat consumption 2015
$heat_{2015}$ = annual heat consumption 2015 for individual house

The standardisation is followed by a linear regression, where the dependent variable = actual heat consumption, and the independent variable = theoretical heat consumption/building characteristics. This linear regression is conducted for two reasons: 1. To determine which part of the variance in energy consumption for heating can be explained by the available building characteristics (AB) in the database; and 2. To make the buildings comparable. The regression coefficients are used to normalise the heating consumption, which makes the buildings comparable even though they have different building characteristics.

$$Nheat_{2010} = c_{2010} + b_{2010}\cdot heat_{2010} \tag{2}$$

$$Nheat_{2015} = c_{2015} + b_{2015}\cdot adjheat_{2015} \tag{3}$$

$Nheat_t$ = normalised heat consumption year t
c_t = constant, result from linear regression year t
b_t = coefficient, result from linear regression year t

Then, the relative heat consumption is determined, which is the normalised heat consumption of 2015 divided by the normalised heat consumption of the year 2010 (Eq. (4)). When the relative heat consumption is close to one it means the heat consumptions of 2010 and 2015 are similar, whereas values of lower than one means the heat consumption of 2015 was lower than in 2010. Further, a figure higher than one means the heat consumption of 2015 was higher than in 2010. To make the data useful for further comparison, the natural logarithm

of the relative consumption is calculated.

$$LRC = \ln\left(\frac{Nheat_{2015}}{Nheat_{2010}}\right) \qquad (4)$$

LRC = logarithm of relative heat consumption

This makes the variance of relative heat consumption of movers and stayers comparable. However, to determine how much influence the movers and stayers have on the variance, first the maximum possible variance has to be determined. This maximum possible variance is determined by adding up the variance of movers in 2010 and 2015[3]. This would be the variance if the consumption level of each house in the second period were totally unrelated to its own level in the first period. Because the logarithm of relative heat consumption is used also the variance of the logarithm of heat consumption for 2010 and 2015 should be used.

$$Var_{max} = \sigma^2_{2010} + \sigma^2_{2015} \qquad (5)$$

Var_{max} = maximum variance
σ^2_t = log variance year t

The following assumptions are crucial for understanding how to define which part of the variance in heating consumption is due to occupants and which part is due to the building characteristics. This study assumes that the heat consumption in houses with the same occupant(s) (stayers) for the two periods would result in a higher correlation of heat consumption between those periods than that in houses with changed occupant(s) (movers). This assumption is made because occupants are expected to have a rather stable heating consumption pattern over time, for example, due to energy consumption practices and comfort expectations that gets embodied and 'carried' from one situation to the next [34,35]. Energy consumption practices referto routinized forms of behaviour that occupants perform in their everyday life, and although such practices have some continuity over time, they are also in constant change, for example in relation to new material surroundings [36,37]. Therefore, occupants are expected to change consumption patterns over time, especially when moving into a new house. Thus, this study distinguishes between two types of changes over time. The first type relates to houses where the occupants do not move, which is expressed in the variance of the logarithm relative heating consumption of the stayers in this research. To these occupants the changes will be referred to as '*changes in heating consumption of the same occupants over time*' (SO). The second type relates to houses where the occupants change because new occupants move in (movers). It is expected that the practices performed by the previous (in 2010) and the new occupants (in 2015) have some similarities because they are performed in more or less the same material surroundings. However, it is also expected that the heating consumption in the 'movers' group changes over time because the occupants in the house are new due to the interaction between the practices that the occupants 'carry' with them and the new material surroundings of the occupants, resulting in completely different consumption patterns. These changes are referred to as '*changes in heating consumption due to new occupants*' (NO). Finally, the linear regression is demonstrated on the variances due to '*available building characteristics*' (AB). For the Dutch case, theoretical heat consumption was available, and for the Danish case, the characteristics are mentioned in Table 2. However, the '*available building characteristics*' (AB) in the databases are probably not the only physical characteristics that explain part of the variance in energy consumption among houses. It is expected that there will be other physical aspects that account for the variance of heat consumption, which will be indicated by the maximum variance in heat consumption. Based on these assumptions, the variance in heat consumption can be explained as follows:

---

[3] Based on the law of propagation of variance of uncorrelated factors.

$\sigma^2[LRC^{stayers}]$	variance in heating consumption of 'stayers': due to changes in heating consumption of the same occupants over time (SO)
$\sigma^2[LRC^{movers}]$	variance in heating consumption of 'movers' due to changes in heating consumption of the same occupants over time (SO) and variance due to changes in heating consumption due to new occupants moving into the house (NO)
$\sigma^2[Var_{max}]$	maximum variance in heating consumption, when everything is different compared to the previous period. Due to changes in heating consumption of the same occupants over time (SO) **and** changes in heating consumption due to new occupants moving into the house (NO) **and** change of physical characteristics that are not taken into account in the linear regression model (Ph)

Following these assumptions, it is possible to calculate how much of the variance is due to 'changes in heating consumption of the same occupants over time' (SO), 'changes in heating consumption due to new occupants' (NO), and 'Physical characteristics that are not taken into account in the linear regression analyses' (Ph). Additionally, there are the results of the linear regression, which indicates how much of the variance can be explained by the building characteristics that are taken into account in the linear regression (AB).

$$SO = \frac{\sigma^2[LRC^{stayers}]}{\sigma^2[Var_{max}]} \cdot (1 - R^2) \qquad (6)$$

$$NO = \frac{\sigma^2[LRC^{movers}] - \sigma^2[LRC^{stayers}]}{\sigma^2[Var_{max}]} \cdot (1 - R^2) \qquad (7)$$

$$Ph = \frac{\sigma^2[Var_{max}] - \sigma^2[LRC^{movers}]}{\sigma^2[Var_{max}]} \cdot (1 - R^2) \qquad (8)$$

$AB = R^2$ of linear regression

To investigate whether the influence of the occupant changes for different type of building characteristics, exactly the same procedure on a split file per building characteristics category is conducted. When the entire procedure is conducted for every building characteristics and each category, the differences per building category characteristics can be compared. The categories we investigated are as follows:

1. Energy label (Dutch data)
2. Construction year (Dutch and Danish data)
3. Building type (Dutch data)
4. Heating system (Dutch and Danish data)
5. Ventilation system (Dutch data)

5. Results

This section presents the results of the different analyses. It starts by showing the general results for both databases, and also describe the intermediate steps. These results are followed by the results per building characteristic. The first building characteristic that is explored is the energy label, followed by the construction period, dwelling type, type of heating system, and type of ventilation system. Depending on data availability, the analyses are executed either on both databases or on the Dutch database.

5.1. General results (full dataset)

As described in the method section, first the heating consumption for 2015 is standardized (Eq. (1)). The results are presented in Table 3, which indicate that the coefficients of variances of 2010 and 2015 are similar, which means that the spread of the consumption is equal for both years.

After this, a linear regression for 2010 and 2015 is conducted, with actual heat consumption as a dependent variable. The independent variables that are used for the regressions are different for the Dutch and the Danish cases due to data availability. For the Dutch case the energy performance of a house which is often referred to as "theoretical heating consumption" is used. The theoretical heating consumption is

Table 3
Standardising heating consumption.

		Full sample (N = 373,582) The Netherlands	Full sample (N = 512,393) Denmark
2010	Mean [kWh]	13,963.7	19,284.3
	Standard deviation	5,969.3	7,672.0
	Coefficient of variance	0.427	0.398
2015	Mean [kWh]	9,909.1	16,267.4
	Standard deviation	4,379.2	6,365.2
	Coefficient of variance	0.441	0.391
2015 adjusted[*]	Mean [kWh]	13,963.7	19,319.9
	Standard deviation	6,158.4	7,559.6
	Coefficient of variance	0.441	0.391

[*] 2015 values multiplied by the ratio of the means.

calculated based on the building characteristics, using the method described in ISSO-publication 82 [38], with the main aim to determine the energy performance certificate of Dutch dwellings (because the theoretical energy consumption is based on all available building characteristics available in the database). For the Danish case, the parameters indicated in Table 2 are used. With this regression it can be determined how much of the variance in heating consumption can be explained by the available building characteristics (average R^2 of regression models). The regression results indicate that the "theoretical heating consumption" explains (on average) 22.7% of the variance in heating consumption for the Dutch case, and 28.2% for the Danish case. The results of the regression (constant and B coefficient) are also used to correct for the building characteristics (Eqs. (2) and (3)). The regression results can be found in the Appendix A in Tables 7–9. After correcting the heating consumption for building characteristics, the results in Table 4 demonstrate (as expected) that the variance and means for both years and for movers and stayers are close.

To identify how the heating consumption of 2010 and 2015 in the movers and stayers groups relate to each other, the relative heating consumption is calculated. This is the heating consumption of 2015 divided by the results for 2010. A natural logarithmic value is used to make the data useful for further comparison (Eq. (4)). A comparison of the natural logarithmic relative heating consumption for movers and stayers with each other shows that the variance differs between movers and stayers. This is an indication that (as assumed) the correlation of heating consumption of stayers between one year and another is higher than the correlation of houses with different occupants (Table 5).

Now the relative heating consumption for movers and stayers is known, the linear regressions show how much of the variance can be explained by the available building characteristics. Next, the maximum possible variance in heating consumption is defined for the occupant, and building characteristics that were not the same over the years. This will enable determining how much of the variance is explained by the physical characteristics that were not available in the database (which are the characteristics not considered in previously- conducted linear regressions). This is achieved by adding the variance of the heating consumption in 2010 from the movers group together with the variance in heating consumption in 2015. For reasons of comparison, the natural

logarithmic variance in heating consumption is used (Table 6).

Following Sonderegger's method, it is assumed that the maximum variance of heating consumption is the sum of three factors:

1. "**changing heat consumption over time of the same occupants**" (SO): time-dependent variable for the ith house
2. "**changing heat consumption due to new occupants moving into the house**" (NO): of the occupant of the ith house, independent of time
3. "**Physical characteristics that are not taken into account in the linear regression analysis because they were not available in the database**" (Ph) of the ith house, time independent.

The variance of relative heating consumption of movers is the sum of two factors:

1. "**changing heat consumption over time of the same occupants**" (SO): time-dependent variable for the ith house
2. "**changing heat consumption due to new occupants moving into the house**" (NO): of the occupant of the ith house, independent of time

Finally, the variance of relative heating consumption of stayers is:

1. "**changing heat consumption over time of the same occupants**" (SO): time-dependent variable for the ith house

Based on these assumptions it can be calculated which part of the variance is caused by which factor. However, it should be remembered that the available building characteristics have been corrected by using the linear regression results. Eqs. (6)–(8) show how the amount of influence of each parameter is calculated. The results are shown in Fig. 2. For the Dutch case: 28% of the variance can be explained by changes in heating consumption due to new occupants over time (NO); 22.6% by changes in heating consumption of the same occupants over time (SO); 29.9% by physical characteristics not available in the database (Ph); and 19.5% by the building characteristics that were available in the databases (AB). For the Danish case: 33.7% of the variance is explained

Table 4
Normalised heating consumption for movers and stayers in the Netherlands and Denmark.

		Stayers (Netherlands) Sample (N = 254,056)	Movers (Netherlands) Sample (N = 121,326)	Stayers (Denmark) Sample (N = 389,890)	Movers (Denmark) Sample (N = 122,503)
2010	Mean	98.855 ± 0.076	101.513 ± 0.115	99.380 ± 0.053	102.377 ± 0.101
	Sd	38.211 ± 0.054	40.155 ± 0.082	33.324 ± 0.038	35.324 ± 0.071
	Cv	0.387 ± 0.001	0.396 ± 0.001	0.335 ± 0.000	0.345 ± 0.001
2015	Mean	101.034 ± 0.080	96.944 ± 0.118	100.306 ± 0.053	99.212 ± 0.096
	Sd	40.411 ± 0.057	40.982 ± 0.083	33.217 ± 0.038	33.473 ± 0.068
	Cv	0.400 ± 0.001	0.423 ± 0.001	0.331 ± 0.000	0.337 ± 0.001

Sd = standard deviation; Cv = coefficient of variance. Error standard deviation was estimated by Sd/sqrt(2 N), error the mean Sd/sqrt N and error of coefficient of variation is error Sd/mean.

Table 5
Relative heating consumption of stayers and movers in the Netherlands and Denmark.

	Stayers (Netherlands) Sample (N = 254,056)	Movers (Netherlands) Sample (N = 121,326)	Stayers (Denmark) Sample (N = 389,890)	Movers (Denmark) Sample (N = 122,503)
	LRC	LRC	LRC	LRC
Mean	0.011 ± 0.001	−0.066 ± 0.002	0.0102 ± 0.001	−0.030 ± 0.001
Standard deviation	0.384 ± 0.001	0.574 ± 0.001	0.379 ± 0.000	0.450 ± 0.001
Variance	0.147 ± 0.049	0.329 ± 0.018	0.143 ± 0.042	0.203 ± 0.030

Table 6
Logarithm normalised heating consumption of movers in the Netherlands and Denmark.

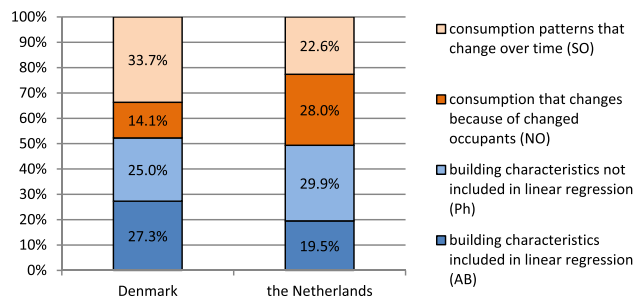|		Movers (Netherlands) LNG Sample (N = 121,326)	Movers (Denmark) LNG Sample (N = 122,503)
2010	Mean	4.529 ± 0.001	4.559 ± 0.001
	Standard deviation	0.473 ± 0.001	0.399 ± 0.001
	Variance	0.224 ± < 0.001	0.159 ± < 0.001
2015	Mean	4.461 ± 0.001	4.530 ± 0.001
	Standard deviation	0.547 ± 0.001	0.393 ± 0.001
	Variance	0.299 ± < 0.001	0.154 ± < 0.001

Fig. 2. Comparison influence building characteristics and occupants on variance energy consumption – Denmark and The Netherlands.

by changing heating consumption patterns of the same occupants over time (SO); 14.1% by changing heating consumption patterns due to new occupants (NO); 25% by physical characteristics that were not available in the database (Ph); and 27.3% by available building char-

acteristics (AB). The use of different prediction variables for the linear regression that determines the influence of available building characteristics explains why there are different percentages for the categories: "*available building characteristics*" and "*other physical characteristics*" for the Dutch and the Danish case. However, for occupant behaviour, large differences were also found between the Dutch and Danish cases. A possible explanation for this could be the origin of the data. The Dutch data is from the social housing sector, while the Danish data contains data from the homeowner-occupied sector. These aspects are addressed more in depth in the discussion section. Nevertheless, both analyses indicate that approximately 50% of the variance is due to occupant behaviour, and the other 50% is due to physical characteristics. These results are different when compared to the results of Sonderegger. This is understandable if our hypothesis that *the amount of influence of the occupant on residential heating consumption is also dependent on the building characteristics of the house they live in* is true. To test this, the same analysis on different groups of the sample in the next sections is conducted. The results are discussed per building characteristic; and depending on data availability, the analyses are conducted on both the Dutch and Danish samples.

5.2. Results per energy label

Executing the same analysis per energy label shows that occupants (changing heating consumption over time (SO) + changing heating consumption due to new occupants (NO)) have on average more influence percentage wise on the variance of energy-efficient houses than on energy-inefficient houses (Fig. 3). This finding is in accordance with the assumptions in previous studies (e.g. [28]). However, this conclusion is only true if we compare dwellings with at least two label steps difference, e.g. the influence of the occupant is on average larger for a B Label dwelling than for an A Label dwelling. Further, it has to be taken into consideration that the variance of buildings with an energy-in-
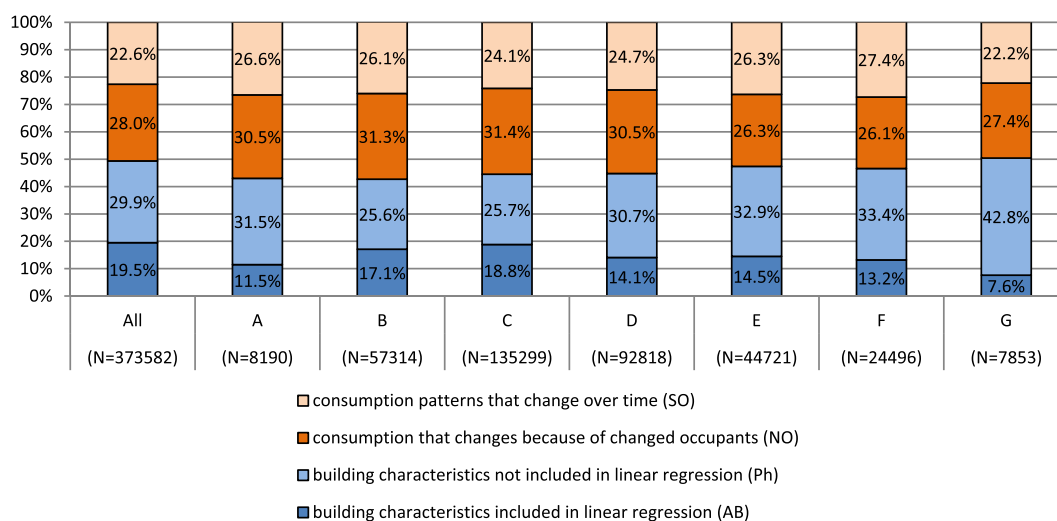

Fig. 3. Comparison of influence of building characteristics and occupants on variance energy consumption – Dutch data energy label.
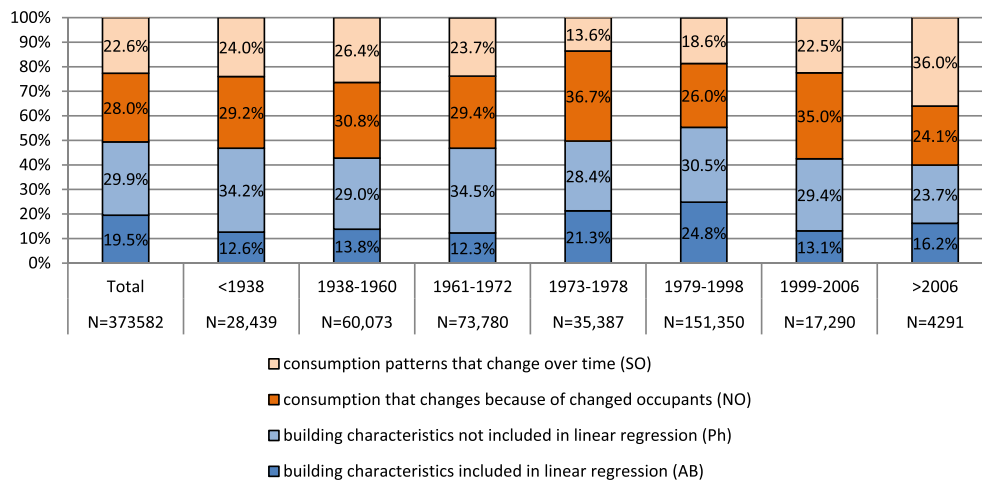
Fig. 4. Comparison of influence of building characteristics and occupants on variance energy consumption – construction year Dutch data.

efficient label is higher than the variance in energy-efficient buildings. This means that if one looks at the physical units, the influence of the occupant is higher for energy inefficient houses, but also the influence of building characteristics is higher for energy-inefficient houses (see Appendix A Fig. 10 for results physical units).

5.3. Results per construction year

An analysis of the construction year confirms our previous results in the analysis of the energy label. Figs. 4 and 5 indicate that in more recently built buildings (which are in most cases more energy-efficient than older buildings) a larger percentage of the variance is caused by occupants, while for older buildings the physical characteristics appear important for explaining the variance. However, especially for the Dutch case, this pattern is less clear than for the energy label results. A possible explanation is that very old buildings are more likely to be renovated than newer buildings. The construction period 1979–1998 forms an exception for both countries and shows a relatively low influence of the occupant. A possible explanation is that those buildings are not renovated yet, while buildings built before 1979 might be more frequently renovated and buildings built after 1999 were initially

already built significantly more energy-efficient.

Fig. 5 shows that the available building characteristics (AB) tend to capture a larger part of the variation in newer buildings, and physical characteristics (Ph) a smaller part. Especially in very new buildings, occupant behaviour seems important for explaining variations across the years.

5.4. Results per building type

Regarding the building type (building types defined in EPISCOPE are used [39]), Fig. 6 indicates that occupants (changing heating consumption over time (SO) + changing heating consumption due to new occupants (NO)) explain a larger percentage of the variance for multi-family houses (common staircase with galleries, common staircase no gallery, maisonette) than for single-family houses (detached houses, semi-detached houses, end houses and terraced houses). Possible explanations for this could be that small changes in consumption patterns are more effective in multi-family houses than in single-family houses, because of the relatively smaller floor area of those dwellings. For example, opening a window in a small room will have more effect on thermal climate than opening a window of similar size in a larger room.
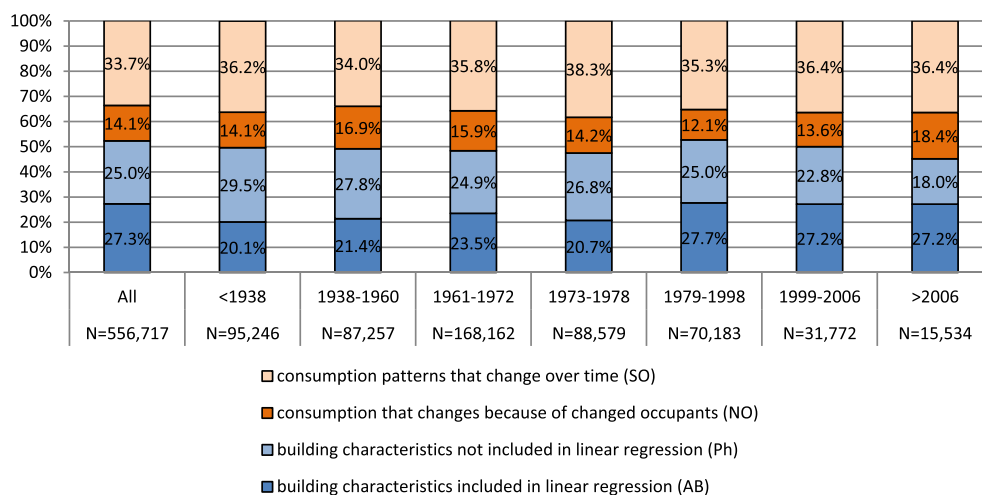

Fig. 5. Comparison of influence of building characteristics and occupants on variance energy consumption – construction year Danish data.
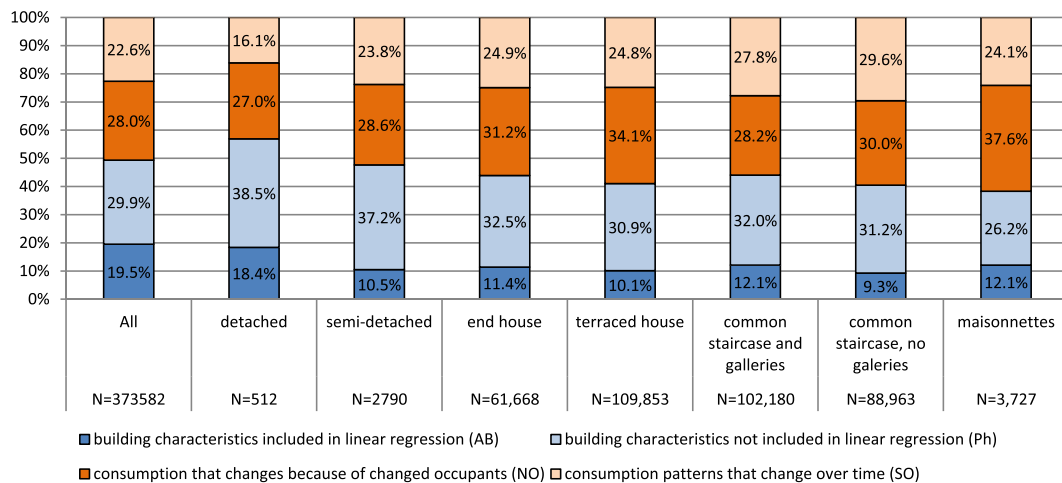
Fig. 6. Comparison of influence of building characteristics and occupants on variance energy consumption – Dutch data dwelling type.
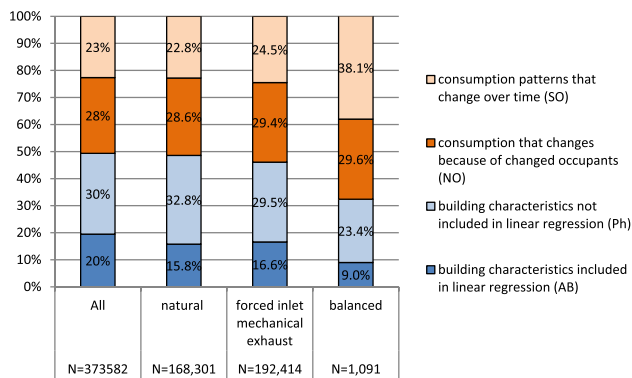

Fig. 7. Comparison of influence of building characteristics and occupants on variance energy consumption – Dutch data ventilation system.

This would also explain why the terraced houses do not show differences with the other multi-family houses, because from the single family houses they have, on average, the smallest floor area.

5.5. Results per type of ventilation system

The comparison of the three different ventilation systems in Fig. 7 indicates that the influence of the occupant is larger for houses with a

balanced ventilation system compared to houses with a natural or forced inlet mechanical exhaust ventilation system. This is expected, because houses with a balanced ventilation system often make use of heat recovery systems. To make optimal use of such a system, all air that enters and leaves the building should go through this system. However, occupants are still able to open windows. Opening the windows means the air does not pass the heat recovery system, which will lead to extra heat losses. Opening windows when a heat recovery system is installed will therefore have a larger effect than in houses where no heat recovery system is installed. Further, balanced ventilation systems are primarily installed in energy-efficient buildings. In Fig. 3 it was already demonstrated that energy-efficient buildings are relatively more sensitive to occupant behaviour compared with energy-inefficient buildings.

5.6. Results per type of heating system

Finally, the heating systems are compared. Because of the differences in the databases, the compared categories are different for the Dutch and Danish cases. For the Dutch case, different gas heating systems are compared. The results of the Dutch case (Fig. 8) indicate (contrary to previous findings) that on average relatively energy-efficient installations are less sensitive to occupant behaviour than energy-inefficient systems. However, the most energy-efficient condensing boiler is an exception and the differences are relatively small, and
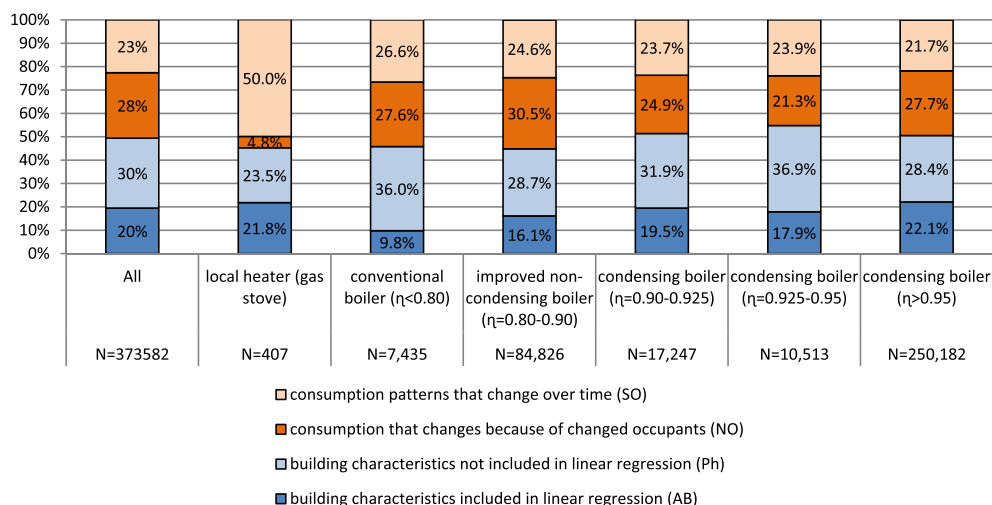

Fig. 8. Comparison of influence of building characteristics and occupants on variance energy consumption – Dutch data heating system.

therefore no conclusion can be drawn from this comparison. Furthermore, the figure shows that the consumption patterns that change over time (SO) are significantly higher for houses with a local heater (gas stove). One could expect that this is due to the relatively small sample of the local heater, however if we study the error of the variances the results seem still reliable (error of ± 1%). This is interesting, because the operation of boiler systems are more or less the same, but the local gas heaters have a different operating system. Therefore, these results could indicate that different operation systems cause differences in behaviour.

For the Danish case, a comparison was made between houses with gas heating and district heating systems. The results indicate, in particular, that the share of consumption that changes, because of changed occupants, is lower for houses with a district heating system compared to houses that are heated by gas (see Fig. 9).
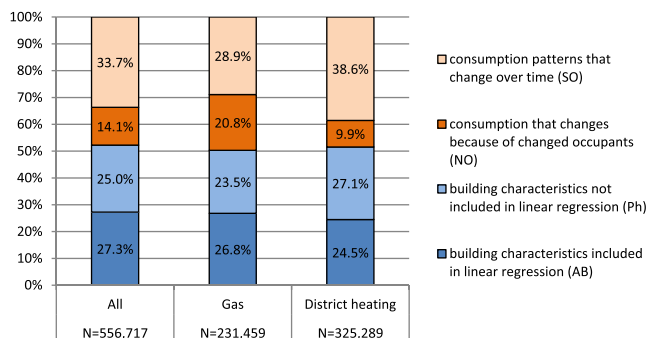

Fig. 9. Comparison of influence of building characteristics and occupants on variance energy consumption – Danish data heating system.

6. Discussion

One of the main advantages of this study compared to previous studies is that this study could make use of two big datasets that included housing data over a six-year period. Using longitudinal data in residential heating consumption research presents significant potential for evaluating the effect of policy changes, newly installed technologies and renovations. Further, analyses on this topic have seldom been conducted based on two large datasets from two different countries (the Netherlands and Denmark).

There are some significant differences between the Dutch and the Danish datasets that should not be neglected. The most important difference is that the Dutch database contains multi- and single-family social rental houses, while the Danish dataset contains private detached houses. Several studies have shown that there is a difference between tenant and homeowner behaviour. Moreover, it could be expected that the building type would influence the results, because in multi-family housing one apartment can be heated from the other. This implies that the energy consumption in an apartment might also change when the neighbours change. This effect is not shown in the analysis separately. If this is the case, then the change due to change of neighbours is included in the change in occupant consumption patterns over time. Despite the differences, both databases indicated that occupants are responsible for half of the residential heating consumption and the building characteristics for the other half. Further, other values calculated from the datasets seemed to be remarkably close together. The difference might be reflected in the distribution of occupant consumption patterns. The results show that the percentage explained by moving occupants is relatively higher for the Dutch dataset (28%) compared to the Danish dataset (14.1%). This suggests that the consumption patterns of the moving Dutch occupants differ more from the consumption patterns of

the previous occupants, compared to the Danish occupants. This could be due to house buyers exhibiting more similarities in consumption patterns with the previous owners, compared to new tenants with previous tenants. This could be the case because occupant characteristics of Dutch social housing tenants are very diverse, while the houses show more similarities and all have a low rental price compared to the owner-occupied housing stock.

One of the uncertainties in this study is the choice of using the data from 2010 and 2015. As Sonderegger [10] mentions in his study, it is expected that the variance in heating consumption among stayers increases over the years. However, it is expected that the variance will proportionally increase in time, because of the limited number of decisions that can be taken, the workings of peer pressure, and other 'stabilising influences'. In his paper, Sonderegger assumes that equilibrium will be achieved after six years, which supports our choice of years. However, he also states that his assumption awaits confirmation by further research. Accordingly, this is an uncertainty that should be taken into consideration.

7. Conclusions

This research investigated the influence of building characteristics and occupants on the variances in residential energy consumption. Therewith this study contributes to a better understanding of the energy performance gap and better interpretation of residential energy modelling and forecasting results. This is one of the first studies towards the influence of building characteristics and occupants on actual residential heating consumption on such a large scale with data from two different countries, which is seldom seen in the field.

This paper showed that variations in residential heating consumption across the years of Dutch social housing can be explained by occupants (49%), the Dutch energy simulation model (theoretical consumption) (20%), and by other physical characteristics that are not taken into account in the building simulation model (32%). For the Danish case, the results showed that 48% of the variation in residential heating consumption can be explained by occupants, 27% by the building characteristics mentioned in Table 2 and 25% by other physical characteristics. These results suggest that approximately half of the variation in residential heat can be ascribed to differences between buildings and approximately half of the variation to differences in occupant behaviour. These results were found by using an existing method (suggested by Sonderegger in 1978) with new and strongly improved data. This enabled comparisons of national contexts (The Netherlands and Denmark), of different types of heat supply (district heating and natural gas), different housing formats (social housing and private single-family houses), and different building types (detached and multi-family).

The results show that approximately half of the variance could be attributed to buildings and half to occupants. However, the follow-up analysis per building characteristic showed that the influences of the occupant are dependent on the building characteristics of the building. For example, the influence of occupants is larger for energy-efficient houses than for energy-inefficient houses. This is demonstrated in both comparisons of houses with different energy labels, and the analysis of houses built in a different period for the Dutch and the Danish cases. The results also show that the influence of occupants is dependent on the type of building installations in the house. For example, the occupant consumption patterns seem more important when the house has a local gas stove as a heating system than when the house has a gas boiler. Further, the influence of occupants is different, depending on the type of house.

The results of this research suggest that, on average, occupants significantly influence the variance in energy among buildings.

Moreover, the magnitude of this influence is dependent on building characteristics, because some buildings are more sensitive to occupant consumption patterns than others. This is an important insight, because this indicates that building simulations will not be able to predict actual heating consumption correctly and accurately if occupant consumption patterns are considered. Although the results indicated that the influence of occupants is almost as important as the influence of building characteristics on residential heating consumption, thermal renovations will remain an important measure for reducing residential heating consumption. This is because deep thermal renovations (if correctly executed) usually result in an energy reduction for heating. Regarding occupant behaviour, more research is needed to determine the extent that occupant consumption patterns can be influenced to reduce residential energy consumption.

The results also indicate that there is still a relatively large number of physical characteristics that cause variance in heating consumption. More research is needed to determine the nature of these physical characteristics. If more is known about these parameters, they could be used to improve building simulation models. The high influence of

occupants also suggests that it is not useful to aim for a perfect simulation model for one specific building, especially when the occupant behaviour is unknown. However, one can aim for a simulation model that shows the average heating consumption of a larger group of buildings.

This paper is one of the first studies to make use of large longitudinal databases in the field of residential heating consumption. It has already demonstrated the importance of this type of data for the field. Longitudinal databases that contain residential heating consumption data present significant potential for evaluating the effect of policy changes, newly installed technologies, and renovations.

Acknowledgements

Appendix A

See Tables 7–9 and Fig. 10.

Table 7
Linear regression results of the Dutch sample year 2010 and 2015 (AB).

Model	Unstandardized coefficients 2010		t	Sig.
	B	Std. Error		
Constant	27,224.01	79.84	341.07	< 0.01
Theoretical gas consumption [MJ]	19.03	0.0070	314.022	< 0.01

*R^2 0.210, dependent variable gas consumption 2010.

Model	Unstandardized coefficients 2015		t	Sig.
	B	Std. Error		
Constant	28,246.89	83.88	803.153	< 0.01
Theoretical gas consumption [MJ]	19.24	0.07	0.517	< 0.01

*R^2 0.180, dependent variable adjusted gas consumption 2015.
(AB) Average explanation of available building characteristics is (0.217 + 0.186)/2 = 0.2015 = 20.15%.

Table 8
Linear regression results of the Danish sample year 2010 and 2015.

Model	Unstandardized coefficients 2010		t	Sig.
	B	Std. Error		
Constant	− 37,990.73	240.366	− 158.05	< 0.01

*R^2 0.2860, dependent variable gas consumption 2010

Model	Unstandardized coefficients 2015		t	Sig.
	B	Std. Error		
Constant	− 38,546.99	241.192	− 159.82	< 0.01

*R^2 0.2595, dependent variable adjusted gas consumption 2015
(AB) Average explanation of available building characteristics is (0.286 + 0.2595)/2 = 0.27275 = 27.28%.

Table 9
Coefficients per parameter of the linear regression of the Danish sample year 2010 and 2015.

Model	Heat2010		Adjheat2015	
	Coef.	SE	Coef.	SE
Area (logarithmic transformed)	11,420.611[***]	55.956	11,757.833[***]	56.149
Gas (1 = Yes)	3.727,314[***]	18.492	1,024.326[***]	18.556
Number of rooms	173.442[***]	10.298	212.825[***]	10.333
Wood-stove (1 = Yes)	−1,316.447[***]	23.932	−1,329.944[***]	24.015
Attic floor (1 = Yes)	−715.892[***]	27.752	−944.257[***]	27.847
Basement (1 = Yes)	3,025.429[***]	24.130	3,344.441[***]	24.213
Building period (ref. "Before 1938")				
1938–1960	−508.627[***]	34.381	−587.511[***]	34.499
1961–1972	−1,737.207[***]	36.120	−1,859.880[***]	36.244
1973–1978	−2,954.989[***]	40.084	−3,170.666[***]	40.222
1979–1998	−5,149.254[***]	43.136	−5,040.772[***]	43.284
1999–2006	−5,937.958[***]	53.440	−6,018.504[***]	53.624
After 2006	−7,816.063[***]	66.363	−8,027.991[***]	66.591
Roof material (Ref. "Fibercement")				
Cement stone	346.156[***]	29.141	19.578	29.241
Tile	1,204.509[***]	25.175	1,068.295[***]	25.262
Other material	1,513.594[***]	35.536	1,609.229[***]	35.658
Exterior wall material (Ref. "Bricks")				
Wood	−1,963.962[***]	71.393	−1,480.936[***]	71.638
Concrete	364.676[***]	46.569	329.333[***]	46.729
Other material	−565.596[***]	91.204	−503.470[***]	91.517
Constant	−37,990.729[***]	240.366	−38,546.992[***]	241.192
R2	0.286		0.260	
Number of observations	512,393		512,393	

Note:
[*] $p < 0.1$.
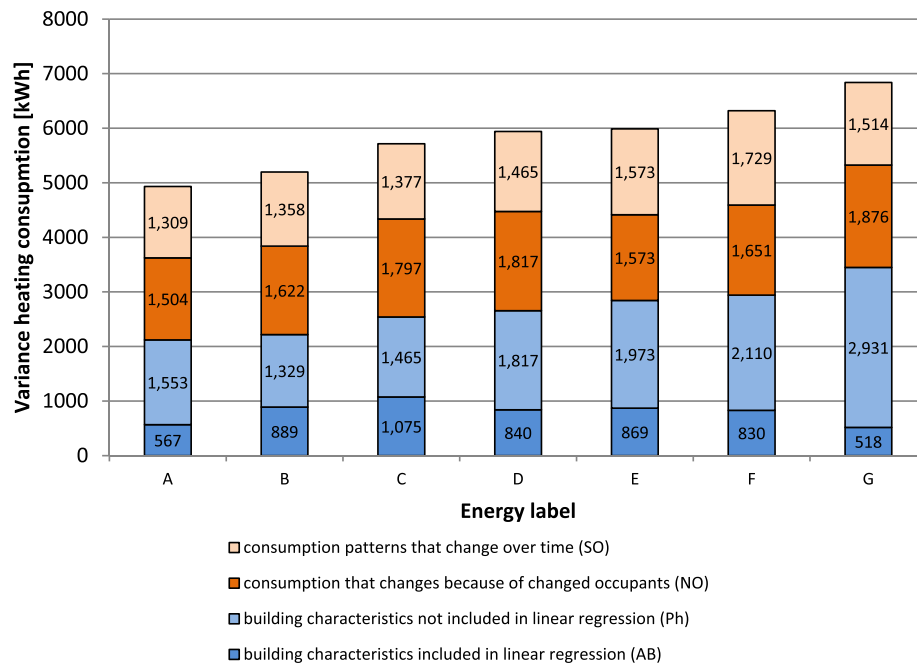[**] $p < 0.05$.
[***] $p < 0.01$.

Fig. 10. Comparison of influence of building characteristics and occupants on variance energy consumption using physical units instead of percentages – Dutch data energy label.

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apenergy.2019.05.078.

References

[1] EC. EU energy in figures, in statistical pocketbook 2018, E. Commission, Editor. 2018, EU Publications Office: Luxembourg. p. 268.

[2] Balaras CA, et al. Empirical assessment of calculated and actual heating energy use in Hellenic residential buildings. Appl Energy 2016;164:115–32.

[3] Menezes AC, et al. Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. Appl Energy 2012;97:355–64.

[4] O'Neill BC, Chen BS. Demographic determinants of household energy use in the United States. Population Dev. Rev. 2002;28:53–88.

[5] Steemers K, Yun GY. Household energy consumption: a study of the role of occupants. Build. Res. Inf. 2009;37(5–6):625–37.

[6] Yohanis YG, et al. Real-life energy use in the UK: How occupancy and dwelling characteristics affect domestic electricity use. Energy Build 2008;40(6):1053–9.

[7] Delzendeh E, et al. The impact of occupants' behaviours on building energy analysis: A research review. Renew Sustain Energy Rev 2017;80:1061–71.

[8] Gram-Hanssen K. Efficient technologies or user behaviour, which is the more important when reducing households' energy consumption? Energ Effi 2013;6(3):447–57.

[9] Aydin E, Kok N, Brounen D. The rebound effect in residential heating; 2013. p. 33.

[10] Sonderegger RC. Movers and stayers: The resident's contribution to variation across houses in energy consumption for space heating. Energy Build 1978;1(3):313–24.

[11] Guerra Santin O, Itard L, Visscher H. The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. Energy Build 2009;41(11):1223–32.

[12] Huebner GM, et al. Explaining domestic energy consumption – The comparative contribution of building factors, socio-demographics, behaviours and attitudes. Appl Energy 2015;159(Supplement C):589–600.

[13] Brounen D, Kok N, Quigley JM. Residential energy use and conservation: Economics and demographics. Eur Econ Rev 2012;2012(56):931–45.

[14] Verhallen TMM, Raaij WFv. Household behavior and the use of natural gas for home heating. J Consumer Res 1981.

[15] Hansen AR, Gram-Hanssen K, Knudsen HN. How building design and technologies influence heat-related habits. Build Res Inf 2018;46(1):83–98.

[16] Gill ZM, et al. Low-energy dwellings: the contribution of behaviours to actual performance. Build Res Inf 2010;38(5):491–508.

[17] Spataru C, Gillott M, Hall MR. Domestic energy and occupancy: a novel post-occupancy evaluation study. Int J Low-Carbon Technol 2010;5(3):148–57.

[18] Guerra-Santin O, Itard L. Occupants' behaviour: determinants and effects on residential heating consumption. Build Res Inf 2010;38(3):318–38.

[19] Hansen AR. The social structure of heat consumption in Denmark: New interpretations from quantitative analysis. Energy Res Social Sci 2016;11(Supplement C):109–18.

[20] Estiri H. Building and household X-factors and energy consumption at the residential sector: A structural equation analysis of the effects of household and building characteristics on the annual energy consumption of US residential buildings. Energy Econ 2014;43:178–84.

[21] Chen J, Wang X, Steemers K. A statistical analysis of a residential energy consumption survey study in Hangzhou, China. Energy Build 2013;66:193–202.

[22] Druckman A, Jackson T. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model. Energy Policy 2008;36(8):3167–82.

[23] Jeeninga H, Uyterlinde M, Uitzinger J. Energieverbruik van energiezuinige woningen : effecten van gedrag en besparingsmaatregelen op de spreiding in en de hoogte van het reële energieverbruik. ECN-C-01-072. 2001, [Petten]: Energieonderzoek Centrum Nederland.

[24] Kaza N. Understanding the spectrum of residential energy consumption: A quantile regression approach. Energy Policy 2010;38(11):6574–85.

[25] Vringer K, Blok K. The direct and indirect energy requirements of households in the Netherlands. Energy Policy 1995;23(10):893–910.

[26] Alberini A, Filippini M. Response of residential electricity demand to price: The effect of measurement error. Energy Econ 2011;33(5):889–95.

[27] Alberini A, Gans W, Velez-Lopez D. Residential consumption of gas and electricity in the U.S.: The role of prices and income. Energy Econ 2011;33(5):870–81.

[28] Haas R, Auer H, Biermayr P. The impact of consumer behavior on residential energy demand for space heating. Energy Build 1998;27(2):195–205.

[29] Yu Z, et al. A systematic procedure to study the influence of occupant behavior on building energy consumption. Energy Build 2011;43(6):1409–17.

[30] Belaïd F. Understanding the spectrum of domestic energy consumption: Empirical evidence from France. Energy Policy 2016;92(Supplement C):220–33.

[31] ECN. Energie-nederland and Netbeheer-nederland, energietrends 2014, in energietrends; 2014.

[32] Odyssee-Mure. Netherlands | Energy Profile, in Energy efficiency trends and policies. odyssee-muree; 2018.

[33] Fuentes E, Arce L, Salom J. A review of domestic hot water consumption profiles for application in systems and buildings energy performance analysis. Renew Sustain Energy Rev 2018;81:1530–47.

[34] Hansen AR. 'Sticky' energy practices: The impact of childhood and early adulthood experience on later energy consumption practices. Energy Res Social Sci 2018;46:125–39.

[35] Strengers Y, Maller C. Adapting to 'extreme' weather: mobile practice memories of keeping warm and cool as a climate change adaptation strategy. Environ Planning A: Economy Space 2017;49(6):1432–50.

[36] Gram-Hanssen K. Understanding change and continuity in residential energy consumption. J Consumer Culture 2011;11(1):61–78.

[37] Rinkinen J, Jalas M. Moving home: houses, new occupants and the formation of heating practices. Build Res Inf 2017;45(3):293–302.

[38] ISSO. Energieprestatie advies woningen, in 82.3, ISSO, Editor. Rotterdam; 2011.

[39] Filippidou F, Nieboer NET. Residential building typology. EPISCOPE 2013 2013 [cited 2019 20-3-2019]; classification scheme for residential buildings — construction year classes — pictures of example buildings]. Available from: http://episcope.eu/building-typology/country/nl/.